Selective Nontarget Inhibition in Multiple Object Tracking (MOT)

Correspondence to:

Zenon Pylyshyn
Center for Cognitive Science,
Psychology Bldg Addition, Busch Campus,
Rutgers University, New Brunswick,
152 Frelinghuysen Road
Piscataway, NJ 08854-8020

Email: zenon@ruccs.rutgers.edu

In Press: *Visual Cognition* (2008) VC BA 04 06R

Acknowledgments. This research was supported by NIH Grant 1R01 MH60924. The authors are grateful to Jeremy Wolfe and Andrew Hollingworth for their comments as reviewers of an earlier version of this manuscript..

Selective Nontarget Inhibition in Multiple Object Tracking (MOT)

Zenon W. Pylyshyn, H. Haroutioun Haladjian, Charles E. King & James E. Reilly

Rutgers Center for Cognitive Science

Rutgers, The State University of New Jersey, New Brunswick. NJ

Abstract

We previously reported that in the Multiple Object Tracking (MOT) task, which requires tracking several identical targets moving unpredictably among identical nontargets, the nontargets appear to be inhibited, as measured by a probe-dot detection method. The inhibition appears to be local to nontargets and does not extend to the space between objects – dropping off very rapidly away from targets and nontargets. In the present three experiments we show that (1) nontargets that are identical to targets but remain in a fixed location are not inhibited and (2) moving objects that have a different shape from targets are inhibited as much as same-shape nontargets, and (3) nontargets that are on a different depth plane and so are easily filtered out are not inhibited. This is consistent with a task-dependent view of item inhibition wherein nontargets are inhibited if (and only if) they are likely to be mistaken for targets.

Introduction

The Multiple Object Tracking (MOT) task has been widely used to study the properties of visual attention. In the MOT task (described in Pylyshyn, 2001; Pylyshyn & Storm, 1988), a subset of simple figures (*targets*) is briefly distinguished at the beginning of a trial. Then these targets, now identical to the other items on the screen, travel in an unpredictable manner among the other items, sometimes with collision-avoiding repulsion, and sometimes in smooth trajectories unconstrained except for “bouncing” off the edges of the display.

Observers are extremely adept at this task under a surprising range of conditions (for a review see Pylyshyn, 2003, chapter 5). The present study concerns the question: Why are targets not confused more often with nontargets, especially when they come close together? Exploring why this should be the case led us to hypothesize that nontargets might be kept from interfering with the tracking task by a process of inhibition.

Pylyshyn (2006) presented evidence that there is inhibition on nontargets and not in the space between objects. The experiments used probe dots that were presented during a tacking trial or during an identical trial in which no tracking was required. The stimulus sequence is illustrated in Figure 1.

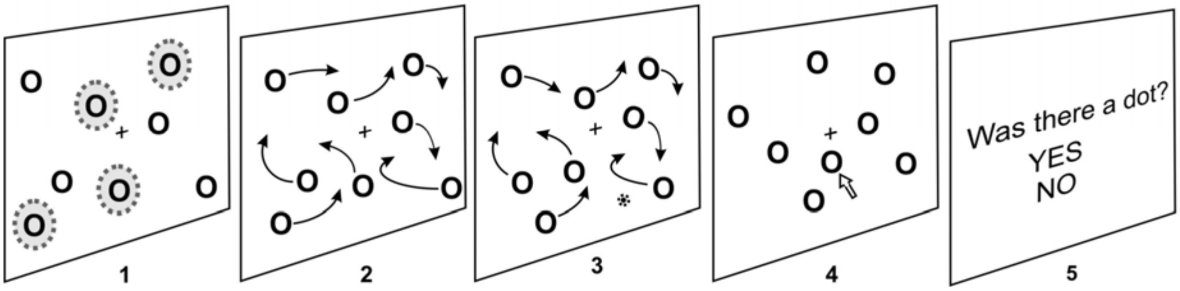

Figure 1. Illustration of experiments exploring inhibition during MOT. During the trial a small probe dot occurs on half the trials (shown as an asterisk in panel 3), When the trial ends, observers use a mouse to select the targets (panel 4). Then they indicate whether there had been a probe on that trial (panel 5).

In a series of experiments we found that detection of probes on nontargets is significantly worse than either on targets or in empty space, but detection of probes on targets is not significantly different from the detection of probes in empty space. We controlled for factors other than tracking by comparing probe detection performance during tracking with probe detection performance during a probe-monitoring task which involved no tracking. We found that that the detection of probes is impaired only on nontargets during tracking. Probe detection just 1.35° away from target or a nontarget was not significantly different from detection in empty space.

One reason why nontargets might be inhibited (suggested by Watson & Humphreys, 1997, 1998) is that inhibition is deployed whenever it is relevant to the task at hand. On this assumption one would expect that the degree of inhibition might be greatest when it is most needed, e.g., when targets are difficult to distinguish from nontargets. We test this idea using dual tasks: MOT and probe dot detection. The dot-probe detection task, which has been used by (Watson & Humphreys, 1997) as well as others (Ogawa, Takeda, & Yagi, 2002; Olivers, Watson, & Humphreys, 1999; Theeuwes, 2004) assumes that performance in detecting a small faint dot at a particular location serves as a measure of attentional enhancement or inhibition at that location. The first experiment tests whether a nontarget that has the same shape as a target, but does not move during a tracking trial, is inhibited. The second experiment asks whether a nontarget that has a different shape from a target is inhibited. The third experiment asks whether a feature known to be capable of preattentively segregating sets of objects (namely stereo disparity) eliminates inhibition on those objects.

Experiment 1

This study examines whether static nontargets, which are easily distinguished from targets, are inhibited relative to identical moving nontargets.

Method

Materials and apparatus

MOT Task. The experiment was programmed using the VisionShell© graphics libraries (Comtois, 1999) and was presented on an iMac computer. The circles in the tracking task consisted of white outline rings (luminance 55.8 cd/m^2) with dark interiors, displayed on a dark background. They were 47 pixels or 2.7 degrees of visual angle in diameter with outer rings 2 pixels (approximately 0.12°) thick. The experiment used 12 identical circles: 4 targets, 4 moving nontargets and 4 identical nontargets that remained in a randomly chosen fixed location throughout the trial.

In computing object trajectories we adopted a repulsion technique to keep the circles from colliding. The algorithm for computing item trajectories was the same as used in previous studies (e.g., Scholl & Pylyshyn, 1999). The moving circles were assigned random initial locations as well as horizontal and vertical velocity components chosen at random between -3 to +3 pixels/frame (with frames lasting 17.1 ms). These could be incremented or decremented on each frame by a single step or left unchanged with probability 0.90. The circles were prevented from getting too close by a "repulsion" between circles as well as the edges of the display. In the resulting motion no circle ever came nearer than about 4° to another circle. We estimated the average scalar speed by examining the frame-by-frame record of locations of a sample sequence of 1176 frames and determined that the mean speed was 11.6 deg/s (Standard Error = .08).

Probe Detection Task. While tracking targets, observers were also required to monitor the occurrence of a small dot that occur anywhere in the display on half the trials, and to indicate at the end of the trial whether there had been a dot on that trial. The forced-choice response was made by clicking on one of two boxes on the screen, as shown in Figure 1. The probes were distributed equally often at the centers of targets, moving nontargets, static nontargets as well as at a randomly chosen location in an empty region at least 2 diameters from the center of any circle. The probes were white squares measuring 6 x 6 pixels (approximately 0.35° x 0.35°) with a luminance of 22.9 cd/m^2 displayed for 128 ms.

Procedure. The 4 different probe location conditions (Empty space, Target, Moving Nontarget, or Static Nontarget) were randomized and grouped into 2 blocks of 120 trials, half of which had no probes. The first block was the nontracking control condition, which was identical to the tracking condition except for the requirement of tracking targets and identifying them at the end of the trial. It was presented prior to the tracking condition in order to discourage observers from adopting a tracking strategy out of habit. Each trial was 5 seconds long and the experiment lasted about 75 minutes.

Participants Twenty Rutgers undergraduates participated as part of their course requirements or for remuneration.

Results

Probe detection performance, measured as percent of probes correctly detected, is show in Figure 2 for both the tracking and the nontracking (control) condition. An ANOVA of probe detection scores showed that the overall difference between control and tracking scores was not significant, F(1,19)=0.85, MS=.007, p>0.37, but the effect of probe location was significant, F(3,57)=45.3, MS=.47, p<0.001, and the interaction of probe location and the control-tracking factor was also significant; F(3,57)=5.00, MS=.045, p<0.003. A planned comparison t-test showed that the control-tracking difference was significant only for the Moving Nontarget condition (t=2.35, df=19, two-tailed p<0.03); the difference between probe detection in the control condition and in the tracking condition on the static nontarget was not significant (t=.84, df=19, two-tailed p>0.41). (Note that there is no distinction between targets and nontargets in the nontracking control condition. For purposes of the analysis and to accommodate the equal-n and equal-variance assumption of the analyses, the program that randomly selected objects to be targets or nontargets in the experimental condition also arbitrarily classified them in the same way in the control condition.)

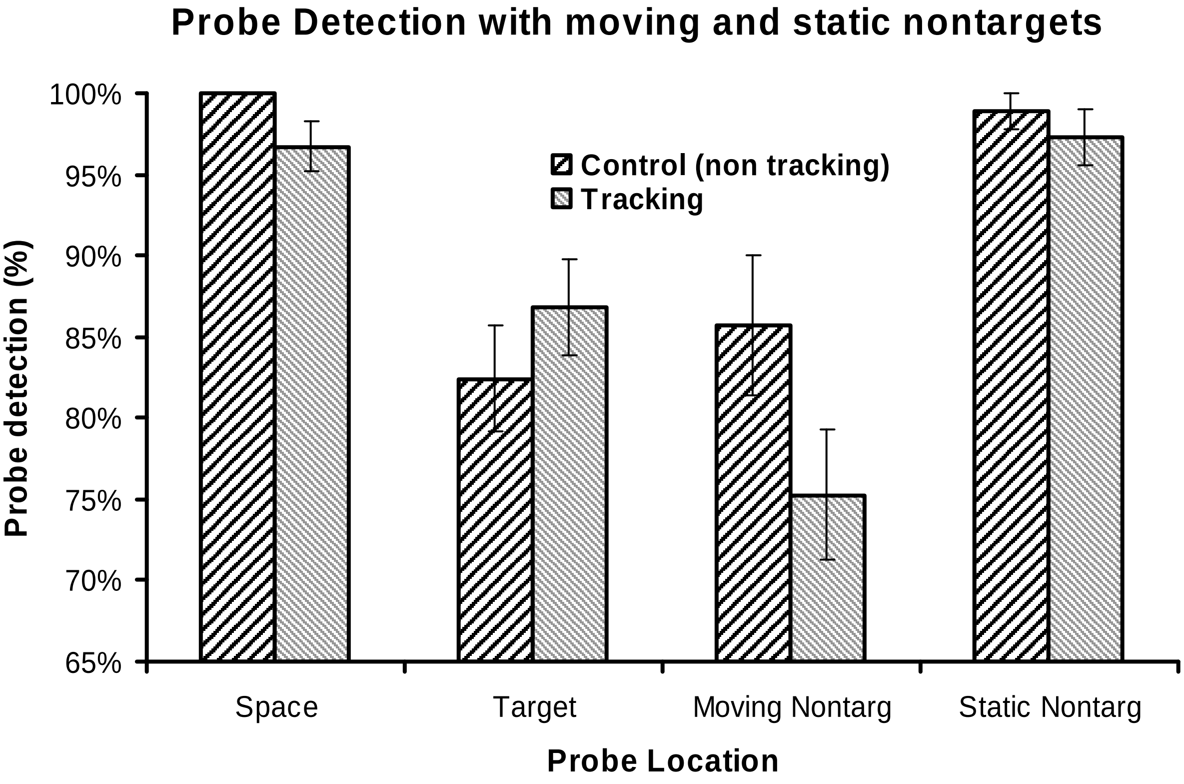

Figure 2. Results of Experiment 1. Probe detection performance at different locations. The difference between the nontracking control and the tracking condition is significant only on moving nontargets.

We also analyzed the tracking performance and found absolutely no difference in tracking as a function of where the probe occurred, or whether there had been a probe or not (the overall mean was 94.2%, Standard error = 0.5; repeated measures ANOVA yielded F(4,76)=0.3, SS=0.00025, p>0.87).

Discussion

The results of experiment 1 show that static nontargets are not inhibited. Because static nontargets are not easily confused with moving targets these results lend support to the hypothesis that easily-distinguished nontargets are not inhibited because they do not need to be for purposes of the tracking task.

It should be noted that the static nontargets are not only highly distinguishable from the moving nontargets, they are also different in a number of other ways. For example, because of their salience their fixed locations they may skew the distribution of attention. This possibility is consistent with the finding that even in the nontracking control condition, probe detection on the static circles is significantly better than on moving circles and equal to the “empty space” condition.

Experiment 2

If inhibition serves to keep targets from being confused with nontargets (i.e., if inhibition is applied in a top-down manner in response to task requirements, as suggested by Watson & Humphreys, 1997), then nontargets that differ in appearance from targets would not be inhibited. Experiment 2 tests this hypothesis by making half the nontargets a different shape from targets.

Method

Materials and apparatus

The materials and procedure were the same as in Experiment 1 except that all objects moved throughout the 5 second trial and half of the moving nontargets were squares (with dimensions equal to the diameter of the circles) rather than static circles There were 4 circular targets, 4 circular nontargets and 4 square nontargets.

Procedure. Same as in Experiment 1.

Participants Nineteen Rutgers undergraduates participated either as part of their course requirements or for remuneration. One participant was eliminated based on a poor score (<40%) on probe detection in the control condition.

Results

Results are shown in Figure 3. An ANOVA of the probe detection scores showed that probe detection in the Control condition differed significantly from that in the Tracking condition, F(1,17)=7.89, MS=.065, p<0.01, the 4 probe locations differed significantly, F(3,51)=22.3, MS=0.194, p<0.000, and the interaction between these two factors was significant, F(3,51)=7.08, MS=.041, p<0.000. A planned comparison t-test showed that the differences between Control and Tracking was significant both on Circular nontargets (t=2.94, df=17, p<0.01) and on Square nontargets (t=3.51, df17, p<0.002). The difference was not significant on the Targets (t=1.45, df=17, p>.16) nor in empty space (t=1.50, df=17, p>0.15). Analysis of tracking performance again showed no difference in tracking with different probe locations (including no probe condition): The mean was 92.1% (Standard Error 1.27); ANOVA yielded F(3,68)=1.78, MS=0.00009, p>0.95.

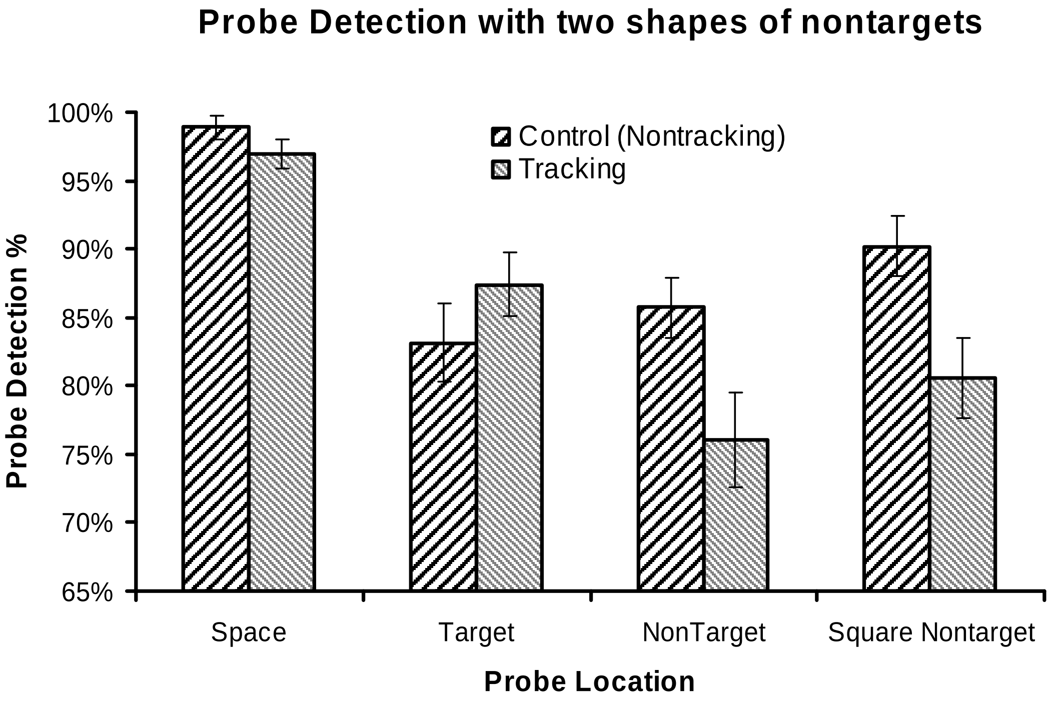

Figure 3. Probe detection score for probes in 4 different locations, shown when observers were tracking and when they were simply monitoring for probe occurrence..

Discussion

The results of experiment 2 showed a different pattern from that of experiment 1. As before, the regular nontargets (which have the same shape as targets) are inhibited while empty space and targets are not inhibited. But in contrast with experiment 1, where we found that easily-distinguishable static nontargets were not inhibited, here we find that different-shaped moving nontargets are inhibited just as much as regular targets.

An obvious counter to this result is that the difference between a circle and a square is insufficient to automatically exclude one of them from being in contention during tracking. Although the two shapes are very different, we already know that the shape of objects is often not encoded in the course of tracking (Bahrami, 2003; Scholl, Pylyshyn, & Franconeri, 1999). Thus the different-shape condition does not provide a strong test of the hypothesis that only objects that might be confused with the targets are inhibited. In order to better test this hypothesis we designed an experiment that uses a property known to be capable of perceptually separating a set of objects into subsets, namely stereo disparity.

Experiment 3

A number of studies have shown that stereo depth is computed early by the visual system (as shown, for example, by random-dot stereograms, Julesz, 1971). Moreover, detection of stereo disparity and other features appears to proceed independently. For example, whereas the conjunction of pairs of features cannot be detected in parallel, the conjunction of stereo disparity and features such as color or shape can be can be carried out in parallel in a visual search task (Nakayama & Silverman, 1986). Thus a difference in stereo disparity might allow some nontargets to be kept distinct, thus obviating the need for inhibition.

Experiment 3 tests the hypothesis that when half the nontargets are distinguished by their stereo disparity, then the nontargets that have the same disparity as the targets (are perceived to be on the same depth plane) will be inhibited, whereas the other nontargets (perceived to be on a different depth plane) will not be inhibited.

Method

Materials and apparatus

The materials and procedure were similar to those of Experiment 2, although the apparatus was different since all our experiments are now being run on a PC programmed in Matlab® and PsychToolbox®. Stimuli were viewed through Crystaleyes3® stereo goggles. The targets and half the nontargets were show as being on the “front plane” of the stereo display and the remaining 4 nontargets are shown as being on the “back plane” (this was done by making the retinal disparity between these two conditions 0.96 degrees). The probe dots were shown with the same disparity as the circles in which they appeared. To shorten the duration of the experiment we omitted the nontracking controls, whose role was primarily to control for the possible difference in lateral masking between the open space condition and the other conditions since the critical comparison in this case is between probe detection on the nontargets in the front plane and those in the back plane.

Procedure. Same as in Experiment 1.

Participants Nineteen Rutgers undergraduates participated as part of their course requirements or for remuneration. Three participants were eliminated based on a their poor score (<50%) on probe detection.

Results

Results of Experiment 3 are shown in Figure 4. The effect of probe location was significant (df=3, MS=0.087, F=4.9, p<.005). Only nontargets on the same plane as the targets were inhibited. The difference between probe detection on nontargets on a different plane from the targets and those on the same plane as targets was significant by a planned comparison paired t-test (t=2.784, df=18, p<.01). No other pairs were significant. This finding is consistent with the hypothesis that objects that might be confused with the targets are inhibited.

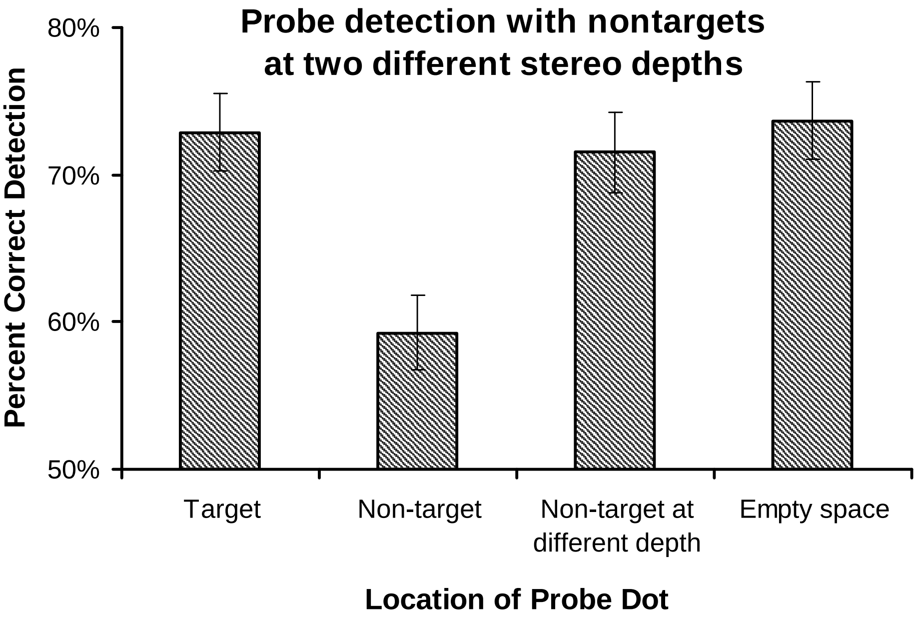

Figure 4. Probe detection score for probes in 4 different locations. Only detection of probes on nontargets at the same depth as targets showed inhibition.

Summary and General Discussion

In an earlier study (Pylyshyn, 2004) we used a probe-dot detection technique to show that probes on nontargets tend to be detected less well than probes either on targets or in the empty space between objects. This raised the question of the conditions under which inhibition is used to distinguish nontargets. In the present study three experiments examined the hypothesis that observers use inhibition in order to keep targets distinct from nontargets only when the nontargets cannot be preattentively distinguished from targets. In particular, we examined the task-specific use of inhibition to filter out nontargets distinguished by shape, motion and stereo disparity.

In the earlier paper we speculated that all and only moving objects in the field of view that are not being tracked might be inhibited. But the present experiments examined a simpler explanation for the particular pattern of nontarget inhibition observed in MOT, namely that in tracking targets, inhibition is applied to objects that might interfere with the task – i.e., nontarget objects that could be mistaken for targets. In the first experiment we found no inhibition of static nontargets that are easily distinguished from moving targets. This is consistent with the finding that moving items can be treated as a group and can be filtered out in visual search, regardless of their direction of motion (McLeod, Driver, Dienes, & Crisp, 1991), and can be inhibited as a group (Ogawa et al., 2002; Tipper, Brehaut, & Driver, 1990).

Contrary to expectations, however, Experiment 2 found inhibition of nontargets that differed from targets in their shape (circle vs square). While these different shapes appear to be relatively easy to discriminate, the difference between circle and square may not be sufficiently salient to serve as a cue for filtering out moving nontargets. Moreover it is known that the shape of objects is typically not encoded in MOT (Bahrami, 2003; Scholl & Pylyshyn, 1999). In order to test the stronger hypothesis that *preattentively* discriminable nontargets need not be inhibited we sought another dimension that, like motion, is known to be detected preattentively.

A property that allows preattentive separability of a set of objects is stereo disparity. Nakayama & Silverman (1986) showed that whereas conjunctions of features tend to slow down visual search, when one of the conjuncts is stereo disparity the conjoined features can be easily detected in parallel. Thus the finding in Experiment 3, that nontargets at a different depth plane from targets are not inhibited supports the view that inhibition is applied when the task requires it and the preattentional mechanism allows it.

The hypothesis that irrelevant but potentially disruptive distractors can be inhibited in MOT has far reaching ramifications. If it can be made more precise with further study, it would help explain the ubiquitous human skill of selecting and keeping track of moving objects in a world filled with moving distractors, such as encountered in team sports and in navigating through the everyday world of pedestrian and vehicular traffic.

References

Bahrami, B. (2003). Object property encoding and change blindness in multiple object tracking. *Visual Cognition, 10*(8), 949-963.

Comtois, R. (1999). VisionShell PPC Software Library.

Julesz, B. (1971). *Foundations of Cyclopean Perception*. Chicago: Univ. of Chicago Press.

McLeod, P., Driver, J., Dienes, Z., & Crisp, J. (1991). Filtering by movement in visual search. *Journal of Experimental Psychology: Human Perception & Performance*, *17*(1), 55-64.

Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature, 320*(6059), 264-265.

Ogawa, H., Takeda, Y., & Yagi, A. (2002). Inhibitory tagging on randomly moving objects. *Psychological Science*, *13*(2), 125-129.

Olivers, C. N. J., Watson, D. G., & Humphreys, G. W. (1999). Visual marking of locations and feature maps: Evidence from within-dimension defined conjunctions. *Quarterly Journal of Experimental Psychology A, 52A*(3), 679-715.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition, 80*(1/2), 127-158.

Pylyshyn, Z. W. (2003). *Seeing and visualizing: It's not what you think*. Cambridge, MA: MIT Press/Bradford Books.

Pylyshyn, Z. W. (2004). Some puzzling findings in multiple object tracking (MOT): I. Tracking without keeping track of object identities. *Visual Cognition, 11*(7), 801-822.

Pylyshyn, Z. W. (2006). Some puzzling findings in multiple object tracking (MOT): II. Inhibition of moving nontargets. *Visual Cognition, 14*(2), 175-198.

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

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology, 38*(2), 259-290.

Scholl, B. J., Pylyshyn, Z. W., & Franconeri, S. L. (1999). When are featural and spatiotemporal properties encoded as a result of attentional allocation? *Investigative Ophthalmology & Visual Science, 40*(4), 4195.

Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review, 11*(1), 65-70.

Tipper, S. P., Brehaut, J. C., & Driver, J. (1990). Selection of moving and static objects for the control of spatially directed action. *Journal of Experimental Psychology: Human Perception & Performance, 16*(3), 492-504.

Watson, D. G., & Humphreys, G. W. (1997). Visual marking: prioritizing selection for new objects by top-down attentional inhibition of old objects. *Psychological Review, 104*(1), 90-122.

Watson, D. G., & Humphreys, G. W. (1998). Visual marking of moving objects: A role for top-down feature-based inhibition in selection. *Journal of Experimental Psychology: Human Perception and Performance, 24*(3), 946-962.