Multiple object tracking is based on scene, not retinal, coordinates

Geniva Liu, Erin Austen, Kelly Booth,
Brian Fisher, Mark Rempel, James T. Enns
University of British Columbia

Address all correspondence to:
J.T. Enns
Department of Psychology
University of British Columbia
2136 West Mall
Vancouver, BC, Canada V6T 1Z4
jenns@psych.ubc.ca
604-822-6634 (Voice)
604-822-6923 (Fax)

Abstract

This study tested whether multiple-object tracking—the ability to visually index objects based on their spatial-temporal history—is scene-based or image-based. Initial experiments showed equivalent tracking accuracy for objects in 2D and 3D motion. Subsequent experiments manipulated the speed of object motion independently from the motion speed of the scene as a whole. Results showed that tracking accuracy was influenced by object speed but not scene speed. This held whether the scene underwent translation, zoom, rotation, or even a combination of all three motions, which we termed the 'wild ride.' A final series of experiments interfered with observers' ability to see a coherent 3D scene: observers tracked objects moving at different speeds (multiple speeds reduces scene coherence) or tracked objects moving at identical retinal speeds but in distorted 3D space. These manipulations reduced tracking accuracy, confirming that tracking is accomplished using a scene-based (allocentric) frame of reference.

Multiple object tracking is based on scene, not retinal, coordinates

An important task of the visual system is to keep track of objects as they move through space. Whether the observer is an air traffic controller tracking airplanes on a radar screen, or an athlete tracking team members and opposing players on a field, there is a need to maintain a visual index for objects that are changing in their spatial location over time. It has been shown that human observers can track up to four or five randomly moving objects with fairly good accuracy (Pylyshyn & Storm, 1998; Scholl, 2001; Yantis, 1992). Tracking performance is high even when the tracked objects are identical to untracked objects in all respects other than their motion paths, pointing to a tracking ability that is based solely on spatiotemporal history of the objects.

The ability to track multiple objects has been used to investigate the role of perceptual organization in tracking (Yantis, 1992), whether attention can be deployed in depth as readily as it can be deployed across vertical and horizontal space (Viswanathan & Mingolla, 1998) and the nature of visual object representations (Scholl & Pylyshyn, 1999). However, there is a fundamental question concerning tracking that is not yet well understood. In what frame of reference are objects being tracked? Are the visual indices or 'pointers' that observers use to track objects pointing to locations on a retinotopic map (a coordinate system with respect to the retina), or are they pointing to locations in an allocentric map (a coordinate system with respect to the scene)? Although there is not yet any research on this question, there are good reasons to suspect that either one of these options may be correct.

One reason to suspect a retinal frame of reference is because the entire human visual system is organized at the physiological level in a retinotopic fashion. When neurons in one visual area of the brain (e.g., area V1) communicate with neurons in other areas (e.g., V5 or temporal lobe), they tend to maintain a strict spatial correspondence. This means that neurons in different visual areas that are responding to the same object in a given visual field location are automatically linked, simply by virtue of their common reference to the same visual field location (Van Essen et al, 2001). A mechanism that was designed to 'keep a finger' on an object as it moved over time would simply have to track the changing neural activity in one of these retinotopically organized visual areas.

Yet there are equally compelling reasons to suspect that tracking is accomplished using a reference frame tied to locations in the world rather than in the eye. One such reason comes from an examination of eye movements. Saccades, those high-speed ballistic eye movements that are made from one location to another, are referenced to stationary environmental landmarks rather than to specific retinal coordinates. This is evident when small changes are made to the locations of saccadic targets while the eye is en route to the target; the eye automatically corrects for these changes in location even when observers are unaware that the target has moved (Deubel, Bridgeman & Schneider, 1998). Smooth pursuit eye movements are also linked to environmental rather than retinal locations, as can be seen when a moving object is tracked while simultaneously rocking ones' head back and forth (Raymond, Shapiro & Rose, 1984). Studies of change blindness in scene perception tell a similar story. Large changes made to a scene during a brief interruption often go unnoticed by observers, provided that the overall gist and layout of the scene remains intact (Henderson & Hollingworth, 2002). These and many other psychophysical studies suggest that visual perception is geared toward registering the position of objects in the environment rather than registering objects with respect to their retinal location (Fecteau, Chua, Franks & Enns, 2001; Liu, Healey & Enns, 2003; Li & Warren, 2000).

The goal of the present study was to determine whether multiple object tracking is based on retinal or allocentric coordinates. Our approach began with the longstanding observation that tracking accuracy varies systematically with object speed: objects moving at a slower speed are generally tracked more accurately than objects moving at a higher speed (Pylyshyn & Storm, 1988; Yantis, 1992). However, in these studies retinal motion and scene motion are confounded. In the present study we varied the speed of object motion relative to the center of the scene (allocentric speed) separately from the speed of motion of the scene relative to the viewing frame (retinal speed). The prediction for a retinal-based tracking mechanism is that tracking accuracy should vary directly as a function of the speed with which the objects transit the eye, regardless of their relative speed of movement within the scene. However, if tracking is based on an allocentric frame of reference, then accuracy should vary most directly with the speed of objects within the scene, and retinal speed should not matter.

Overview of Experiments

In Experiment 1, the tracking accuracy for objects moving within the confines of a two-dimensional (2D) rectangle was compared with objects moving within a depicted three-dimensional (3D) box. In both conditions, the speed of the moving objects was varied to determine the sensitivity of tracking accuracy to changes in retinal speed. The results showed that objects could be tracked equally well in both situations, with a small tendency for tracking to be even more accurate in the 3D display. Most critically for the remaining experiments, tracking accuracy declined systematically with increases in object speed.

In Experiment 2, tracking accuracy for objects within the 3D box was measured while the box as a whole underwent a ‘wild ride,’ consisting of dynamic and simultaneous translations in the picture plane, rotations in depth around the vertical plane, and dilations and contractions in depth. That is, in addition to varying the relative speed of objects within the 3D box, the motion of the whole box varied in a complex way. Yet the results showed clearly that tracking accuracy was unaffected by these global variations in scene motion. Only the motion of the objects relative to the scene as a whole influenced tracking accuracy.

In Experiment 3 most of the pictorial support for the 3D box was removed from the display, in order to see to what extent the perception of a stable scene depended on the wire frame and the checkerboard ground plane that had been used to convey the layout of the scene. The results showed that tracking accuracy was unaffected by the removal of these cues to the third dimension. This suggested that the movement of the objects themselves, within the confines of the depicted 3D box, were sufficient to provide the ‘structure from motion’ necessary to perceive the layout of the 3D scene.

In Experiments 4 and 5 the allocentric tracking hypothesis was tested by attempts to reduce the perceived coherence of the 3D structure. In Experiment 4, the objects to be tracked moved at two different speeds within the same scene, thereby sharply reducing both the coherence of the 3D scene and tracking accuracy. Scene coherence was reduced in Experiment 5 by projecting the image of the scene onto the junction of two dihedral surfaces. Even though the retinal projection for the observer in this condition was identical to the conditions in which tracking accuracy had been high (Experiments 2 and 3), tracking accuracy was reduced along with the reduced coherence of the scene. Taken together, these results provide strong support for the view that multiple object tracking is accomplished using an allocentric frame of reference.

Experiment 1: Baseline Tracking Performance

The purpose of Experiment 1 was to establish several important baseline measurements for the experiments that followed. First, because the displays in all the subsequent experiments depicted objects moving in a 3D scene, we sought to compare tracking accuracy in 2D and 3D displays as directly as possible. Previous studies have reported that multiple object tracking is not impaired in accuracy when objects disappear briefly as they pass behind occluding surfaces (Scholl & Pylyshyn, 1999). Studies using the additional cue of binocular disparity have reported improved tracking accuracy relative to control displays without this cue (Viswanathan & Mingolla, 2002). Tracking accuracy is also improved when the moving objects are distributed across two planes in depth rather than when they are moving in only a single plane (Viswanathan & Mingolla, 2002). Our goal in Experiment 1 was therefore to provide as rich a 3D environment as possible, using only pictorial and motion cues for depth, and to compare tracking under these conditions with the 'standard' case of tracking on a 2D screen.

To manipulate motion relative to an allocentric frame in the present study, we depicted the objects moving within a 3D box defined by a wire frame and a checkerboard floor, as shown in Figure 1. To help reinforce the perception of 3D motion we added the depth cue of dynamic changes in relative size. When objects were closest to the viewer they subtended 0.5° of visual angle and when they were farthest away they subtended

0.8° degrees. At intermediate depth locations their size varied smoothly between these extreme values.

------------- Insert Figure 1 about here ----------------

Our second goal was to establish tracking accuracy when the objects in the scene were moving at various rates of speed. In Experiment 1, all the objects moving in a display moved at the same speed, but the rate of speed on any given trial was 1°, 2°, or 6° per second. This was a sufficiently large variation in speed to have a large influence on tracking accuracy.

Our third goal was to measure the decline in tracking accuracy as the number of objects was increased, thereby allowing us to obtain a stable measure of tracking 'capacity' for every condition that was tested (Pashler, 1988). In Experiment 1, a total of 16 moving objects were present in each display, but the number designated as targets varied randomly from 2, 4, or 6. This turned out to be a large enough range to observe tracking accuracy that was near perfect in some cases and near chance in others.

Method

Participants. 25 undergraduate students (X females) with mean age = x years participated in an hour-long session in exchange for extra-course credit. All reported normal or corrected-to-normal vision. 12 participated in the 2D condition and 13 participated in the 3D condition.

Apparatus. Displays and data collection for all experiments were controlled by a Dell Pentium III 533MHz computer running custom software written in C++ using Open GL for 3D graphics. Participants were seated at a viewing distance of 57cm from a 19" Sony Trinitron monitor (resolution 1024 x 768) that had a viewable area of 35° (width) by 26° (height).

Stimuli and Procedure. Moving objects consisted of 16 small white airplanes seen within a viewing frame, as shown in Figure 1. In the 2D condition, the frame was a 2D rectangle subtending 20° (wide) x 16° (high) of visual angle. In the 3D condition, the frame was a depicted 3D rectangular box, drawn in white on a black background, with an aspect ratio of 285 pixels horizontally (X), 360 pixels in depth (Y) and 285 pixels vertically (Z). This corresponded to approximately 16° (X) x 20° (Y) x 16° (Z) of visual angle when each dimension was viewed from an orthogonal vantage point. The floor of the rectangular box was depicted in gray, overlaid with a black square grid. The airplanes subtended 0.65° of visual angle in the 2D condition and 0.5°- 0.8° in the 3D condition. In the 3D condition, the objects and frame were drawn to create a camera angle of 45° to the X-Y plane of the frame.

It is important to note that although observers tracked small airplane shapes in this experiment, the shape of the objects to be tracked had no influence on performance. We confirmed this with out own tests comparing airplanes with dots of similar size, and others have found the same result (Horowitz, personal communication).

At the beginning of each trial, 16 stationary objects were randomly positioned onscreen. After 1 s, a subset of 2, 4, 6, or 8 objects was each surrounded by green marking circles. The marking circles flashed off and on four times at 200 ms intervals and then remained onscreen for another 2 s. This designated the target set that observers were to track through a period of motion. The marking circles then disappeared and all objects began to move at a constant speed of 1, 2, or 6°/s for a duration of 10 seconds.

Objects moved in a straight line in a randomly chosen direction until they reached the edge of the frame. Upon meeting the frame edge, the objects' trajectory was changed so that it appeared to bounce off the edge (2D) or the wall (3D) and continue on a trajectory consistent with the physics of a billiard ball of constant velocity. Object boundaries were allowed to intersect from the perspective of the viewer for both 2D and 3D conditions. However, in the 3D condition objects were not allowed to occupy the same region of 3D space. At the end of the 10-s period of continuous motion, all of the objects stopped moving and a green circle surrounded one object. This circle flashed briefly four times and then remained onscreen until the participant responded.

On half the trials this probe surrounded a target object (one of the objects to be tracked); on the other half it surrounded a nontarget object.

The observer's task was to indicate whether the probed object was part of the original target set. Observers pressed the 'z' key for target objects and the '/' key for nontarget objects. Correct responses were followed by a centrally presented '+' symbol, incorrect responses by a '-' symbol, and no response by a '0' symbol. At the end of each block of trials, a message displayed the percentage of errors for the most recent block of trials and participants were prompted to initiate the next block when ready. Observers were instructed to respond accurately but to guess when uncertain.

Observers performed 20 practice trials in each condition before formal testing began. The 2D and 3D conditions each consisted of 144 trials (3 object speeds x 4 sets of objects to be tracked x 2 probe types x 6 repetitions). The order of conditions was randomized throughout the experiment. The testing session was divided into three blocks of 48 trials, with a self-paced break between each block. Observers were instructed to maintain fixation at the center of the display throughout the trials but eye movements were not monitored; they have been shown not to affect tracking accuracy (Scholl & Pylyshyn, 1999).

Results

Tracking accuracy was analyzed using two different dependent measures. First, tracking accuracy was examined with an analysis of variance (ANOVA) on the between-observers factor of Display (2D, 3D) and the within-observer factors of Object Speed (1°, 2°, 6° per second) and Target Number (2, 4, 6). Trials involving 8 targets were excluded from the analysis because accuracy was near the chance response level of 50% in all conditions (less than 60% correct). This analysis revealed the broad effects of the experimental factors on tracking accuracy.

Second, tracking accuracy was assessed using a measure of 'capacity' (K) adapted from Pashler (1988). This measure was originally developed as a quantitative estimate of the number of items held in memory during a change-detection task. The formula is: K = [

NT * (pHits – pFA)] /(1-pFA), where K = capacity, NT is the number of items to be tracked, pHits is the proportion of hits and pFA is the proportion of false alarms. The upper limit of K is bounded by the number of objects that the observer is asked to track and so trials in which Target Number = 2 were excluded in order to avoid artificial deflation of K.

Accuracy. Mean proportion correct is shown in Figure 2. Tracking accuracy decreased as object speed increased from 1°/s to 6°/s, F(2, 44) = 70.50, p < .001. Accuracy also decreased as number of targets increased from 2 to 6, F(2, 44) = 142.47, p < .001. Finally, there was a marginally significant interaction of Display x Target Number, F(2, 44) = 3.08, p < .06, indicating that accuracy remained higher as number of objects increased in the 3D than in the 2D displays objects.

Capacity. Capacity decreased as object speed increased, from 1°/s (K = 3.66) to 2°/s (K = 2.53) to 6°/s (K = 1.60), F(2, 46) = 25.52, p < .001. No other effects were significant.

------------- Insert Figure 2 about here ---------------

Discussion

Experiment 1 replicated previous findings that tracking accuracy is impaired by increases in both object speed and number of objects tracked (Pylyshyn & Storm, 1988; Yantis, 1992). As measured by capacity (K), the number of items observers could track successfully ranged from almost four items when objects were moving slowly to only one and a half objects on average when objects were moving rapidly.

These results also established that the addition of several pictorial depth cues (wire frame, checkerboard floor) and dynamic changes in relative size did not negatively affect tracking accuracy. If anything, there was a trend for improved tracking in the large target sizes in the 3D condition. Thus, whereas previous studies have reported that occlusion cues and binocular disparity can enhance tracking performance (Viswanathan & Mingolla, 2002), the present findings showed that this benefit was only weakly present for the pictorial 3D cues employed here.

Experiment 2: Tracking during a 'Wild Ride'

The next step was to manipulate the speed of the objects in the 3D scene independently of their speed on the retina. The speed of the 3D scene was manipulated by applying three motion transformations to the 3D scene as a whole. These included translation (back and forth movement of the box horizontally across the screen), rotation (a swiveling of the box about its central vertical axis), and zoom (movement of the box both toward and away from the viewer). Applying all three of these motion transformations to the box of moving objects had the effect of making the box swing, swoop and rotate wildly while the observer attempted to track the target objects inside the box. Two example video frames from these motion sequences are shown in Figure 3.

The speed of the objects relative to the box boundaries was varied in a similar way to Experiment 1. That is, while the box itself was undergoing a complex path of motion, the objects inside it were all moving at 1°, 2°, or 6°/s, relative to an imaginary viewer that was viewing the box from 57 cm and keeping a constant viewing angle on the box.

------------- Insert Figure 3 about here ---------------

We must also note that prior to collecting data in this condition, we tested tracking accuracy in each of the scene motion conditions individually (translation, rotation, zoom). None of these conditions varied significantly from one another in their tracking accuracy, and none of them were significantly different from the 3D condition in Experiment 1. It is for this reason that only the condition in which all three transformations of scene motion were applied simultaneously is presented here in detail.

Method

Participants. 17 undergraduate volunteers (X females, mean age = x years) participated in an hour-long session in exchange for course credit. All reported normal or corrected-to-normal vision.

Stimuli, Design and Procedure. Displays consisted of 16 objects moving within a depicted 3D rectangular box and were identical to the 3D condition in Experiment 1 with the following modifications. In this and all subsequent experiments, the moving objects were 3D spheres that appeared as discs from any given vantage point. Observers were asked to track only 2, 4, or 6 objects, since Experiment 1 had shown that accuracy for 8 objects was very poor. Finally, because the greatest differences on tracking accuracy were observed with the most extreme object speeds, only speeds of 1°/s and 6°/s were tested.

Scene motion transformations. The wire frame box containing the moving objects in this experiment underwent motion involving simultaneous changes in translation, rotation, and zoom. Considered singly, each transformation was as follows: Translation involved moving the box horizontally across the screen (X-axis motion). The speed of translation was measured by taking the left-right distance traversed on the screen over time. Rotation involved moving the box around its vertical or Z-axis. On half the trials, the box rotated clockwise for the duration of the tracking episode, on the remaining half it rotated counterclockwise. Rotation speed was measured as the polar angle of rotation around the Z-axis over time. Zoom involved expanded and contracted the size of the box and its contents proportionately, which is retinally equivalent to moving the box closer and further from the observer. Zoom speed was measured as the change in size (in degrees of visual angle) over time.

For all types of transformations, when the box changed directions from left to right, clockwise or counterclockwise rotation, or moved from near to far, the speed of the box changed in a sinusoidal manner, so that the box did not bounce sharply from one direction to another.

Scene motions were classified as no motion, slow, and fast. For no-motion, the box remained static while objects moved within the confines of the box as they had in Experiment 1 (translation: 0°/s, rotation: 0°/s, zoom: 0°/s). For slow motion the speeds were: translation: 2.4°/s, rotation: 2°/s, zoom: 1.6°/s. For fast motion they were: translation: 3.4°/s, rotation: 4°/s, zoom: 3.25°/s. These three scene motion conditions were presented randomly and equally often within each block of trials. There were a total of 144 trials (3 scene speeds x 2 object speeds x 3 target numbers x 2 probe types x 4 repetitions).

Results

Accuracy. Mean proportion correct is shown in Figure 4. ANOVA indicated that tracking accuracy decreased as object speed increased from 1°/s to 6°/s, $F(1, 16)= 93.95$, $p < .001$. Accuracy also decreased as target number increased from 2 to 6, $F(2, 32)= 46.52$, $p < .001$. No other effects were significant, including the factor of Scene Speed, $F(2, 32) < 1$.

Capacity. Capacity decreased as object speed increased, from 1°/s = 3.64 to 6°/s = 2.03, $F(1, 20) = 32.45$, $p < .0001$. No other effects were significant, including the factor of Scene Speed, $F(2, 40) < 1$.

------------- Insert Figure 4 about here ---------------

Discussion

Experiment 2 revealed two main results. First, regardless of whether the 3D box of moving objects remained stationary or was moving, tracking was influenced by increases in object speed and number of objects tracked. Faster object motion and a larger number of target objects both reduced tracking accuracy. Second, adding slow or fast motion to the entire scene caused no additional impairment in tracking accuracy. This is inconsistent with tracking being accomplished with a retinotopic frame of reference, since adding scene motion to the object motions resulted in both (1) a marked increase in the retinal motions of many of the objects being tracked for large portions of the tracking episode, and (2) a marked increase in the variability of the retinal motions of the various objects in the box. Given the sensitivity of tracking accuracy to variations in object speed already shown in both Experiments 1 and 2, it is surprising therefore that these additional variations in speed caused by scene motion had no measurable influence. We can only conclude that object tracking is not based on retinal coordinates, but rather that it is based on the speed of objects relative to the boundaries of the box, whether they are stationary or moving.

We must also note that the excellent tracking accuracy for objects inside the moving box cannot be attributed to observers smoothly pursuing the center of the box with their gaze, so as to leave the motion of the objects on the retina of the observer roughly equal in the stationary and moving conditions. This can be ruled out because only one of the three motion transformations (translation) even lends itself to the possibility of maintaining equal retinal motion through smooth pursuit. It is certainly possible to track the center of the box when it is simply moving back and forth in a predictable pattern across the screen. However, the situation is very different for rotation, where fixating the center of the box would result in objects generally speeding up their retinal motion as they passed by the center of gaze and generally slowing down their retinal motion as the moved to the outside edges of the box. For zoom, the same strategy of center fixation would result in retinal motions that were slower then usual as the box moved away and then were much faster than usual as the box moved toward the viewer. Our testing of each of these conditions separately (not reported here in order to save space) indicated that all three resulted in patterns of tracking accuracy that were not significantly different from the pattern reported for the ‘wild ride,’ (Figure 3) in which all three motions were applied simultaneously. What makes this even more remarkable is the observation that the maximum retinal motions possible in the ‘wild ride’ were markedly faster even than any of the motions considered separately. Yet, this had no influence on tracking accuracy.

Experiment 3: 3D Structure of the ‘Wild Ride’

Tracking accuracy in Experiment 2 was strongly affected by the speed of motion of individual objects, but was unaffected by the speed of motion of the box in which objects were moving. This points to an allocentric tracking mechanism, one that tracks objects with respect to their position in the environment rather than with respect to the position and viewpoint of the observer.

If tracking is truly allocentric, one way to interfere with it should be to reduce the visual cues supporting the perception of the 3D space in which the objects are moving. In Experiment 2, observers saw objects moving inside the confines of a wire frame box that included a textured floor. These two features may have provided important pictorial cues for the 3D structure of the box in which the objects were moving as well as important cues to the nature of the motion path undertaken by the box as a whole. However, there were also additional 3D cues intrinsic to the moving objects themselves. For example, objects expanded and contracted slightly as they moved, to support the appearance that they were either nearer (larger) or farther (smaller) from the viewer. The moving objects also changed direction every time they encountered the invisible 'walls' of the box. Finally, when the box moved, the moving objects also moved coherently on the screen so that, regardless of the individual path being taken by each object within the box, its motion was also consistent with the box moving as a whole.

In Experiment 3 we stripped away the wire frame and the textured floor from the displays in order to test whether the 3D cues that were intrinsic to the moving objects alone were sufficient to support the perception of a coherent 3D scene. Previous research on the perception of 3D structure from coherent motion indicates that human observers are adept at interpreting volume from the motion of only 3 or 4 points, provided that the points are fixed to the surfaces of a rigid object (Siegel & Andersen, 1988; Treue, Husain, & Andersen, 1991; Ullman, 1979). The differences for the present displays include: (1) the moving objects are not rigidly fixed to any part of the box, (2) there are a total of 16 objects moving independently, and (3) the rigid 'walls' of the box are indicated only by the dynamic change in direction that occurs for the moving objects when they encounter them. It is therefore very much an open question whether allocentric object tracking will still be possible if the wire frame and the textured floor are removed from the displays. If tracking is impaired by the removal of these features, it will suggest that the visible features of the box are essential for establishing a 3D environment in which the objects can be tracked. If tracking is unaffected, we will have to conclude that the 3D structure inherent in the cloud of moving objects is itself sufficient to support the perception of a volume in which the objects are moving.

Method

Participants. 20 undergraduate students (X females, mean age = x years) participated in exchange for course credit. All reported normal or corrected-to-normal vision.

Stimuli, Design, and Procedure. All methodological details were identical to Experiment 2 with the exception that the wire frame and textured floor of the box were not displayed.

Results

Accuracy. Mean proportion correct is shown in Figure 5. ANOVA indicated that tracking accuracy decreased as object speed increased from 1°/s to 6°/s, $F(1, 19)= 121.24$, $p < .001$. Accuracy also decreased as target number increased from 2 to 6, $F(2, 38)= 45.93$, $p < .001$. There was also an interaction of Object Speed x Number of Targets, $F(2, 38)= 4.60$, $p < .02$, reflecting that the decrease in accuracy with number of targets was greater for fast than for slow object motion. No other factors were significant, including the factor of Scene Speed, $F(2, 38) = 2.98$.

Capacity. Capacity decreased as object speed increased, from 1°/s = 3.64 to 6°/s = 2.03, $F(1, 19) = 87.03$, $p < .001$. No other effects were significant, including the factor of Scene Speed, $F(2, 38) = 1.35$.

Accuracy and capacity measures were also compared to those in Experiment 2 with mixed-design ANOVAs to determine if the presence of the wire frame and textured floor had any effect on tracking performance. No factors involving Experiment were significant in either measure, all $Fs < 1.0$.

------------- Insert Figure 5 about here ---------------

Discussion

As in previous experiments, tracking accuracy was still strongly impaired by increases in object speed and in the number of objects to be tracked. As in Experiment 2, tracking accuracy was also unaffected by increases in the speed of the box in which the objects moved. Yet, tracking accuracy was unaffected by the presence versus absence of the wire frame and the textured floor of the box in which the objects moved. This suggests that there is sufficient 3D structure implied by the moving objects themselves to allow for perception of a coherent, albeit moving, 3D environment. Observers are able to recover the structure of this environment solely from the relative motions of the objects.

Experiment 4: Simultaneously tracking objects moving at two speeds

Experiment 4 was conducted in a further effort to discover factors underlying the allocentric tracking of multiple objects. Experiment 3 had already shown that visible cues to the boundaries of the box were not critical to this ability, since excellent tracking accuracy was still possible when the visible cues to the boundaries of the 3D box were eliminated. The factor we considered next was the increase in the variability of motion speeds that occurs when the box is placed into motion. We noted in the discussion to Experiment 2 that one of the remarkable features of allocentric tracking ability, from the perspective of the retinal motions involved, was that adding scene motion to the displays resulted in a marked increase in the variability of the retinal motions of the various objects in the box. Although this had no negative effects on allocentric tracking, we suspected that adding variability to the speeds of individual objects in motion might impair tracking accuracy, especially when the entire box of objects was also in motion. If so, it would confirm that one of the factors assisting the successful tracking of objects in Experiment 3 (no visible boundaries of the box) was the constant speed of motion of all the objects moving inside the box.

In Experiment 4 half the objects inside the box moved at one speed while the other half moved at the other speed (either 1°/s or 6°/s). The target objects to be tracked were also equally divided between slow and fast moving objects. The motion of the box as a whole was varied in the same way as in Experiments 2-3. If multiple speeds of motion generally impair tracking accuracy, then the presence of two speeds of motion should impair tracking even when the box is stationary. However, if multiple speeds impair tracking accuracy only when the box itself is in motion, it would indicate that the constant allocentric speed of motion in previous experiments was a contributing factor to the establishment of a coherent 3D scene in which objects were moving.

Previous studies of multiple object tracking have either had objects within a scene moving at different speeds (Pylyshyn & Storm, 1988; Scholl & Pylyshyn, 1999) or the same speed (Yantis, 1992; Viswanathan & Mingolla, 2002), but none have directly examined these two conditions. Previous studies have also not distinguished between increased variability in allocentric speed versus retinal speed of motion.

Method

Participants. 20 undergraduate students (X females, mean age = x years) participated in exchange for course credit. All reported normal or corrected-to-normal vision.

Stimuli, Design, and Procedure. Displays and other details of the method were identical to Experiment 2, except that all objects, including targets, now moved at one of two speeds inside the box (1°/s and 6°/s). The target probe at the end of the tracking episode therefore also varied randomly between either a slow or fast target.

Results

Accuracy. Mean proportion correct is shown in Figure 6. ANOVA indicated that tracking accuracy did not vary with object speed in this experiment, $F(1, 19) < 1$. This is likely because observers were tracking objects of both speeds on every trial and had to be prepared to respond to either a ‘slow’ or a ‘fast’ probe. However, as in previous experiments, accuracy decreased as target number increased from 2 to 6, $F(2, 38)= 65.72$, $p < .001$.

Most importantly, accuracy decreased as the motion of the scene increased from no motion to slow to fast, $F(2, 38)= 8.92$, $p < .001$. This is the first time this effect has been observed in this study, suggesting that multiple speeds of allocentric motion impairs the perceived structure of the 3D scene. This interpretation is strengthened by an interaction of Scene Speed x Object Speed, F(2, 38)= 3.45, p < .05, which reflects a larger impairment of object speed when the box was in fast motion than when the box was stationary. It is also strengthened by an interaction of Scene Speed x Number Tracked, F(2, 38)= 8.92, p < .001, which reflects an exaggerated decrease in accuracy for larger numbers of targets when the box was in motion. No other effects were significant.

Capacity. Capacity did not vary with object speed, F(1, 19) < 1, but capacity did decrease with increases in scene speed: no motion, K = 3.36; slow motion, K = 2.59; fast motion, K = 2.46; F(2, 38) = 5.48, p < .03. A direct comparison of these capacity estimates with those in Experiment 2 revealed a significant interaction of Experiment x Scene Motion, F(2, 78) = 4.13, p < .02. Whereas capacity was unaffected by scene motion in Experiment 2 (mean K = 2.85) it decreased significantly with increases in scene motion in the present experiment.

------------- Insert Figure 6 about here ---------------

Discussion

These results indicate that an important factor contributing to allocentric object tracking is the perception of a coherent 3D scene. When the perception of this scene was impaired in the present experiment by increasing variability in the speeds of objects in motion, tracking accuracy was reduced. It is notable that the mere existence of two speeds of motion was in itself not detrimental to tracking accuracy. When the box was stationary, tracking accuracy for targets moving at two speeds was comparable to that of previous experiments, where only one speed of motion was present in any given display. Yet, the simultaneous presence of these two speeds of object motion was detrimental to tracking accuracy when the box of moving objects was itself moving. This indicates that the constant speed of object motion is an important cue to the perceived structure of the 3D space in which the objects are moving.

Experiment 5: Tracking Objects in a Non-rigid 3D Space

Experiment 4 suggested that allocentric object tracking depends on the perception of a coherent 3D scene. Experiment 5 put this idea to a further test by examining tracking accuracy when the 3D space in which objects move appears to be unstable. The main manipulation was inspired by the observation that shape and space constancy break down when a scene is viewed from more than one vantage point (Cavanagh, Peters & von Grünau, 1988; Cavanagh & von Grünau, 1989).

The general failure of shape constancy when a scene is viewed from more than one vantage point can be easily illustrated by placing a vertical fold in a ten or twenty dollar bill, centered on the face of the individual on the bill. The fold can be either concave (open book) or convex (book spine) with respect to the viewer. While you are viewing the folded face, slowly rotate the bill around its horizontal axis. You will notice the facial expression of the depicted person changes dramatically as the bill is rotated. Now smooth out the fold in the bill and view the face again, while slowly rotating the bill both horizontally and vertically. There is no longer any change in the expression of the face. Taken together, these two conditions illustrate that shape constancy is readily achieved when an image is viewed from a wide range of vantage points (the smooth bill), even though these vantage points may distort the retinal image of the scene considerably. However, shape constancy breaks down as soon as the scene is viewed from two or more vantage points, or when the image of the scene is projected onto two or more surfaces, as occurs when the bill is folded.

In Experiment 5 we projected the image of the tracking displays onto the convex corner of a dihedral surface, as illustrated in Figure 7. Importantly, we had the observer perform the tracking task on these displays from the same vantage point as the projector, to ensure that the retinal projection in this experiment was roughly equivalent to what it had been in previous experiments. Yet, despite this equivalence at a retinal level, projecting the image in this way had the effect of greatly distorting the perceived structure of the space in which the objects were moving. The distortion included a rubber-like bending of the 3D box and large apparent changes in object speed when objects crossed the folded center of the projection zone. If tracking is based on retinal coordinates, then tracking accuracy should be comparable to that in Experiments 3 and 4. However, if tracking is allocentric, then accuracy should be impaired by the reduction in scene coherence.

------------- Insert Figure 7 about here ---------------

Method

Participants. 14 undergraduate students (X females, mean age = x years) participated in exchange for course credit. All reported normal or corrected-to-normal vision.

Apparatus. Displays were projected onto a convex corner using a NEC MultiSync projector positioned 225 cm from the corner. The convex projection screen consisted of two white foam-core boards (3′ x 5′) connected at a 90° angle. To help reinforce the spatial layout of the projection screen a 1″ wide green ribbon outlined the screen and a 60-watt lamp illuminated the screen from overhead on the right side, casting a noticeable shadow on the left side of the screen. Participants sat with their heads positioned directly underneath the projector, at a viewing distance of 171 cm, to ensure a projection that was nearly equivalent to that obtained when the displays were viewed on a computer screen.

Stimuli, Design, and Procedure. With the exception of the details of projection, the displays and methods were identical to Experiment 2.

Results

Accuracy. Mean proportion correct is shown in Figure 8. ANOVA indicated that tracking accuracy decreased with increases in object speed, $F(1, 13) = 34.65$, $p < .001$, as had been found in Experiments 1-3. Also, as in previous experiments, accuracy decreased as target number increased, $F(2, 26)= 38.07$, $p < .001$. The critical result for this experiment, however, was that tracking accuracy decreased with increases in scene speed, $F(2, 26)= 7.28$, $p < .01$. This finding indicates that reducing the coherence of the

3D scene impairs tracking accuracy even though the retinal speeds of motion are identical to other conditions in which scene speed is not a factor (Experiment 2-3).

This interpretation was strengthened by a marginal interaction of Scene Speed x Object Speed, F(2, 26)= 3.13, $p < .06$, which reflects a larger impairment of object speed when the box was in fast motion than when the box was stationary. It is also strengthened by an interaction of Scene Speed x Number Tracked, F(2, 26)= 2.97, $p < .03$, which reflects an exaggerated decrease in accuracy for larger numbers of targets when the box was in motion. No other effects were significant.

Capacity. Capacity was impaired by both increases in Object Speed (1°/s, K = 3.08; 6°/s, K = 1.65), F(1, 13) = 29.91, $p < .001$, and with increases in Scene Speed (no motion, K = 2.92; slow motion, K = 2.27; fast motion, K = 1.90; F(2, 26) = 5.02, $p < .02$. There was also an interaction of Object Speed x Scene Speed, F(2, 26) = 5.51, $p < .01$, reflecting the fact that differences in Scene Speed were greater when the objects were moving slowly than when they were moving fast. This is likely because accuracy in the fast object condition was already so low (near the chance level of 50%) that the full effects of Scene Speed could no longer be measured.

A direct comparison of these tracking capacity estimates with those in Experiment 2 revealed an interaction of Experiment x Scene Motion, F(2, 66) = 3.78, $p < .03$. Whereas capacity was unaffected by scene motion in Experiment 2 (mean K = 2.85) it decreased significantly with increases in scene motion in the present experiment.

------------- Insert Figure 8 about here ---------------

Discussion

These results confirm that allocentric object tracking depends on the perception of a coherent 3D scene. Scene perception was impaired in this experiment by projecting the image of the moving objects onto a convex corner, a manipulation that is known to sharply impair the recovery of the 3D structure of a scene. At the same time, the retinal relations among the moving objects were preserved by this manipulation. Notably, projection of this kind had little if any influence on tracking objects when the box was stationary. However, it sharply reduced tracking accuracy when the box was in motion. These results are all consistent with the idea that allocentric object tracking depends on the perception of a coherent 3D scene.

General Discussion

The question addressed by this study was whether multiple object tracking is based on retinal or allocentric spatial coordinates. Although there has been much interest in the capacity of human vision to keep track of a small number of objects in a dynamically changing environment (Pylyshyn & Storm, 1988; Scholl & Pylyshyn, 1999; Yantis, 1992), the question of the spatial frame of reference in which object position is indexed and updated has not been a research focus to date. As a first step in addressing this issue, the present study asked whether object tracking is based on a retinotopic frame or whether it is referenced to the scene, which we refer to as an allocentric frame of reference.

Our approach to addressing this question was based on the fact that tracking accuracy generally declines as the speed objects to be tracked is increased. However in the studies used to establish this finding, the motion of the objects on the retina (assuming a fixed gaze) and the motion of the objects in the scene had not been decoupled. We uncoupled them in the present study by designing an object tracking environment in which the motion of the objects in a scene could be varied independently of their motion on the retina. This was accomplished by moving the entire scene of moving objects as a whole in complex ways across the viewing screen. We were then able to ask whether the motion of the scene as a whole had an influence on tracking accuracy over and above the influence of the motions of the objects within the scene.

The clear answer we obtained after measuring the influence of three different scene motions, including translation, rotation in depth, and zoom, both alone and in combination, was that motion of the scene as a whole had negligible effects on tracking accuracy. In short, the number of objects that could be tracked successfully was not different when the depicted 3D box in which the objects moved was stationary on the screen versus when it moved in complex ways on the screen. This is an important result, both for our understanding of the human visual system and for the application of current knowledge about human vision to the problems of human-machine interaction (HCI). In the discussion that follows, we will address implications of this finding for each of these areas in turn.

Implications of Allocentric Tracking for Human Vision

A primary implication of the finding that object tracking is allocentric is that tracking must therefore depend on continual visual interaction with a perceived environment. This leads to numerous questions such as: Which visual cues are used to perceive the layout or structure of this environment? How richly and for how long is that environment represented in the visual system? To what extent is visual interaction with the environment based on past experience versus being based only on the most recently updated information? These and many other questions relevant to the conclusion of allocentric tracking will now have to be addressed.

In the present series of experiments we only began to address these issues. The experiments reported here focused primarily on whether the stability of the perceived 3D environment was important. We first looked at whether any cues from the environment itself needed to be explicitly presented. When we removed all external visual support for the scene in Experiment 3, which included the wire frame box and the textured floor, we found (1) that tracking accuracy was still allocentric (not significantly affected by complex motion of the box as a whole) and (2) that tracking accuracy was comparable to that obtained when these scene ‘supports’ were visible. This means that the structural cues remaining in the moving objects themselves, which included changes in object size consistent with relative distance and constant motion trajectories that changed only when the invisible walls of the box were encountered, were sufficient to allow observers to maintain the perception of a coherent 3D environment.

Our second approach involved presenting all the visible scene structure (wire frame and textured floor) but reducing the coherence of the object motions internal to the box by allowing objects to move at one of two different speeds (Experiment 4). This had no negative effect on tracking accuracy when the 3D environment was stationary, as would be expected based on previous studies (Pylyshyn & Storm, 1988). However, the presence of two different object motions led to a marked reduction in tracking accuracy when the box as a whole was in motion. Since the perceived coherence of the 3D scene was also much reduced in this condition it suggested that scene stability or coherence was a critical factor in sustaining high levels of tracking accuracy.

This hypothesis was put to a direct test in Experiment 5 where scene coherence was reduced by projecting its image onto the junction of two dihedral surfaces. Despite the fact that the retinal projection was still identical to previous conditions in which tracking accuracy was high (Experiments 2 and 3), tracking accuracy was again reduced when the scene was placed into motion under these conditions. The apparent plastic transformations of the scene structure that occurred when the box moved resulted in a significant decrease in tracking accuracy. Taken together, these results converge on the conclusion that multiple object tracking is accomplished using allocentric spatial references rather than a retinal frame of reference. These spatial references can be fairly abstract, as seen in Experiment 3, where all explicit evidence regarding the box was removed, but they do need to be stable, as indicated by the results from Experiments 4 and 5.

Having established that object tracking is allocentric, it is of interest to consider the perspective that this gives to some of the extant findings regarding object tracking in the literature. Consider first the finding that whether or not observers are permitted to move their eyes during a tracking episode has no effect on tracking accuracy (Pylyshyn & Storm, 1988). Maintaining high levels of tracking accuracy when eye movements are permitted would require an additional mental operation if tracking is fundamentally retinotopic. However, for an allocentric mechanism, eye movements are not an impediment to tracking, unless the periods between eye fixations become so long that they exceed the capacity of the system to maintain an updated representation of the environment. In fact, eye movements may even improve tracking accuracy, especially if they benefit the perception of the stable environment in which the objects are moving by, for instance, providing greater spatial detail of the environment.

We can also consider the finding that tracking accuracy remains high despite the presence of surfaces in the scene that temporally occlude vision of the objects to be tracked (Scholl & Pylyshyn, 1999). Again, an allocentric tracking mechanism would only be disturbed by occluding objects if they interfered with either the ability to maintain a representation of the object (if the occluded period exceeded the temporal capacity of the system) or if they interfered with the perception of a stable environment. Again, to the extent that visual occlusion is a rich cue to the 3D structure of an environment, there is also the real possibility that visual occlusion could be used to enhance tracking accuracy, through its effect on reinforcing the coherence and stability of the environment in which objects are moving. The finding that the addition of binocular disparity to a display can improve tracking accuracy (Viswanathan & Mingolla, 2002) is completely consistent with this point.

Finally, it is worth considering how the perspective of allocentric tracking may alter one's interpretation of the finding that tracking becomes very difficult when the entities to be tracked undergo plastic transformations or are deformable 'substances' rather than rigid objects (VanMarle & Scholl, 2003). This work was originally presented to show that the visual tracking mechanism can receive as inputs only objects that are 'rigid' and 'cohesive.' The present study raises the possibility that the critical stability may lie in the environment in which the objects are moving, not necessarily in the nature of the items being tracked per se. For example, consider a tracking display in which the surface in the scene is a non-uniform landscape (e.g, a surface of hills and valleys, a crumpled newspaper, an unmade bed). A subset of 'soldiers' roaming at random across this landscape should be very easy to track, despite the fact that the retinal projections of these objects would undergo considerable plastic transformation, as they are progressively made visible and occluded from the vantage point of the observer and as they change in their apparent distance. This is because both these objects and the environment in which they move can be interpreted as rigid and stable.

However, now consider the possibility that the 'soldiers' randomly shrink and expand in length as they make their way over the landscape. The stable object hypothesis (VanMarle & Scholl, 2003) would predict that tracking accuracy would now be reduced, since the objects are no longer rigid. Yet, the stable environment hypothesis (present study) would allow for the possibility of successful tracking under these circumstances, since the locations of the moving objects are still unambiguously visible in a stable 3D environment. In fact, an extremely strong version of the stable environment hypothesis would predict that even only the projections of cast shadows of rigid objects, flying overhead the non-uniform landscape, would still permit successful tracking of the shadows, even though this would make for an extremely plastic set of 'objects.' Control conditions could include the same object deformations in the absence of an interpretable landscape. Experiments like this clearly need to be conducted in order to determine the nature of the representations used in object tracking. The main benefit of the present finding, that tracking is allocentric under some circumstances, is to clarify and focus the questions that need next to be addressed.

Implications of Allocentric Tracking for Human Machine Interaction

Further Questions

On which visual cues is the perception of this environment established?
How richly and for how long is that environment represented in the visual system?
To what extent is visual interaction with the environment based on past experience versus being based only on the most recently updated information?

References

Cavanagh, P., & von Grünau, M. (1989). 3-D objects that appear nonrigid during rotation. Investigative Ophthalmology and Visual Science Supplement, 30, 263.

Cavanagh, P., Peters, S. C., & von Grünau, M. (1988). Rigidity failure and its effect on the Queen. Perception, 17, A27.

Deubel, H., Bridgeman, B., & Schneider, W. X. (1998). Immediate post-saccadic information mediates space constancy. Vision Research, 38, 3147-3159.

Fecteau, J.H., Chua, R., Franks, I., & Enns, J. T. (2001). Visual awareness and the on-line modification of action. Canadian Journal of Experimental Psychology, 55, 106-112.

Henderson, J.M., & Hollingworth, A. (2002). Eye movements, visual memory, and scene representation. In M.A. Peterson & G. Rhodes (Eds.), Analytic and holistic processes in the perception of faces, objects, and scenes. New York: JAI/Ablex.

Li, L., & Warren, W. H. (2000) Perception of heading during rotation: Sufficiency of dense motion parallax and reference objects. Vision Research, 40, 3873-3894.

Liu, G., Healey, C. G, & Enns, J. T. (2003). Target detection and localization in visual search: A dual systems perspective. Perception & Psychophysics, 65, 678-694.

Pashler, H. (1988). Familiarity and the detection of change in visual displays. Perception & Psychophysics, 44 , 369-378.

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

Raymond, J. E., Shapiro, K. L., & Rose, D. J. (1984). Optokinetic backgrounds affect perceived velocity during ocular tracking. Perception & Psychophysics, 36, 221–224.

Scholl, B.J. (2001). Objects and attention: The state of the art. Cognition, 80, 1–46.

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

Scholl, B.J., Pylyshyn, Z.W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple-object tracking. Cognition, 80, 159–177.

Siegel, R.M. & Andersen, R.A. (1988). Perception of three-dimensional structure from two-dimensional motion in monkey and man. Nature, 331, 259-261.

Treue, S., Husain, M. & Andersen, R. (1991). Human perception of structure from motion. Vision Research, 31, 59-75.

Ullman, S. (1979). The interpretation of structure from motion. Proceedings of the Royal Society of London B, 203, 405-426.

Van Essen, D.C., Lewis, J.W., Drury, H.A., Hadjikhani, N., Tootell, R.B.H, Bakircioglu, M., Miller, M.I. (2001). Mapping visual cortex in monkeys and humans using surface-based atlases. Vision Research, 41, 1359–1378.

vanMarle, K., & Scholl, B. J. (2003). Attentive tracking of objects vs. substances. Psychological Science, 14, 498-504.

Viswanathan, L. and Mingolla, E. (2002). Dynamics of attention in depth: Evidence from multi-element tracking. Perception, 31, 1415-143.

Viswanathan, L., & Mingolla, E. (1998). Attention in depth: Disparity and occlusion cues facilitate multi-element visual tracking (Technical Report No. CAS/CNS-98-012). Boston University Center for Adaptive Systems and Department of Cognitive and Neural Systems.

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

Figure Captions

Figure 1. Experiment 1 compared tracking accuracy for objects moving in either a depicted two-dimensional (A) or three-dimensional environment (B).

Figure 2. Mean proportion correct tracking accuracy in Experiment 1. Error bars represent +/- 1 SE. The numbers beside each accuracy function refer to K, an estimate of tracking capacity in which the units are "total number of objects" tracked successfully (Pashler, 1988).

Figure 3. Displays in Experiment 2 consisted of a depicted 3D wire frame box with a textured floor. In addition to the objects (spheres) inside the box moving while observers attempted to track them visually, the entire box could translate across the screen, swivel around the vertical center of the box, and zoom in and out. When these transformations all occurred at the same time we referred to it as the 'wild ride.'

Figure 4. Mean proportion correct tracking accuracy in Experiment 2. Error bars represent +/- 1 SE. The numbers beside each accuracy function refer to K, an estimate of tracking capacity in which the units are "total number of objects" tracked successfully (Pashler, 1988).

Figure 5. Mean proportion correct tracking accuracy in Experiment 3. Error bars represent +/- 1 SE. The numbers beside each accuracy function refer to K, an estimate of tracking capacity in which the units are "total number of objects" tracked successfully (Pashler, 1988).

Figure 6. Mean proportion correct tracking accuracy in Experiment 4. Error bars represent +/- 1 SE. The numbers beside each accuracy function refer to K, an estimate of tracking capacity in which the units are "total number of objects" tracked successfully (Pashler, 1988).

Figure 7. In Experiment 5 displays were projected onto a convex corner in order to test whether tracking accuracy is influenced by the coherence of the 3D scene. The display from the observer's perspective is shown in (A). A bird's eye view of the setup is shown in (B).

Figure 8. Mean proportion correct tracking accuracy in Experiment 5. Error bars represent +/- 1 SE. The numbers beside each accuracy function refer to K, an estimate of tracking capacity in which the units are "total number of objects" tracked successfully (Pashler, 1988).

Figure 1

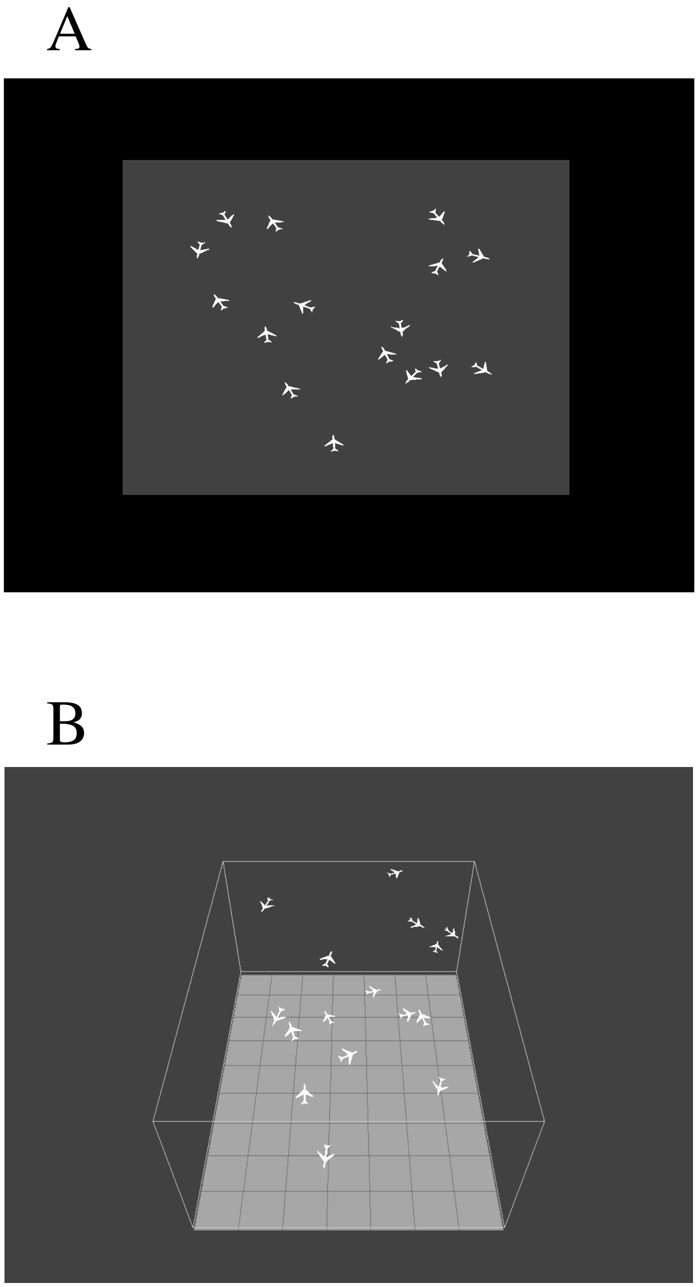


Figure 2

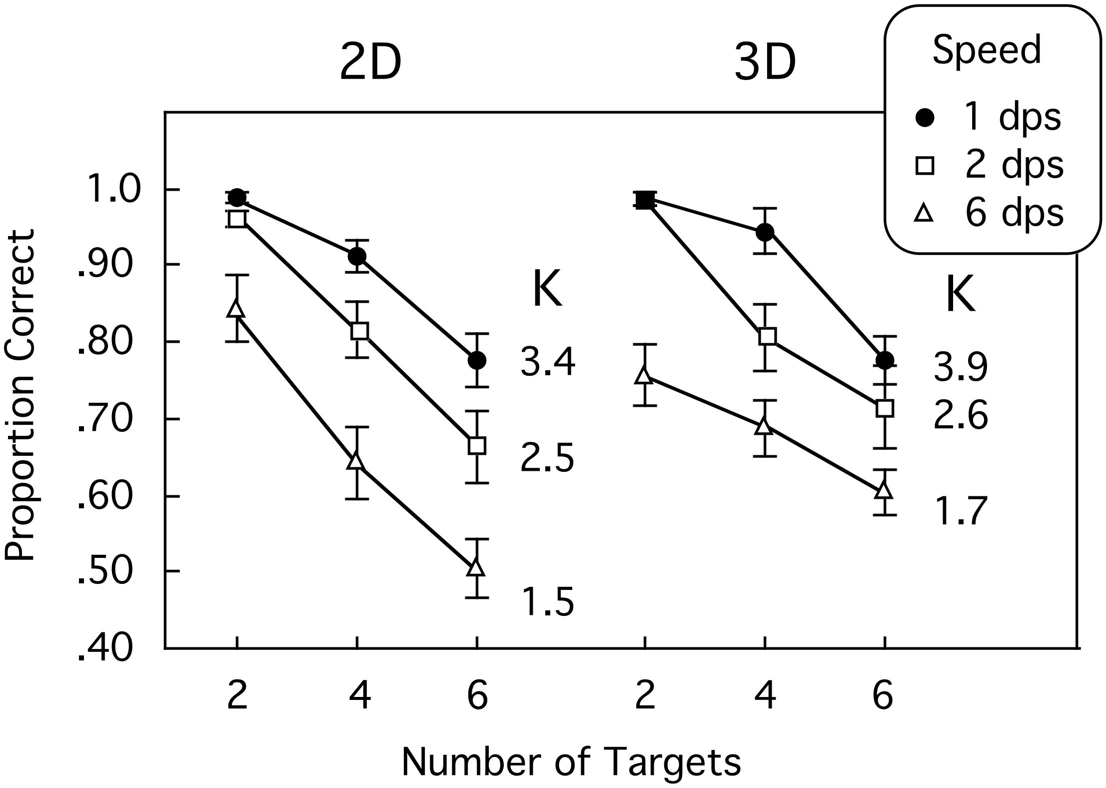

2D
3D
Speed
● 1 dps
□ 2 dps
△ 6 dps
Proportion Correct
1.0
.90
.80
.70
.60
.50
.40
K
3.4
2.5
1.5
K
3.9
2.6
1.7
2
4
6
2
4
6
Number of Targets

Figure 3

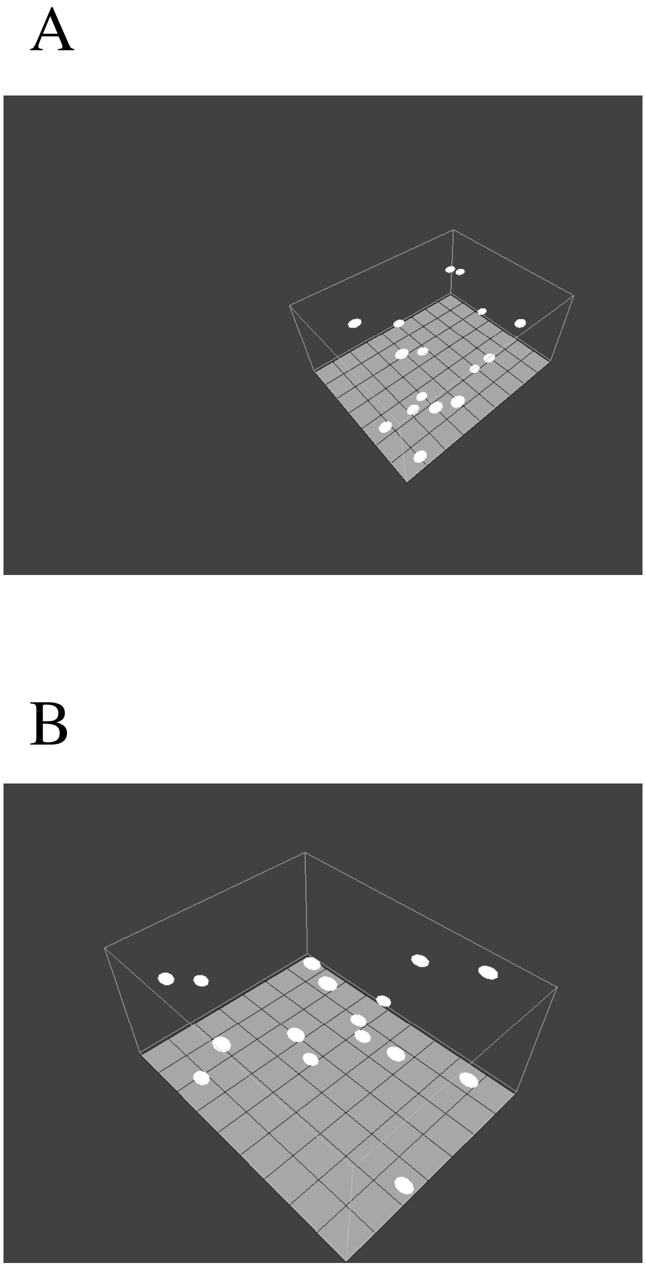


Figure 4

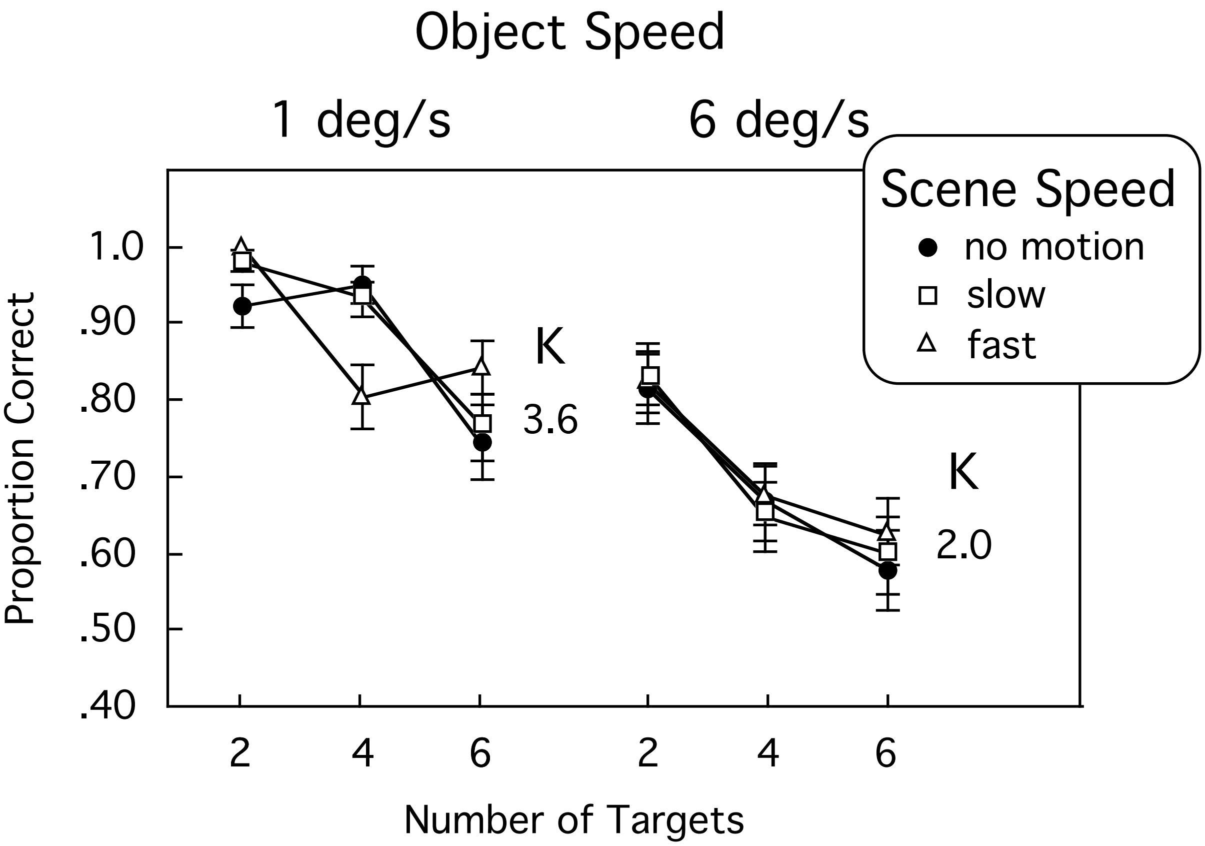

Object Speed
1 deg/s
6 deg/s
Scene Speed
● no motion
□ slow
△ fast
K
3.6
K
2.0
1.0
.90
.80
.70
.60
.50
.40
Proportion Correct
2
4
6
2
4
6
Number of Targets

Figure 5

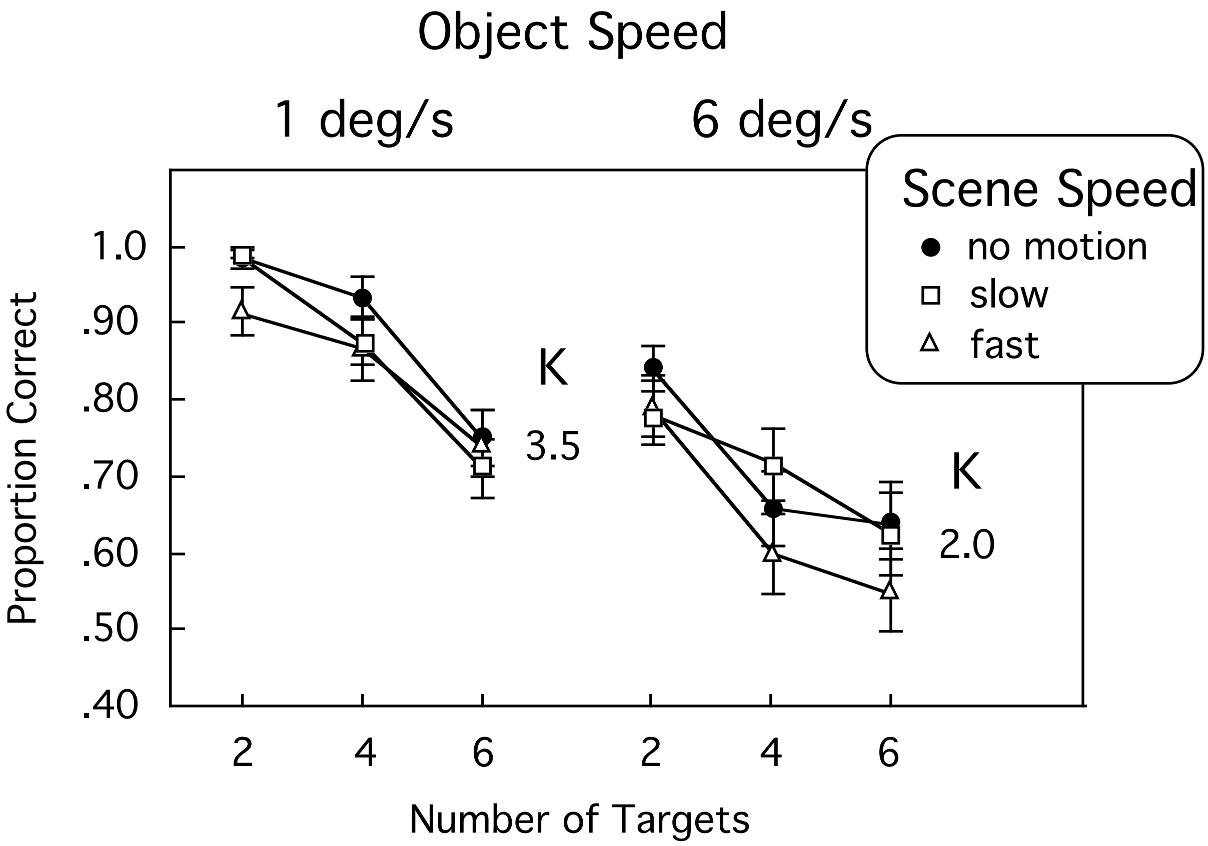

Object Speed
1 deg/s
6 deg/s
Scene Speed
● no motion
□ slow
△ fast
Proportion Correct
1.0
.90
.80
.70
.60
.50
.40
K
3.5
K
2.0
2
4
6
2
4
6
Number of Targets

Figure 6

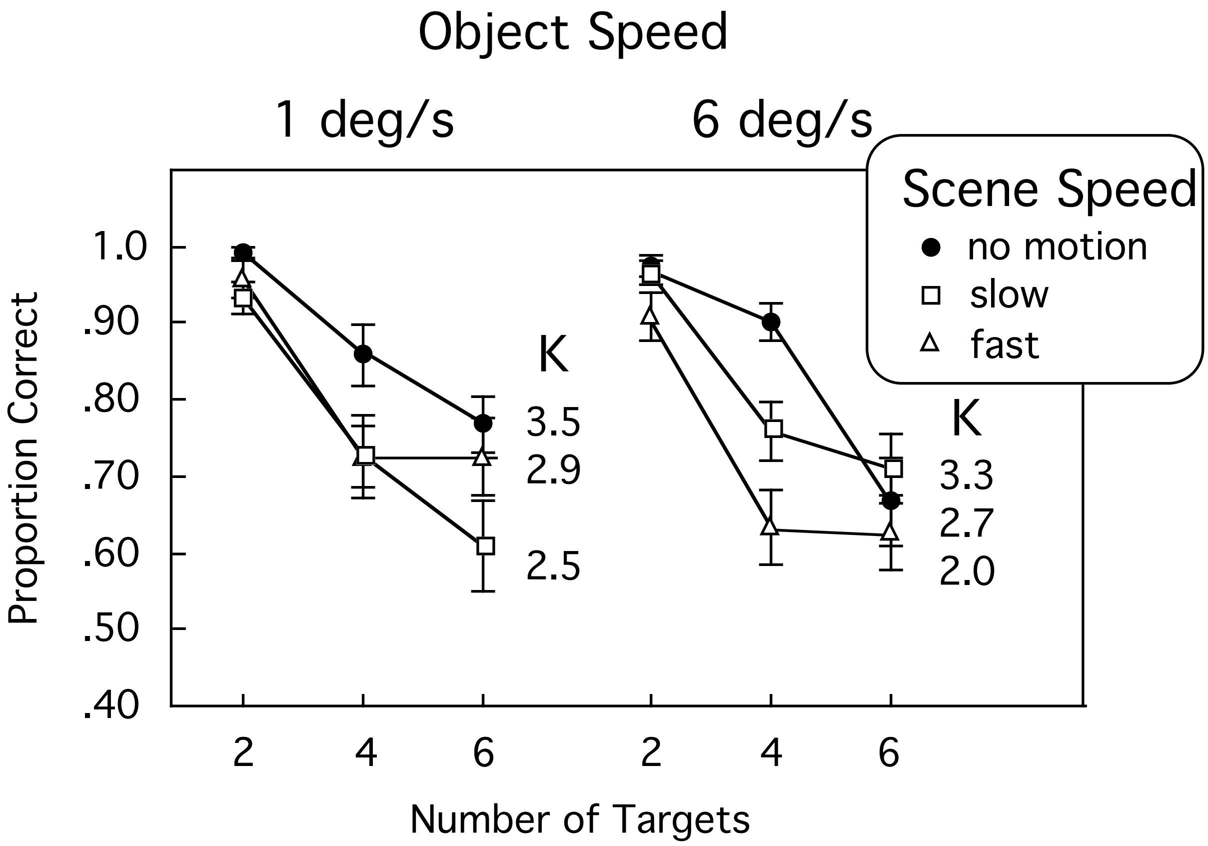

Object Speed
1 deg/s
6 deg/s
Scene Speed
● no motion
□ slow
△ fast
Proportion Correct
1.0
.90
.80
.70
.60
.50
.40
K
3.5
2.9
2.5
K
3.3
2.7
2.0
2
4
6
2
4
6
Number of Targets

Figure 7

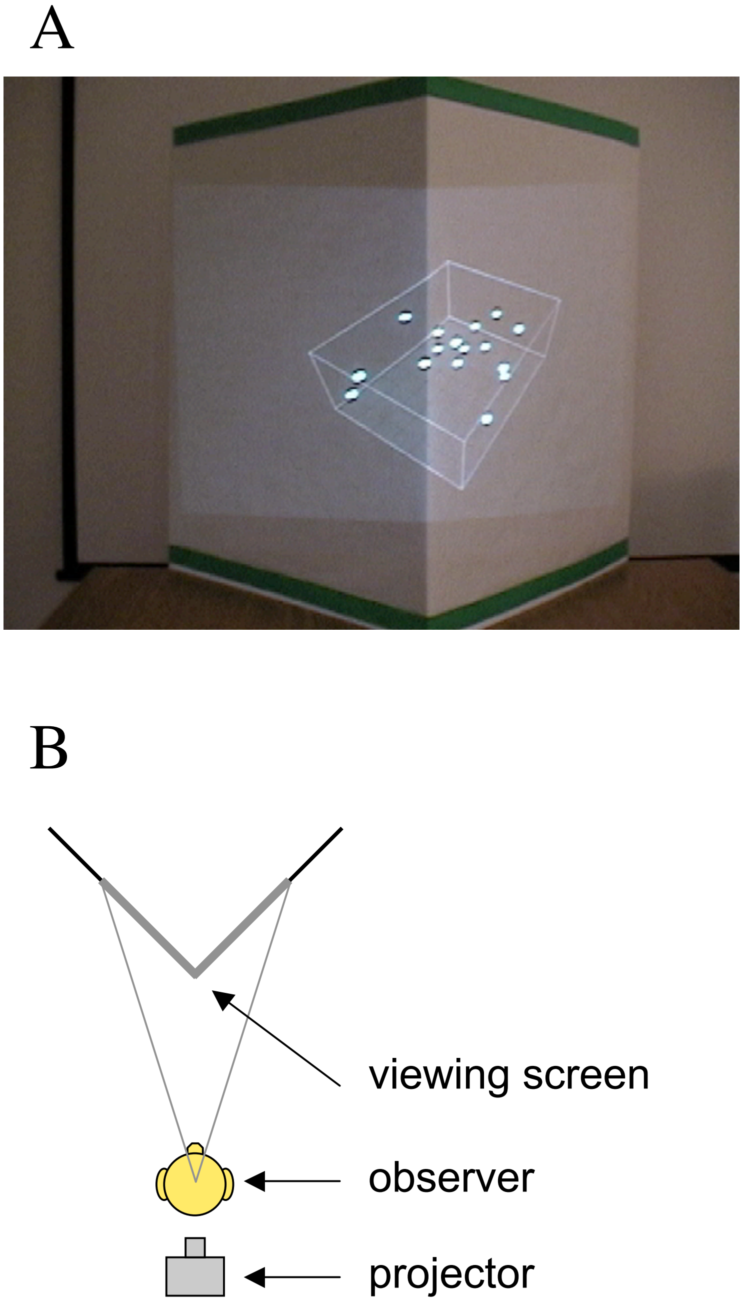
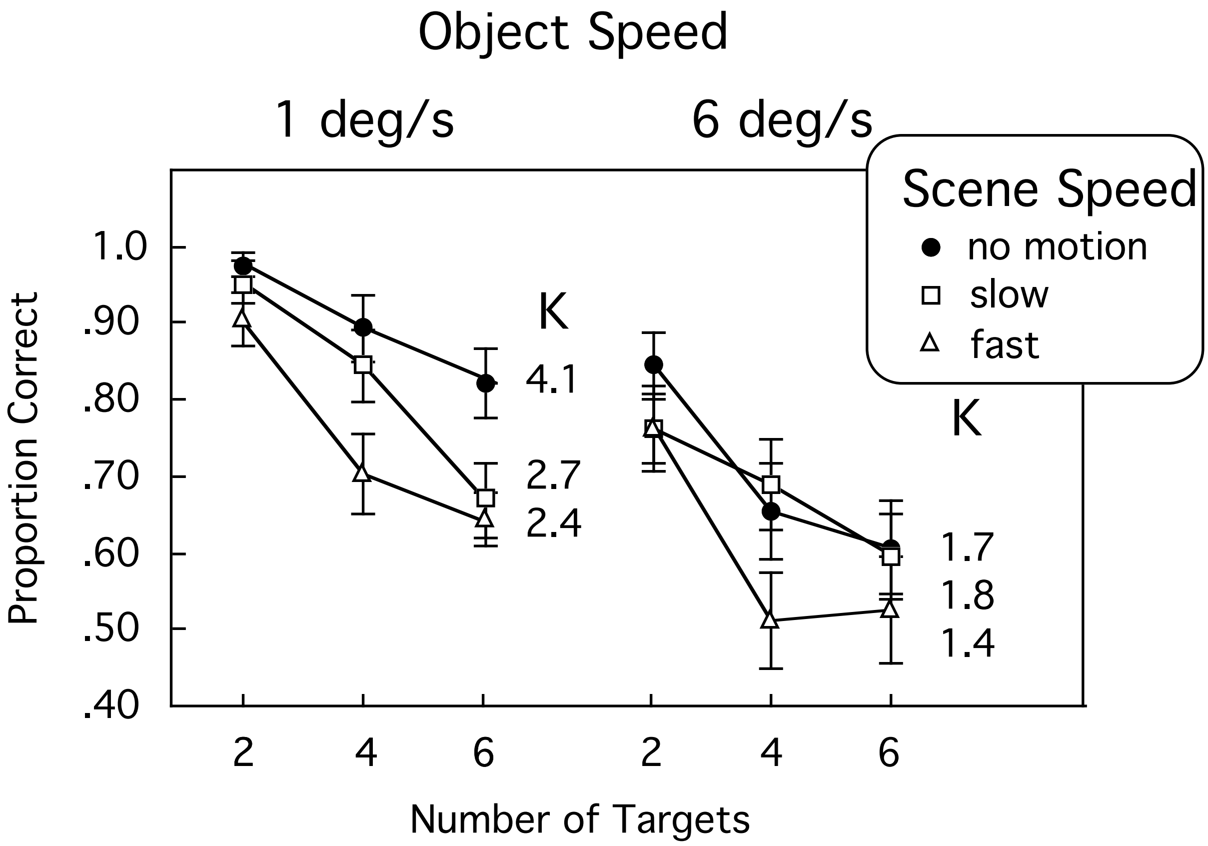


Figure 8


Object Speed
1 deg/s
6 deg/s
Scene Speed
● no motion
□ slow
△ fast
K
4.1
2.7
2.4
K
1.7
1.8
1.4
Proportion Correct
1.0
.90
.80
.70
.60
.50
.40
2
4
6
2
4
6
Number of Targets