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.,

**Abstract**

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
increased by 4-fold. If speed were the limiting factor, then per
display magnified by 4-fold, which consequently had speeds
tracked a set of objects in a small tracking display and in a
2008; see also Intriligator & Cavanagh, 2001). Participants
One previous study supports this possibility (Franconeri et al.,
lead to lower tracking capacity? We suggest that increasing
is tracked independently, there should be a constant upper
limit on speed for each object, but there should be no interac
tions 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 pre
dict 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 interac
tions 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 per-
formance 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 interac-
tions. 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 per-
formance should be primarily limited not by object speed, but by
the cumulative distance the objects travel. If objects in a dis-
play moved at 10°/s for 10 s, and then the same animation
were played in “fast forward,” running in one half or one quar-
ter 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 Psy-
chophysics 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 cir-
cles (0 cd/m²; diameter = 8 pixels) were presented on a white
(~70 cd/m²) 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 reported forgetting known targets when asked to click on all 6 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 (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 values 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
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...
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


