



Published in final edited form as:

*Cognition*. 2008 July ; 108(1): 201–209. doi:10.1016/j.cognition.2007.11.008.

## Eye movements during multiple object tracking: Where do participants look?

Hilda M. Fehd\* and Adriane E. Seiffert

Vanderbilt University, Department of Psychology, 111 21st Avenue South, Nashville 37203, United States

### Abstract

Similar to the eye movements you might make when viewing a sports game, this experiment investigated where participants tend to look while keeping track of multiple objects. While eye movements were recorded, participants tracked either 1 or 3 of 8 red dots that moved randomly within a square box on a black background. Results indicated that participants fixated closer to targets more often than to distractors. However, on 3-target trials, fixation was closer to the center of the triangle formed by the targets more often than to any individual targets. This center-looking strategy seemed to reflect that people were grouping the targets into a single object rather than simultaneously minimizing all target eccentricities. Here we find that observers deliberately focus their eyes on a location that is different from the objects they are attending, perhaps as a consequence of representing those objects as a group.

### Keywords

Tracking; Eye movements; Attention

### 1. Introduction

We often need to keep track of several objects at once. If you are walking on a crowded sidewalk with your family, you will strive to keep track of them and not confuse them with any passersby. Or, if you are watching your favorite sports team play, you will be attending to the players as they move to determine which team has the advantage. The ability to track multiple objects in this way has fascinated vision researchers for almost 20 years (Cavanagh & Alvarez, 2005; Pylyshyn & Storm, 1988). Yet, one aspect of tracking multiple objects has yet to be explored – the position of eye gaze during tracking.

Eye movements made during tracking are of interest because, although it is not necessary to move one's eyes to attend to an item moving in the periphery (Verstraten, Cavanagh, & Labianca, 2000), patterns of eye movements may reveal common strategies that lead to successful tracking. How people are able to successfully track multiple objects at once is still a question of debate. The initial theory posited by Pylyshyn and Storm (1988) was that the

---

\* Corresponding author. E-mail address: hilda.m.fehd@vanderbilt.edu (H.M. Fehd).

**Publisher's Disclaimer:** This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues. Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit: <http://www.elsevier.com/copyright>

mind has a limited number of pre-attentive visual indexes (see also Pylyshyn, 2001), which represent the locations of targets and are dynamically updated as the targets move. Focal attention can access information about targets by using visual indexes as pointers, but is allocated to only one location at a time. An alternative account is that multiple foci of attention can be allocated simultaneously to multiple locations (Cavanagh & Alvarez, 2005). Attention could be divided amongst representations of multiple salient locations either independently or through grouping the targets together as one object (Yantis, 1992). These theories do not make explicit predictions about eye gaze position during tracking, however, the strategies revealed from measuring eye movements may be more consistent with one theory than another.

Participants may continually saccade from target to target, exposing them all to brief periods of high visual resolution. Landry, Sheridan, and Yufik (2001) have looked at eye movements of participants during a tracking task that simulated air-traffic control in which participants were asked to monitor moving objects for potential collisions. They found that participants made more saccades between targets of a potential collision than to other targets that were not in danger. This evidence suggests that making eye movements to targets during tracking helps participants keep track of them. The strategy of saccading from target to target may be more consistent with the visual index theory of tracking, as the theory describes that attention is serially allocated to them one at a time (Pylyshyn, 2001), which may drive eye movements to do the same.

Another possibility, however, is that participants may focus on a central location while tracking. Participants might look towards a point in the center of the target array in an attempt to minimize the eccentricity of each of the targets. This is likely to help tracking because visual acuity limitations make it more difficult to discriminate peripheral targets from distractors. In addition, the ability to individuate two nearby items, or attentional resolution, falls off steeply with eccentricity (Intriligator & Cavanagh, 2001). Thus, attempting to reduce the eccentricity of targets should aid in tracking them. Alternatively, participants might focus centrally because they are attending to all the targets as a group. Previous investigations of multiple object tracking have shown that tracking performance is improved when participants employed the strategy of mentally grouping the multiple targets into a single polygon and tracking the contorting “virtual” object as a whole (Yantis, 1992). Perhaps if people conceive of the targets as forming an object, then they may look at the center of the object formed by the targets. Eye movement experiments have shown that when participants make a saccade to an object in their periphery, they saccade to roughly the center of the object (Kowler, 1995; Vishwanath & Kowler, 2003). The strategy of focusing the eyes centrally may be consistent with the multi-focal theory of attention, because gaze would not be biased towards any one target if attention were divided amongst several objects at once.

## 2. Methods

### 2.1. Participants

Seventeen participants (9 females; aged 20–33) from Vanderbilt University participated in this experiment following the procedures for the protection of human participants defined in the APA Code (2002). Two participants' data were excluded from analysis due to signal loss by the eyetracking equipment that caused insufficient eye movement data to be acquired (see Section 3).

### 2.2. Apparatus

Eye movements were monitored using an Applied Systems Laboratory EYE-TRAC 6000 (ASL, Bedford, MA, USA) running at 120 Hz. Participants used a chinrest and headrest to sit 38.5 cm from the computer monitor. Stimuli were created with Matlab for OS X and the

Psychophysics toolbox (Brainard, 1997; Pelli, 1997). The visual display was generated by a Macintosh G4 driving a Sony Trinitron Multiscan E540 monitor.

### 2.3. Stimuli

Stimuli were 8 red dots presented within a white square frame on a black background. Each dot subtended  $2.1^\circ$  of visual angle and the frame was  $36.3^\circ$  by  $36.3^\circ$ . Green rings,  $3.0^\circ$  in diameter, were used to designate targets. Randomized starting positions were constrained to prevent dots from overlapping with each other or the bordering frame. Each dot moved in a random Brownian-like motion constrained so that each dot moved on average 1.8 pixels per frame ( $\sim 15^\circ/s$ ).

### 2.4. Procedure

Each participant completed one 50 min session containing 70 experimental and 6 practice trials. At the beginning, and after every 5 trials, the eyetracking system was calibrated using a 17-point calibration. To maintain calibration on every trial, participants kept their eyes focused on a central dot for 2 s at the beginning of each trial. Once the array of 8 dots appeared inside the frame, participants were allowed to move their eyes freely, and did so, for the remainder of the trial. Starting at the onset of the dot array, green rings designating the target(s) appeared on either 1 or 3 of the dots for 3 s. Typically, participants saccaded between targets during this cue period. Cues were removed and the dots remained stationary for another 500 ms. Dots moved for 3 s, then participants selected each target with the mouse. A high or low tone provided feedback for each correct and incorrect selection, respectively. Half of the trials had only one target, and half had three targets.

## 3. Results

Percent correct was defined as the number of trials in which all targets were correctly identified divided by the total number of trials. The average percent correct was significantly higher (99.8%) for 1-target trials than for 3-target trials (93.1%), ( $t(14) = 4.9, p < .05$ ). Trials were selected for eye movement analysis if all targets were identified correctly and less than 10% of the eye movement data was lost due to errors with the equipment, calibration, or participants' motion (such as blinks and head motion). Data from two participants were removed on the basis that their number of excluded trials was greater than 30%, indicating an unreliable eyetracking signal for these individuals. The average percentage of excluded trials for the rest of the participants was 12.3%.

Looking at the trial-by-trial data for each participant, 3-target trials seemed to fall into three different types. First, there were trials in which eye gaze stayed in approximately the same place for the duration of the motion period (Fig. 1A). Second, there were trials in which gaze seemed to pursue the overall motion of the three targets (Fig. 1B). Finally, there were trials in which gaze tended to jump from the vicinity of one target to another (Fig. 1C). The tendency to follow one of the above patterns was quantified by comparing the eye gaze position to the position of each dot.

We conducted a location competition analysis to determine the location to which gaze was closest. In this procedure, each dot was assigned a weight of zero at the beginning of the trial, and weights were then adjusted after each frame such that the dot closest to the fixated position received an increase in its weight by 7 while the remaining 7 dots received a decrease in weight of 1. The dot with the highest weight value on a given frame was then considered the winner for that frame. Summing the total time that each dot was the winner across all frames and averaging across trials, we measured the average percentage time that gaze was directed towards each dot. The advantage of this analysis is that accumulating weights provide a history

that is resistant to frame-to-frame noise in the data. Similar results were found with analysis of the proportion of time the fixation was within a specified window around each dot.

The average percentage of time in 1-target trials (Fig. 2A) that the target was the winning dot was 96.0%, which was significantly higher than the time the seven distractors were winners (0.0–0.5%;  $t(14)s > 58.37$ ,  $p < .01$  for all comparisons). The average percentage of time that a target was the winning dot in 3-target trials (Fig. 2B) was 8.5%, which was also significantly higher than the time distractors won (1.0–2.2%;  $t(14)s > 2.7$ ,  $p < .01$  for all comparisons). This was also significantly less than the percentage of time targets were winners in 1-target trials ( $t(14) = 41.83$ ,  $p < .01$ ). We examined whether participants looked at the center of the triangle formed by the targets, by including the centroid as one of the competitors (Fig. 2B). The centroid is the intersection of the medians of a triangle and is the same as the average position of the vertices in a triangle and the center of mass. The average time that the centroid won was 65.7%, which was significantly higher than the time any of the targets were winners ( $t(14)s > 8.38$ ,  $p < .01$  for all comparisons). We tested the generality of this finding in a follow-up experiment using 3, 4, or 5 targets in an array of 10 dots. Participants looked significantly more at the centroid (41.6–42.4%) than at each of the targets when tracking 3 (10.7%), 4 (9.3%), and 5 (8.1%) targets ( $t(15)s > 6.84$ ,  $p < .01$  for all comparisons). The average of the coordinates of the targets is not the same position as the centroid for 4 or 5 targets and did not account for the position of eye gaze as well as the centroid ( $t(14)s > 6.6$ ,  $p < .01$  for both comparisons). Thus, when multiple objects are tracked, more time is spent looking towards the center of the target array than at each target individually.

Participants may have been looking at the center of the target array because they perceived it as a virtual object or because they were attempting to minimize target eccentricities. We examined two different points of minimum eccentricity, one that minimized the maximum eccentricity of any one target (labeled “max”) and one that minimized the average eccentricity of all the targets (labeled “avg.”). Looking at only the frames where these points differed from the centroid by more than one degree, we found that average distance to eye gaze was less from the centroid than from both eccentricity minimizing points (4.1° vs. 4.8°,  $t(14) = 6.1$ ,  $p < .01$  for max; 4.1° vs. 5.4°,  $t(14) = 6.0$ ,  $p < .01$  for avg.). Further, using these points as competitors in place of the centroid in the competitive analysis revealed that they were winners less often than the centroid (28.4% for max and 28.6% for avg. vs. 65.7% for centroid). An additional experiment, which kept the centroid distinct from the minimizing eccentricity points on every trial, also confirmed that participants look closer to the centroid than to the points that minimize eccentricity (5.5° vs. 6.5°,  $t(15) = 6.2$ ,  $p < .01$  for max; 5.4° vs. 7.7°,  $t(15) = 7.2$ ,  $p < .01$  for avg.). Though the averages of these two minimum eccentricity definitions predicts a point that is very close to the location of the centroid, it is difficult to determine whether it is more computationally costly to calculate this sort of average of averages or the properties of the virtual object formed by the targets. We call on the principle of parsimony to conclude that people are more likely looking towards the center of mass, the centroid, of the target dots during multiple object tracking.

It should be noted that our data did produce one clear indication of eccentricity's influence on eye gaze position during tracking. On average, participants biased their gaze by  $\sim 1.2^\circ$  so that the centroid appeared in the lower visual field, where attentional resolution has been shown to be higher (Intriligator & Cavanagh, 2001). Position of gaze, then, was affected by the attentional, rather than visual, resolution of the targets.

## 4. Discussion

In this study, we investigated eye gaze position during multiple object tracking. Participants were not given instructions on where to look during the task. Results indicate that participants

tended to look more at the center of the triangle formed by the 3 targets than at any of the targets individually. The center-looking strategy seems to be a result of grouping the targets into a single object (Yantis, 1992), rather than reflecting an attempt to minimize the average or maximum target eccentricities.

Previous research has shown that eye movements during an air-traffic control task were made mostly between potential collision targets (Landry et al., 2001). This strategy differs from the pattern of results seen here, likely because of task differences. In Landry and colleagues' experiment, the task was to report potential collisions between targets, while in the current experiment only current locations of targets needed to be attended. It seems likely that requiring people to make fine-tuned discriminations of the tracked targets also would increase fixation on the targets. However, the current observations indicate that when the only task is to keep track of locations of multiple objects, center-looking is a preferred strategy.

The observation that participants look towards the center of the target array during multiple object tracking is more consistent with a multi-focal attention account of tracking than the visual index theory. Pre-attentive visual indexes that guide a single attentional focus to one location at a time would more likely produce eye movements in which targets are fixated serially. Participants in our study fixated near targets only a small percentage of the time. Attention divided simultaneously amongst several locations may lead to eye movements that attempt to satisfy all locations at once and thus pull gaze to a central location. Consistent with this account, we also observed that participants kept more of the target array in the lower visual field, possibly to take advantage of its greater attentional resolution.

We have also suggested that people look at the centroid because it reflects the conception of those targets as parts of a single virtual object. Although grouping targets in this way would not provide a mechanism to accurately track these independently moving objects, people may do it. Previous work shows that grouping targets into a single virtual object improved multiple object tracking (Yantis, 1992). There is also evidence to suggest that attention tends to concentrate near the center of an attended object (Alvarez & Scholl, 2005), so fixation may have been located centrally because attention was concentrated there. This possibility, however, would imply that attention was directed to the location of fixation, which was usually not the location of the targets in this task. That attention is directed towards fixation is supported by previous work showing that central distractors are more distracting than peripheral distractors during peripheral covert attention (Beck & Lavie, 2005; Goolkasian, 1981), and that extinguishing a fixation point leads to shorter saccade latencies to peripheral targets (Fischer & Breitmeyer, 1987; Fischer & Weber, 1993; Pratt, Lajonchere, & Abrams, 2006; Saslow, 1967). However, other evidence shows that the bias to attend at fixation can be overcome (Linnell & Humphreys, 2004). If it were the case that there was some attention directed towards fixation during tracking, it would decrease the amount of attention available for each attended target in the periphery. As such, participants are choosing to keep their gaze towards the center of the virtual object even when the computation to maintain this mental spatial relationship might be costly and it would benefit them to look directly at the targets to obtain more accurate location information. In future research, we will determine whether participants continue to look at the object center even when it may be disadvantageous, for example, when a distractor dot is presented at the centroid.

One implication of this work is that participants can accurately maintain the representation of an object's location without constant foveation. Though you may no longer see a pen that you set behind you, you still have a mental representation of its location and would be able to pick it back up without looking towards it. Here we find that observers deliberately focus their eyes on a location that is different from the objects they are attending, perhaps as a consequence of representing those objects as a group. The ability to mentally represent the world without

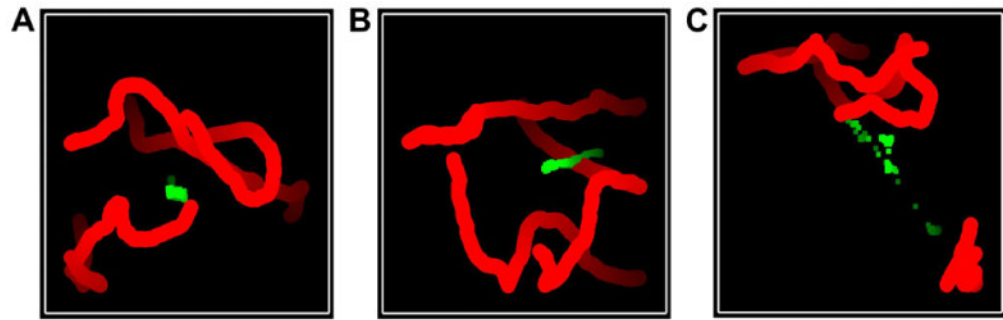
directly fixating attended objects allows us to interact successfully with our complicated surrounding environment.

## Acknowledgements

We would like to thank the editor and two anonymous reviewers for their helpful comments. This work was supported by the Vanderbilt Vision Research Center and NIH R01-EY014984.

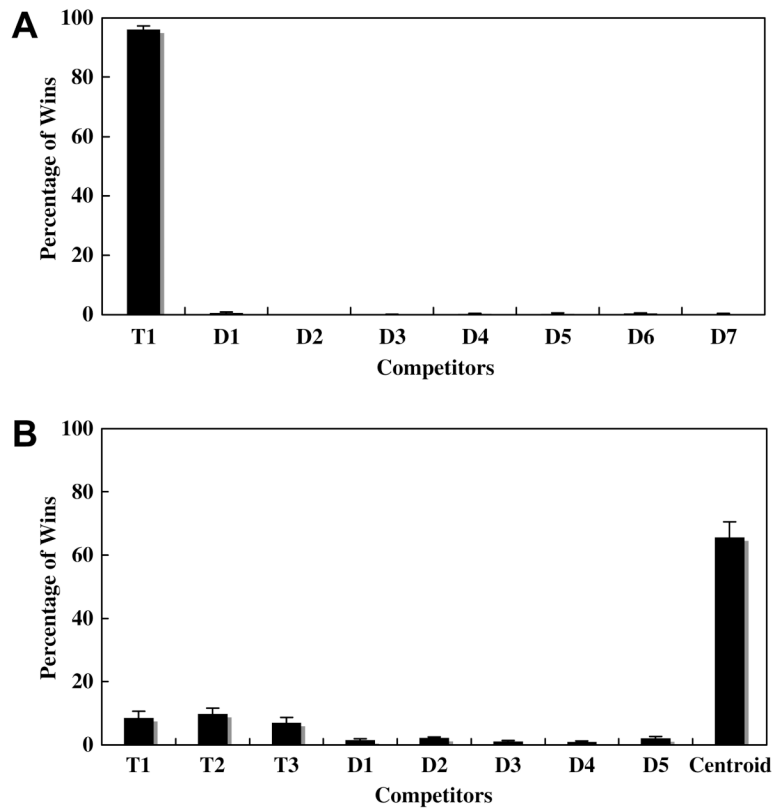
## References

- Alvarez GA, Scholl BJ. How does attention select and track spatially extended objects? New effects of attentional concentration and amplification. *Journal of Experimental Psychology: General* 2005;134(4):461–476. [PubMed: 16316286]
- APA Code of Ethics. 2002. Published online by the American Psychological Association at <http://www.apa.org/ethics/code2002.html>
- Beck DM, Lavie N. Look here but ignore what you see: Effects of distractors at fixation. *Journal of Experimental Psychology: Human Perception and Performance* 2005;31(3):592–607. [PubMed: 15982133]
- Brainard DH. The Psychophysics Toolbox. *Spatial Vision* 1997;10(4):433–436. [PubMed: 9176952]
- Cavanagh P, Alvarez GA. Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences* 2005;9(7):349–354. [PubMed: 15953754]
- Fischer B, Breitmeyer B. Mechanisms of visual attention revealed by saccadic eye movements. *Neuropsychologia* 1987;25(1A):73–83. [PubMed: 3574652]
- Fischer B, Weber H. Express saccades and visual attention. *Behavioral and Brain Sciences* 1993;16(3):553–567.
- Goolkasian P. Retinal location and its effect on the processing of target and distractor information. *Journal of Experimental Psychology: Human Perception and Performance* 1981;7(6):1247–1257. [PubMed: 6458649]
- Intriligator J, Cavanagh P. The spatial resolution of visual attention. *Cognitive Psychology* 2001;43(3):171–216. [PubMed: 11689021]
- Kowler, E. Eye movements. In: Kosslyn, SM.; Oshersohn, DN., editors. *Visual cognition*. Cambridge, MA: MIT Press; 1995. p. 215-255.
- Landry SJ, Sheridan TB, Yufik YM. A methodology for studying cognitive groupings in a target-tracking task. *IEEE Transactions on Intelligent Transportation Systems* 2001;2(2):92–100.
- Linnell KJ, Humphreys GW. Attentional selection of a peripheral ring overrules the central attentional bias. *Perception and Psychophysics* 2004;66(5):743–751. [PubMed: 15495900]
- Pelli DG. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision* 1997;10(4):437–442. [PubMed: 9176953]
- Pratt J, Lajonchere CM, Abrams RA. Attentional modulation of the gap effect. *Vision Research* 2006;46(16):2602–2607. [PubMed: 16524610]
- Pylyshyn Z. Visual indexes, preconceptual objects, and situated vision. *Cognition* 2001;80(1–2):127–158. [PubMed: 11245842]
- Pylyshyn Z, Storm R. Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision* 1988;3(3):179–197. [PubMed: 3153671]
- Saslow MG. Latency for saccadic eye movement. *Journal of the Optical Society of America* 1967;57(8):1030–1033. [PubMed: 6035297]
- Verstraten FAJ, Cavanagh P, Labianca AT. Limits of attentive tracking reveal temporal properties of attention. *Vision Research* 2000;40(26):3651–3664. [PubMed: 11116167]
- Vishwanath D, Kowler E. Localization of shapes: Eye movements and perception compared. *Vision Research* 2003;43(15):1637–1653. [PubMed: 12798146]
- Yantis S. Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology* 1992;24(3):295–340. [PubMed: 1516359]



**Fig. 1.**

Traces of targets positions and the point of fixation during 3 trials. The trajectories of targets (red dots) and the point of fixation (green squares) are shown in three example 3-target trials (A–C). Distractor dots are not represented in these graphs. Time course of the trial is represented by the brightening of the color, such that the locations of the dots and eye fixation at the beginning of the trial are shown in dark colors and the locations at the end of the trial are shown in bright colors. Typically eye movements followed a pattern of either (A) staying in roughly the same place throughout the trial, (B) pursuing the general motion of the targets, or (C) saccading rapidly between targets.



**Fig. 2.**

Percentage of wins for each dot. A location competition analysis was conducted that classified a dot as a winner if it was the closest to eye gaze at a given point in time. The percentage of time that each competitor was the winner is shown for (A) 1-target and (B) 3-target trials with the centroid included. Targets (T1, T2, and T3) won the competitive analysis more often than distractors (D1–D7), though the centroid won most often.