Research Report

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science
21(7) 920–925

DOI: 10.1177/0956797610373935
http://pss.sagepub.com
SAGE

Tracking Multiple Objects Is Limited Only by Object Spacing, Not by Speed, Time, or Capacity

S.L. Franconeri, S.V. Jonathan, and J.M. Scimeca
Northwestern University

Abstract

In dealing with a dynamic world, people have the ability to maintain selective attention on a subset of moving objects in the environment. Performance in such multiple-object tracking is limited by three primary factors—the number of objects that one can track, the speed at which one can track them, and how close together they can be. We argue that this last limit, of *object spacing*, is the root cause of all performance constraints in multiple-object tracking. In two experiments, we found that as long as the distribution of object spacing is held constant, tracking performance is unaffected by large changes in object speed and tracking time. These results suggest that barring object-spacing constraints, people could reliably track an unlimited number of objects as fast as they could track a single object.

Keywords

multiple-object tracking, MOT, crowding, surround inhibition, divided attention

Received 9/8/09; Revision accepted 12/8/09

In daily visual life, objects in the world shift drastically across the retina as their positions move relative to the observer's field of view. Despite these dynamic changes, these objects must be continuously selected if they are to be monitored, compared, or encoded in memory. To explore the ability to maintain attention on more than one object at a time, researchers often rely on the multiple-object-tracking (MOT) task. This task requires observers to mentally track a set of target objects moving among featurally identical distractor objects (Pylyshyn & Storm, 1988); the challenge is similar to tracking a single cup as a street magician shuffles it rapidly among other identical cups.

Performance in this task reveals several limits on tracking abilities. First, there is a limit on capacity, the number of objects that can be tracked concurrently. Initially, many results suggested that this limit was four objects (Intriligator & Cavanagh, 2001; Pylyshyn & Storm, 1988; Yantis, 1992), but later work demonstrated that with some methodological changes, tracking capacity could reach eight or nine objects (Alvarez & Franconeri, 2007). Second, tracking capacity is reduced when objects move more quickly (Alvarez & Franconeri, 2007). Third, there is a limit from object spacing, with tighter spacing leading to lower performance (Franconeri, Lin, Pylyshyn, Fisher, & Enns, 2008; Intriligator & Cavanagh, 2001; Pylyshyn, 2004; Tombu & Seiffert, 2008).

These limits must be taken into consideration in models of how the visual system might concurrently track multiple objects. Past accounts have explained these limits by positing a set number of trackers (Pylyshyn & Storm, 1988), a variable number of trackers (Alvarez & Franconeri, 2007), or memory for the global shape created by the targets (Yantis, 1992). Here, we propose a more parsimonious account that predicts the limits on capacity, the influence of speed on tracking capacity, and object-spacing limits using known limitations of the visual system. We suggest that there is no limit on the number of trackers, and no limit per se on tracking capacity. Instead, tracking is accomplished in parallel for an unlimited number of objects at once. Such a system could be implemented by local and independent neural circuits that maintain a local activation peak for a tracked object while inhibiting nearby objects (Koch & Ullman, 1985; Pylyshyn, 2000). The important limit for this mechanism would arise from two types of spatially modulated interactions among objects. First, because a locus of spatial attention is known to have a suppressive surround (Bahcall & Kowler, 1999; Cutzu & Tsotsos, 2003; Hopf et al.,

Corresponding Author:
S.L. Franconeri, Northwestern University, Department of Psychology, 2029 Sheridan Rd., Evanston, IL 60208
E-mail: franconeri@northwestern.edu

2006; Tsotsos, Culhane, Wai, Davis, & Nuflo, 1995), tracked targets should inhibit each other if they are within a critical distance, creating noise in the selection region. Recent studies using MOT tasks have shown that decreased target-target spacing impairs MOT performance (Carlson, Alvarez, & Cavanagh, 2007; Pylyshyn, 2004; Shim, Alvarez, & Jiang, 2008). Second, moving targets closer to distractors would increase the likelihood that the selection regions for the targets would fail to exclude nearby distractors (Intriligator & Cavanagh, 2001; Pylyshyn, 2004).

Object-spacing limits can explain capacity limits in MOT because the capacity for independently selecting *static* locations is also eight or nine objects, and capacities for static locations diminish rapidly as the selected objects are placed closer together (Franconeri, Alvarez, & Enns, 2007). However, at first glance, the object-spacing account does not appear to predict the influence of speed on tracking capacity. If each object is tracked independently, there should be a constant upper limit on speed for each object, but there should be no interaction between that speed limit and the number of tracked objects. Within this framework, why should increased speed lead to lower tracking capacity? We suggest that increasing speed increases the number of close interactions among objects. One previous study supports this possibility (Franconeri et al., 2008; see also Intriligator & Cavanagh, 2001). Participants tracked a set of objects in a small tracking display and in a display magnified by 4-fold, which consequently had speeds increased by 4-fold. If speed were the limiting factor, then performance should have dropped dramatically. However, because the impact of object spacing should not change with screen magnification (e.g., Toet & Levi, 1992), there should have been little difference in the distribution of close interactions. Accuracy levels were highly similar across the two conditions, a result suggesting that object speed affected MOT performance only through its impact on the distribution of interactions among the objects. However, because this display-scaling manipulation may have altered other aspects of the tracking display, such as the spatial-frequency profile of the moving objects, more evidence is needed to support the object-spacing account.

The study we report here provides direct evidence that object spacing is the root cause of limits on MOT performance. If the critical factor limiting performance is the number of times that objects pass too closely to one another, then performance should be primarily limited not by object speed, but by the cumulative distance the objects travel. If objects in a display moved at 10°/s for 10 s, and then the same animation were played in "fast forward," running in one half or one quarter of the time, the cumulative distance covered by each object would not change, and the distribution of object interactions would be identical. The object-spacing account predicts that performance should be identical in these two conditions, despite the large changes in object speed. In Experiments 1 and 2, we tested MOT performance under a variety of object-speed and tracking-time combinations, chosen so that a given cumulative traveled distance was paired with widely varying speeds. Experiment 1 tested four combinations of speed and time, including a variety of cumulative distances. Experiment 2 replicated and extended our results using six combinations.

Method

Participants

Twenty-three observers participated in Experiment 1, and 24 in Experiment 2, in exchange for course credit or payment. Some participants were removed from the analysis (3 in Experiment 1, 5 in Experiment 2) because they were not able to track objects with at least 75% accuracy in the shortest-cumulative-distance condition.

Stimuli and apparatus

The experiments were run on Intel Macintosh computers using MATLAB 7.6 (The MathWorks, Natick, MA) and Psychophysics Toolbox Version 3 (Brainard, 1997; Pelli, 1997). Figure 1 illustrates the stimulus display. Participants sat approximately 50 cm from a 15-in. Viewsonic monitor (640 × 480 resolution) running at 120 Hz. Cumulative distances are reported in pixels (1° ≈ 18 pixels). On each trial, 12 black circles (0 cd/m^2; diameter = 8 pixels) were presented on a white (~70 cd/m^2) background. Targets and distractors were paired (paired circles were separated by 55 pixels in the case of the center pairs and by 110 pixels in the case of the corner pairs), and the members of each pair always remained 180° apart as they orbited an imaginary center point. The center points for the four outer pairs were on the corners of an imaginary square 300 pixels wide, and the center points for the two middle pairs were 60 pixels above and below the fixation point.

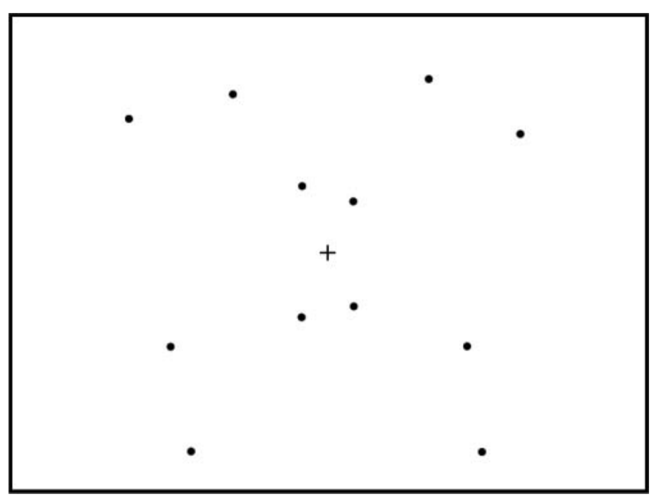

Fig. 1. Illustration of the tracking displays used in Experiments 1 and 2. Each of six pairs of black circles revolved around a center point, changing directions randomly and independently. The targets to be tracked were cued in red for 2 s at the beginning of a trial.

Object pairs revolved around their center points in a clockwise or counterclockwise pattern, always at a set speed and with instantaneous transitions in direction. Revolution speed was between 0.167 and 1.6 revolutions per second, and the duration of the tracking task was from 1.5 to 6 s. Object pairs randomly and independently changed the direction of their revolution (clockwise, counterclockwise), completing at least 0.1 and at most 2 revolutions before changing direction; the timing of each direction change was chosen randomly from a rectangular distribution.

Procedure

Subjects were given strict fixation instructions. At the start of a trial, all 12 circles appeared on the screen and began moving, with targets cued in red for 2 s, and then participants tracked the 6 targets for the designated time period. In Experiment 1, objects slowed exponentially over the final 0.5 s of the tracking period, whereas in Experiment 2, the objects stopped abruptly. The participant then heard a voice cue to click on the targets within the "top" or "bottom" three pairs of objects. We used this partial report because participants in pilot experiments reported forgetting known targets when asked to click on all 6 objects. After participants selected all targets they knew or chose to guess, they pressed the space bar, and the computer selected a target, with 50% (chance) accuracy, for any remaining pair within the partial report. Object speeds and tracking times for all conditions are shown in Figures 2a and 2b. Each condition was presented in a separate block, with block order randomized across subjects, and each block included 20 (Experiment 1) or 18 (Experiment 2) trials. Each subject received 5 practice trials using the condition with the shortest cumulative distance. Experiment 1 lasted approximately 40 min, and Experiment 2 lasted approximately 45 min.

Results and Discussion

Figure 2a depicts accuracy rates for Experiment 1. Accuracy values were submitted to an analysis of variance (ANOVA), which revealed a main effect of condition, $F(3, 57) = 88.0$, $p < .001$, $\eta_p^2 = .822$. Accuracy was highest for the shortest-cumulative-distance condition (M = 92.2%). Specifically, accuracy was significantly higher in this condition than in the two medium-cumulative-distance conditions (Ms = 74.5% and 74.2%), both $t(19)$s > 8.2, ps < .001, ds > 2.20. Accuracy did not differ between the latter two conditions, $t(19) < 1$, but accuracy was higher in each of these conditions than in the longest-cumulative-distance-condition (M = 61.2%), both $t(19)$s > 6.9, ps < .001, ds > 1.49.

Figure 2b depicts accuracy rates for Experiment 2. An ANOVA revealed a main effect of condition, $F(5, 90) = 40.0$, $p < .001$, $\eta_p^2 = .688$. Accuracy was highest for the shortest-cumulative-distance-condition (M = 86.4%), all $t(18)$s > 4.6, ps < .001, ds > 1.67. There were no significant differences in accuracy among the three medium-cumulative-distance conditions with speeds of 0.4, 0.8, and 1.2 revolutions per second (Ms = 71.5%, 74.5%, and 71.7%, respectively). The largest-cumulative distance condition had lower accuracy (M = 58.4%) than all other conditions, all $t(18)$s > 4.9, ps < .001, ds > 1.11. The condition with the highest speed, and medium cumulative distance, showed a medium accuracy level (M = 66.0%) that was slightly lower than accuracy in two of the other conditions with equivalent cumulative distance, both $t(18)$s > 2.7, ps < .02, ds > 0.71, but not the third, $t(18) = 1.8$, $p = .09$, $d = 0.62$. Accuracy was again best captured by differences in cumulative distance traveled, not speed or time.

One result in Experiment 2 seems at first incongruous with the object-spacing account. In the condition with the highest speed (right-most bar in Fig. 2b), cumulative distance was the medium cumulative distance, used in three additional conditions, yet accuracy was slightly (M = 6.6%) lower than in those conditions. Is this evidence for an influence of speed on tracking capacity even when distance is controlled? We think not. Even when all objects are tracked in parallel with independent speed limits, any individual object is still subject to the independent speed limit; thus, for example, faster targets will move farther during eyeblinks, and therefore are more likely to be lost. (See Norman & Bobrow, 1975, for a related dissociation, between *data-limited* and *resource-limited* processes, that is also discussed in Alvarez & Franconeri, 2007). Critically, this type of speed limit should not interact with the number of objects tracked.

Therefore, in a separate control experiment (N = 8) using identical displays, we asked participants to track only two objects. We had them track two objects instead of one to prevent them from using eye movements. Accuracy in tracking two objects should be about the same as accuracy in tracking one object as long as the objects are located in separate visual hemifields (Alvarez & Cavanagh, 2005). The targets were always drawn from the two diagonally opposite corner pairs, and both targets were reported. There were two conditions, with equal cumulative distances: slow (0.4 revolutions per second) but long (6 s) and fast (1.6 revolutions per second) but short (1.5 s). Performance was more accurate in the slow condition (M = 93.7%) than in the fast condition (82.7%), $t(7) = 3.6$, $p = .017$. This result suggests that the speed impairment observed in the highest-speed condition in Experiment 2 was not due to an influence of speed on tracking capacity, but rather was due to a main effect of speed.

Figure 3 depicts the results of both main experiments (accuracy levels for 10 conditions) as a function of cumulative distance, speed, and tracking time. The figure shows that cumulative distance best accounts for the variance in tracking accuracy. The logarithmic relationship (seen also in Alvarez & Franconeri, 2007) is likely due to the diminishing impact of distance on accuracy at greater distances. If close object spacing results in an unrecoverable target loss, the impact of spacing (or any other factor that impairs tracking) should be lower when more targets have already been lost. In contrast to the cumulative-distance panel in Figure 3, the speed panel shows

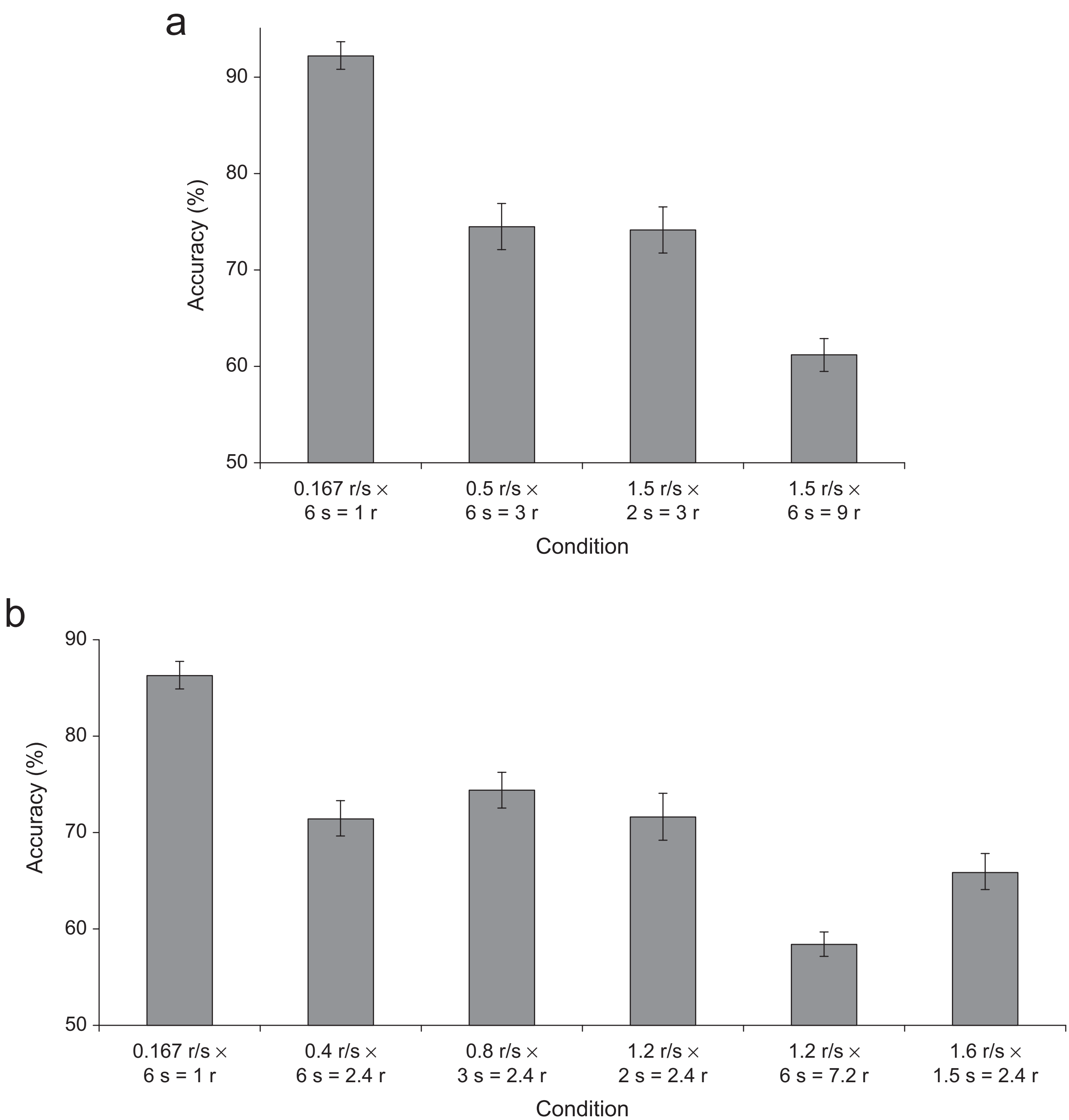

Fig. 2. Tracking accuracy in (a) Experiment 1 and (b) Experiment 2. For each condition, the revolution speed (r/s = revolutions/second), tracking duration, and number of revolutions (cumulative distance) during a trial is indicated. Note that one of the medium-cumulative-distance conditions in Experiment 2 employed a speed that was significantly faster than the speed used in the condition with the longest cumulative distance. Error bars represent standard errors.

roughly constant accuracy levels across a 4-fold difference in object speed. Note that the few points in this panel that seem to indicate a relationship between accuracy and speed (the two highest and the two lowest) confound extreme cumulative distances. The time panel in this figure shows no relationship between time and accuracy. (See the Supplemental Material available online for evidence that the impairment associated with cumulative distance is related to object-spacing effects.)

Conclusions

Across 10 combinations of object speed and tracking time, we found that cumulative distance traveled was by far the best predictor of tracking accuracy. Although this variable may also affect some yet-unknown factor related to tracking performance, this result is consistent with an account according to which all limits on multiple-object tracking have their origin

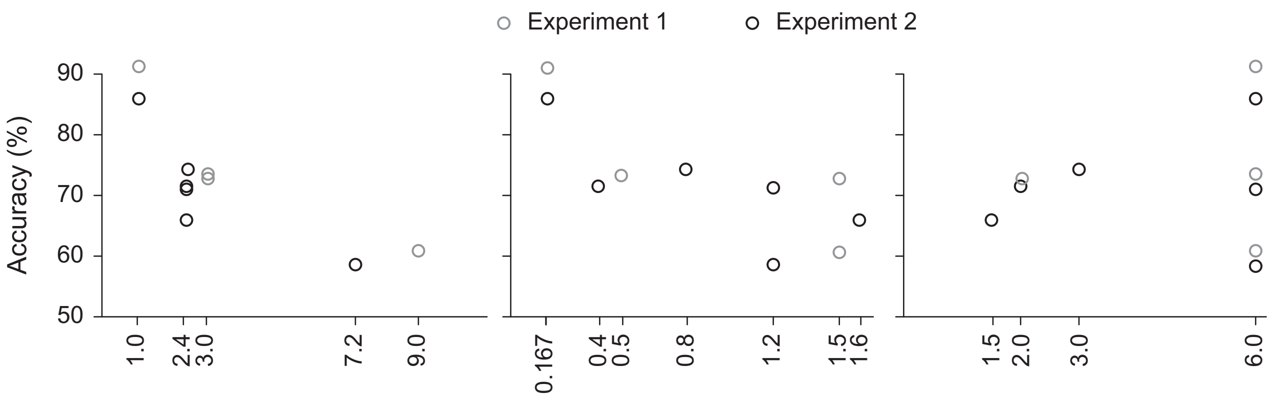

Fig. 3. Tracking accuracy for the 10 conditions in Experiments 1 and 2. Accuracy is graphed as a function of cumulative distance (left panel), speed (middle panel), and tracking time (right panel).

in object-spacing constraints among the tracked objects. This single parsimonious explanation can predict many previous results in MOT tasks, including the following:

- When target-surround inhibition is blocked by placing tracked objects in separate hemifields or quadrants such that object interactions are eliminated or reduced (Chakravarthi & Cavanagh, 2009; T. Liu, Jiang, Sun, & He, 2009), tracking performance is fully or partially independent for those objects (Alvarez & Cavanagh, 2005; Carlson et al., 2007).
- If the speed of all objects in a tracking display is increased by translating or scaling the display as a whole, which does not change relative object spacing, performance is unaffected (Franconeri et al., 2008; Intriligator & Cavanagh, 2001; G. Liu et al., 2005).
- Observers are as successful at tracking many moving objects as they are at tracking mixed collections of moving objects and static locations (Howe, Cohen, Yair, & Horowitz, 2009).
- Asking participants to track for longer periods of time, which increases cumulative distance traveled, impairs accuracy (Oksama & Hyönä, 2004).
- Video-game training can improve MOT performance (Green & Bavelier, 2003), perhaps by tightening participants' spatial resolution for object interactions (Green & Bavelier, 2007).
- Distractors that pass closer to targets can experience more inhibition (as measured by probes on objects; Doran & Hoffman, 2010), and such increased competition can tighten tuning of target location representations (Iordanescu, Grabowecky, & Suzuki, 2009).
- Constraining the global virtual polygon created by the target objects to remain convex leads to better performance (Yantis, 1992). This result could be explained by the fact that this constraint should serve to keep targets farther apart.

The object-spacing account provides a concrete mechanism for limits on object tracking and moves beyond redescriptions of those limits that label the tracking process as "resource dependent" or "requiring attention." Instead, it presents a simple and falsifiable hypothesis of the limits underlying the ability to track multiple objects at once. This account implies that barring object-spacing constraints, people could reliably track an unlimited number of objects as fast as they could track a single object. We hope that future work will test whether this explanation alone can account for all limits in the ability to maintain selection of multiple objects in the environment.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Supplemental Material

Additional supporting information may be found at http://pss.sagepub.com/content/by/supplemental-data

References

Alvarez, G.A., & Cavanagh, P. (2005). Independent resources for attentional tracking in the left and right visual hemifields. *Psychological Science*, *16*, 637–643.

Alvarez, G.A., & Franconeri, S.L. (2007). How many objects can you track?: Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, *7*(13), Article 14. Available from http://www.journalofvision.org/content/7/13/14

Bahcall, D.O., & Kowler, E. (1999). Attentional interference at small spatial separations. *Vision Research*, *39*, 71–86.

Brainard, D.H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Carlson, T.A., Alvarez, G.A., & Cavanagh, P.C. (2007). Quadrantic deficit reveals anatomical constraints on selection. *Proceedings of the National Academy of Sciences, USA*, *104*, 13496–13500.

Chakravarthi, R., & Cavanagh, C. (2009). Bilateral field advantage in visual crowding. *Vision Research*, *49*, 1638–1646.

Cutzu, F., & Tsotsos, J.K. (2003). The selective tuning model of visual attention: Testing the predictions arising from the inhibitory surround mechanism. *Vision Research*, *43*, 205–219.

Doran, M.M., & Hoffman, J.E. (2010). Target enhancement and distractor suppression in multiple object tracking. *Visual Cognition*, *18*, 126–129.

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

Franconeri, S.L., Lin, J., Pylyshyn, Z.W., Fisher, B., & Enns, J.T. (2008). Multiple object tracking is limited by crowding, but not speed. *Psychonomic Bulletin & Review*, *15*, 802–808.

Green, C.S., & Bavelier, D. (2003). Action video games modify visual selective attention. *Nature*, *423*, 534–537.

Green, C.S., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of attention. *Psychological Science*, *18*, 88–94.

Hopf, J.M., Boehler, C.N., Luck, S.J., Tsotsos, J.K., Heinze, H.J., & Schoenfeld, A.M. (2006). Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proceedings of the National Academy of Sciences, USA*, *103*, 1053–1058.

Howe, P.D., Cohen, M., Yair, P., & Horowitz, T.S. (2009, May). *Distinguishing between parallel and serial accounts of multiple object tracking*. Paper presented at the annual meeting of the Vision Sciences Society, Naples, FL.

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

Iordanescu, L., Grabowecky, M., & Suzuki, S. (2009). Demand-based dynamic distribution of attention and monitoring of velocities during multiple-object tracking. *Journal of Vision*, *9*(4), Article 1. Available from http://www.journalofvision.org/content/9/4/1

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

Liu, G., Austen, E.L., Booth, K.S., Fisher, B., 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.

Liu, T., Jiang, Y., Sun, X., & He, S. (2009). Reduction of the crowding effect in spatially adjacent but cortically remote visual stimuli. *Current Biology*, *19*, 127–132.

Norman, D.A., & Bobrow, D.G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7*, 44–64.

Oksama, L., & Hyönä, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Visual Cognition*, *11*, 631–671.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

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

Pylyshyn, Z.W. (2004). Some puzzling findings in multiple object tracking: I. Tracking without keeping track of object identities. *Visual Cognition*, *11*, 801–822.

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

Shim, W.M., Alvarez, G.A., & Jiang, Y.V. (2008). Spatial separation between targets constrains maintenance of attention on multiple objects. *Psychonomic Bulletin & Review*, *15*, 390–397.

Toet, A., & Levi, D.M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, *32*, 1349–1357.

Tombu, M., & Seiffert, A.E. (2008). Attentional costs in multiple object tracking. *Cognition*, *108*, 1–25.

Tsotsos, J.K., Culhane, S.W., Wai, Y.L., Davis, N., & Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence*, *78*, 507–547.

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