Journal of Experimental Psychology:
Human Perception and Performance
2014, Vol. 40, No. 2, 502–524

0096-1523/14/$12.00 DOI: 10.1037/a0035245

The Aperture Capture Illusion: Misperceived Forms in Dynamic Occlusion Displays

Evan M. Palmer
Wichita State University

Philip J. Kellman
University of California, Los Angeles

Visual illusions can reveal unconscious representations and processes at work in perception. Here we report a robust illusion that involves the misperception of moving, partially occluded objects. When a dynamically occluded object is seen through 2 misaligned apertures, the object appears misaligned in the direction of the apertures, creating the *Aperture Capture Illusion*. Specifically, when part of a dynamically occluded object disappears behind an occluding surface and then another part of the object comes into view immediately afterward, the 2 parts appear misaligned in the direction of the offset of the apertures through which they were seen. This illusion can be nulled: Separating the 2 object parts to increase the time interval between their appearance produced the percept of alignment. The ability to null the illusion in this manner demonstrates that dynamically occluded regions of moving objects continue to persist in perceptual awareness but, we argue, are perceived to move at a slower velocity than visible regions. We report 7 experiments establishing the existence of the illusion and ruling out several classes of explanation for it. We interpret the illusion and the ability to nullify it within the context of Palmer, Kellman, and Shipley's (2006) theory of spatiotemporal object formation.

Keywords: dynamic occlusion, motion, illusion, contour interpolation, shape integration

One of the remarkable accomplishments of the human visual system is that it generates representations of complete objects despite partial input. As a result of occlusion, the visible regions of objects in the environment are often fragmented in the retinal projection. Perceiving whole objects, then, requires perceptual processes that connect visible regions across gaps in the input to achieve accurate representations of unity and shape. When objects or observers move, the visible regions of objects can change dramatically over time, greatly complicating the requirements of object formation. The system must deal with fragmentation, not only across space but over time as well. Dynamic occlusion situations of this sort are pervasive and challenging stimulus situations, yet our visual system typically handles them with ease (Palmer, Kellman, & Shipley, 2006).

Consider the perceptual situation of looking through a thick hedge and seeing a woman walk down the street (Figure 1; *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/Walking_People.mov*). Because of the branches and leaves, only bits and pieces of the woman are visible at any given moment. Yet when viewing the movie from which these frames were taken, one perceives not a jumble of disconnected bits of hair, facial features, clothing and books, but a unitary, solid object—a woman—walking down the street. How does the visual system achieve this extraordinary feat of spatiotemporal object formation?

Evan M. Palmer, Department of Psychology, Wichita State University; Philip J. Kellman, Department of Psychology, University of California, Los Angeles.

Correspondence concerning this article should be addressed to Evan M. Palmer, Department of Psychology, Wichita State University, 1845 North Fairmount, Wichita, KS 67260-0034. E-mail: evan.palmer@wichita.edu

Using a laboratory analogue of the situation in Figure 1, Palmer et al. (2006) studied spatiotemporal object formation under conditions of multiple apertures and found highly accurate perception from fragmented displays under certain conditions. To account for their data, they proposed the theory of *spatiotemporal relatability*, which describes the geometric relations between visual fragments in time and space that support object formation and they also proposed mechanisms for accomplishing the required integration of information. Specifically, they introduced the notion of a *dynamic visual icon* (DVI) that represents continuously moving objects after they have gone out of sight. The choice of the term *dynamic visual icon* reflects both the visual icon's capacity to maintain representations of previously seen visual inputs and also to extrapolate stored fragments' positions over time during occlusion. According to spatiotemporal relatability, currently visible and recently occluded object fragments are unified into perceptual wholes via edge completion processes that obey the geometry of *relatability* (Kellman & Shipley, 1991) within an object-centered, distal reference frame (Shipley & Cunningham, 2001). Within the DVI representation, both visible and occluded object fragments contribute to shape perception, even though they might not be simultaneously visible. Thus, the DVI enables perception of dynamically occluded objects by collecting and integrating shape information over time, as bits and pieces of the object become visible through apertures in occluding surfaces.

The reasons for theorizing a special representational type—the DVI – may be found in the earlier work (Palmer et al., 2006). Here we review some of the rationale and consider a range of findings, including some since the DVI notion was initially proposed, that clarify and provide evidence for this theoretical construct. The theory of spatiotemporal relatability proposes that the DVI can be understood as having two aspects: *persistence* and *position updat-*

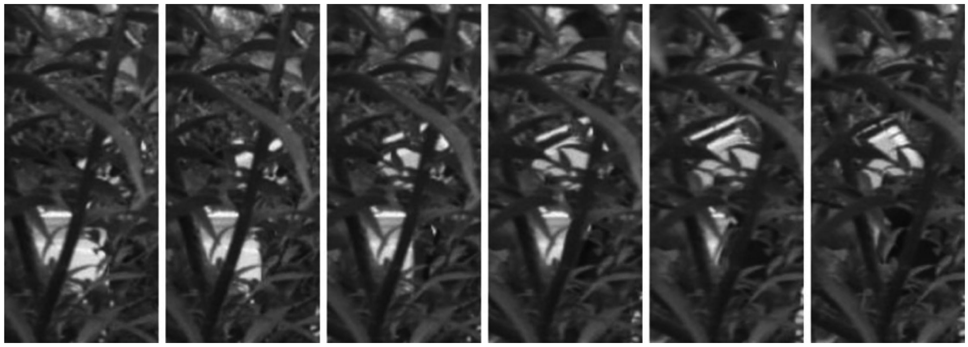

Figure 1. Six frames from a movie of a woman walking from right to left (see *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/Walking_People.mov*), as seen through a thick hedge. As a result of occlusion and the walker's motion, the visible regions of the woman change dramatically over time.

ing. The DVI builds upon previous work indicating mechanisms of storage that persist after visual input has ceased. Perceptual processing is extended in time and can operate on both current and recently received inputs. Sperling (1960) demonstrated the existence of a visual buffer that retains shape information after a display is extinguished and allows observers to make judgments about visual stimuli after they are no longer visible. Neisser (1967) labeled this short-term visual buffer the *visual icon*, and connected it with other such temporally extended perceptual buffers like *echoic memory* in the auditory modality. Since that time, numerous researchers have studied the properties of the visual icon and catalogued its many aspects (e.g., Banks & Barber, 1977; Clark, 1969; Demkiw & Michaels, 1976; Turvey & Kravetz, 1970; Von Wright, 1968).

A visual icon, as described by Sperling and others, characterized as a high-capacity, short-term perceptual buffer would provide the persistence of information required in spatiotemporal interpolation. However, conceptions of a visual icon have typically been retinotopic in preserving information at locations where it appeared. The position updating that is implicated in spatiotemporal interpolation, and some other visual phenomena, requires both persistence and a changing representation of object position. As persisting visual representations may be individuated based on their properties (Coltheart, 1980), we label the required storage a DVI, to distinguish it from representations involving static persistence. Whether this is a unique representation, or whether the DVI simply incorporates a previously unsuspected property of representations described earlier (such as the notions of visible persistence or informational persistence described by Coltheart, 1980), is an issue to which we return after considering other research and the experiments below.

The role of the DVI in spatiotemporal object formation may be simply stated. When an object fragment becomes visible through gaps in an occluding surface, the visual system may have inadequate information to determine much about the object. Other fragments of the same object may be visible earlier or later, and some means is needed to connect these. Finally, for moving objects, or for stationary objects viewed by a moving observer, the spatial position at which a fragment appears is typically a poor guide to its spatial relations to other sequentially appearing object parts. How can the visual system acquire fragmentary information across space and time and use it to achieve coherent representations of objects and surfaces?

In stationary displays, strong geometric constraints, formalized as spatial relatability, govern which contour fragments connect across gaps to form complete illusory or occluded objects (Kellman, Garrigan, & Shipley, 2005; Kellman & Shipley, 1991). Palmer et al. (2006) reasoned that the same spatial geometry might apply to spatiotemporal object formation as well. Connecting the spatial case to the spatiotemporal case might produce a unified spatiotemporal view of object formation, of which the more frequently studied static, two-dimensional situation is seen as a limiting case (Kellman & Shipley, 1991; Palmer et al., 2006). This unification can be accomplished in a straightforward way using two assumptions. First, visible parts persist as perceptual representations for a short time after going out of sight. Second, the visual system updates the position of these stored fragments based on velocity signals acquired when they were visible.

Of the two functions of the DVI necessary to explain spatiotemporal object formation, persistence has been often studied, whereas the position updating aspect is relatively unexplored. If the perception of dynamically occluded objects is possible because of a DVI, then the positions of occluded fragments within the buffer must be updated over time so that when new parts of the object become visible, they can be united with occluded fragments at the position *they would occupy if they were visible.* Perhaps position extrapolation mechanisms such as those proposed by Nijhawan (2002) to explain the flash-lag illusion play a role in anticipating an object's future positions behind an occluding surface. For expanded discussion of classic evidence for position updating after occlusion, the basis for inferring such a mechanism, and its relation to visual storage and position updating mechanisms that have been previously proposed, see Palmer, Kellman and Shipley (2006, pp. 514–515, 517–519, and 536–539).

In summary, the notion of a DVI proposes that, for a short time after occlusion, there is perceptual equivalence between visible and occluded regions of a dynamically occluded object. Support for this notion requires evidence that dynamically occluded objects continue to be perceptually processed even though they are not physically visible (cf., Michotte, Thinès, & Crabbé, 1964). Below, we review some recent evidence in support of this claim.

Evidence for a Dynamic Visual Icon

Palmer et al. (2006) provided support for the notion of a dynamic visual icon in a variety of ways. Most specifically, they showed that displays fulfilling the criteria of spatiotemporal relatability led to perception of complete objects in both occluded and illusory displays, as evidenced by objective performance advantages in a task facilitated by object formation. One experiment in Palmer et al. (2006) provided specific information about the power of the DVI representation. In a static exposure condition, observers saw three fragments of a static object through three large apertures in an occluding surface. The three fragments were visible for either 80 ms or 440 ms and the participants' task was to identify which shape configuration they saw among two alternative configurations, differing only in the horizontal alignment of one of the three pieces of the object. Two regions of the object were always occluded, making this a spatial interpolation task. Participants performed better in the 440-ms than the 80-ms static exposure condition, as would be expected. In the dynamic exposure condition, the object moved behind an occluding surface with three rows of four small vertical apertures each, and observers performed the same identification judgment. The apertures were horizontally and vertically misaligned, making this a spatiotemporal interpolation task—no single frame of the animation provided enough shape information to allow perceptual completion of the object. Importantly, even though it took 440 ms for the object to traverse the width of the occluding surface, the total physical exposure time of the dynamic object fragments (the total time that any give pixel of the object image was visible through the apertures) was just 80 ms, the same as the shorter static exposure condition. However, participants' performance on the shape identification task in the dynamic exposure condition was significantly better than the 80-ms and more similar to the 440-ms static exposure condition. This suggests that perceptual processing of dynamically occluded objects is better than would be expected based on the raw physical exposure times of the fragments, providing evidence for the existence of the DVI.

Keane, Lu, and Kellman (2007) supported the notion of a DVI for perceiving dynamically occluded objects used a spatiotemporal classification image paradigm (Gold & Shubel, 2006). In this study, noise-corrupted, dynamically occluded objects translated laterally in front of several inducing elements and participants attempted to classify the images as either "fat" or "thin," an objective experimental task. The authors demonstrated that dynamically occluded objects yield classification images consistent with contour interpolation between regions visible at different times and places, regardless of whether the inducing elements were real or illusory, as long as the contours obeyed the parameters of spatiotemporal relatability. In a fragmented condition that violated spatiotemporal relatability, little or no edge interpolation was observed. Keane et al. (2007) concluded that their spatiotemporal classification images support the notion of active contour interpolation between spatiotemporally misaligned regions, consistent with the notion of the DVI.

Further evidence for a DVI that actively represents the position of moving objects behind an occluding surface comes from Flombaum, Scholl, and Pylyshyn (2008). They performed a series of multiple object tracking experiments in which moving disks sometimes passed behind occluding surfaces during the tracking phase. Scholl and Pylyshyn (1999) previously established that target disks can be accurately tracked through occlusion in a multiple object tracking task, provided that appropriate optical cues to progressive occlusion are provided to the observer (i.e., deletion/accretion of the disk rather than implosion/explosion). In the Flombaum, Scholl, & Pylyshyn (2008) experiment, observers tracked disks that could occasionally become occluded and also performed a secondary task of reporting whenever they saw a dim gray probe dot presented somewhere in the tracking field. The probe could appear on or near an unoccluded disk, on or near an occluded disk (on top of the occluder in the location where the disk would be at that moment), or on an occluding surface itself. Observers were overall more accurate at identifying the appearance of the probe when it appeared on a tracked object than a distractor object, even when the target was occluded at the time. Surprisingly, Flombaum, Scholl, and Pylyshyn (2008) found that participants were more accurate at identifying probes presented on occluded targets than unoccluded targets, overall. They termed this the "Attentional High-Beams" effect and argued that observers allocate more attentional resources for processing the positions of occluded than visible objects. Thus, the Flombaum et al. (2008) data suggest that observers actively represent the position of temporarily occluded moving objects, consistent with the proposal of a DVI.

Neurophysiological evidence for the continued representation of occluded objects during multiple object tracking comes from evoked response potential studies by Drew and Vogel (2007). Previous work with evoked response potentials established that the amplitude of a sustained contralateral activity emanating from the posterior parietal cortex is directly related to the number of objects being tracked in a multiple object tracking task (Drew & Vogel, 2006), allowing researchers to determine whether one, two, or three disks were being actively tracked at any given moment. Because observers can track disks through occlusion (Scholl & Pylyshyn, 1999), Drew and Vogel (2007) asked whether observers continue to actively represent the disks during occlusion or whether they stop tracking disks at the moment of occlusion and reacquire them as soon as they reappear. Sustained contralateral activity during the moments of occlusion established that observers continued to actively track occluded disks even though they were not physically visible.

Several functional MRI experiments have shown stronger cortical activity for dynamically occluded objects that obey naturalistic optical transformation rules than those that do not. Shuwairi, Curtis, and Johnson (2007) identified a network of cortical regions in the extrastriate visual cortex that were more active in response to objects that underwent accretion and deletion transformations (Gibson, Kaplan, Reynolds, & Wheeler, 1969) before occlusion than objects that shrank at the occlusion boundary. Yi et al. (2008) found stronger cortical responses in the fusiform face area for dynamically occluded pictures of faces that followed continuous as opposed to discontinuous motion trajectories. It seems that the human perceptual system is designed to track dynamic objects through occlusion, provided that the perceptual cues to occlusion match those that objects exhibit in the real world.

Studying the Dynamic Visual Icon

Palmer et al. (2006) proposed the notion of the dynamic visual icon, but their focus was on object formation in circumstances resembling those of ordinary perception, with information appearing through multiple apertures. As Figure 2 shows, their displays had laterally translating objects specified by appearance of fragments through small rectangular apertures (dynamic occluded objects; *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/Orig_DO_Paradigm_STD.mov*) or by interruption of small rectangular inducing elements (dynamic illusory displays; *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/Orig_DO_Paradigm_STD_ILL.mov*).

Here, our primary focus is the DVI itself. Assuming that a DVI allows the perceptual system to continue to represent moving objects after occlusion, can we say more about it other than it supports accurate perception? For instance, how well does it work under a range of circumstances? Is it always accurate? To assess these questions, we reduced the stimulus situation in Figure 2 to the simplest possible case: two apertures, spatially separated vertically and horizontally. Lateral misalignment ensures, for laterally translating object parts, that perception requires processing across both space and time. Vertical separation ensures that visible parts must be perceptually integrated across gaps where no shape information is specified. As will become clear, these conditions produce a robust visual illusion that, despite differences in the accuracy of perception from previously studied conditions, provides strong evidence for the hypothesized mechanism of a DVI. The results converge on the idea that under simple conditions, the position updating function is systematically nonveridical.

Experiment 1A and 1B: Form Perception in Minimal Dynamic Occlusion Displays

The displays used in these experiments represent the most basic case for studying spatial relations in dynamically occluded object perception—the case in which an object moves behind an occluding surface and two nonoverlapping portions of it are seen through two apertures at different locations in the visual field. In this situation, the shape information from the object is discontinuous in both space and time and contour interpolation processes must be engaged to unite the two separated fragments into a perceptual whole. In the displays shown in Figure 3, two apertures are horizontally arranged so that the rightmost boundary of one aligns vertically with the leftmost boundary of the other. The apertures are also separated in the vertical dimension, so that object formation from fragments seen in the two apertures in such a display would require interpolation across gaps. The two apertures reveal the top and bottom regions of the object, but the middle 1/3 of the object is never seen. Such an arrangement allows the examination of the position-updating process within the dynamic visual icon. Participants never saw the unoccluded rod and thus had to create a representation of the alignment of the two pieces of the rod by spatiotemporally integrating the fragments visible through the two apertures. Referring to Figure 3, the top right edge of the rod becomes occluded first and must be tracked behind the display before it can be integrated with the bottom right edge of the rod, once it appears. The observer must maintain a representation of the occluded region of the object and update its position over time behind the occluding surface to correctly perceive the alignment of the rod.

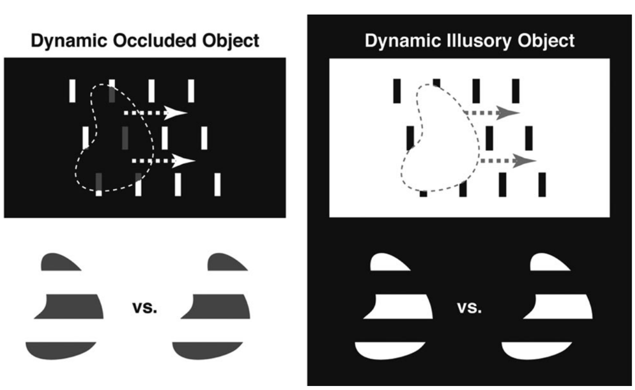

Figure 2. Schematic depiction of the dynamic occluded and dynamic illusory object displays used by Palmer, Kellman, and Shipley (2006). Participants saw fragments of dynamic objects either through apertures (left; *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/Orig_DO_Paradigm_STD.mov*) or on top of similarly shaped elements (right; *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/Orig_DO_Paradigm_STD_ILL.mov*). After the motion display, participants chose which of two shape configurations they saw (bottom).

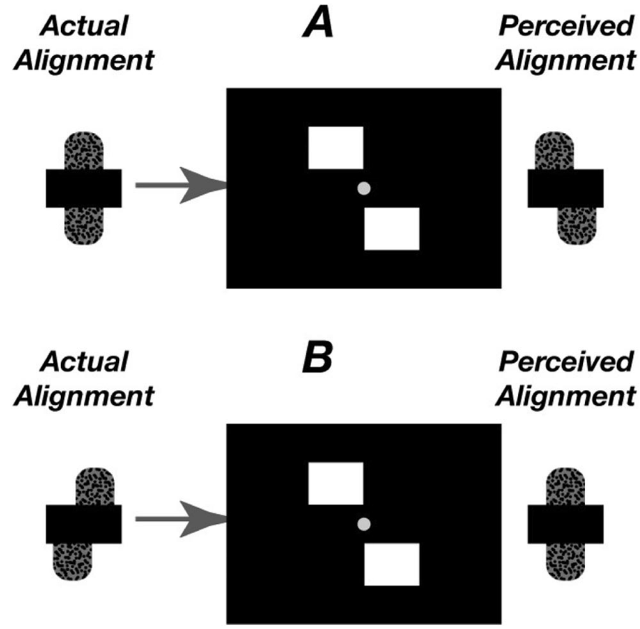

Figure 3. Stimuli used in Experiment 1. With the exception of the width of the black occluding surface with two apertures (which is about 37% of the width used in the experiments), all aspects of the stimuli, including illusion magnitudes, are drawn to scale. The illusion magnitude shown is for the 510-arcmin/sec speed in Experiment 1A. A) An aligned rod seen moving behind an occluder with misaligned apertures appears misaligned in the direction of the apertures (*http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/ACIllusion-Aligned.mov*). B) A rod that is misaligned in the direction opposite of the apertures appears to be aligned when it passes behind the occluder (*http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/ACIllusion-Misaligned.mov*). Note that the rod was red with black speckles and that participants never saw the unoccluded rod in the experiment.

Pilot work with these displays revealed a robust illusion: when an aligned rod moves behind an occluding surface with two misaligned apertures, the rod appears broken and misaligned in the same direction as the offset of the apertures (see Figure 3A and *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/ACIllusion-Aligned.mov*). Moreover, it appeared that the perceived misalignment of the rod could be nulled by repositioning the two pieces (Figure 3B and *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/ACIllusion-Misaligned.mov*). In Experiment 1A and 1B we investigated this illusion systematically, specifically focusing on quantifying the magnitude of the illusion and examining its dependence on stimulus velocity.

Method

Experiments 1 through 7 used the same apparatus, procedure, and data analyses, along with highly similar stimuli and design. We describe these aspects in detail for Experiment 1 and indicate minor variations from this general framework in Experiments 2 through 7.

Participants. Twelve UCLA undergraduates participated in Experiment 1A and 10 UCLA undergraduates participated in Experiment 1B in partial fulfillment of course requirements for an introductory psychology class. All participants gave informed consent, reported normal or corrected-to-normal vision, and were naïve to the purposes of the experiment. Two participants in Experiment 1A were excluded because they failed to meet an objective standard for task compliance (described in the Dependent Measures and Data Analysis section below).

Apparatus. Stimuli were presented and responses recorded using a 400-MHz Apple Macintosh G3 with a 17" (diagonal; 43.1 cm) monitor at a resolution of 1024 × 768 pixels (76.9 pix/in; 30.2 pix/cm) and a refresh rate of 85 Hz. A chinrest was used to stabilize participants' heads at a distance of 115 cm from the monitor, such that each pixel subtended one minute of visual angle. Responses were gathered with an Apple Macintosh extended keyboard. The experiment was programmed in the MacProbe programming language (Hunt, 1994), which controlled the timing and presentation of the stimuli.

Stimuli. In each display, a red rod moved horizontally behind a black occluding surface containing two apertures (as in Figure 3). The red rod was 90 arcmin tall by 30 arcmin wide and had black speckled texture (to help disambiguate its motion direction within each aperture). The black occluding surface was 600 arcmin wide by 160 arcmin tall, and the apertures were each 40 arcmin wide by 35 arcmin tall. There was a 30 arcmin vertical gap between the two apertures, meaning that the middle 1/3 of the red rod was never seen. The innermost edges of the apertures were aligned so that the inner left edge of one was aligned with the inner right edge of the other. The red rod translated laterally behind the occluding surface at one of three velocities (chosen from a counterbalanced set of conditions—see below). The top and bottom portions of the rod were fully visible only in succession, because of the arrangement of the apertures within the occluding surface. A fixation dot was presented in the middle of the occluding surface and participants were instructed to keep their gaze focused on it, but eye movements were not monitored.

The animations were generated by displaying one frame at every screen refresh, and by displacing the position of the rod by a fixed number of pixels on each frame. For the slowest velocity in Experiment 1A, the rod was moved by two pixels in the direction of motion every screen refresh (85 Hz), for a translation velocity of two arcmin per 11.77 ms or 170 arcmin/sec. Likewise, the two other velocities in Experiment 1A were produced by displacing the rod by four or six pixels per screen refresh, producing velocities of 340 and 510 arcmin/sec, respectively. In Experiment 1B, the rod was displaced by one, three, or five pixels per screen refresh, yielding translation velocities of 85, 255, and 425 arcmin/sec, respectively. This method of animation produced motion that appeared smooth and continuous, while at the same time precisely controlling the velocity of the rod.

Design. We used an adjustment ("nulling") procedure in which each participant watched a misaligned rod move horizontally behind the occluding surface and then adjusted the top piece of the rod to make it appear aligned with the bottom piece. Once the participant perceived the rod as appearing aligned as it passed through the apertures, the trial ended.

The three independent variables manipulated in these experiments were the motion direction (leftward or rightward) and velocity (170, 340, and 510 arcmin/sec in Experiment 1A or 85, 255, and 425 arcmin/sec in Experiment 1B) of the rod, as well as the relative placement of the apertures. The apertures were arranged so that the top window was either to the right or to the left of the bottom window (forming either a "rightward" or "leftward" aperture configuration, respectively). Two trials were run for each combination of window configuration, motion direction, and translation velocity, one with the top piece of the rod to the right of the bottom piece (a "rightward" rod configuration), and the other with the top piece to the left (a "leftward" rod configuration). Each participant saw only three of the motion velocities, depending on whether they were in Experiment 1A or 1B, and the order of the 24 trials was randomized for each participant.

Procedure. Participants received instructions on how to perform the experiment and then completed two practice trials. The practice phase was used to introduce them to the adjustment method and data from these trials were discarded. Next, participants completed 24 experimental trials, which were presented in random order from a counterbalanced set. Participants were given two minute-long breaks during the experimental phase. The existence of the illusion was not revealed to the participants until the end of the experiment.

On each trial, the top and bottom pieces of the rod were misaligned by between 5 and 10 arcmin, chosen randomly. Half of the time the top piece was to the left and half of the time to the right of the bottom piece. Participants watched a motion sequence in which the misaligned rod translated once horizontally, either rightward or leftward on each trial, behind the apertures. A stationary, aligned version of the rod (with the middle third occluded) was always visible at the bottom of the screen as a reference for the alignment they were seeking to achieve. After the motion sequence, the participant adjusted the position of the top piece of the rod (using designated keys on the keyboard) to make it look more aligned. The motion sequence was repeated, with the adjusted rod alignment, and the participant continued to enter adjustments until the rod appeared aligned as it passed behind the occluder. Once the participant felt the top and bottom pieces of the rod were aligned, they pressed the 'M' key on the keyboard (for "Matched"). Afterward, the same rod alignment was shown once more so the participant could double-check that it indeed appeared aligned to them. At this point, they could either adjust it more, or press ‘M’ again. Each trial was ended when the participant pressed ‘M’ twice in succession.

Dependent measures and data analysis. In this and the following experiments, participants were excluded from analysis if 25% or more of their final alignments of the rod were the exact same as the initial random alignment at the beginning of the trial. This occurred infrequently, and was typically associated with subjective reports from the experimenters of participant noncompliance (e.g., speeding through the task to finish quickly).

The final position of the top piece of the rod, relative to the bottom piece, that appeared aligned to each participant was recorded for later analysis. Final alignments of the rod that resulted in a rightward configuration (the top piece to the right of the bottom piece) were recorded as positive numbers, and leftward configurations were recorded as negative numbers. The two final alignments for each condition were averaged, creating a 4 × 3 matrix of final alignment values consisting of four aperture and motion configurations (left or right motion crossed with leftward or rightward apertures) for the three velocities. The values plotted on the graphs are the means of participants’ final alignments.

Results

The results of Experiment 1A and 1B appear in Figure 4. The graphs plot participants’ mean adjusted positions of the top object fragment relative to the bottom fragment. As these adjustments reflect the positioning required to nullify the illusion, the illusion magnitude may be considered to be a perceived relative misalignment in the direction opposite to the adjustment. In the data graphs, it is apparent that illusion magnitude increases with velocity. It is also apparent that the spatial configuration of the apertures determines the direction of the illusion, regardless of motion direction. These observations were confirmed by the analyses.

Experiment 1A. The final positions of the top piece of the rod that participants judged as being aligned with the bottom piece were submitted to a 3 × 2 × 2 (Velocity × Aperture Configuration × Motion Direction) within subjects analysis of variance (ANOVA). The analyses revealed a main effect of aperture configuration, $F(1, 9) = 49.14$, $p < .0001$, indicating that participants consistently judged the moving rod to be aligned when the top piece was positioned to the right of the bottom piece for leftward aperture configurations, and to the left of the bottom piece for rightward aperture configurations. Additionally, the ANOVA revealed an interaction of velocity by aperture configuration, $F(2, 18) = 34.43$, $p < .0001$, reflecting the fact that as the velocity of the rod increased, so did the magnitude of the final alignments, with rightward aperture configurations causing participants to position the top piece of the rod to the left, and vice versa for leftward aperture configurations (see Figure 4). Neither the main effect of motion direction ($p > .10$) nor velocity ($F < 1$) were significant in this analysis. There were no other significant main effects or interactions (all $p > .10$).

The lack of a velocity main effect appears to be attributable to the fact that although the absolute magnitudes of the final alignments increased with velocity, half of the alignments were negative and half were positive, thus canceling out in the final analysis. Therefore, to better understand the effect of velocity, a 3 × 2 × 2

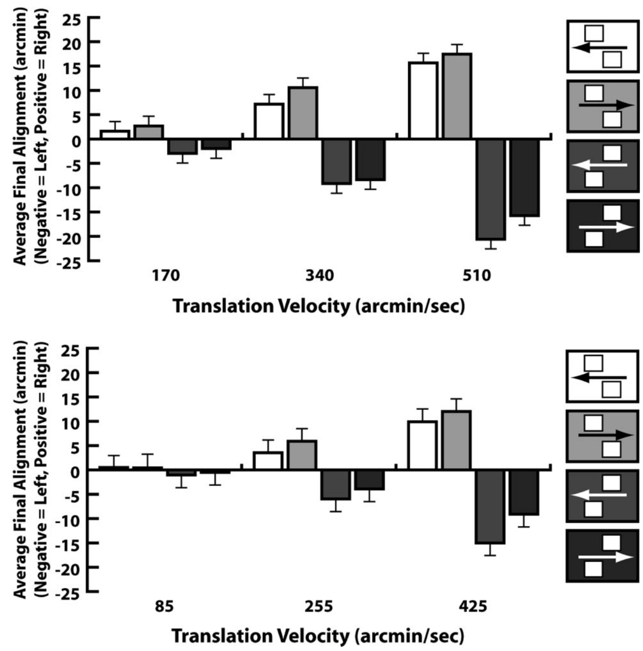

Figure 4. Final adjusted position of top piece of the rod, relative to the bottom piece, by stimulus velocity, motion direction, and aperture configuration for Experiment 1A (top) and 1B (bottom). Participants adjusted the top object fragment to appear aligned with the bottom. Positive alignment values indicate rightward displacement and negative alignment values indicate leftward displacement. Error bars are within-subjects confidence intervals (Loftus & Masson, 1994).

(Velocity × Aperture Configuration × Motion Direction) within-subjects ANOVA was run on the absolute values of the final alignments of the rod to prevent the window configuration scores from canceling each other out. This analysis revealed a main effect of velocity, $F(2, 18) = 34.94$, $p < .0001$, indicating that the final reported position of the top piece of the rod became more misaligned from the bottom piece as the velocity of the rod increased. The magnitude of this effect appears to account fully for the interaction of aperture configuration by velocity in the earlier analysis. There were no other significant main effects or interactions in this analysis (all $p > .10$).

Based on the results of the analyses conducted so far, we can view the primary effects in the data more simply by collapsing across motion direction and aperture configuration. Figure 5A depicts the overall unsigned magnitude of the final alignment of the rod. This presentation of the data shows that as velocity increased, illusion magnitude increased. Planned comparisons were conducted on the unsigned magnitudes of the final alignments between the three velocities. These confirmed that the overall illusion magnitude for the 510 arcmin/sec condition was greater than the 340 arcmin/sec condition, $t(18) = 2.17$, $p = .044$, which was in turn greater than the 170 arcmin/sec condition, $t(18) = 3.72$, $p = .0016$. The best-fitting linear function that describes the data was calculated by using least-squares estimation and yielded a slope of .044 arcmin/(arcmin/sec) and a y intercept of -5.58 arcmin ($r^2 = .99$). The slope of .044 arcmin/(arcmin/sec) reduces to .044 sec, which represents the extra amount of time per velocity unit that the first-seen piece of the rod needed to appear

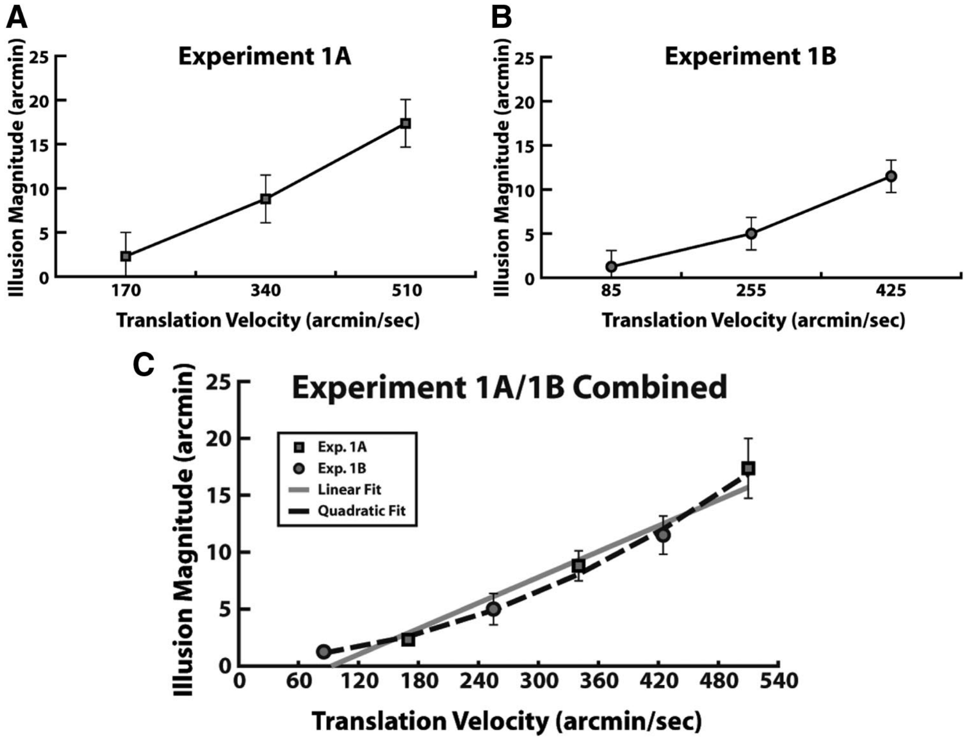

Figure 5. Unsigned magnitude of the illusion as a function of velocity for Experiment 1. A) Data for Experiment 1A. B) Data for Experiment 1B. C) Combined data for Experiment 1A and 1B, along with best-fitting linear and quadratic functions. The data represent absolute values of final alignments, averaged across motion direction and aperture configuration. Error bars in A and B represent within-subjects confidence intervals (Loftus & Masson, 1994), whereas error bars in C are standard error of the mean.

ahead of its aligned position to look aligned with the second piece of the rod. Thus, as the object velocity gets faster, the necessary time between appearances of the pieces of the rod increases in order for the two pieces to appear aligned.

Experiment 1B. The structure of the analyses for Experiment 1B was the same as for Experiment 1A. A within-subjects ANOVA detected main effects of aperture configuration, $F(1, 9) = 26.08$, $p < .001$ and motion direction, $F(1, 9) = 9.23$, $p < .05$, and an interaction of velocity by aperture configuration, $F(2, 18) = 43.28$, $p < .0001$. There were no other significant effects in this analysis (all $p > .05$).

As with Experiment 1A, we analyzed absolute values of the final alignments to better understand the effect of velocity on these data. This analysis indicated a main effect of velocity, $F(2, 18) = 34.34$, $p < .0001$, and planned comparisons established that the illusion was greater for the 425 than the 255 arcmin/sec velocity, $t(18) = 3.03$, $p < .005$, which in turn was larger than the 85 arcmin/sec condition, $t(18) = 4.69$, $p < .0001$. The ANOVA on unsigned misalignment magnitudes also returned significant interactions of aperture configuration by motion direction, $F(1, 9) = 8.16$, $p < .05$, and velocity by aperture configuration by motion direction, $F(2, 18) = 4.05$, $p < .05$. There were no other significant main effects or interactions in this analysis (all $F < 1$).

An index of the overall unsigned magnitude of the illusion as a function of velocity for Experiment 1B was created by collapsing the data across motion direction and aperture configuration (see Figure 5B). The best-fitting linear function that describes these data was calculated by using a least-squares estimation and yielded a slope of .030 sec and a y intercept of -1.8 arcmin ($r^2 = .98$).

The results of Experiment 1 indicate that the overall magnitude of the illusion increases monotonically with object velocity. Furthermore, the perceived misalignment of the rod is a function of the configuration of the apertures, not the motion direction of the rod. The data from Experiment 1A and 1B were combined to yield a function that describes the overall magnitude of the illusion for the six velocities tested across the two experiments (Figure 5C). The best fitting linear function for the combined data was determined using least-squares estimation, and yielded a slope of .039 sec with a y intercept of -3.72 arcmin ($r^2 = .96$). A slightly better fit of the data was achieved with a quadratic function that was determined using least-squares estimation to be $y = 0.000055x^2 + 0.0058x - 0.01$ ($r^2 = .99$). This indicates that a curvilinear fit may be slightly better for these data because it captures the flattened illusion magnitudes at lower velocities better than the linear fit.

Discussion

Experiment 1A and 1B provided clear evidence for a robust perceptual illusion: When a dynamically occluded object is perceived through two nonoverlapping and misaligned apertures, observers consistently misperceive the spatial alignment of the visible regions of the object to be distorted in the direction of the apertures. Thus, we refer to this illusion as the *aperture capture illusion*. The misperceived alignment can be nulled by shifting one of the pieces of the object relative to the other along the same axis as the object's motion trajectory. The end result is that faster moving objects needed larger spatial misalignments to appear perceptually aligned when seen through these apertures. The aperture capture illusion stands in stark contrast to the robust, spatially accurate perception of object form that comes from perceiving a dynamically occluded object through many apertures (Palmer et al., 2006; Experiment 5, below).

What causes the aperture capture illusion? We cannot answer this question from this experiment alone, but we can rule out one possible explanation. A potential explanation for the aperture capture illusion is that it represents a failure of temporal resolution by the visual system: if the appearance of both pieces of the rod through both windows happens close enough in time, then perhaps the visual system treats these two events as one event (similar to Chun & Potter's, 1995 theory of the attentional blink and Di Lollo, Hogben, & Dixon's, 1994 theory of temporal integration), and perceives the rod as misaligned in the direction of the apertures. After all, the top and bottom portions of the rod *did* appear through the apertures at misaligned locations in the visual field, so a misperception of form in the direction of the apertures is perhaps not that surprising.

The temporal resolution hypothesis, however, is disconfirmed by the fact that the illusion can be nulled. By misaligning the top and bottom pieces in the direction opposite to the displacement of the apertures, the two object fragments can be made to appear aligned, under all of the conditions tested in this experiment. By misaligning the rod in the direction opposite to the apertures, one is effectively *increasing* the amount of time between the appearance of the top and bottom of the rod. However, the perception that one has in this case is that the rod looks *more aligned* rather than less, despite the fact that the time between the presentations of the two pieces of the rod has been increased. Therefore, the source of the illusion does not appear to be a failure of temporal resolution by the visual system.

Beside the preceding possibility, how else might we understand the aperture capture illusion? As discussed above, the theory of Spatiotemporal Relatability proposed by Palmer et al. (2006) to explain the perception of dynamically occluded objects suggests a framework for understanding this illusion. Perhaps position updating mechanisms that allow object formation from currently and previously visible fragments do not always operate veridically. In impoverished situations, such as this two aperture display, the visual system may underestimate the velocity of the rod after a piece of it passes behind the occluding surface.

To present this account more clearly, it is useful to distinguish between two sorts of object velocities in these displays. We will refer to the perceived velocity of a fully visible object as its *real velocity* and the perceived velocity of an occluded (yet persisting) object as its *occlusion velocity*. Thus, using this new terminology, one principle that might explain the illusion is that occlusion velocity is perceived to be slower than real velocity. Such an error would lead to inaccurate position updating of the occluded regions and incorrect perception of the physical relationship between previously seen and later seen object regions, leading to the illusion described here. Importantly, this explanation for the illusion also explains how the illusion could be nulled by misaligning the pieces of the rod in the direction opposite to the offset of the windows since that would give the occluded fragment a "Head Start" to compensate for its slower occlusion velocity. Underestimation of velocity could arise from the following: a) improper perception of real velocity within the leading aperture (the aperture in which the rod initially appears), b) inaccurate maintenance of occlusion velocity after occlusion, or c) inaccurate perception of real velocity of the second appearing rod piece.

With regard to the first error, given that the size of the apertures was the same for all velocities, it follows that the amount of time that the rod was visible through the apertures decreased as a function of velocity. Thus, the illusion may have arisen because the visual system did not have enough time or space to extract an accurate motion signal from the rod at faster velocities, and consequently did not update the position of the rod behind the occluder correctly. This hypothesis is addressed in Experiment 2.

Experiment 2: Inaccurate Motion Signals Before Occlusion?

Experiment 1A and 1B established that the magnitude of the aperture capture illusion increases as the velocity of the rod increases. One possible explanation for the illusion is an inaccurate initial representation of the velocity of the rod when it is seen within the leading aperture. Given that the size of the apertures was held constant for all velocities, the amount of time that the rod was visible within the first aperture decreased as the velocity of the rod increased. If accurate velocity perception is dependent upon adequate exposure time, then the real velocity of the figure may have been underestimated at higher velocities with lower exposure durations.

A constant aperture size not only limits the amount of time available to extract a motion signal for higher velocities, but also limits the space available. Van de Grind, Koenderink, and Van Doorn (1986) showed that the minimum spatial interval on the retina necessary to perceive motion accurately increases with stimulus velocity. Given this notion, it is possible that the aperture sizes in our experiments did not allow enough space for accurate registration of faster velocities. If this possibility is true, then the illusion could be caused by underestimation of object velocity for the first appearing fragment. If so, a larger leading aperture in our displays might allow more accurate extraction of a velocity signal from the stimulus and eliminate the illusion.

In Experiment 2, we varied the size of the leading aperture along with the velocity of the rod such that approximately the same exposure time was achieved for the three velocities tested (see Figure 6 and *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/AC_Exp2_2.mov* for the 170-arcmin/sec velocity, *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/AC_Exp2_4.mov* for the 340-arcmin/sec velocity, and *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/AC_Exp2_6.mov* for the 510-arcmin/sec velocity, each with a 350-ms exposure duration). Besides testing exposure time between velocities, we also tested it within velocity by using three different exposure durations for each rod velocity. If exposure time of the leading aperture affects the magnitude of the illusion, then the illusion should be the same strength between velocities with the same exposure duration and should vary within each velocity as exposure duration varied.

We increased only the size of the leading aperture because we were interested in evaluating the effect of improved velocity

Figure 6. Depiction of the occluders used in Experiment 2. The size of the leading aperture was increased in proportion to the velocity of the rod to ensure that the top rod fragment was always visible for approximately the same amount of time.

information before the first part of the rod became occluded. We assumed that the process of integrating the occluded and visible regions of the rod could begin immediately upon appearance of the second part of the rod (this assumption was tested directly in Experiment 7).

Method

Unless otherwise stated, methods for this experiment were identical to those in Experiment 1.

Participants. Thirty-four UCLA undergraduates participated in the experiment in partial fulfillment of course requirements for an introductory psychology class. All participants gave informed consent, reported normal or corrected-to-normal vision, and were naïve to the purposes of the experiment. Each participant was assigned to one of three experimental conditions, for a total of 10 participants in each condition. Four participants were excluded for failing to fulfill the objective standard of task compliance.

Stimuli. The occluder was modified in this experiment so that the leading aperture (the aperture in which the rod was seen first) was larger than the trailing aperture (the aperture in which the rod was seen last). While the trailing aperture's width remained 40 arcmin (the same as Experiment 1), the leading aperture size was adjusted according to the velocity of the rod to yield exposure durations of approximately 272, 350, and 428 ms (see Table 1). The exposure duration of 350 ms was achieved precisely for all three velocities, whereas the other two exposure durations were closely approximated within the timing limits imposed by the refresh rate of the monitor.

Design. This experiment employed a 3 (Velocity of Rod) $\times$ 2 (Aperture Configuration) $\times$ 2 (Motion Direction) $\times$ 3 (Leading Aperture Size/Exposure Duration) mixed design. The three velocity conditions were run between subjects.

Results

Varying exposure durations had little effect on illusion magnitude. As in Experiment 1, the magnitude of the illusion strongly increased with the velocity of the rod. However, equalized exposure durations did not lead to equalized illusion magnitudes. Taken together, these two findings demonstrate that the velocity of the rod and not the length of time it is seen through the leading aperture is what determines illusion magnitude.

These findings were confirmed by the analyses. The final alignments of the top versus bottom pieces of the moving rod were submitted to a 3 $\times$ 2 $\times$ 2 $\times$ 3 (Velocity Group $\times$ Aperture Configuration $\times$ Motion Direction $\times$ Exposure Duration) mixed ANOVA with the first factor as a between-subjects variable and repeated measures on the other three factors. Analyses revealed a main effect of aperture configuration, $F(1, 27) = 33.72$, $p < .0001$, and a significant interaction of aperture configuration $\times$ velocity group, $F(2, 27) = 4.50$, $p < .05$. There were no other significant main effects or interactions in this analysis (all $p > .10$).

As in the previous two experiments, a 3 $\times$ 3 (Velocity $\times$ Exposure Duration) mixed ANOVA was run on the unsigned illusion magnitude data, with velocity as a between-subjects variable (see Figure 7). The analysis detected a main effect of velocity, $F(2, 27) = 15.06$, $p < .001$, with participants in the 510 arcmin/sec condition exhibiting larger illusion magnitudes than participants in the 340 arcmin/sec condition, $t(9) = 2.55$, $p = .031$, who in turn showed a larger illusion magnitude than the 170 arcmin/sec group, $t(9) = 3.23$, $p = .010$. The ANOVA did not reveal any other significant main effects or interactions (all $p > .10$, observed power for the main effect of exposure duration was 0.295).

Table 1
Aperture Sizes and Exposure Durations in Experiment 2

Translation velocity of rod	Exposure space and time in leading aperture: Short	Medium	Long
170 arcmin/sec	47 arcmin *274 ms*	60 arcmin *350 ms*	73 arcmin *426 ms*
340 arcmin/sec	93 arcmin *271 ms*	120 arcmin *350 ms*	147 arcmin *428 ms*
510 arcmin/sec	160 arcmin *272 ms*	180 arcmin *350 ms*	200 arcmin *428 ms*

Discussion

The results of Experiment 2 indicate that the illusion is not related to the time or space available for registration of the motion of the rod within the leading aperture. The increasing illusion magnitude at higher velocities in Experiment 1 was not attributable to use of a common aperture size. Figure 7 shows that velocity predicts the magnitude of the illusion independently of exposure duration.

The apertures used in this experiment equated the exposure duration of the rod within the leading window, ensuring that time of exposure was not correlated with velocity. The size of the leading aperture for the fastest velocity was dramatically increased (from 40 arcmin to 200 arcmin), ensuring that motion perception circuits had enough space over which to properly encode the velocity of the stimulus within the leading aperture.

Experiment 2 addressed an important family of possible accounts of the illusion relating to inaccurate initial perception of the real velocity of the object. There are, however, other possible contributing factors to accurate velocity perception besides the spatial and temporal intervals in which velocity information is gathered. One such variable involves the possibility that motion signals are not veridical in the absence of a textured background.

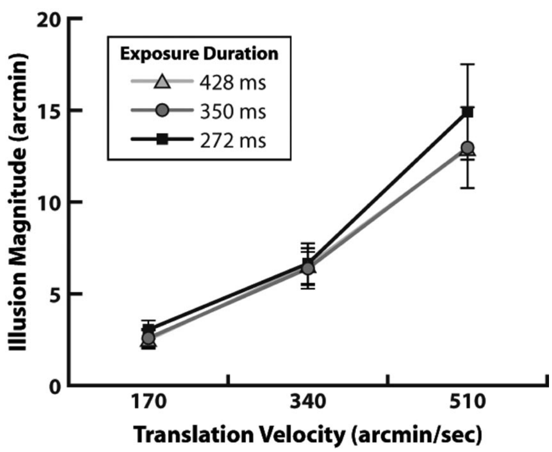

Figure 7. Illusion magnitude in Experiment 2 as a function of translation velocity of rod and exposure duration in leading aperture. Illusion magnitude depends primarily on velocity, not exposure duration. Error bars indicate ± one standard error of the mean.

The representation of the velocity of a fragment that has become occluded depends on the velocity registered while it is visible; a manipulation that improved registration of the visible fragment's velocity might decrease the illusion. Another factor not yet explored is the possible role of eye movements for registering velocity in these displays. Experiment 3 investigated the effect of adding background texture to the displays (which provides accretion/deletion information that may improve accurate motion tracking), and Experiment 4 examined the effect of smooth pursuit eye movements tracking the rod as it moved behind the occluding surface.

Experiment 3: Accretion/Deletion Cues

The results of the previous experiments indicate that if the aperture capture illusion is the result of an underestimation of velocity within the leading aperture, the cause of this underestimation is not a lack of time or space to extract a veridical motion signal. What other sources of information could influence the accurate extraction of the rod's velocity? It is known that motion sensitivity is enhanced when motion occurs relative to visible background features (Wallach, 1959). Motion in front of a textured background also provides an additional cue to depth and motion–accretion and deletion of background texture elements (Gibson et al., 1969) and might serve to improve the extraction of an accurate velocity signal within the leading aperture.

Accordingly, in this experiment we added a random dot texture to the background to provide accretion/deletion cues as the rod was seen through the apertures. It seemed possible that this extra source of motion information might improve participants' perception of the rod's real velocity, and therefore reduce or eliminate the illusion.

Method

Participants. Eleven UCLA undergraduates participated in the experiment in partial fulfillment of course requirements for an introductory psychology class. All participants gave informed consent, reported normal or corrected-to-normal vision, and were naïve to the purposes of the experiment. One participant was excluded for failure to meet the objective standard of task compliance.

Stimuli. In this experiment, the rod had squared corners rather than the rounded corners used in the previous experiments. This was done to ensure that there was no white space in the image of the rod that would cause improper accretion/deletion cues since the extra white space in the corners of the rod image file caused the textured dot background to seem as if it is being covered and uncovered by a surface slightly larger than the rod itself. To solve this problem, the corners of the bitmap were filled out to create a rod with squared corners rather than rounded corners (as in Figure 8 and *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/AC_Exp3_4.mov*). A pilot study with three subjects found no difference in performance on a version of Experiment 1 with squared versus rounded corners.

The occluder in this experiment was identical to the one used in Experiment 1, except that a blue dot background texture was visible through the apertures (see Figure 8). The background dot texture was randomly generated on each trial to rule out the possibility that participants were using the placement of the dots as a reference point while aligning the rod.

Results

Adding a textured background within the apertures of the occluding surface did not eliminate the illusion and, in fact, increased it relative to the magnitudes at the same velocities in Experiment 1A. In all other respects, the pattern of results observed in Experiment 1 was also observed here. The major determinant of illusion magnitude was the arrangement of the apertures rather than the motion direction of the rod, and the magnitude of the illusion increased monotonically with rod translation velocity.

A 3 × 2 × 2 (Velocity × Aperture Configuration × Motion Direction) within-subjects ANOVA was run on these data. The analysis revealed a main effect of aperture configuration, $F(1, 9) = 139.07$, $p < .0001$, and a significant interaction of aperture configuration by velocity, $F(2, 18) = 78.63$, $p < .0001$. A 3 × 2 × 2 (Velocity × Aperture Configuration × Motion Direction) within-subjects ANOVA on the absolute values of the data revealed a main effect of velocity, $F(2, 18) = 78.29$, $p < .0001$. No other main effects or interactions were significant in this analysis (all $p > .25$).

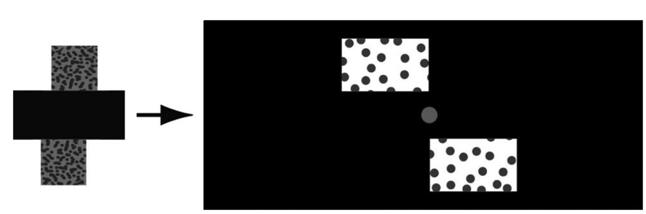

Figure 8. Occluder and dot-textured background used in Experiment 3. The dot texture that was visible through the apertures provided the additional visual cue of accretion/deletion of background texture (see *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/AC_Exp3_4.mov* for a depiction of the rod moving at the 340 arcmin/sec velocity). Note that the rod was red with black speckles and the dots behind the aperture were blue in the original displays.

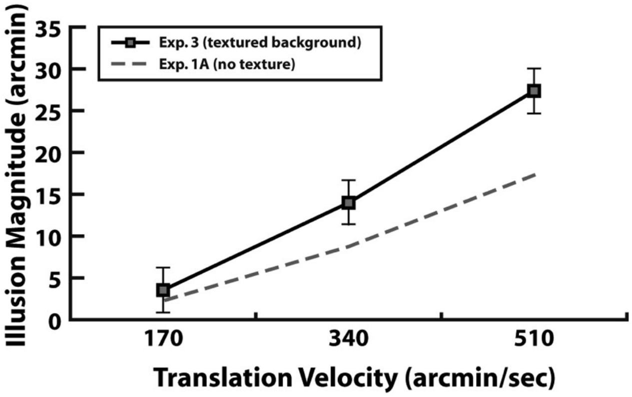

Figure 9. Illusion magnitude as a function of the translation velocity of the rod in Experiment 3. Error bars are within-subjects confidence intervals (Loftus & Masson, 1994).

The unsigned magnitude of illusion is depicted in Figure 9. Planned comparisons between the three velocities confirmed that the overall magnitude of the illusion was smaller in the 170 arcmin/sec condition than in the 340 arcmin/sec condition, $t(18) = 6.28$, $p < .0001$, and larger in the 510 arcmin/sec condition than in the 340 arcmin/sec condition, $t(18) = 4.58$, $p = .0002$. A linear trend analysis of these data using least-squares estimation yielded an estimated slope of .066 sec and a y-intercept of -7.95 arcmin ($r^2 = .99$).

Discussion

This experiment tested whether adding a visual cue that might provide enhanced registration of the true velocity of the rod would decrease or eliminate the illusion. We used a random dot background behind the occluding surface to provide stationary reference features and accretion/deletion information. If the source of the illusion is an initial underestimation of the rod's velocity within the leading aperture, we would expect participants to benefit from enhanced visual information, with the consequence of a decreased illusion magnitude.

Contrary to this prediction, however, the magnitude of the illusion *increased* rather than decreased. In Experiment 3, the slope of the illusion function was 70 ms, significantly larger than the slope 44 ms slope observed in Experiment 1A (with the same sized apertures and velocities), $t(9) = 4.11$, $p < .05$. Additionally, the maximum magnitude of illusion observed in Experiment 3 was 27 arcmin as opposed to 17 arcmin in Experiment 1A, an increase of 10 arcmin, or roughly 1/3 the width of the rod. These two experiments differed only in the presence of a random dot background, and in a minor shape change to the rod (that produced no significant differences in a pilot study without a textured background).

The most likely explanation for the increase in illusion magnitude compared with Experiment 1A is that the addition of a textured background may have made the rod appear to move faster than it would have appeared on an untextured background. Because the magnitude of the illusion increases with velocity, larger illusion magnitudes were observed.

Taken together with the results of Experiment 2, these data so far do not support the hypothesis that the illusion is caused by misperception of the rod's velocity within the leading aperture.

Experiment 4: Smooth Pursuit Eye Movements

The previous two experiments aimed to evaluate the idea that there is some shortcoming in the visual system's ability to extract the real velocity of the rod while it is visible within the leading aperture of the displays. Underestimation of velocity of the first rod fragment might predict the aperture capture illusion observed in the first three experiments. Stimulus manipulations designed to enhance velocity perception for the visible fragment proved unsuccessful at eliminating the illusion.

In Experiment 4, we explored the effect of smooth pursuit eye movements on the illusion, with two objectives. One objective was to provide a different, unambiguous cue to the rod's real velocity. The other objective was to assess whether and to what degree the illusion occurred when the observer made a smooth pursuit eye movement tracking the rod behind the occluder.

To assess both the retinal painting hypothesis and the effect of unambiguous velocity information about the rod, participants in this experiment tracked a small fixation dot that translated across the visual field in lock step with the rod (Figure 10 and *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/AC_Exp4_6.mov*). The tracking dot was displayed in front of the occluding surface at the vertical midpoint of the rod, at a randomly chosen horizontal position within the boundaries of the occluded rod (this ensured that the participants were not able to use the dot itself as a frame of reference for determining the alignment of the rod from trial to trial). During the animation sequence, the dot moved at the same rate as the rod, within the boundaries of the rod, thus allowing the projection over time of aligned bottom and top pieces of the rod onto the retina, if accurate ocular pursuit was maintained. Participants were instructed to maintain fixation on the moving dot while performing the illusion nulling task and the experimenter monitored their eye movements to make sure smooth pursuit tracking of the rod occurred on every trial.

Figure 10. Schematic of displays used in Experiment 4 (see *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/AC_Exp4_6.mov* for an example of the rod moving at 510 arcmin/sec). The fixation dot moved either rightward or leftward during each trial, in concert with the motion direction of the rod. The dot first appeared at its initial position, then a beep sounded, after which the dot traveled horizontally across the occluding surface at the same rate and within the boundaries of the occluded rod. Participants were instructed to track the fixation dot with their eyes as it moved across the screen. Note that the rod was red with black speckles and was never fully visible during the experiment—it is shown here for illustrative purposes.

The retinal painting hypothesis predicts that as observers track the fixation point, the top and bottom portions of the rod will be emblazoned on their retinas in succession, allowing them to accurately perceive the alignment of the rod pieces and thus perceive no illusion. In addition to the retinal painting hypothesis, there is another reason to expect that the moving fixation dot might eliminate the illusion. To the extent that smooth pursuit eye movements indicate the real velocity of an object moving in the environment, we would predict that participants have knowledge, at some oculomotor level, of the rod's velocity. Therefore, the moving fixation dot should provide excellent perceptual evidence of the rod's velocity, within and between both apertures of the occluding surface.

Method

Participants. Twelve UCLA undergraduates participated in the experiment in partial fulfillment of course requirements in an introductory psychology class. All participants reported normal or corrected-to-normal vision, gave informed consent, and were naïve to the purposes of the experiment. Two participants were excluded for failure to meet the objective standard of compliance with the experimental protocol.

Stimuli. In this experiment, the fixation dot moved horizontally across the screen, along with the rod (see Figure 10). The fixation dot was always presented in front of the occluder, but within the boundaries of the rod. The position of the dot was presented at a random location horizontally within 10 pixels of the average position of the rod and vertically exactly between the centers of the top and bottom pieces.

Procedure. The same procedure that was used in Experiment 1 was also used here, except that the experimenter monitored observers' eye movements during the study. The experimenter sat in a chair on one side of the computer screen and monitored observers' eyes to ensure they executed a smooth pursuit eye movement on each trial. If the observer did not execute a smooth pursuit eye movement during a presentation of the stimulus, they were reminded to do so before the next presentation by the experimenter. All participants complied with experiment instructions and executed smooth pursuit eye movement during stimulus presentation and nulling.

Results

Providing a smooth pursuit fixation target that traveled along with the rod as it moved behind the occluding surface did not eliminate the illusion (Figure 11), contrary to the retinal painting hypothesis. Illusion magnitudes were smaller than in the previous experiments, but still reliably greater than zero. The moving fixation manipulation did yield a slightly different pattern of illusion magnitudes, however. In this experiment, with smooth pursuit eye movements, when the rod was seen through the bottom aperture first, illusion magnitudes were reduced (but not eliminated) relative to when the rod was seen through the top aperture first.

These observations were confirmed by the statistical analyses. Rod alignments were submitted to a 3 × 2 × 2 (Velocity × Aperture Configuration × Motion Direction) ANOVA. The analysis revealed a main effect of aperture configuration, $F(1, 9) = 70.92$, $p < .0001$, reflecting the fact that participants chose final alignments for the rod that were misaligned in the direction opposite to the apertures. Unlike previous experiments, there was a main effect of motion direction, with rightward motions yielding slightly negative alignments (rod top to the left of rod bottom) and leftward motions yielding slightly positive alignments (rod top to the right of rod bottom), $F(1, 9) = 16.38$, $p < .005$. There was also a significant interaction of aperture configuration by velocity of the rod, $F(2, 18) = 19.19$, $p < .0001$, attributable to the difference between alignment scores for rightward and leftward aperture configurations increasing along with the velocity of the rod. Finally, motion direction had a significant interaction with velocity, $F(2, 18) = 4.41$, $p < .05$, owing to larger differences between rightward and leftward motion directions at higher velocities. There were no other significant main effects or interactions (all $p > .10$).

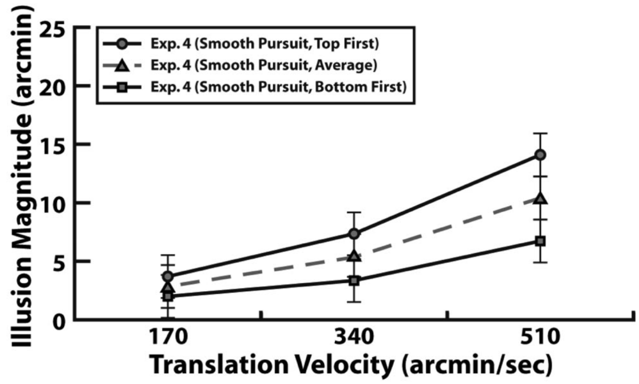

Figure 11. Overall magnitude of the illusion for Experiment 4, plotted as a function of whether the translating rod was visible first through the top aperture or the bottom aperture. Mean illusion magnitudes are also presented. Error bars are within-subjects confidence intervals (Loftus & Masson, 1994).

Absolute values of the final alignments were analyzed with a 3 × 2 × 2 (Velocity × Aperture Configuration × Motion Direction) ANOVA. This analysis revealed a main effect of velocity, $F(2, 18) = 15.62$, $p < .001$, and detected an aperture configuration by motion direction interaction, $F(1, 9) = 16.29$, $p < .005$, with the magnitude of the illusion being smaller when the rod was seen in the lower aperture first, regardless of motion direction. There were no other significant main effects or interactions in this analysis (all $p > .10$).

For the sake of comparison between experiments, an index of the overall magnitude of illusion for each velocity was created (see Figure 11). It is important to note that there was both a main effect of motion direction as well as an interaction of motion direction by aperture configuration in this experiment. Accordingly, we plot data separately for rod configurations that entered the top aperture first, the bottom aperture first, and the average of these two. A linear trend analysis using least-squares estimation was conducted on the average illusion magnitudes and established that the best fitting linear function for these data has a slope of .023 sec and a y intercept of -1.60 arcmin ($r^2 = .95$).

Discussion

The overall magnitude of the illusion in this experiment was less than in Experiment 1A, except at the slowest velocity. This indicates that participants received some benefit from the smooth-pursuit tracking manipulation. However, contrary to the predictions of the retinal painting hypothesis, the illusion was not eliminated. The stimulus arrangement in this experiment was designed to enable observers to take advantage of retinal painting, if at all possible. Participants were given a fixation dot that moved in sync with the rod as it passed behind the occluding surface, along with instructions to track the dot. With accurate ocular pursuit of the moving fixation dot, the two fragments of the rod appearing through the apertures would be perfectly aligned on the retinas when they were aligned in the physical stimulus. The results show that even with support given for appropriate ocular pursuit and constant monitoring by an experimenter to ensure pursuit of the dot, observers could not use retinal painting to properly align the top and bottom halves of the rod. We believe that these results eliminate the possibility that the illusion is attributable to retinal painting.

The magnitude of the illusion was reduced in conditions in which the rod entered the bottom aperture first, regardless of motion direction. This result is interesting in that it suggests more accurate processing for velocity signals in the lower visual field for stimuli that are being tracked with the eyes. Some research suggests that attentional resolution is better in the lower visual field than in the upper visual field. For instance, He, Cavanagh, and Intriligator (1996) used an attentional tracking task in which observers fixated a central dot and then tracked several targets and found that tracking performance was better in the lower visual field. Within the context of the present experiment, the He et al. (1996) result suggests that attentional resources may have been occupied for a shorter time when the rod was visible in the lower visual field first, and then were able to be allocated to the top portion of the rod more quickly due to the higher temporal resolution of attention in the lower visual field.

Additionally, some physiological evidence suggests that there is a larger cortical representation of the lower visual field than the upper visual field for some dorsal extrastriate areas, such as MT/V5 (Maunsell & Newsome, 1987; Maunsell & van Essen, 1983). Previc (1990) argued that the lower visual field is specialized for visuomotor tasks controlled primarily by the dorsal stream, whereas the upper visual field is specialized for object and scene perception tasks controlled primarily by the ventral stream. If this is the case, then portions of the rod that appeared in the lower visual field first may have benefited from a processing advantage for position information that did not occur when the rod appeared in the upper visual field first.

However, it should be noted that the performance advantage for stimuli that were visible in the lower visual field first only occurred in the presence of smooth-pursuit eye movements and was not observed in any of the other experiments reported in this study. Consequently, this effect will not be considered further.

Experiment 5a and 5b: Effects of Multiple Apertures on the Illusion

The idea of erroneous position updating due to underestimated occlusion velocity in representations of occluded fragments is perplexing in at least one respect. Palmer et al. (2006) found results that indicated high precision object formation from spatiotemporally separated object fragments when occluding surfaces had many apertures. Specifically, object formation depended on relatability constraints similar to those in static arrays (Kellman & Shipley, 1991), and discrimination performance for spatial relations of fragments was markedly enhanced under conditions of object formation. Both of these findings imply that the spatial positions of the fragments were very accurately represented in the dynamically occluded displays studied by Palmer et al. (2006). Given that there is little or no positional distortion in the presence of multiple apertures, but large and systematic distortion in the presence of just two apertures, it seems important to explore the effect of the number of apertures on the perceived alignment of the rod.

In this experiment, the overall visible area within the apertures was the same as in Experiments 1, 3, and 4, but the number of apertures was increased (see Figure 12). This was accomplished by dividing the width of each aperture in half each time the number of apertures was doubled. By increasing the number of apertures, while at the same time holding constant the total area of the rod that was physically visible, we were able to specifically examine the effects of repeated exposure of the rod on the final alignments of the top and bottom pieces. A video of the configuration with four apertures is available from *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/AC_Exp5_4.mov*, and a video of the configuration with eight apertures is available from *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/AC_Exp5_8.mov*.

The hypothesis in these experiments is that as the number of apertures increases, the overall magnitude of the illusion should

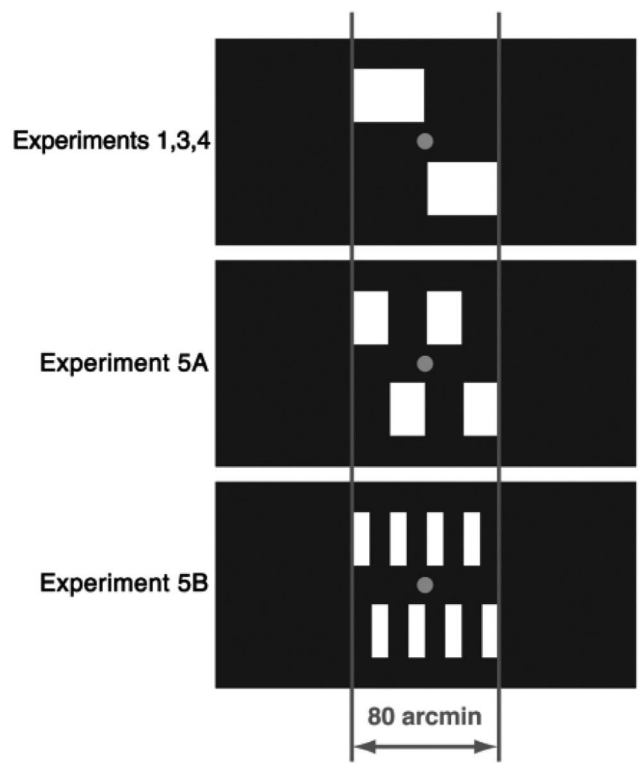

Figure 12. The three occluding surfaces used in Experiments 1–4, 5A, and 5B. The overall area of exposure within the apertures and the horizontal extent of the apertures within all three occluders is the same, but the number of apertures varies.

decrease. We predicted a decrease in the magnitude of the illusion for several reasons: a) There was no illusion observed in the dynamic occlusion experiments reported by Palmer et al. (2006), so manipulations that make the current displays more like those should decrease the illusion, b) If the illusion is the result of an underestimation of occlusion velocity, it should decrease for shorter occlusion episodes with multiple apertures since there will be less time for the underestimated occlusion velocity to have an impact on perception, and c) If the illusion is the result of an underestimation of occlusion velocity, then allowing the visual system to resample the true position of the rod after occlusion (through multiple apertures) could correct or replace erroneous position tracking based on the underestimated velocity of the rod.

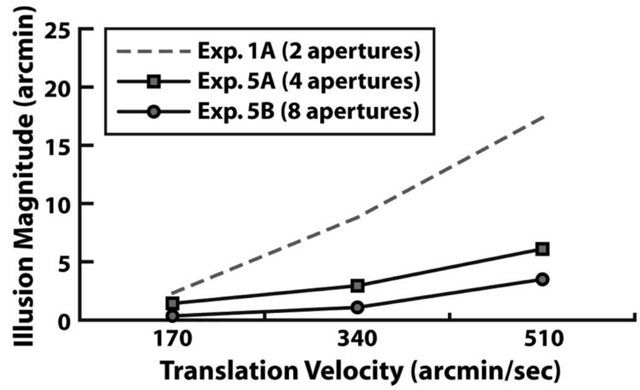

Figure 13. Overall illusion magnitudes for Experiment 5A and 5B compared with Experiment 1A. Dividing two apertures into four or eight apertures decreased illusion magnitudes dramatically. Error bars indicate ± one standard error of the mean and are too small to be seen in this graph.

Method

Participants. Twelve UCLA undergraduates participated in Experiment 5A and 10 UCLA undergraduates participated in Experiment 5B. Participants performed the experiments in partial fulfillment of course requirements for an introductory psychology class. All participants reported normal or corrected-to-normal vision, provided informed consent, and were naïve to the purposes of the experiment. Two participants were excluded from Experiment 5A because they did not meet the objective requirement of task compliance.

Stimuli. The rod traveled behind occluding surfaces that are depicted in Figure 12. In Experiment 5A, the occluding surface had four windows measuring 20 arcmin wide each and in Experiment 5B, the occluding surface had eight windows measuring 10 arcmin wide each. The total width visible through the occluding surface was always 80 arcmin, as in Experiments 1, 3, and 4.

Results

The dividing of two large apertures into either four or eight smaller ones resulted in a dramatic reduction in the size of the illusion. The reduction was greater for displays with eight apertures than for those with four apertures. It seems that many small apertures lead to better perception of a dynamically occluded object than fewer large apertures, even if total exposure time and space is held constant.

Experiment 5A. Final alignments were analyzed with a 3 × 2 × 2 (Velocity × Aperture Configuration × Motion Direction) within-subjects ANOVA, which revealed a main effect of aperture configuration, $F(1, 9) = 38.16$, $p < .0005$, and a significant interaction of velocity by aperture configuration, $F(2, 18) = 23.06$, $p < .0001$. A 3 × 2 × 2 (Velocity × Aperture Configuration × Motion Direction) within-subjects ANOVA on the absolute value data detected a main effect of velocity, $F(2, 18) = 19.20$, $p < .0001$. There were no other significant main effects or interactions (all $p > .15$).

An index of the unsigned magnitudes of final alignments was created by collapsing the full absolute value dataset across motion direction and aperture configuration (see Figure 13). Planned comparisons of these data established that the 170 arcmin/sec condition yielded significantly lower overall magnitudes of final alignments when compared to the 340 arcmin/sec condition, $t(18) = 2.43$, $p = .026$. Additionally, the overall magnitude of the final alignments in the 510 arcmin/sec condition was significantly greater than the overall magnitudes of the final alignments for the 340 arcmin/sec condition, $t(18) = 2.86$, $p = .011$. The best-fitting linear function for these data, as determined by a least-squares fit, has a slope of .014 sec and a y intercept of -1.19 arcmin ($r^2 = .96$).

Experiment 5B. A 3 × 2 × 2 (Velocity × Aperture Configuration × Motion Direction) within-subjects ANOVA of the raw alignment data revealed main effects of aperture configuration, $F(1, 9) = 34.48$, $p < .0005$, and motion direction, $F(1, 9) = 11.51$, $p < .01$, as well as a significant interaction of velocity by aperture configuration, $F(2, 18) = 64.69$, $p < .0001$ (see Figure 13). The absolute value data were analyzed with a 3 × 2 × 2 (Velocity × Aperture Configuration × Motion Direction) within-subjects ANOVA. The analysis detected a main effect of velocity, $F(2, 18) = 31.51$, $p < .0001$, and significant interactions of aperture configuration by motion direction, $F(1, 9) = 9.28$, $p < .05$, and velocity by aperture configuration by motion direction, $F(2, 18) = 9.28$, $p < .005$. There were no other significant main effects or interactions in this analysis (all $p > .05$).

An index of the unsigned magnitudes of the final alignments was created by collapsing the absolute value data across motion direction and aperture configuration (see Figure 13). Planned comparisons of index of the unsigned magnitudes of the final alignments revealed that the overall magnitude of the final alignments increased as a function of velocity, with the 170-arcmin/sec condition yielding lower scores than the 340 arcmin/sec condition, $t(18) = 2.99$, $p < .01$, which in turn was lower than the 510 arcmin/sec condition, $t(18) = 3.95$, $p < .001$. A linear trend analysis of these data determined that the best fitting linear function using the least-squares estimation technique has a slope of .009 sec and a y intercept of -1.39 arcmin ($r^2 = .90$).

Discussion

The displays with many small apertures in Experiment 5A and 5B yielded more accurate judgments of rod alignment than the displays with only two apertures in Experiments 1, 3, and 4. This improvement in alignment accuracy occurred even though both the total amount of visible area through the apertures and the overall horizontal extent of the visible regions were the same across experiments. These findings indicate that repeated exposure of the top and bottom pieces of the rod through many small apertures provides more accurate position information to the observer than a single exposure of the top and bottom pieces only once each through two large apertures. This may be attributable to the visual system having several opportunities to sample the velocity of the rod when there are multiple apertures, and also to the smaller occluded regions between apertures over which the observer must maintain a representation of the rod behind the occluder. Another possibility is that each pair of apertures induced an illusion of misalignment in the opposite direction from the last, thus canceling each other out in the final percept.

To evaluate the effect of the number of apertures on the perceived alignment of the dynamically occluded rod, the data from Experiments 1A, 5A, and 5B were submitted to a 3 × 3 (Number of Apertures × Velocity) ANOVA, with the first factor as a between-subjects variable. This analysis revealed main effects of number of apertures, $F(2, 27) = 22.50$, $p < .0001$, and velocity, $F(2, 54) = 65.15$, $p < .0001$, indicating that the overall amount of illusion was indeed smaller for displays with more apertures, and that the amount of illusion increased as a function of the velocity of the rod in all three experiments (see Figure 13). Additionally, the analyses detected a significant interaction of velocity by number of apertures, $F(4, 54) = 16.29$, $p < .0001$, reflecting the fact that the slopes of the illusion function decreased as the number of apertures increased.

The finding that many small, closely spaced apertures allow for more accurate perception of dynamically occluded objects clarifies how the illusion may exist at the same time spatiotemporal object formation in ordinary vision works well. Arguably, multiaperture displays are ecologically more common than displays with just two large, misaligned apertures. In situations where an observer sees an object moving through intervening foliage, for example, there are typically many gaps in the leaves through which the object regions project to the eyes (see Figure 1). Also, the size of the occluded regions between these gaps in foliage is often small. Perception of dynamically occluded objects under such real-world circumstances appears to be quite accurate, as suggested by ordinary experience and by empirical studies (Palmer et al., 2006).

One important difference between single, relatively large apertures (Experiments 1–4) and multiple, relatively small apertures (Experiment 5) is that in the latter case, the visual system has several opportunities to resample the dynamically occluded object whereas in the former case there is just one opportunity for sampling each piece. When a dynamically occluded object is seen only once and then must be continuously represented behind an occluding surface, the representation of its occluded position is less accurate. Specifically, the data so far suggest that the occluded position is perceived as not being as far in the direction of motion as it should be, consistent with occlusion velocity being slower than real velocity.

How would newly acquired samples of a dynamically occluded object be integrated with representations of the same object regions already in the dynamic visual icon? It seems reasonable to assume that the visual system would favor physically specified position information over perceptually interpolated position information whenever possible. Therefore, we suggest that although the visual system is able to continuously represent the position of occluded fragments of objects, it abandons this strategy whenever a region reappears from behind an occluding surface. This notion is consistent with findings from Keane and Pylyshyn (2006), who showed that objects were not better tracked when they disappeared and then reappeared at their extrapolated trajectory in a multiple object tracking task, and Franconeri, Pylyshyn, and Scholl (2012) and Scholl and Nevarez (2002), who demonstrated that people do not seem to notice when an object reappears too soon from behind an occluding surface in a multiple object tracking task. Although observers clearly track the position of objects behind occluders in multiple object tracking (Flombaum et al., 2008), when occluded disks reappear, the particular trajectories that the objects "should have been" following while they were occluded or invisible are abandoned in favor of the new, more accurate position information (Franconeri et al., 2012). The fact that observers do not notice the early reappearance of an object from behind an occluder suggests that physically visible object position information "writes over" any other position information being generated in the DVI.

Experiment 6: Effects of Increased Occlusion Distance on the Illusion

Experiment 5 suggested that the occlusion distance over which position updating must be applied is a major determinant of the strength of the illusion. If this is true, then participants should experience a greater illusion with increased distance between apertures. Experiment 6 evaluated this hypothesis.

The design of this experiment was identical to that of Experiment 1A, with the exception that a 20-arcmin gap between apertures was added to the display (see *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/ACIllusionWideGap2.mov*). If the illusion is the result of underestimated occlusion velocity, then adding more space (and time) over which the misperception of position can apply should increase the magnitude of the illusion.

Method

Participants. Ten UCLA undergraduates participated in Experiment 6 in exchange for partial fulfillment of course requirements for an introductory psychology class. All participants reported normal or corrected-to-normal vision, provided informed consent, and were naïve to the purposes of the experiment.

Stimuli. The occluding surfaces were the same as Experiment 1A, except that the two apertures were 20 arcmin apart (as opposed to 0 arcmin apart as in previous experiments; *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/ACIllusionWideGap2.mov*). The total visible area through the apertures was 80 arcmin, as in Experiments 1, 3, 4, and 5.

Results

As Figure 14 shows, increasing the distance between apertures in the occluding surface resulted in larger illusion magnitudes, relative to Experiment 1A. Analyses of final alignments of the top and bottom pieces of the rod confirmed this observation.

The final reported alignments of the rod were submitted to a 3 × 2 × 2 (Velocity × Aperture Configuration × Motion Direction) within-subjects ANOVA. As in previous analyses of the illusion, the analysis revealed a main effect of aperture configuration, $F(1,$

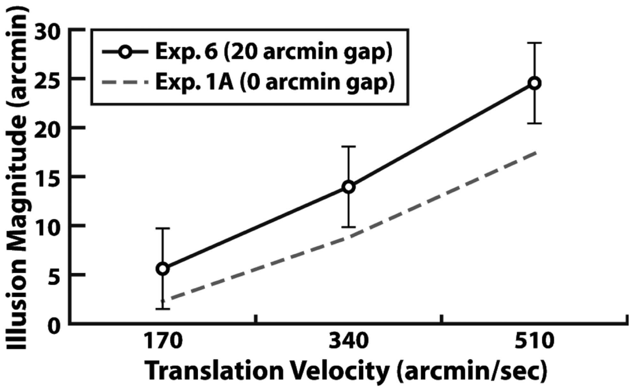

Figure 14. Overall magnitude of the illusion in Experiment 6 (with Experiment 1A data included for reference). Increasing the gap between apertures in the occluding surface caused an increase in the final alignments of the top piece of the rod relative to the bottom piece. Error bars are within-subjects confidence intervals (Loftus & Masson, 1994).

9) = 37.23, $p < .0005$, and a significant interaction of velocity by aperture configuration $F(2, 18) = 23.75$, $p < .0001$, but no other significant effects (all $p > .10$).

To test for a main effect of velocity, the absolute values of the final alignments were submitted to a 3 × 2 × 2 (Velocity × Aperture Configuration × Motion Direction) within-subjects ANOVA. This analysis detected a main effect of velocity, $F(2, 18) = 22.77$, $p < .0001$, indicating that the overall magnitude of the final alignments increased as a function of the velocity of the rod. There were no other significant main effects or interactions in this analysis (all $p > .10$).

Planned comparisons between the three velocities confirmed that the 510 arcmin/sec condition was faster than the 340 arcmin/sec condition, $t(18) = 10.50$, $p < .0001$, which was in turn faster than the 170 arcmin/sec condition, $t(18) = 8.30$, $p = .0006$ (see Figure 14). A linear trend analysis of these data using least squares estimation showed that the best-fitting linear function had a slope of .056 sec and a y-intercept of −4.22 arcmin ($r^2 = .99$).

Discussion

The overall magnitude of the illusion in this experiment was larger than Experiment 1A at all three translation velocities (see Figure 14). The increased illusion magnitude can be attributed to the greater distance between apertures in the occluding surface, as this was the only difference between the displays in the two experiments.

It is interesting to note that the participants' responses to the greater delay between appearances of the rod through the apertures (result from the larger gap size) was to increase the misalignment of the rod, which only delayed the presentation of the top and bottom portions even further. This is more evidence that participants are not aligning the rod based on a strategy of minimizing the timing of the appearance of the two pieces of the rod within the apertures. Larger temporal misalignments are perceived as smaller spatial misalignments, which is consistent with continued representation of the rod behind the occluding surface at a slower velocity.

Even with the increased spacing in Experiment 6, participants were still able to find a configuration of the rod's pieces that appeared aligned as it moved behind the occluder. Taking the temporal offset of both the rod and the apertures into consideration by calculating the length of time between the occlusion of the first edge of the rod in the leading aperture and the disocclusion (appearance) of the first edge of the rod in the trailing aperture, the largest delay in appearance between the two pieces of the rod that appeared aligned to observers after adjustment was 161.6 ms in the 340 arcmin/sec velocity condition. Thus, we can estimate that a perceptual representation of the occluded portion of the rod can be maintained for at least 160 ms.

In sum, Experiment 6 demonstrated that an increase in the distance between apertures causes a corresponding increase in the illusion. Participants' response to the increased delay between appearances of the rod was to increase the delay even further by choosing larger misalignments between the top and bottom portions of the rod. Analysis of the timing of the rod's appearances through the two apertures indicates that a perceptual representation of the rod behind the occluding surface can be maintained for at least 160 ms.

Experiment 7: A Direct Test of the Underestimated Occlusion Velocity Hypothesis

The evidence gathered so far suggests that the illusion is attributable to an underestimation of the rod's occlusion velocity. This inference has mostly come from ruling out alternative explanations. Experimental manipulations meant to improve the extraction of real velocity information within the leading aperture did not eliminate the illusion. However, manipulations that decreased the occluded area behind which the rod traveled before reappearing (Experiment 5A and 5B) did decrease the magnitude of the illusion. Nonetheless, the data gathered so far are only suggestive of underestimated occlusion velocity. For this hypothesis to be supported, a more direct demonstration is needed.

We developed a new set of displays to assess the slower occlusion velocity hypothesis directly. The displays used two apertures that were aligned along one side, with one aperture longer than the other (see Figure 15). An interesting feature of these occluding surfaces is that both the top and bottom portions of the rod are simultaneously visible through the apertures for a short time. Therefore, if participants can confine their perceptual processing to the time and place where both portions of the rod are visible, they should exhibit highly accurate performance.

The critical manipulation in this experiment is whether the disappearance of the rod pieces from the two apertures is staggered or simultaneous. When the top portion of the rod becomes occluded before the bottom portion (the "staggered exit" condition; Figure 15A and *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/AC_Exp7_Staggered.mov*), we predicted that it would appear misaligned in the direction of the shorter aperture because the successive disappearance of the rod from the windows allows for the hypothesized slower occlusion velocity to apply to the top piece of the rod. However, when the rod leaves both apertures simultaneously (the "simultaneous exit" condition; Figure 15B and *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/AC_Exp7_Simultaneous.mov*), we predicted that the rod would not appear misaligned because the occlusion velocity would

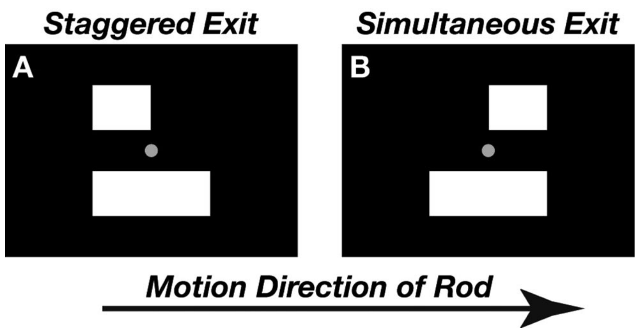

Figure 15. Aperture configurations used in Experiment 7. A) The staggered exit condition (see *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/AC_Exp7_Staggered.mov*). B) The simultaneous exit condition (see *http://webs.wichita.edu/depttools/depttoolsmemberfiles/AttentionLab/Video_Files/AC_Exp7_Simultaneous.mov*). Note that for leftward motion directions, the assignment of occluding surfaces to conditions would be reversed.

apply to both portions of the rod at once, and they would appear to slow down together.

Note that this paradigm provides a direct test of slowed occlusion velocity because representation of the updated positions of a hidden part is not strictly necessary to judge the relations of the rod parts in the staggered exit condition. Both parts in that condition appear simultaneously. Thus, it was completely possible that tests of perceived alignment in this new paradigm would show no systematic error. On the other hand, the representation of a partly occluded fragment may be obligatory, and its perceived position after occlusion may influence perception. If this latter possibility was correct, we expected that the paradigm would provide useful information about the dynamic representation of moving, occluded regions.

Method

Participants. Eleven UCLA undergraduates participated in the experiment in partial fulfillment of course requirements for an introductory psychology class. All participants reported normal or corrected-to-normal vision, provided informed consent, and were naïve to the purposes of the experiment. One participant was excluded from the analysis for failing to achieve the objective criterion of task compliance.

Stimuli. The occluding surface had two apertures, with the top aperture measuring 40 arcmin wide and the bottom aperture measuring 80 arcmin wide. The top and bottom apertures were always aligned on either their left edge (as in Figure 15A) or on their right edge (as in Figure 15B). The same 30-arcmin wide rod that was used in all previous experiments was also used here.

Design. This experiment used a 3 × 2 × 2 (Velocity × Motion Direction × Aperture Exit Order) design. The aperture exit order could either be staggered (the top portion of the rod became occluded before the bottom portion, as in Figure 15A) or simultaneous (the top and bottom portions of the rod became occluded at the same time, as in Figure 15B). Two trials for each of the 12 condition combinations were tested, with one trial having an initial alignment with the top piece of the rod at a random position to the left of the bottom piece, and the other with the top piece of the rod at a random position to the right of the bottom piece.

Results

Figure 16 shows the main results of Experiment 7. Little, if any, systematic misperception of alignment occurred when visible parts exited the apertures at the same time. However, strong illusion effects consistent with the occlusion velocity hypothesis appeared when the visible parts appeared together and aligned but then exited the apertures at different times.

These effects were confirmed by the analyses. Alignment data were submitted to a 3 × 2 × 2 (Velocity × Aperture Configuration × Aperture Exit Order) within-subjects ANOVA. This analysis revealed a main effect of Aperture Configuration, $F(1, 9) = 60.92$, $p < .0001$, with the leftward aperture configurations yielding slightly leftward alignments and the rightward aperture configurations yielding slightly rightward alignments. There were also interactions of velocity by aperture configuration, $F(2, 18) = 53.23$, $p < .0001$, aperture configuration by aperture exit order, $F(1, 9) = 28.86$, $p < .0005$, and velocity by aperture configuration by aperture exit order $F(2, 18) = 14.69$, $p < .0005$. These interactions appear to be attributable to the influence of the aperture configuration on the appearance of the rod, combined with increasing overall magnitudes of final alignments as a function of velocity and larger final alignments for rods that had a staggered aperture exit order than for rods that had a simultaneous aperture exit order. The test for a main effect of velocity was not significant ($p = .29$). The data for the final alignments of the rod are plotted in Figure 16.

A 3 × 2 × 2 (Velocity × Aperture Configuration × Aperture Exit Order) within-subjects ANOVA was conducted on the absolute value data and detected main effects of velocity, $F(2, 18) = 86.20$, $p < .0001$, and aperture exit order, $F(1, 9) = 22.55$, $p < .005$, with simultaneous exits producing lower illusion magnitudes than staggered exits. Finally, there was a significant velocity by

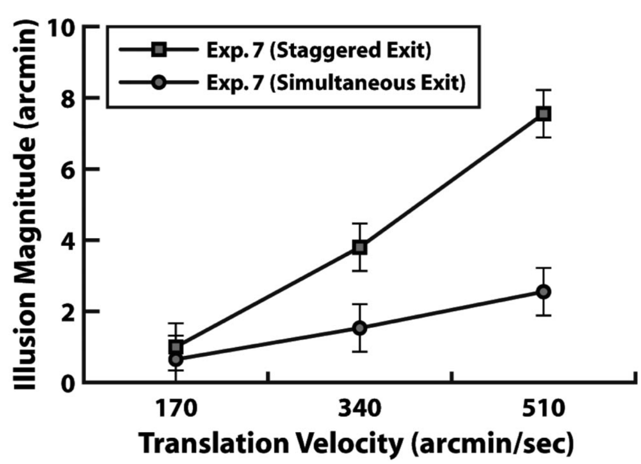

Figure 16. Illusion magnitudes in Experiment 7, split by whether both pieces of the rod became occluded simultaneously or in succession. Participants showed the illusion for the staggered aperture exit condition, but not for the simultaneous aperture exit condition. Error bars are within-subjects confidence intervals (Loftus & Masson, 1994).

aperture exit order interaction, $F(2, 18) = 9.25$, $p < .005$, with the overall magnitudes of the final alignments increasing as a function of velocity for the staggered, but not simultaneous, aperture exit order condition (see Figure 16). There were no other significant main effects or interactions in the analysis (all $p > .10$).

Discussion

In this experiment, a significant and robust illusion was observed when the top piece of the rod became occluded before the bottom piece (the staggered exit condition), but not when both pieces became occluded simultaneously (the simultaneous exit condition). These two very different results occurred with the same aperture configurations and velocities, and depended on which direction the rod traveled behind the occluder. In this case, the cause of the illusion was the occlusion of one portion of the rod before the other. Furthermore, the illusion was observed in the staggered exit condition even though both the top and bottom portions of the rod were visible at the same time during part of the display sequence.

In connection to the flash-lag illusion, Maus and Nijhawan (2009) examined the perceptual alignment of moving objects when one object disappeared and reported a similar illusion to the one described here. In their displays, two aligned rods translated laterally across the visual field and one rod abruptly disappeared while the other continued to move. Participants reported seeing the two rods as misaligned at the time of the disappearance. Maus and Nijhawan explain their findings by appealing to extrapolated motion mechanisms for continuously moving objects (e.g., Nijhawan, 2002). They suggested that the abrupt disappearance of the rod caused a "stop" signal to be transmitted to the visual system, thus canceling the perception of that object at its extrapolated position, but not canceling the perception of the continuously moving object being at an extrapolated position. Consequently, the two rods appeared misaligned when they were, in fact, aligned.

Referring to earlier reports of the aperture capture illusion (Palmer & Kellman, 2001, 2002, 2003), Maus and Nijhawan (2009) suggest that their theory may be related to the aperture capture illusion, though they did not provide specifics. The displays used in this experiment certainly do bear a similarity to the displays used by Maus & Nijhawan in that two pieces of a stimulus are seen moving together and then one disappears (becomes occluded in the present experiment), resulting in the perception of misalignment between the two pieces. However, it is not clear how to relate Maus and Nijhawan's (2009) explanation to the experiments already reported in which the two pieces of the rod were not visible simultaneously but were instead seen through two nonoverlapping apertures (i.e., Experiments 1–4, and 6). Furthermore, the motion extrapolation account of the illusion proposed by Maus & Nijhawan depends on the transmission of a "stop" signal to the visual system, but with disappearance of an object part through gradual occlusion, there is nothing that would qualify as a "stop" signal in the current displays. Moreover, if a "stop" signal causes the last position of the rod that disappeared to be seen accurately (rather than in an extrapolated position), then why is the illusion able to be nulled?

Although phenomena involving abrupt disappearance studied by Maus and Nijhawan (2009) cannot directly explain the experimental results here, we believe there is an important connection. Maus & Nijhawan's explanation of their phenomena involve two components: a motion extrapolation mechanism and a mechanism that pegs an object's position when it abruptly disappears (or changes direction). The first component—motion extrapolation—is the connection among these visual phenomena. In their displays, according to their theory, motion extrapolation is going on even while an object is fully visible. Our experiments indicate that extrapolation goes on even when the object becomes gradually occluded; that is, there is a storage mechanism that preserves object information and extrapolates its position. The Maus & Nijhawan work illustrates an interesting feature of a DVI mechanism, namely that a representation of a persisting, spatially shifting object is actually operating even while the object is fully visible. In contrast, as Maus and Nijhawan (2009) suggest, their illusion does not have much to do with the fact that extrapolated motion under occlusion is slower than the prior velocity signals would predict. Their displays did not contain occlusion, and in any case, their illusion would still occur if any motion extrapolation (even if slowed) gets compared with a transient signal produced by the abrupt disappearance of another object or object part.

In the discussion of Experiment 1, we suggested that the illusion might be the result of one of three errors by the visual system: a) inaccurate registration of the motion of the first appearing object piece within the leading aperture (i.e., slower real velocity), b) underestimation of the first disappearing object piece's motion after occlusion (i.e., slower occlusion velocity), or c) inaccurate registration of the later appearing object piece in the later aperture. Experiments 2–4 ruled out the first possibility, whereas the results of Experiment 5 were consistent with the second and third possibility. Experiment 7 can rule out the third possibility: It is not possible that the illusion was caused by inaccurate perception of the later appearing piece of the object through an aperture because there *was no* later appearing object piece. The bottom portion of the rod was fully visible during the motion sequence, yet when the top piece of the rod was occluded before the bottom piece, a robust illusion occurred. Only the second of our three possible sources of error leading to the illusion—underestimated occlusion velocity—remains viable.

The results of this experiment are consistent with the hypotheses that visible surfaces that become occluded continue to be represented, that the positions of moving, occluded surfaces are updated in the representation, and that this aspect of the representation, which we label for convenience "occlusion velocity" is slower than real velocity. Slower occlusion velocity means that interpolated positions of occluded fragments lag behind what would be expected if their true velocity were veridically extracted and used for position updating. When accurately perceived, moving visible regions are perceptually combined with slower-moving occluded regions, the illusion results.

Our findings indicate that the illusion occurs when an occluded region of a moving object must be perceptually integrated with a visible region. Experiment 7 makes an important additional point that was only implicit in the earlier experiments. It appears that the integration of visible and hidden regions of an object is obligatory and their influence on perception unavoidable. Subjects performed a task that could have been done without any contribution from the hidden region, yet the data showed that they were influenced by it. Combining currently available information with spatiotemporally extrapolated visual representations of previously obtained information may be a basic characteristic of visual perception.

Additionally, the results of this experiment provide further evidence that the illusion is not caused by an inability to extract a proper motion signal within a short aperture, because the same aperture sizes caused the illusion in the staggered exit condition but not in the simultaneous exit condition: the only difference between these two conditions was the order in which the top piece of the rod became occluded relative to the bottom piece.

To our knowledge, the only remaining explanation for the illusion that is consistent with these results is that occluded regions of a moving object remain perceptually available for a short time after their disappearance, but under these circumstances, their extrapolated movement is slower velocity than visible regions. Consequently, in Experiment 7, when the top portion of the rod becomes occluded before the bottom portion, its position behind the occluder is inaccurately updated, causing it to be perceived as misaligned in the direction of the shorter aperture. This account of the illusion explains why adjusting the top piece of the rod to be misaligned in the direction opposite to the offset of the apertures allows the pieces to appear aligned.

General Discussion

The experiments reported here investigated a powerful perceptual illusion in which a moving, aligned rod that is seen through two (or more) misaligned apertures appears misaligned in the direction of the apertures (Experiment 1A and 1B). This illusion appears to be caused by an underestimation of the velocity of the rod after occlusion that is not attributable to an inaccurate estimation of velocity within the leading aperture of the displays (Experiments 2, 3, and 4). Consistent with this notion, if the spacing between apertures is increased, the magnitude of the illusion also increases (Experiment 6). Additionally, the underestimation of velocity occurs only after part of the object becomes occluded (Experiment 7), but can be alleviated through repeated exposure of the rod through multiple apertures which allows for the true velocity of the rod to be resampled (Experiment 5).

Eye movements may play a role in the illusion, though tracking the object as it moved behind the occluder did not eliminate the illusion, contrary to the retinal painting hypothesis (Experiment 4). However, in the one condition involving smooth pursuit eye movements, we observed a different pattern of data than the other experiments in which participants were instructed to maintain steady fixation. In Experiment 4, stimulus conditions in which the rod was seen in the bottom aperture first yielded lower illusion magnitudes than those in which the rod was seen in the top aperture first, regardless of motion direction. The interaction of eye movements and dynamically occluded object perception warrants further study.

These results can be fit with a quantitative model that allows one to estimate the relative speed of occlusion velocity versus real velocity (Palmer, 2003). However such a model involves additional data relating to the time course of the aperture capture illusion, which is beyond the scope of the current article. Data about the time course of the aperture capture illusion and a quantitative model of occlusion velocity will be described in a forthcoming study.

Implications for the Dynamic Visual Icon

To explain perception of dynamically occluded objects, Palmer, Kellman, and Shipley (2006) argued that the particular spatial and temporal relationships involved required a special kind of representation. It is known that previously presented information persists for short intervals, as reflected in the notion of iconic storage, proposed by Neisser (1967) to characterize persistence effects discovered by Sperling (1960; see also Coltheart, 1980). Palmer et al. (2006) proposed that the persistence and position updating of dynamically occluded fragments not currently available in the stimulus requires a different kind of representation, which they called the *dynamic visual icon* (DVI).

The reasons for postulating a different representation come from the properties that have been described previously for persisting visual representations. Some have argued that iconic visual storage depends on persistence of activity in photoreceptors (e.g., Sakitt, 1975, 1976). Whether or not the persistence is mediated by photoreceptors, this characterization is similar to perhaps the most common understanding of the visual icon. Julesz (1971) expressed the idea by saying, “The ‘short-term visual memory’ of Sperling (1960) is a detailed texture memory, but fades out in .1 sec like the afterglow of a CRT and is merely an afterimage” (Julesz, 1971, p. 103). Any storage mechanism of this type would be unable to incorporate spatial transformations as appear to occur in a DVI representation of previously available object parts.

Based on experimental evidence that the visual icon was relatively unaffected by variables that should influence photoreceptor persistence, Adelson and Jonides (1980) argued that the icon is postretinal. Typical suggestions about a postretinal icon, however, also involve preserving the spatial pattern as it appeared, perhaps in terms of detectors for edges (Adelson & Jonides, 1980). Coltheart (1980) reviewed a great deal of research on visual icon representations and concluded that there are actually three forms in which a visual stimulus may be considered to exist after it is physically extinguished: neural persistence, visible persistence, and informational persistence. He argued that these can be separated based on the effects of different variables on different persistence phenomena. Coltheart suggested that informational persistence is defined by the partial report procedures of Sperling (1960) and Averbach and Coriell (1961). This store is not based on photoreceptor persistence but is “a decaying store of visual information” (Coltheart, 1980, p. 188) lasting on the order of 300 ms. One of its key properties, as in the foundational experiments that established this type of representation, is that information can be accessed by the location at which it appeared in a prior display.

The effects in our study seem to require a representation that combines some properties of the “informational persistence” and “visible persistence” notions discussed by Coltheart (1980). He made a strong case that partial report results, such as those of Sperling (1960), require an iconic memory representation corresponding to informational persistence, because these effects do not show two characteristics of visible persistence effects. One is the inverse duration effect, such that visible persistence decreases for longer stimulus presentations, and the other is the inverse luminance effect, such that visible persistence effects are shorter following higher luminance displays. Our informal observations suggest that it is unlikely that the aperture capture illusion would show these inverse effects, although we have not carried out formal studies. On the other hand, Coltheart and others identify a "visible persistence" mechanism with the phenomenal experience of seeing. The current results and related work suggest that recently occluded object fragments are part of phenomenal experience in an amodal sense of seeing (Kanizsa, 1979; Michotte, Thinès, & Crabbé, 1964). Participants in our studies experience a vivid impression of misalignment of object parts despite the fact that one part is not simultaneously present in the display with the other. Such results suggest that in an important sense, the scope of perceptual reality—what is being *seen*—is not limited to information momentarily available.

It is clear that further investigation will be useful in clarifying the properties of postretinal storage mechanisms in vision. Following Coltheart (1980), it is especially useful to distinguish possible representations by their properties and by the variables that affect them. The most important impetus for postulating a DVI representation is that virtually all of the previously identified storage mechanisms encode the spatial position of information as it appeared in the display. Referring to partial report paradigms, for example, Coltheart (1980) notes that the spatial locations items occupied in the display must be part of iconic memory, because one can access them through a spatial cue after stimulus offset. Such an encoding of fixed spatial positions is deeply inherent in notions of visible persistence, which have often been argued to involve retinotopic information.

The DVI as described by Palmer, Kellman, and Shipley (2006) and in relation to the current findings must be a representation that not only preserves information for a short duration but transforms it spatially. The duration of visible persistence was estimated by Coltheart (1980) to be 300 ms or less based on the partial report superiority effect in Sperling's (1960) work. Experiment 6 of the present work provided an estimate of DVI duration to be about 160 ms. Another estimate of the duration of the DVI comes from Experiment 2 of Palmer et al. (2006). In those displays, the dynamically occluded object was physically visible through apertures in the occluding surface for only 80 ms, but participant performance indicated that the effective exposure duration was somewhere between 250 and 350 ms (Palmer et al., 2006, Table 1). Subtracting the actual exposure duration of 80 ms, we can reason that the duration of the DVI representation was in the range of 170–270 ms. Thus, estimates of the timecourse of the DVI representation are consistent with the timecourse of the visible persistence representation, but the properties of these representations appear to differ. The DVI has some properties of informational persistence and the phenomenal character of visible persistence, but unlike typical descriptions of these representational formats, it involves an active extrapolation, and experience, of position of moving fragments over time. It is also possible that future work will resolve some apparent discrepancies and indicate that the DVI we have labeled and characterized is the same representation that underlies partial report in the Sperling procedure or other visual persistence phenomena that, in static presentations, preserve spatial location information. If so, characterizing a DVI representation may prove useful in attaining a more unified theoretical account and highlighting the transformational properties of a crucial visual storage mechanism.

Much of the value of a theoretical construct, such as the DVI, lies in connecting data from seemingly different paradigms and accounting for these data using a common underlying mechanism. The dynamic occlusion paradigm used in earlier work and the illusion studies here involved different kinds of stimuli and methods. Both, however, converge in implicating a dynamic form of iconic visual storage. In the dynamic occlusion work, a representation for preserving and spatially updating object information was invoked to explain findings of superior discrimination performance for spatiotemporally relatable fragments (Palmer, Kellman, & Shipley, 2006). The current work reveals that spatial updating in such a representation is nonveridical and varies with a number of stimulus variables, as measured by nulling of an illusion of misalignment between object parts that are not simultaneously present. As indicated earlier, the DVI may also account for the motion extrapolation component in the explanation of flash-lag phenomena suggested by Maus and Nijhawan (2009).

Implications for Spatiotemporal Object Formation

Visible regions of a dynamically occluded object transmit contour, surface, depth, and motion information to the observer, allowing him or her to use this information to constrain and construct later representations of object shape. When parts of a visible scene become occluded because of object or observer motion, information from previously visible regions continues to be phenomenally available for a short time in the DVI representation. Thus, both visible and occluded shape information can be combined to form visual units whose perceived shape may be more than is physically specified at any given time or place. In particular, the bounding contours of objects residing in these representations can be interpolated across gaps in space via the Kellman and Shipley (1991) relatability process, and over time using the Palmer et al. (2006) persistence and position-updating components of the DVI.

According to the framework of spatiotemporal relatability, the illusion reported here can be conceptualized as an error in the position-updating process: occluded regions of objects remain perceptually available for a short time after their disappearance, but their perceived velocity behind the occluding surface is underestimated. Consequently, when visual unit formation proceeds on visible and occluded shape representations, the position of fragments within the DVI representation are mistakenly perceived as being nearer to their last visible position than they really are. In the case of a rod moving behind an occluding surface with two misaligned apertures, the rod is always perceived as misaligned in the same direction as the offset of the apertures. This illusion occurs whenever one portion of an object becomes occluded before another, and when there are no other apertures available to resample the object's true velocity.

However, in situations in which a dynamically occluded object is perceived through an occluding surface with many apertures (e.g., through foliage), repeated exposures of the object through the apertures minimize the illusion so that it has no discernible perceptual effect. It is only in sparse displays with just a few large, misaligned apertures that the illusion reliably occurs. Occluding surfaces with multiple apertures, such as those in Experiment 5B, allow the visual system to accurately represent the position of hidden portions of dynamically occluded objects. Informal observation indicates that the ecological conditions leading to dynamic occlusion displays typically have many closely spaced apertures, suggesting that spatiotemporal relatability is well-suited to the demands of everyday perception.

There is some indication (from Experiment 5) that the visual system prioritizes physically visible object information over information in the DVI. This is may be because though we are able to (and routinely do) track the positions of objects behind occluding surfaces, we also prioritize information from objects when they are visible over when they are occluded. There are some results in the multiple object tracking literature suggesting this kind of looseness with trajectories of occluded objects (e.g., Franconeri et al., 2012; Keane & Pylyshyn, 2006; Scholl & Nevarez, 2002) that seem at odds with our apparently good ability to perceive dynamically occluded objects which requires good representation of occluded trajectories (Palmer et al., 2006; Experiment 5). Perhaps once previously occluded object pieces reappear, all bets are off, and the newly observed position "writes over" any inaccurate (but better than nothing) representations that were being used previously.

Why Is Occlusion Velocity Underestimated?

It seems strange that the perceptual processes that enable extremely accurate perception of dynamically occluded objects under normal circumstances would be subject to such a strong illusion under the minimal circumstances investigated here. However, as is the case with most illusions, the perceptual processes that lead to the illusion may, in fact, be optimized for a different set of circumstances than the ones encountered in the laboratory. We can think of two hypotheses that would account for the underestimation of occlusion velocity by appealing to perceptual strategies that may work well under normal conditions.

The first hypothesis is that underestimated occlusion velocity results from the continued activity of motion-processing mechanisms after occlusion (similar to the notion of neural persistence described by Coltheart, 1980). According to this explanation, neurons that encode the velocity of a stimulus do not discontinue their activity immediately after an object becomes occluded. Rather, they maintain a slightly elevated firing rate for a short time so that if the object comes back into view, their response to its velocity will already be primed. Additionally, given that accurate perception of the global motion of a dynamically occluded object necessitates the interaction of motion processing mechanisms between apertures to overcome the aperture problem (Liden & Pack, 1999; Lorenceau & Shiffrar, 1992, 1999), the priming of motion units that specify particular directions and speeds of motion in the recent past may be important for constraining new motion information and overcoming the local ambiguity of motion information within apertures (Watamaniuk & McKee, 1995). Thus, the underestimation of occlusion velocity occurs because motion units maintain a heightened firing rate due to their priming, but are not firing as fast as they would to a visible stimulus.

The second possible explanation for the underestimation of occlusion velocity is that motion processing in the visual system relies on Bayesian estimation processes that take into account the evidence for a particular motion stimulus at a given time, as well as prior information about motion in the environment (e.g., Ascher & Grzywacz, 2000; Weiss, Simoncelli, & Adelson, 2002). Given that the majority of objects in the environment are stationary, such a model would assume that slower velocities are more common than faster velocities and that perceptual impressions of velocity contain noise (Weiss et al., 2002). With these two assumptions, any degradations in visual information about object velocity would lead to greater influence of the slow velocity priors, resulting in perceptions of velocities as being slower than they really are. Weiss et al. (2002) used exactly such an approach to model velocity perception for low-contrast objects, showing that this model explains why people perceive low-contrast objects to be moving more slowly than they really are (e.g., Stone, Watson, & Mulligan, 1990). When an object moves behind an occluding surface, the quality of information available about its velocity is degraded, so such a model might also explain slower perceived velocities in the dynamic occlusion velocities reported here.

These two explanations are not incompatible; rather, they focus on different levels (Marr, 1982). The first explanation relates to properties of the mechanisms that carry out motion processing. The second starts with ecological (or in Marr's often misinterpreted term) computational considerations (i.e., what kinds of motions are out in the world to be perceived?). For effectively functioning perceptual systems, we would expect substantial coherence between the facts at these levels. The explanations may be considered different if the motion mechanisms account is interpreted as relating to unavoidable or incidental features of the kinds of motion detectors we possess, whereas the Bayesian account of weighting motion priors more heavily as a region becomes occluded provide a more purposeful explanation of why motion mechanisms may behave a certain way. Of course, the motion priors explanation could be implemented in other ways.

A final comment on these potential explanations is that the motion priors idea could be argued to be somewhat loosely fitting here, or at least other ideas about priors are possible. A prior based on the overall distribution of stationary and moving objects in the world does not seem to be the most relevant reference class for moving objects that pass behind occluders. Such objects have already been registered as being in motion, so factoring in the existence of many other stationary objects may or may not be considered a good idea. A more specific description for the prior would be the distribution of object velocity for short durations after occlusion, given previous velocity information from immediately preceding unoccluded viewing. Intuitively, one might think the best estimate of the object's motion for a short duration after occlusion would be its last observed velocity. These are issues that are ripe for consideration in further development of potential Bayesian explanations of the data indicating that occlusion velocity in simple situations is slower than real velocity.

Regardless of the cause of the underestimated occlusion velocity, several points seem clear. First, observers naturally and habitually continue to represent the shapes and motion trajectories of moving objects that become occluded. The relevant representation, which we have called the DVI, allows for currently visible and currently occluded regions of moving objects to be united together into perceptual wholes, enabling observers to collect and integrate shape information over time. Second, this process of dynamically occluded object perception, which works well in most real-world situations, can produce a powerful illusion when one piece of a moving object becomes occluded before another piece and is not seen again. In such situations, position updating of the occluded piece is flawed due to underestimated occlusion velocity, leading to the piece seeming to be closer to the aperture in which it was last seen through than it should be.

In closing, we note that the present research suggests a broad and deeply interesting point about perception in general, one that may be evident in other work but that is particularly salient here. Our conscious experience of the world is extended in time and depends in part on mental representations, not simply the stimulus information of the moment. In the aperture capture illusion, the perceived relation of object parts is misperceived. In more favorable dynamic occlusion situations, the dynamic visual icon allows accurate perception of coherent objects from fragments that, due to occlusion and motion, are discontinuous in both space and time. Both kinds of results arise from the visual system combining currently incoming information about some object parts with information about other parts that are not currently available in the stimulus, but are preserved and transformed in perceptual representations. Our experience of the world comes from processes that combine the currently seen with the previously seen to connect them across gaps in both space and time.

References

Adelson, E. H., & Jonides, J. (1980). The psychophysics of iconic storage. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 486–493. doi:10.1037/0096-1523.6.3.486

Ascher, D., & Grzywacz, N. M. (2000). A Bayesian model for the measurement of visual velocity. *Vision Research, 40,* 3427–3434. doi:10.1016/S0042-6989(00)00176-0

Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal, 40,* 309–328. doi:10.1002/j.1538-7305.1961.tb03987.x

Banks, W. P., & Barber, G. (1977). Color information in iconic memory. *Psychological Review, 84,* 536–546. doi:10.1037/0033-295X.84.6.536

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 109–127. doi:10.1037/0096-1523.21.1.109

Clark, S. E. (1969). Retrieval of color information from preperceptual memory. *Journal of Experimental Psychology, 82,* 263–266. doi:10.1037/h0028135

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics, 27,* 183–228. doi:10.3758/BF03204258

Demkiw, P., & Michaels, C. F. (1976). Motion information in iconic memory. *Acta Psychologica, 40,* 257–264. doi:10.1016/0001-6918(76)90029-9

Di Lollo, V., Hogben, J. H., & Dixon, P. (1994). Temporal integration and segregation of brief visual stimuli: Patterns of correlation in time. *Perception & Psychophysics, 55,* 373–386. doi:10.3758/BF03205295

Drew, T., & Vogel, E. K. (2006). *An electrophysiological measure of multiple object tracking: Correlations with visual working memory*. Paper presented at the Cognitive Science Meeting for Interdisciplinary Learning (CSAIL), Hood River, OR.

Drew, T., & Vogel, E. K. (2007). *An electrophysiological measure of tracking through occlusion*. Paper presented at the Vision Sciences Society Annual Meeting, Sarasota, FL.

Flombaum, J. I., Scholl, B. J., & Pylyshyn, Z. W. (2008). Attentional resources in visual tracking through occlusion: The high-beams effect. *Cognition, 107,* 904–931. doi:10.1016/j.cognition.2007.12.015

Franconeri, S. L., Pylyshyn, Z. W., & Scholl, B. J. (2012). A simple proximity heuristic allows tracking of multiple objects through occlusion. *Attention, Perception, & Psychophysics, 74,* 691–702.

Gibson, J. J., Kaplan, G. A., Reynolds, H. N., & Wheeler, K. (1969). The change from visible to invisible. *Perception & Psychophysics, 5,* 113–116. doi:10.3758/BF03210533

Gold, J. M., & Shubel, E. (2006). The spatiotemporal properties of visual completion measured by response classification. *Journal of Vision, 6,* 5. doi:10.1167/6.4.5

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383,* 334–337. doi:10.1038/383334a0

Hunt, S. M. J. (1994). MacProbe: A Macintosh-based experimenter's workstation for the cognitive sciences. *Behavior Research Methods, Instruments & Computers, 26,* 345–351. doi:10.3758/BF03204643

Julesz, B. (1971). *Foundations of cyclopean perception*. Oxford, England: University of Chicago Press.

Kanizsa, G. (1979). *Organization in vision: Essays on Gestalt perception*. New York, NY: Praeger.

Keane, B. P., Lu, H., & Kellman, P. J. (2007). Classification images reveal spatiotemporal contour interpolation. *Vision Research, 47,* 3460–3475. doi:10.1016/j.visres.2007.10.003

Keane, B. P., & Pylyshyn, Z. W. (2006). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function. *Cognitive Psychology, 52,* 346–368. doi:10.1016/j.cogpsych.2005.12.001

Kellman, P. J., Garrigan, P., & Shipley, T. F. (2005). Object interpolation in three dimensions. *Psychological Review, 112,* 586. doi:10.1037/0033-295X.112.3.586

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology, 23,* 141–221. doi:10.1016/0010-0285(91)90009-D

Lidén, L. H., & Pack, C. C. (1999). The role of terminators and occlusion cues in motion integration and segmentation: A neural network model. *Vision Research, 39,* 3301–3320. doi:10.1016/S0042-6989(99)00055-3

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1,* 476–490. doi:10.3758/BF03210951

Lorenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research, 32,* 263–273. doi:10.1016/0042-6989(92)90137-8

Lorenceau, J., & Shiffrar, M. (1999). The linkage of visual motion signals. *Visual Cognition, 6,* 431–460. doi:10.1080/135062899395055

Marr, D. (1982). *Vision: A computational approach,* San Francisco, CA: Freeman & Co.

Maunsell, J. H. R., & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience, 10,* 363–401. doi:10.1146/annurev.ne.10.030187.002051

Maunsell, J. H. R., & van Essen, D. C. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *The Journal of Neuroscience, 3,* 2563–2586.

Maus, G., & Nijhawan, R. (2009). Going, going, gone: Localizing abrupt offsets of moving objects. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 611–626. doi:10.1037/a0012317

Michotte, A., Thinès, G., & Crabbé, G. (1964). *Les complements amodaux des structures perceptive*. Institut de psychologie de l'Université de Louvain.

Neisser, U. (1967). *Cognitive psychology*. New York, NY: Appleton, Century, Crofts.

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends in Cognitive Sciences, 6,* 387–393. doi:10.1016/S1364-6613(02)01963-0

Palmer, E. M. (2003). *Spatiotemporal relatability in the perception of dynamically occluded objects*. University of California, Los Angeles.

Palmer, E. M., & Kellman, P. J. (2001). The aperture capture effect: Misperceived forms in dynamic occlusion displays [Abstract]. *Journal of Vision, 1,* 381. doi:10.1167/1.3.381

Palmer, E. M., & Kellman, P. J. (2002). Underestimation of velocity after occlusion causes the aperture-capture illusion [Abstract]. *Journal of Vision, 2,* 477. doi:10.1167/2.7.477

Palmer, E. M., & Kellman, P. J. (2003). (Mis)Perception of motion and form after occlusion: Anorthoscopic perception revisited [Abstract]. *Journal of Vision, 3,* 251. doi:10.1167/3.9.251

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. doi:10.1037/0096-3445.135.4.513

Previc, F. H. (1990). *Functional specialization in the lower and upper visual fields in humans: Its ecological origins and neurophysiological implications* (No. USAFSAM-JA-88–43). School of Aerospace Medicine, Brooks Air Force Base, Texas.

Sakitt, B. (1975). Locus of short-term visual storage. *Science, 190,* 1318–1319. doi:10.1126/science.1198117

Sakitt, B. (1976). Iconic memory. *Psychological Review, 83,* 257. doi:10.1037/0033-295X.83.4.257

Scholl, B. J., & Nevarez, H. G. (2002). *Why so slow? The role of speed discontinuities in maintaining object persistence through occlusion*. Paper presented at the Annual Meeting of the Psychonomic Society, Kansas City, KS.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology, 38,* 259–290. doi:10.1006/cogp.1998.0698

Shipley, T. F., & Cunningham, D. W. (2001). Perception of occluding and occluded objects over time: Spatiotemporal segmentation and unit formation. In T. F. Shipley & P. J. Kellman (Eds.), *From fragments to objects: Segmentation and grouping in vision* (pp. 557–585). Amsterdam, The Netherlands: Elsevier. doi:10.1016/S0166-4115(01)80038-8

Shuwairi, S. M., Curtis, C. E., & Johnson, S. P. (2007). Neural substrates of dynamic object occlusion. *Journal of Cognitive Neuroscience, 19,* 1275–1285. doi:10.1162/jocn.2007.19.8.1275

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74,* 1. doi:10.1037/h0093759

Stone, L. S., Watson, A. B., & Mulligan, J. B. (1990). Effect of contrast on the perceived direction of a moving plaid. *Vision Research, 30,* 1049–1067. doi:10.1016/0042-6989(90)90114-Z

Turvey, M. T., & Kravetz, S. (1970). Retrieval from iconic memory with shape as the selection criterion. *Perception & Psychophysics, 8,* 171–172. doi:10.3758/BF03210198

van de Grind, W. A., Koenderink, J. J., & Van Doorn, A. J. (1986). The distribution of human motion detector properties in the monocular visual field. *Vision Research, 26,* 797–810. doi:10.1016/0042-6989(86)90095-7

Von Wright, J. M. (1968). Selection in visual immediate memory. *The Quarterly Journal of Experimental Psychology, 20,* 62–68. doi:10.1080/14640746808400128

Wallach, H. (1959). The perception of motion. *Scientific American, 201,* 56–60. doi:10.1038/scientificamerican0759-56

Watamaniuk, S. N. J., & McKee, S. P. (1995). Seeing motion behind occluders. *Nature, 377,* 729–730. doi:10.1038/377729a0

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience, 5,* 598–604. doi:10.1038/nn0602-858

Yi, D. J., Turk-Browne, N. B., Flombaum, J. I., Kim, M. S., Scholl, B. J., & Chun, M. M. (2008). Spatiotemporal object continuity in human ventral visual cortex. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 105,* 8840–8845. doi:10.1073/pnas.0802525105

Received December 17, 2012
Revision received October 1, 2013
Accepted October 2, 2013 ■