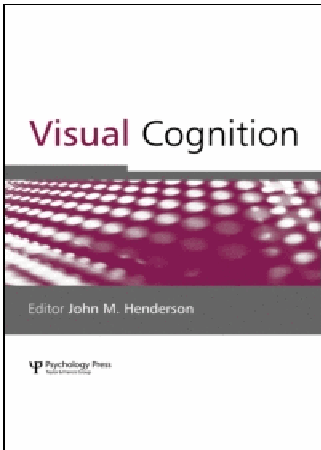


Visual Cognition
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683696

# An eye movement analysis of multiple object tracking in a realistic environment

Gregory J. Zelinsky [a]; Mark B. Neider [b]
[a] Department of Psychology, Stony Brook University, Stony Brook, NY, USA
[b] Beckman Institute, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, USA

First Published on: 01 April 2008

To cite this Article: Zelinsky, Gregory J. and Neider, Mark B. (2008) 'An eye movement analysis of multiple object tracking in a realistic environment', Visual Cognition, 16:5, 553 — 566

To link to this article: DOI: 10.1080/13506280802000752
URL: http://dx.doi.org/10.1080/13506280802000752

VISUAL COGNITION, 2008, 16 (5), 553–566

Psychology Press
Taylor & Francis Group



# An eye movement analysis of multiple object tracking in a realistic environment

Gregory J. Zelinsky
*Department of Psychology, Stony Brook University, Stony Brook, NY, USA*

Mark B. Neider
*Beckman Institute, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, USA*

To study multiple object tracking under naturalistic conditions, observers tracked 1–4 sharks (9 in total) swimming throughout an underwater scene. Accuracy was high in the Track 1–3 conditions (>92%), but declined when tracking 4 targets (78%). Gaze analyses revealed a dependency between tracking strategy and target number. Observers tracking 2 targets kept their gaze on the target centroid rather than individual objects; observers tracking 4 targets switched their gaze back-and-forth between sharks. Using an oculomotor method for identifying targets lost during tracking, we confirmed that this strategy shift was real and not an artifact of centroid definition. Moreover, we found that tracking errors increased with gaze time on targets, and decreased with time spent looking at the centroid. Depending on tracking load, both centroid and target-switching strategies are used, with accuracy improving with reliance on centroid tracking. An index juggling hypothesis is advanced to explain the suboptimal tendency to fixate tracked objects.

Many everyday tasks, such as navigating a busy city street or watching a sporting event on television, require the coordinated contribution of a deceptively large number of visual and cognitive processes. Features must be bound into objects and these objects must be segregated from a background (e.g., Neider & Zelinsky, 2006b; Prinzmetal, 1981; see Shipley & Kellman, 2001, for a review). Indexing and memory processes must then link the

Please address all correspondence to Gregory J. Zelinsky, Department of Psychology, Stony Brook University, Stony Brook, NY 11794-2500, USA.
E-mail: Gregory.Zelinsky@stonybrook.edu

This work was supported by the National Science Foundation under Grant IIS-0527585 to GJZ. We thank Andrei Todor for creating the shark motion sequences and for help with data analysis, and Todd Horowitz, Brian Scholl, and Jon Flombaum for comments on an earlier draft of this paper.


http://www.psypress.com/viscog DOI: 10.1080/13506280802000752

different views of an object so as to experience coherent object motion (e.g., Kahneman, Treisman, & Gibbs, 1992; Pylyshyn, 1989; see Scholl, 2001, for a review). Finally, attention-related processes must prioritize and tag these objects so as to confine computations to only the task relevant items (Alvarez & Scholl, 2005; Cavanagh & Alvarez, 2005; see also Ullman, 1984).

The field of visual cognition devoted to understanding these interrelated processes is tracking, and one of the principle paradigms used to study tracking in the laboratory is the multiple object tracking (MOT) task (Pylyshyn & Storm, 1988; Yantis, 1992). Using this task, researchers have made great progress in understanding the types of object transformations and visual interruptions through which we can, and cannot, track (e.g., Scholl & Pylyshyn, 1999; vanMarle & Scholl, 2003), and more generally the factors affecting the creation of perceptual objects (e.g., Scholl & Pylyshyn, 1999; Scholl, Pylyshyn, & Feldman, 2001).

Our work builds on these earlier efforts by asking *how people track*—what strategies are used to track multiple moving targets and how do these strategies impact tracking performance? To answer these questions we supplement an error analysis of tracking performance with analyses of gaze behaviour. Our general approach parallels the use of oculomotor variables to study visual search. Although people can often search accurately without making eye movements, we ordinarily do make eye movements as we search, and analyses of these movements have revealed a great deal about underlying search processes (Chen & Zelinsky, 2006; Neider & Zelinsky, 2006a, 2006b; Zelinsky, 1996; Zelinsky, Rao, Hayhoe, & Ballard, 1997; Zelinsky & Sheinberg, 1997). Similarly, although accurate tracking may be possible without eye movements (e.g., Pylyshyn & Storm, 1988), people tracking freely may elect to move their eyes, and the information to be gleaned from these behaviours remains largely unexplored.[1] Our goal is to analyse these eye movements, to the extent that they occur, to observe tracking behaviour as observers are actually engaged in tracking. We will attempt to define tracking strategies in terms of oculomotor measures, and evaluate the use of these strategies on a trial-by-trial basis.

## METHOD

We explored MOT using sharks moving in a rendered underwater scene. Like the simpler stimuli upon which the tracking literature was built, sharks are visually very similar, and the motion of any one can be independent from the

[1] Landry, Sheridan, and Yufik (2001) analysed eye movements in a mock air traffic control task, but to date there have been no journal publications describing eye movements in the context of the MOT paradigm. However, several investigators have independently started to engage this topic, and this work has recently been presented at meetings of the Psychonomics and Vision Sciences Societies (Doran, Hoffman, & Scholl, 2006; Fazl & Mingolla, 2007; Fehd & Seiffert, 2007; Zelinsky, Neider, & Todor, 2007).

others. An underwater scene is also free of any contextual constraints (e.g., confinement to roadways; see also Neider & Zelinsky, 2006a; Torralba, Oliva, Castelhano, & Henderson, 2006) that might complicate the interpretation of real-world tracking. However, shark tracking is closer to real-world applications in that our sharks are perceived as moving in depth as well as changing position on a screen (see also Liu et al., 2005, and Viswanathan & Mingolla, 2002). They are also featurally more complex than dots, providing information that might be used by a tracking process. For these reasons we believe our tracking task describes a useful middle ground between traditional laboratory experiments and tracking in the real world.

Nine identical sharks swam for 20 s throughout an underwater scene. Stimuli were created, animated, and rendered using Autodesk's 3D Studio Max. Shark movement was pseudorealistic, and confined to the swim volume illustrated in Figure 1. The events comprising a typical trial are illustrated in Figure 2.[2] The task was standard MOT; indicate whether a probed shark was one of the targets designated prior to the motion sequence. Stimuli were presented in colour, and DirectDraw functionality was used to display the 600 frames comprising each motion sequence at 30 frames/s. Eye position was recorded throughout using an EyeLink II eyetracker, sampling at 500 Hz (with chinrest).

Task difficulty was manipulated across four target conditions; observers had to track one, two, three, or four of nine sharks over the 20 s motion sequence. We also manipulated target occlusion. In the no-occlusion condition none of the target sharks were occluded during their motion, nor did they occlude any nontarget shark. In the occlusion condition at least one target shark occluded, or was occluded by, at least one other shark.[3] The occlusion manipulation was a within-subjects variable; the 112 trials per observer were evenly divided into occlusion and no-occlusion conditions. Target condition was a between-subjects manipulation, with seven observers participating in each of the four conditions (28 participants in total).

## RESULTS AND DISCUSSION

### How well did people track?

Tracking accuracy decreased with the number of targets, 98.9%, 96.6%, 92.4%, and 77.6%, in the Track 1–4 conditions, respectively, $F(3, 24) = 11.43$, $p < .001$. This decrease was most precipitous between the Track 3 and Track

[2] See also http://www.psychology.sunysb.edu/gzelinsky-/movies/mot/Track_4_trial.avi for a movie of a typical Track 4 trial.

[3] Occlusions were defined as bounding box intersections in the image plane occurring over at least three consecutive frames.

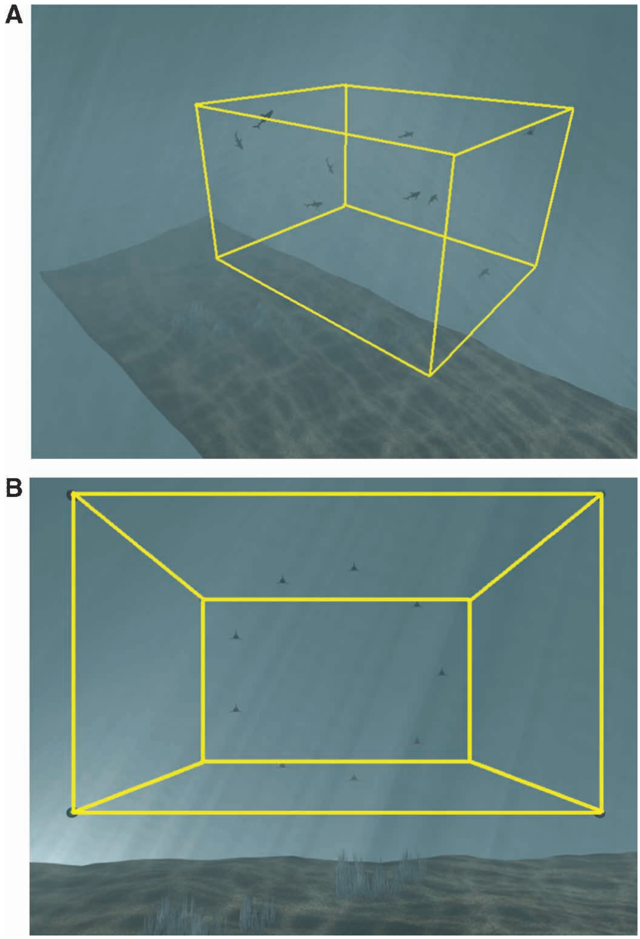


**Figure 1.** (A) A delineation of the swim volume. This volume subtended a visual angle of $25^{\circ} \times 15^{\circ}$, which corresponded to a virtual space of 100 (width) $\times$ 60 (height) $\times$ 80 (depth) m. Individual sharks subtended a maximum and minimum visual angle of $1.13^{\circ}$ and $0.48^{\circ}$, respectively, depending on the shark's perceived depth in the scene. Except for constraints preventing sharks from colliding with each other and bouncing abruptly off of the virtual boundaries of the swim volume, shark trajectories were random and shark velocity was a constant 15 m/s, corresponding to a maximum velocity of $1.13^{\circ}$/s in the image. (B) The swim volume as viewed from the observer's perspective. Note that it did not intersect the ground plane, meaning that sharks were always viewed against a background of water. To view this figure in colour, please see the online issue of the Journal.

4 conditions, $t(12) = 2.67$, $p = .02$. Such an abrupt drop in accuracy suggests a limit on the number of objects that can be simultaneously tracked, which in our task was about three sharks. Table 1 groups the accuracy data by occlusion condition. Although Track 3 accuracy was lower with occlusions,

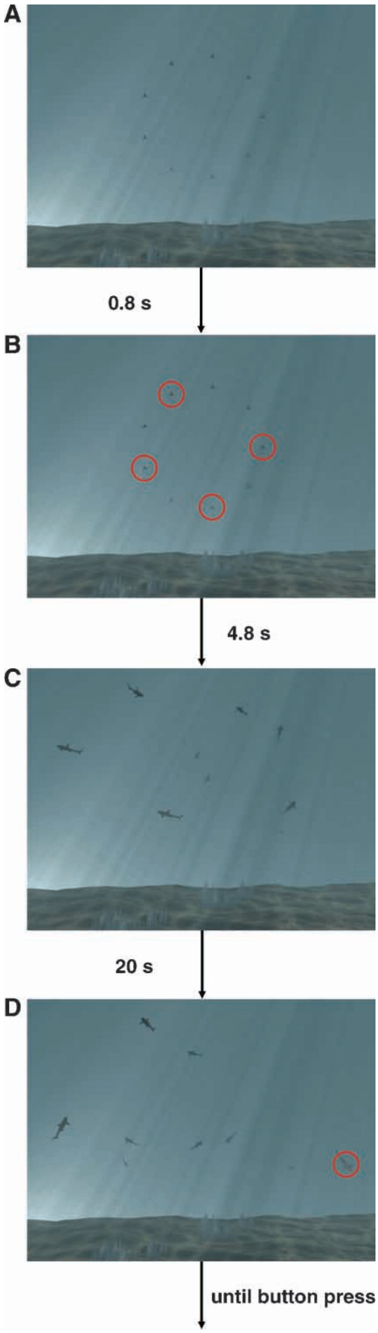


**Figure 2.** (A) Each trial began with the presentation of a target frame, showing 9 sharks arranged into a circle and oriented so as to have visually identical shapes. (B) One to four target sharks were then designated by flashing red rings. Target positions were balanced, as much as possible, over the 9 target frame positions in the 4 target conditions. (C) All sharks swam throughout the scene for 20 s. (D) Upon reaching the end of the motion sequence (frame 600), objects froze in position and a red ring was drawn around one of the sharks. On half of the trials this probed shark was one of the targets; on the other half of the trials a nontarget shark was probed. The observer's task was to indicate whether this probed object was, or was not, one of the targets. The probe display remained visible until the judgement, and observers were instructed to respond as accurately as possible without regard for time. There was no accuracy feedback. To view this figure in colour, please see the online issue of the Journal.

TABLE 1
Percentage correct tracking accuracy as a function of number of targets and occlusion condition

| | *Number of targets* | | | |
|---|---|---|---|---|
| | *1* | *2* | *3* | *4* |
| No occlusion | 98.6 (0.7) | 98.0 (0.8) | 95.4 (1.5) | 79.3 (5.9) |
| Occlusion | 99.3 (0.4) | 94.8 (2.0) | 89.0 (2.1) | 75.9 (5.3) |

Values in parentheses indicate one standard error of the mean (SEM).

$t(6) = 2.45$, $p = .05$, this effect was small (6%). There were no significant effects of occlusion in the other tracking conditions, $t(6) \leq 1.73$, $p \geq .14$. We interpret these patterns as being broadly consistent with the MOT literature; observers tracked successfully through occlusions, but good tracking performance was limited to a small number of targets (Scholl et al., 2001).

## What strategies did people use?

To explore how people tracked in our task we considered the use of three potential strategies. In the *look-at-one strategy*, observers would track one target with their gaze, while keeping track of the other targets in their peripheral vision. This strategy predicts that the maximum time spent looking at one target under Track 2–4 conditions should roughly equal the time spent looking at the single target under Track 1 conditions. In the *target switching strategy*, observers would keep switching their gaze between the tracked sharks. Exclusive use of this strategy predicts that the total time spent looking at targets under Track 2–4 conditions should be comparable to the total time on target from the Track 1 condition. A switching strategy also predicts that the total time on targets should be longer than the total time spent looking at nontargets or the background. Lastly, observers using a *centroid strategy* would keep their gaze on the averaged spatial position of the targets, and not distribute their gaze between the targets. Similar to the switching strategy, a centroid strategy predicts that the total time on centroid under Track 2–4 conditions should roughly equal the time on target from the Track 1 condition, and be longer than the total time spent looking at targets in the Track 2–4 conditions.

To begin teasing apart tracking strategies we first calculated the maximum time spent looking at any one of the targets (Table 2). Predictably, the maximum time on target was quite high in the Track 1 condition; on average, observers held their gaze on the target for 18.7 s out of each 20 s

TABLE 2
Average maximum time (in s) spent looking at any one target during the motion sequence of a tracking trial, by target condition

| *Number of targets* | | | |
|---|---|---|---|
| *1* | *2* | *3* | *4* |
| 18.7 (0.9) | 5.6 (0.7) | 4.7 (0.3) | 5.3 (0.8) |

Values in parentheses indicate one standard error of the mean (SEM).

trial.[4] We found a very different pattern in the Track 2–4 conditions, where observers looked at any one target for only 5.2 s, on average. These data suggest that observers used a look-at-one strategy only in the minimal task of tracking one target; this strategy was not used for multiple target tracking.

To assess the use of switching and centroid tracking strategies in our task we first assigned each eye position sample to one of three events: Looking at the centroid, looking at a target, or looking at a nontarget. These assignments were made using a shortest distance rule, meaning that if the x,y gaze coordinate on a given sample was closest to the geometric centroid of the target shark configuration, that sample would be counted as a look to the centroid. Using this rule, each of the 10,000 gaze samples during a 20 s motion sequence could be uniquely assigned to a centroid, target, or nontarget event.

Table 3 plots the average time during the 20 s motion sequence devoted to each of these three events. Consistent with the centroid tracking strategy, the total time spent looking at the centroid in the Track 2 condition was longer than the gaze time devoted to either the targets, $t(6) = 3.0$, $p < .05$, or the nontargets, $t(6) = 7.13$, $p < .001$.[5] This centroid preference disappeared in the Track 3 condition.[6] Observers tracking three targets looked at the centroid and targets about equally, $t(6) = 0.42$, $p > .1$, and both were preferred relative to the nontargets, $t(6) = 2.67$, $p = .05$. However, under Track 4 conditions the cumulative time spent looking at the targets was longer than the total time on centroid, $t(6) = 2.62$, $p < .05$. Indeed, gaze time on the centroid in this condition no longer differed from the nontarget viewing time, $t(6) =$

[4] A representative Track 1 trial can be viewed at: http://www.psychology.sunysb.edu/gzelinsky-/movies/mot/Track_1_data.avi The target shark in this movie is indicated by a dynamic yellow box displayed around the object. The instantaneous eye position is indicated by the red dot.

[5] A representative Track 2 trial illustrating centroid tracking can be viewed at: http://www.psychology.sunysb.edu/gzelinsky-/movies/mot/Track_2_data.avi The geometric centroid of the target configuration is indicated by the green dot.

[6] http://www.psychology.sunysb.edu/gzelinsky-/movies/mot/Track_3_data.avi

TABLE 3
Average total time (in s) spent looking at the centroid, targets, and nontargets during the motion sequence of a trial, by target condition

| | *Tracking strategy* | | |
|---|---|---|---|
| *Target condition* | *Centroid* | *Targets* | *Nontargets* |
| Track 2 | 9.4 (1.0) | 7.4 (1.0) | 3.3 (0.3) |
| Track 3 | 7.7 (0.9) | 8.3 (0.6) | 4.1 (0.4) |
| Track 4 | 4.7 (0.9) | 10.3 (1.3) | 5.1 (0.7) |
| Track 4—lost target | 7.1 (0.5) | 8.7 (0.7) | 4.3 (0.5) |
| Track 4—no lost targets | 2.2 (0.4) | 11.9 (1.1) | 6.0 (0.8) |

Values in parentheses indicate one standard error of the mean (SEM).

0.44, $p > .1$. This pattern reflects movement to a switching strategy, characterized by a frenetic shifting of gaze between targets.[7]

## Can ''lost targets'' be identified during tracking?

The previously described patterns suggest a movement from a centroid strategy at low tracking loads to a target switching strategy at higher tracking loads, but might this apparent strategy shift be an artifact of how we defined a centroid? Recall that the centroid computation assumed that all of the targets were tracked, but this is not always the case, particularly under Track 4 conditions where one or more targets might often be lost. Rather than abandoning a centroid tracking strategy under high loads, perhaps observers in the Track 4 condition instead abandoned one of the tracked sharks, and tracked the centroid of the remaining three.

To evaluate this possibility we first define an oculomotor marker to identify targets not tracked during a trial. We define such a *lost target* as one inspected by gaze for less than 350 ms during 20 s of motion, under the assumption that actively tracked targets, even those tracked using a centroid strategy, will be inspected for at least this long during a trial.[8] Supporting this assumption we found that lost target trials were relatively rare (about 10%) in the highly accurate Track 2–3 conditions. However, using the same criterion a lost target was identified in 52% of the Track 4 trials. We further validate this measure by showing a relationship between lost targets and tracking accuracy as a function of probe. Given the paucity of cases with fewer targets, this

[7] http://www.psychology.sunysb.edu/gzelinsky-/movies/mot/Track_4_data.avi

[8] A deliberately short criterion was chosen so as to capture cases in which targets are lost early in a trial, thereby creating a cleaner separation between lost target and no lost target groups. However, in exploring the data we found that qualitatively similar group differences are obtained if this criterion is allowed to vary from 1 to 1000 ms.

analysis is restricted to the Track 4 condition. When the probed shark was inspected for longer than 350 ms during tracking (i.e., not lost), accuracy was high (81%). Accuracy was also high (79%) when there was a lost target, but a different target was probed. However, when the probed target was lost during tracking, accuracy dropped to near chance performance (55%). We can therefore be reasonably confident that our measure of a lost target is a meaningful indicator that a given target was not successfully tracked.

Having defined a valid oculomotor marker for lost targets, we grouped the correct Track 4 data into lost target trials (i.e., at least one of the four targets were lost) and no lost target trials, and recomputed centroid, target, and nontarget gaze events. Importantly, lost targets were now excluded from the centroid computation, meaning that centroids in the lost target trials were typically based on only three targets. These data appear in the bottom two rows of Table 3. Analyses of lost target trials showed a roughly even split between centroid and target tracking strategies, $t(6) = 1.32$, $p > .1$, similar to what was reported for the Track 3 condition. This is what would be expected if observers were treating a Track 4 trial with a lost target the same as a Track 3 trial. More central to the current discussion are the no lost target data, where we see a clear preference to look at targets rather than the four-target centroid, $t(6) = 6.68$, $p < .001$. The time spent looking at the centroid in fact decreased by more than half after removing lost targets, relative to the original Track 4 data. This would not be expected if observers were using a centroid tracking strategy; when we improved our estimate of the centroid we found less evidence for centroid tracking. The strategy shift from centroid tracking with two targets to target tracking with four targets is therefore real, and is not an artifact of how we defined the target centroid.

## Which tracking strategy is best?

If a strategy is beneficial to tracking, one would expect fewer errors with more engagement in that strategy. This was clearly not the case for target tracking (Figure 3A). Indeed, we found the opposite relationship; the longer observers spent looking at targets, the worse their tracking performance, $r = .62$, $p < .01$. Separate within-condition analyses revealed correlations of .78, .84, and .37 for the Track 2–4 conditions, respectively. In contrast to target tracking, evaluation of the centroid strategy yielded evidence for a tracking benefit (Figure 3B). The longer observers spent looking at the centroid, the lower their error rates, $r = -.82$, $p < .001$. Again, separate analyses suggest similar relationships within each tracking condition (.85, .88, and .75 for the Track 2–4 conditions, respectively). We also computed Pashler (1988) *k*s for the Track 4 observers who could be clearly classified as either target-trackers or centroid-trackers, and found capacity estimates of 1.7 and 3.7 sharks,

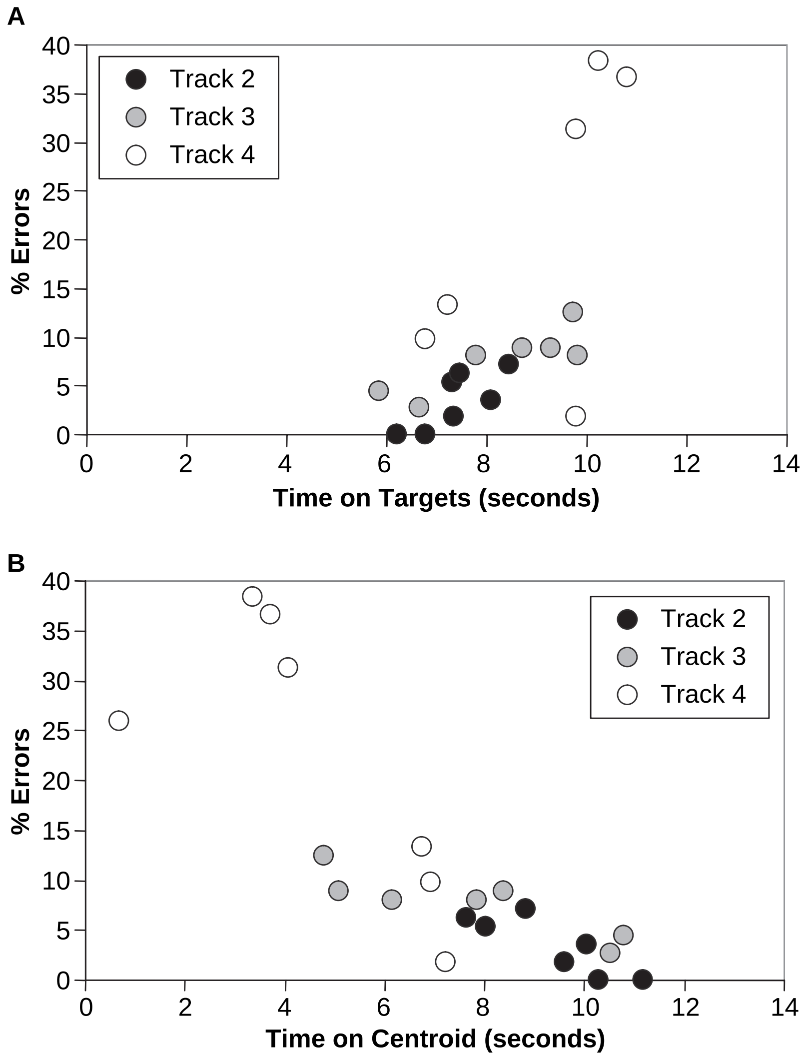


**Figure 3.** Errors for each observer in the Track 2–4 conditions, plotted as a function of (A) the total time spent looking at the targets, and (B) the total time spent looking at the centroid.

respectively.[9] To the extent that observers used a centroid strategy, their tracking ability was good; to the extent that they attempted to keep their gaze on individual targets, their tracking ability was comparatively poor.

[9] Capacity estimates were not calculated for the Track 1–3 conditions, as these might be artificially limited by the small number of objects to track.

## GENERAL DISCUSSION

This study explored how people move their gaze during a free-viewing MOT task. Compared to other visual-cognitive tasks, oculomotor measures have been underutilized in the study of MOT. Motivating this relative neglect may be an assumption that gaze measures will not reveal core processes used in MOT, or perhaps the belief that this effort would be wasted due to the eyes moving minimally in a tracking task. As demonstrated in this report, both of these assumptions are misguided; gaze certainly does move during tracking, and these movements can reveal a wealth of information about the tracking process.

In the current study this information took the form of three contributions. First, using gaze analyses to identify and quantify distinct tracking strategies, we showed that observers preferred a centroid strategy when tracking two targets, but shifted to a target switching strategy when tracking four targets. However, whereas this shift suggests a clear load-dependency on tracking strategy, it is also true that the time devoted to any one strategy was short compared to the time spent looking at the single target in the Track 1 condition, suggesting considerable strategy mixing. Rather than rigidly relating a particular tracking strategy to a particular load condition, tracking strategy might be better conceptualized as a continuum, with observers fluidly moving from one strategy to another depending on the momentary demands of the tracking task.

Second, in addition to identifying strategies, eye movements can be used to identify targets that are lost during tracking. We found that if a person failed to look at a given target while tracking, they were likely to misreport this object as a nontarget. The significance of identifying lost targets extends beyond the current context, which was to confirm the movement away from centroid tracking under high loads. Using our gaze-based method, one can obtain direct evidence for a specific lost target on a specific trial, making it unnecessary to infer the proportion of lost targets based on capacity estimates from aggregated error data. This specificity is important, as it enables a finer-grained understanding of the specific factors contributing to track failures.

A third finding from our study is that tracking tends to be accurate when observers look at the centroid, but relatively error prone when gaze shifts back-and-forth between targets. This is what one would expect if observers were grouping the positions of multiple targets, then tracking the deformations of this composite *virtual object* (Yantis, 1992). Our work extends this tracking theory in two respects. First, if the centroid tracked by our observers corresponds to the centre of this virtual object, it may be possible to recover the shape of this object from eye position, thereby indicating, moment-by-moment, the targets that are being tracked. Second, our work qualifies this theoretical perspective by showing that centroid tracking tends to break down under high tracking loads. One reason for this breakdown may be that virtual object tracking becomes less reliable with larger numbers of targets. For example, with four targets, one will often be contained within the triangle formed by the other three. To the extent that observers are biased to perceive simple shapes, this may lead to errors. Alternatively, a limitation may exist on the number of targets that can be grouped into a virtual object. Assuming idiosyncratic grouping constraints, observers might adopt the strategy of abandoning one or more targets until their personal limit is reached. Our current data cannot distinguish between these possibilities.

These explanations address why observers tended to shift away from centroid tracking under high loads, but neither explains why they often adopted the suboptimal strategy of looking at individual targets. We believe that the decision of which tracking strategy to use, and when to switch from one to the other, depends on the observer's own criterion for abandoning targets. As already noted, if an observer is willing to tolerate a lost target or two, they could conceivably continue using a centroid strategy even under high tracking loads. But what if they are unwilling to let go of a target?

We speculate that our evidence for a target switching strategy reflects an observer's effort to *not* abandon targets by juggling indices to tracked objects. It is widely believed that there exists a capacity limit on the number of indices that are available for tracking (e.g., Kahneman et al., 1992; Pylyshyn & Storm, 1988; see Cavanagh & Alvarez, 2005, for a review). However, it may be possible to momentarily exceed this limitation by dynamically reshuffling a small number of indices over a larger target set.[10] For example, assuming that an observer can comfortably track only three targets simultaneously in our task, she may temporarily deallocate an index to target $x$ so as to maintain track on the other three. Then, based on low-level visual features and a representation of $x$'s trajectory (e.g., Fencsik, Klieger, & Horowitz, 2007), target $x$ can be reindexed at a later moment in time, accompanied by the temporary deallocation of the index to one of the other tracked targets. Note also that this framework does not require the existence of "hidden indices"; during the momentary deallocation of an index to a target, track on that target is lost. By combining *index juggling* with a memory for a target's motion, it is therefore possible to temporarily achieve supercapacity tracking; $m+n$ targets are tracked despite $m$ indices. If one then assumes that reindexing an object occurs most efficiently when that object is fixated by gaze, this might explain the tendency to look at targets under high loads.

[10] Such an explanation dovetails nicely with our postexperiment questioning of observers, who often reported a feeling of panic when tracking four targets.

Future work will test the assumptions of index juggling, focusing on the relationship between gaze and the reindexing of moving objects, and on the delay over which trajectory memory can serve this reindexing operation. Individual differences in tracking will also be studied, particularly with respect to a person's tolerance for lost targets, as this may ultimately determine their tracking strategy.

## REFERENCES


Alvarez, G. A., & Scholl, B. J. (2005). How does attention select and track spatially extended objects? New effects of attentional concentration and amplification. *Journal of Experimental Psychology: General*, *134*, 461–476.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, *9*, 349–354.

Chen, X., & Zelinsky, G. J. (2006). Real-world visual search is dominated by top-down guidance. *Vision Research*, *46*, 4118–4133.

Doran, M. M., Hoffman, J. E., & Scholl, B. J. (2006). The role of eye fixations in amplification and concentration effects during MOT. *Abstracts of the Annual Meeting of the Psychonomic Society*, *11*, 134.

Fazl, A., & Mingolla, E. (2007). Consistency of eye movements during multiple object tracking. *Journal of Vision*, *7*(9), 902a.

Fehd, H. M., & Seiffert, A. E. (2007). Eye movements during multiple object tracking. *Journal of Vision*, *7*(9), 999a.

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

Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*, 175–219.

Landry, S. J., Sheridan, T. B., & Yufik, Y. M. (2001). A methodology for studying cognitive groupings in a target-tracking task. *IEEE Transactions on Intelligent Transportation Systems*, *2*(2), 92–100.

Liu, G., Austen, E. L., Booth, K. S., Fisher, B. D., Argue, R., Rempel, M. I., & Enns, J. T. (2005). Multiple-object tracking is based on scene, not retinal, coordinates. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 235–247.

Neider, M. B., & Zelinsky, G. J. (2006a). Scene context guides eye movements during visual search. *Vision Research*, *46*, 614–621.

Neider, M. B., & Zelinsky, G. J. (2006b). Searching for camouflaged targets: Effects of target-background similarity on visual search. *Vision Research*, *46*, 2217–2235.

Pashler, H. (1988). Familiarity and visual change detection. *Perception and Psychophysics*, *44*, 369–378.

Prinzmetal, W. (1981). Principles of feature integration in visual perception. *Perception and Psychophysics*, *30*, 330–340.

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

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

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

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

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

Shipley, T., & Kellman, P. (Eds.). (2001). *From fragments to objects: Grouping and segmentation in vision* (Advances in Psychology Series, Vol. 130). New York: Elsevier.

Torralba, A., Oliva, A., Castelhano, M., & Henderson, J. M. (2006). Contextual guidance of attention in natural scenes: The role of global features on object search. *Psychological Review*, *113*, 766–786.

Ullman, S. (1984). Visual routines. *Cognition*, *18*, 97–159.

VanMarle, K., & Scholl, B. J. (2003). Attentive tracking of objects versus substances. *Psychological Science*, *14*, 498–504.

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

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

Zelinsky, G. J. (1996). Using eye saccades to assess the selectivity of search movements. *Vision Research*, *36*, 2177–2187.

Zelinsky, G. J., Neider, M. B., & Todor, A. (2007). Multi-object tracking in a realistic 3D environment. *Journal of Vision*, *7*(9), 895a.

Zelinsky, G. J., Rao, R., Hayhoe, M. M., & Ballard, D. H. (1997). Eye movements reveal the spatio-temporal dynamics of visual search. *Psychological Science*, *8*, 448–453.

Zelinsky, G. J., & Sheinberg, D. (1997). Eye movements during parallel-serial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 244–262.

*Manuscript received September 2007*
*Manuscript revised January 2008*
*First published online April 2008*