Splitting attention reduces temporal resolution from 7 Hz for tracking one object to <3 Hz when tracking three

Alex O. Holcombe
School of Psychology, University of Sydney, Sydney, New South Wales, Australia

Wei-Ying Chen
School of Psychology, University of Sydney, Sydney, New South Wales, Australia

Overall performance when tracking moving targets is known to be poorer for larger numbers of targets, but the specific effect on tracking’s temporal resolution has never been investigated. We document a broad range of display parameters for which visual tracking is limited by temporal frequency (the interval between when a target is at each location and a distracter moves in and replaces it) rather than by object speed. We tested tracking of one, two, and three moving targets while the eyes remained fixed. Variation of the number of distracters and their speed revealed both speed limits and temporal frequency limits on tracking. The temporal frequency limit fell from 7 Hz with one target to 4 Hz with two targets and 2.6 Hz with three targets. The large size of this performance decrease implies that in the two-target condition participants would have done better by tracking only one of the two targets and ignoring the other. These effects are predicted by serial models involving a single tracking focus that must switch among the targets, sampling the position of only one target at a time. If parallel processing theories are to explain why dividing the tracking resource reduces temporal resolution so markedly, supplemental assumptions will be required.

Keywords: attention, tracking, motion


Introduction

Visual analysis begins with massively parallel processing. In the first cortical areas, neurons dedicated to small regions across the visual field extract information rapidly and with high temporal resolution. Signals from these automatic and preattentive processes are thought to underlie perception of information restricted to brief temporal intervals (for a review see Holcombe, 2009). For example, motion detectors integrate over only a few dozen milliseconds, allowing us to perceive the motion of gratings drifting faster than 20 Hz (Burr & Ross, 1982). Binocular disparity detectors perform at similarly fast rates, providing us with depth information available only at narrow timescales of less than 50 ms (Morgan & Castet, 1995). Low-level temporal blurring of the signals from the retina may impose many of these fast (above 10 Hz) limits.

Perceptual judgments with poorer temporal resolution are often associated with computations occurring late in the visual system or requiring focused attention (Holcombe, 2009). In these cases the limiting factor is not temporal blur of the image. It may instead be a slow maximum rate of central processing.

Synchrony judgments with alternating stimuli typically have slow limits of between 4 and 10 Hz. One example is judging whether two widely-spaced lights are flickering in synchrony or in alternation (Rogers-Ramachandran & Ramachandran, 1998; Battelli et al., 2001; Aghdaee & Cavanagh, 2007). He, Cavanagh, and Intriligator (1996) suggested that this limit is caused by the temporal resolution of attentional individuation. Fujisaki and Nishida (2010) also theorize that slow synchrony judgments are limited by central factors but attribute it to the limitations of a when pathway, distinct from a what pathway that yields even slower limits when the features must be identified.

As Fujisaki and Nishida (2010) highlighted, synchrony tasks requiring identification of stimulus features such as color do indeed yield slower temporal thresholds of 5 Hz or less (Holcombe & Cavanagh, 2001; Clifford, Holcombe, & Pearson, 2004; Holcombe & Judson, 2007; Holcombe & Cavanagh, 2008; Fujisaki & Nishida, 2010; Holcombe, Linares, & Vaziri-Pashkam, 2011) except in some cases where features are spatially superposed (Holcombe & Cavanagh, 2001; Holcombe, 2009). For example, with a patch alternating between red and green and an adjacent grating alternating between leftward tilted and rightward tilted,
above six frames per second (3 Hz) participants are unable to judge whether the rightward-tilted patch is synchronous with the red or with the green (Holcombe & Cavanagh, 2001; Fujisaki & Nishida, 2010).

Tracking an object with attention (rather than with the eyes or the hand) is another process with a slow limit. In their pioneering study, Verstraten, Cavanagh, and LaBianca (2000) found limits of between 4 and 8 Hz. Like the synchrony judgments reviewed, attentional tracking may be limited by central individuation processes (He, Cavanagh, & Intriligator, 1996; Holcombe, 2009) or the when pathway (Fujisaki & Nishida, 2010).

While tracking’s temporal frequency limit has been investigated in only a single paper (Verstraten et al., 2000), its capacity limits have been tested extensively (e.g., Pylyshyn & Storm, 1988; Alvarez & Cavanagh, 2005; Alvarez & Franconeri, 2007; Bettencourt & Somers, 2009). Under demanding conditions (such as high object speeds), capacity is severely restricted to just a few objects (Alvarez & Franconeri, 2007; Vul, Frank, Tenenbaum, & Alvarez, 2010; Holcombe & Chen, 2012). Some theorists have suggested that the limiting process is identical to attentional selection (Cavanagh & Alvarez, 2005) or immediately precedes it (Pylyshyn & Storm, 1988; Pylyshyn, 1989). If the tracking process is indeed tightly linked to attentional selection, then the temporal limits of tracking will also have implications for attentional tasks more generally.

**Attentional resources and temporal resolution limits**

What causes the slow temporal limit on synchrony and tracking tasks? One possibility is that it is a fixed limitation on central processing, as would occur if a low-pass temporal filter intervened between midlevel and high-level processing. Note that such a filter would not blur signals like white and black together to make grey, but rather it would combine the successive representations of white and black together so that they were still available but could not be individuated into the distinct times they occurred (Van de Grind, Grusser, & Lunkenheimer, 1973; Holcombe, 2009).

Another possibility is that the temporal limit is not fixed but instead reflects the amount of attentional resource devoted to a target object, with less resource yielding a slower temporal limit. When all of an attentional network is devoted to processing a target, processing may be much faster or temporally more precise than when it is divided among two targets.

A third possibility is that attentional processing is serial, not parallel, so that when multiple targets are presented, the attentional focus must visit each target in turn. For the tracking task used here, this would manifest as a proportional decrease in temporal resolution as number of targets is increased. A related idea has been used to explain the attentional blink, the notion being that while the first target is being processed, there is no resource available for processing the second target (Duncan, Ward, & Shapiro, 1994; Chun & Potter, 1995). However, the culprit in the attentional blink may be one of task switching or control of attentional settings (Di Lollo, Kawahara, Ghorashi, & Enns, 2005; Nieuwenstein, 2006; Olivers, van der Stigchel, & Hulleman, 2007) rather than switching the focus of selection, so the issue of the effect of resource availability on temporal resolution remains unresolved. As multiple object tracking is a simpler task possibly constrained only by the selection process (Pylyshyn, 2001; Cavanagh & Alvarez, 2005), it provides a more direct test.

Here we investigated the effect of the number of targets to track on tracking’s temporal limit. In addition to possibly having implications for the consequences of splitting attention on the performance of attentional tasks, the results will constrain theories of how tracking itself is done.

**Understanding tracking through its temporal limits**

In a typical multiple object tracking (MOT) experiment, several identical objects travel about the screen (Pylyshyn & Storm, 1988). One or more are designated as targets and participants attempt to monitor their positions. After the identical targets and distracters move about the screen for several to a few dozen seconds, they stop and participants are asked to report the position of one or more targets.

Verstraten et al. (2000) found evidence that two distinct temporal limits constrain attentive tracking of a moving object—a speed limit and a temporal frequency limit. These experiments used a single moving target that moved in a circular trajectory centered on fixation. With the target one bar of a two-bar radial grating, Verstraten et al. (2000) found that performance declined to 75% when speed was increased to 1.3 revolutions per second for the three participants tested. Recently we used a similar technique (Holcombe & Chen, 2012) but with moving discs (one target, one distracter), and found a similar speed limit (1.9 rps, based on a more lenient 68% performance criterion) with six participants.

In addition to speed, temporal frequency may also constrain tracking. Temporal frequency is the rate that relevant objects pass a given location. In the two-bar (Verstraten et al., 2000) and Holcombe and Chen (2012) experiments, only two objects were present in a trajectory so the temporal frequency was always low.
When more objects are present in a display, this increases the rate at which objects visit any given location—the temporal frequency. Verstraten et al. (2000) manipulated temporal frequency via variation of target speed and the number of distracters that share the target’s circular trajectory. In such a display, the temporal frequency or local flicker rate is the product of the number of stimuli in the orbital path and the speed of the stimuli.

Verstraten et al. (2000) found that when the temporal frequency exceeded several hertz, tracking performance was poor. They suggested that performance was constrained by a frequency limit between 4 and 8 Hz, in addition to being constrained by the speed limit. Although their evidence was not entirely unequivocal (e.g., they did not correct for change in chance level for different numbers of objects), the results of our tests support their claims.

The limit may reflect the minimum temporal window tracking can access or the frequency at which each target is sampled by tracking processes. It appears that published theories of multiple object tracking have never discussed the issue of its temporal frequency limit (e.g., Kazanovich & Borisuyk, 2006; Oksama & Hyönä, 2008; Ma & Huang, 2009; Vul et al., 2010; Tripathy, Ogen, & Narasimhan, 2011). Yet simple serial theories of tracking make a specific prediction for the effect of target number on temporal frequency limits.

**Serial switching theory and temporal limits**

If tracking is serial, then the interval between updates of the location of any target should increase linearly with the number of targets. This implies that temporal frequency limits should decline with the inverse of the number of targets.

This prediction of serial sampling theory assumes that the time to switch between targets is not affected by the distance between them. Pylyshyn and Storm (1988) instead assumed that attention takes longer to switch among locations that are further apart, as if a physical spotlight must sweep across the visual field. Their calculations suggested a sweeping spotlight of implausibly high speed would be required to explain their data. They therefore adopted a parallel processing theory of tracking. However, many studies since have found that attention does not take longer to shift larger distances (Kwak, Dagenbach, & Egeth, 1991; Shih & Sperling, 2002), although voluntary shift times may be distance-dependent (Hazlett & Woldorff, 2004; Chakravarthi & VanRullen, 2011). Tripathy et al. (2011) have further pointed out that iconic memory may transiently buffer the traces of the moving objects, facilitating target recovery in conventional MOT where objects do not share a common trajectory, allowing serial switching to more plausibly explain tracking data.

To understand the prediction of an inverse effect of target number on temporal frequency limit, imagine that with two targets, the position of each is updated every 200 ms, implying that the focus switches between them every 100 ms. Adding a third target, each target’s position would be expected to be updated every 300 ms and with four targets, 400 ms. If when it comes time to revisit a target, the tracking focus goes to the object closest to the last-recorded position, than tracking will succeed only when distracters are no closer to a previously-registered position than the corresponding target is. Smaller distances between distracters and previously-registered target positions result in a failure of tracking. The failure point is the speed and spacing combination when the distracter has traveled half the distance between it and the target. Because the temporal frequency is the inverse of the time to travel the full distance, the temporal frequency failure point is the reciprocal of twice the time between position updates of a target. If the target position is updated every 200 ms (0.2 s) for example, then the temporal frequency limit is 1 / (0.2 * 2) or 2.5 Hz.

We find that increasing the number of targets has an approximately inverse effect on the temporal frequency limit, which is consistent with the serial theory’s prediction. Note however that in the above discussion we avoided the issue of serial switching theory’s prediction for tracking a single target. If attention samples periodically even a single target at the same rate that it switches between objects (VanRullen, Carlson, & Cavanagh, 2007), then the one-target limiting cycle duration is predicted to lie on the same line as that for additional targets. If with a single target attention instead operates more continuously, then the temporal frequency limit should be better.

Verstraten et al. (2000) only tested with a single target and we wanted to assess the effect of additional targets on the temporal frequency limit. As mentioned above, Verstraten et al. (2000) found evidence not only for a temporal frequency limit, but also for a speed limit. We confirm the existence of both limits here and also document a decline in temporal frequency limit with number of targets. We also find tentative evidence for a decline in the speed limit with number of targets.

The core idea of serial switching does not predict either the existence of a speed limit nor its decline with additional targets. However, a serial theory can accommodate these findings by assuming that the farther an object is from its last recorded location the more likely the spotlight will not reacquire it, even without any ambiguity regarding the nearest object. With higher loads, the target will have traveled farther since the last update, reducing the speed limit.
**Parallel tracking theory and temporal limits**

While the prediction of a load-dependent temporal frequency limit seems integral to serial theories, the prediction of parallel theories is less straightforward, largely because most published parallel theories are vague.

The most popular parallel processing theory is flexible resource theory (Alvarez & Franconeri, 2007) which proposes that tracking is mediated by a finite attentional resource that is distributed among the targets. Targets are processed independently, and the more resource allocated to a target, the better the tracking. But the theory does not specify in what way more resource is beneficial. It might reduce spatial interference by narrowing the tracking foci (Franconeri, in press). It might improve temporal resolution by shortening the temporal selection window or decreasing sampling variability in time, both of which could explain the present results. More resource might also increase the speed at which the tracking foci can move, which can accommodate a decrease in speed limit.

A more specific parallel processing theory with some conceptual similarities to flexible resource theory is the neural assembly oscillator model of Kazanovich and Borisyuk (2006). Although its predictions for temporal resolution are not entirely clear, it does appears to predict a decrease with number of targets, as will be described in the General discussion.

**Spatial interference theory**

Franconeri et al. (Franconeri, Lin, Pylyshyn, Fisher, & Enns, 2008; Franconeri, Jonathan, & Scimeca, 2010; Franconeri, in press) argued that the only reason tracking performance decreases with additional targets is because with more targets, in typical displays there are more occasions when targets are close to each other, resulting in more spatial interference. Holcombe and Chen (2012) investigated this by keeping objects farther apart than the range of documented spatial interference effects like crowding. In these circumstances where spatial interference was presumably avoided, Holcombe and Chen (2012) nevertheless found a large decrease in speed limit when two targets were tracked compared to tracking a single target. Moreover, spatial interference predicts that adding a target reduces performance even at very low speeds, which we did not observe.

Our interest here was to again avoid spatial interference so that temporal limits could be isolated. To achieve this, we once again prevented objects from approaching each other very closely. But to reach high temporal frequencies with reasonable object speeds, we had to present objects closer together than in Holcombe and Chen (2012). We did not need to present them so close that the objects’ crowding zones overlapped (Pelli & Tillman, 2008), but in case spatial interference nonetheless occurred, we included data analyses to probe for it. Only small effects were observed.

**The present experiments**

Two experiments were performed, one involving two rings of objects and the other involving three rings. The two-ring experiment was used to assess the temporal limits on one and on two targets, and the three-ring experiment allowed assessment of the temporal limits on two and on three targets. Including the two target condition in both experiments allowed us to show that the results generalize across stimuli (blobs vs. arc segments) and use of eccentricity scaling (three-ring experiment) versus the lack of it (two-ring experiment).

Both experiments document a temporal frequency limit that declines dramatically with the number of targets, as well as evincing a separate speed limit. There is potential for terminological confusion between the theoretical construct of a speed limit versus what we actually measured—the threshold speed. It is certainly the case that tracking performance falls with increasing object speed, but this need not indicate that a speed limit generally impairs performance. Indeed, for most of the displays tested here, the limitation instead appears to be a temporal frequency limit, because the threshold speed fell with number of distracters sharing the circular trajectories with the target, in a manner (inverse proportion) that indicates performance was limited by temporal frequency. The term “speed limit” will be used only for circumstances where it appears performance is constrained by speed rather than by temporal frequency.

In two animated demonstrations, readers can roughly assess their maximum tracking speed and thereby note the decline with the number of distracters sharing the target’s trajectory, which points to a temporal frequency limit. The animations are provided as supplemental movies here, and can also be viewed on their own webpage here: http://bit.ly/temporalTrackingLimits. In Movie S1, the two targets designated for tracking gradually increase in speed. Many readers will be able to successfully track these objects successfully even at the top, final speed of 0.6 rps. Movie S2 shows the same range of speeds but with more distracter objects added to the array of targets. Although it is still easy to track the objects at the very slow speeds of the beginning of the movie, as the speed increases the targets become untrackable. Our data below indicate that the lower speed threshold in the presence of more distracter objects reflects a temporal frequency limit. This limit is higher if only one target
is tracked, as can be seen by viewing Movie S2 again while ignoring one of the targets.

Our results document a range of circumstances in which temporal frequency is the primary constraint on performance which was not recognized by previous theorists who instead considered only spatial interference and speed (Alvarez & Franconeri, 2007; Bettencourt & Somers, 2009; Vul et al., 2010; Franconeri et al., 2010). The decline in temporal frequency limit with additional targets supports serial processing theories, which predict both the existence of the effect and its size. To remain viable, parallel processing theories must add an accessory assumption such as that temporal precision of the tracking focus decreases substantially with number of targets.

Methods

Participants

Six participants (aged 29–37, five males, two authors) were tested for the two-ring experiment and seven (aged 29–37, six males) for the three-ring experiment. The participants reported normal or corrected-to-normal vision. The protocol was approved by the University of Sydney’s ethics committee in accordance with the Declaration of Helsinki.

Apparatus

Stimuli were presented on a 22-inch Mitsubishi Diamond Pro 2070SB CRT monitor (1024 × 768 resolution) with a refresh rate of 160 Hz controlled by a Mac running a python program that used PsychoPy software (Peirce, 2007). Viewing distance was 57 cm in a dimly lit room, with a chin rest and forehead support to avoid subject head movement.

Stimuli

The two-ring experiment stimuli comprised two concentric rings of objects, and the three-ring experiment had three rings of objects, see Figure 1. For each ring, three, six, nine, or twelve objects were evenly spaced about the circular trajectory, forming a ring centered on fixation (fixation was a white dot of radius 0.1°). The background was black (< 1 cd/m², screen size 41° × 31°). In the two-ring experiment, the objects were blobs with a Gaussian luminance profile (visible diameter 0.8°, luminance: 12 cd/m²). In the three-ring experiment, the objects were arc segments with thickness and length scaled by eccentricity of the three rings (for inner: 0.6° × 0.9°; middle: 1.6° × 1.4°; outer: 4.3° × 2.4°).

For the two-ring experiment, one ring had radius of 2.5° and the other was 5.5°. In the three-ring experiment, the three rings had radii 1.5°, 4.5°, and
These values were chosen to keep the rings well outside each other’s crowding zones (Intriligator & Cavanagh, 2001; Pelli & Tillman, 2008).

Procedure

The participants, who all had experience fixating from previous experiments, were told to maintain fixation on the white dot at the display center. The trial started with the target objects presented in white (167 cd/m²) while the remainder (the distracter objects) were red (12 cd/m²; evoked by red gun only). The targets gradually became red (identical to the distracters) over the initial 0.7 s via a linear ramp through RGB space. The subsequent tracking interval was assigned a random duration between 3 and 3.8 s, see Figure 2.

For the two-ring experiment, the objects of the inner ring initially revolved in the opposite direction from that of the outer ring. For the three-ring experiment, the blobs of the inner and outer rings initially revolved in the opposite direction from that of the middle ring. Their initial angle about the circular trajectory was set randomly on each trial.

In the two-ring experiment, in half of the trials two targets were designated, one in each ring. In the other half of trials, only one target was designated, its ring randomly chosen. In the three-ring experiment, there were two or three targets, each in a randomly-chosen ring, with the constraint that each ring could contain only one target.

All the objects in a given ring always revolved about fixation in the same direction and at the same speed. Each ring was assigned an independent series of reversal times which succeeded each other at random intervals between 1.2 and 2 s. The initial angle of each ring (position of the arcs in the circular trajectory) was set randomly.

At the end of the trial, one ring was indicated with a recording of the first author saying “inner,” “outer,” or “middle.” The participant used the mouse to indicate which of the objects in the corresponding ring was the target. For the three-ring experiment, in one-third of trials participants were prompted to indicate the target in the inner trajectory, in one-third the middle trajectory, and one-third the outer trajectory. For the two-ring experiment, in half of trials participants were prompted to indicate the target in the inner trajectory and in half of trials that in the outer trajectory.

All objects revolved about fixation at the same speed throughout a given trial. A range of speeds (0.05–1.2 rps) was presented. The particular speeds for each condition and person were chosen based on piloting. Each person participated in at least 480 trials of an experiment (two-ring or three-ring), which usually involved two sessions, each shorter than fifty minutes.

Data analysis

The terms threshold speed and threshold temporal frequency will be used for the empirical rates at which performance declines to a certain level in a particular experimental condition. The terms speed limit and temporal frequency limit are reserved for hypothesized constraints on tracking that may be the cause of these thresholds.

To estimate the threshold speeds, plots of speed versus proportion correct were fit by logistic regression. By logistic regression, we mean a least squares generalized linear model with logit as the link function. Because logistic regression is for dependent variables that go from zero to one, proportion correct was rescaled. Its range was assumed to be from chance level to the ceiling imposed by the lapse rate, which allowed rescaling from zero to one (Zychaluk & Foster, 2009). For example, with a 1% lapse rate and chance level of 33% (three objects in each ring, one being a target), the ceiling is 99.33% because at ceiling, on 99% of trials participants respond correctly while in the remaining 1% of trials, participants will respond randomly, giving a correct response 33% of the time (0.99 * 1 + 0.01 * .33 = 0.9933).

After the psychometric function (the logistic) was fit to the data, the speed for threshold performance was
Because the chance rate varied with the number of objects in the ring, a particular performance level such as 75% has a different meaning for different object-number conditions. With three objects in the ring (two distracters and one target), chance level was 33% correct, with six objects 17% correct, nine objects 11% correct, and 12 objects 8% correct. Yet Verstraten et al. (2000) reported only the 75% threshold for all conditions, making their results difficult to interpret. A level more likely to be comparable across conditions is that halfway between the chance level and the ceiling imposed by the lapse rate.

Indeed, rescaling the psychometric function in this way from the ceiling to chance and extracting corresponding points is equivalent to the classic method for correcting performance for guessing. We will refer to this threshold adjusted for chance rate as the "midpoint threshold." We examined both the midpoint threshold and, as Verstraten et al. (2000) did, the 75% threshold. Because the pattern of results was very similar or identical, we report the midpoint threshold and provide the 75% point only in supplementary figures (Figures S1 and S2). So the threshold, unless otherwise specified, will refer to the midpoint threshold at which performance fell to the value midway between the ceiling and chance. This was estimated separately for each participant and condition.

In a separate analysis described later, we allowed the lapse rate to vary across conditions, but this had little effect on the pattern of results.

Results and discussion

As speed or temporal frequency increased, performance declined. We report the threshold speed at which performance falls to a criterion level (see Data analysis section above). The corresponding data and psycho-
metric fits are shown in Figure 3. The experiment program code, raw data, and analysis scripts that generated the graph of Figure 4 are available from the University of Sydney repository (http://hdl.handle.net/2123/8783).

A speed and a temporal frequency limit on tracking one target

The mean threshold speeds for each condition are shown in Figure 4. With only three objects in a ring and only one target (leftmost bar of Figure 4), the mean speed limit across participants was 1.7 rps (similar to the 1.6 rps found by Holcombe & Chen, 2012 with similar stimuli but travel distance equated across speeds). As the number of objects in the two-ring experiment increases from three to 12 (left panel of Figure 4), the threshold speed declines to 0.6 rps (simple linear regression for three to 12: \( b = -0.127, r^2 = 0.89, t(23) = -13.1, p < 0.001 \), for six to 12, \( b = -0.091, r^2 = 0.79, t(17) = -7.7, p < 0.001 \). As we explain below, there is reason to believe the three-objects condition reflects a speed limit while six to 12 objects reflects a temporal frequency limit, so we will sometimes analyze these conditions separately.

A temporal frequency constraint on performance can explain the present pattern of a decline in threshold speed with number of objects. This is because for a particular temporal frequency, the corresponding speed is lower for larger numbers of objects—temporal frequency is the product of speed and number of objects. The temporal frequency limit therefore makes a specific quantitative prediction. If this quantitative prediction is met, when the speeds are plotted in terms of temporal frequency the threshold will be constant rather than declining.

Accordingly, Figure 5 shows the threshold speeds converted to temporal frequencies. An alternative approach is to fit psychometric functions to temporal frequencies rather than speeds, which yielded the same pattern of results, not shown but the analysis script to generate the plot is available at http://hdl.handle.net/2123/8783. For the six, nine, and 12 object conditions of the one-target condition, the mean threshold temporal frequency was fairly constant at 6.9 to 7.1 Hz (simple linear regression indicates a nearly flat line: \( b = 0.042, r^2 = 0.02, t(17) = 0.569, p = 0.577 \), suggesting a temporal frequency limit constrains performance for these conditions.

The threshold temporal frequency for the three-object condition is significantly different, at 5.2 Hz—a repeated-measures ANOVA with number of objects and number of targets as factors shows a number of objects effect, \( F(3, 15) = 20.03, p < 0.001 \), and a post-hoc test indicates the threshold temporal frequency for the three-object condition is significantly lower than that for six (\( p = 0.003 \)), nine (\( p = 0.002 \)), and 12 (\( p < 0.001 \)) object conditions. This is consistent with the evidence of Verstraten et al. (2000) that a speed limit of \( 2.0 \) rps constrained performance when only a few objects shared the trajectory. In other words, with only three objects in the ring, the speed limit was exceeded before the temporal frequency limit was reached. Consider that with only three objects in the ring, 1.7 rps corresponded to 5.1 Hz, below the putative temporal frequency limit of \( 7 \) Hz. But for greater numbers of objects in the ring, the temporal frequency limit of \( 7 \) Hz is reached at speeds below 1.7 rps, meaning that \( 7 \) Hz can constrain performance before...
1.7 rps is hit. Thus it appears that for one target (the only condition considered in this section), the constraints on performance are a 1.7 rps speed limit and 7 Hz temporal frequency limit. When either of these limits was exceeded, tracking failed.

Effect of number of targets on tracking limits

Considering first the two-ring experiment, for every number-of-objects condition in every participant, the threshold speed limit was lower for two targets than for one. Paired t tests comparing two targets to one in each condition yielded a p-value of 0.001 or less, three objects, \( t(5) = 8.77; \) six objects, \( t(5) = 8.1; \) nine objects, \( t(5) = 6.6; \) 12 objects, \( t(5) = 10.5. \) Threshold speeds were also slower for three targets than for two (three-ring experiment) for each number of objects condition, from three to 12, \( t(6) = 15.1, t(6) = 4.4, t(6) = 7.4, t(6) = 7.5, \) all ps less than 0.005.

As with the one-target condition, in the two- and three-target conditions (the three-ring experiment) the threshold speeds decreased markedly with number of objects. Simple linear regression documents this for two targets from six to 12 objects, \( b = -0.063, r^2 = 0.64, t(20) = -5.86, p < 0.001, \) for two targets from three to 12 objects, \( b = -0.094, r^2 = 0.72, t(27) = -8.21, p < 0.001, \) for three targets from six to 12 objects, \( b = -0.044, r^2 = 0.57, t(20) = -4.99, p < 0.001, \) for three targets from three to 12 objects, \( b = -0.061, r^2 = 0.58, t(27) = -5.99, p < 0.001. \) Recall that we perform analyses without the three-object condition as the three-object condition may be constrained by a speed limit whereas it appears the other conditions are not, as documented next.

For six to 12 objects in the three-ring experiment, it appears that temporal frequency limits performance as it did in the two-ring experiment. Thresholds expressed as temporal frequency are relatively constant, with all regressions not statistically significant; for two targets from six to 12 objects, \( b = -0.084, r^2 = 0.05, t(20) = -0.97, p = 0.345, \) for two targets from three to 12 objects, \( b = 0.034, r^2 = 0.01, t(27) = 0.58, p = 0.569, \) for three targets from six to 12 objects, \( b = -0.045, r^2 = 0.04, t(20) = -0.82, p = 0.423, \) for three targets from three to 12 objects, \( b = 0.021, r^2 = 0.01, t(27) = 0.52, p = 0.61. \)

The limiting temporal frequency was lower for the two-targets condition (~4.4 Hz) than for the one-target condition (~7 Hz) (two-ring experiment). With thresholds expressed as temporal frequency, target number was significant according to a repeated-measures ANOVA with number of objects and number of targets as factors, \( F(1, 5) = 110.6, p < 0.001, \) and also when the three-object condition (where performance might be limited by speed not temporal frequency) was excluded from the ANOVA, \( F(1, 5) = 94.773, p < 0.001. \)

The hypothesis that an ~4 Hz limit constrained tracking of two targets (mean 4.4 Hz for six through 12 objects in the two-ring experiment) is corroborated by the two-target condition in the three-ring experiment, which shows a similar temporal frequency threshold despite the use of different stimuli (mean 4.0 Hz).

For the two-target condition, the apparent temporal frequency limit of 4 Hz is so low that the corresponding speed in the three-object condition is lower than the
For the three-ring experiment contrasting three targets with two targets, thresholds are again substantially poorer for the larger target load. The threshold temporal frequencies are 3.9 Hz for two targets and 2.6 Hz for three targets, \( F(1, 6) = 94.385, p < 0.001 \). Excluding the three-object condition that may be limited by speed rather than temporal frequency, the comparison remained significant, \( F(1, 6) = 43.902, p = 0.001 \).

For the three-target case, these data do not reveal whether a speed constraint limits performance. With the temporal frequency limit approximately 2.6 Hz, in the three-object condition this is already exceeded by speeds above 0.87 rps. If the speed limit fell from two to three targets, it would have had to fall below 0.87 rps to constrain performance. As the speed threshold measured was 0.8 rps (2.4 Hz), close to 2.6 Hz, it seems more likely that temporal frequency constrains performance.

Performance is even worse than predicted by the capacity-one model

To put any findings of speed limit differences in perspective, we calculated the predictions of an extreme scenario—that participants can only track one object and must guess on the half of trials for which they track the wrong target. On this model, for the half of trials where a participant tracks the target that is subsequently queried, predicted performance for that speed is provided by the one-target logistic curve fit. On the other half of trials, participants guess and therefore perform at chance. Because according to this model participants guess on half of trials, performance can never exceed the level halfway between chance and the ceiling (reflecting a 1% lapse rate, the ceiling is 99.33% with three objects).

The threshold speeds of the model were calculated in the same way as that of the data—from the psychometric curve, extracting the speed at which performance fell to the midpoint threshold (e.g., 66% for the three-object condition, halfway between 99.33% and 33.33% correct).

Because actual performance is close to 100% at slow speeds (as shown in Figure 3), participants clearly can track more than one object at slow speeds. They are certainly not guessing on 50% of trials. But as speed increases, performance approaches the model prediction. It eventually falls below the prediction, which at first may seem paradoxical. After all, it means that at those speeds participants would have been better off tracking only a single target and guessing on the other. If they had done that, they would have matched the model performance.

Figure 6 shows the empirical threshold speeds for the two-ring experiment along with the threshold speeds of the capacity-one model for the two-target condition.

In every condition, the empirical two-target speed limit is even slower than that predicted by the capacity-one model. A repeated-measures ANOVA indicates the difference between the empirical and predicted speed thresholds is significant, \( F(1, 5) = 46.523, p = 0.001 \).

How is it that performance is worse than this seemingly worst-case model? If participants do not have sufficient tracking resources to track both targets of the two-target condition at high speeds but nonetheless attempt to track both, they can end up with worse performance than if they had only attempted to track one. Splitting the tracking resource apparently results in insufficient resource per target for tracking to succeed, implying that the resource versus performance function (Norman & Bobrow, 1975; Pastukhov, Fischer, & Braun, 2008) is sufficiently steep at high speeds that going from 100% resource to 50% resource causes performance to fall farther than halfway towards chance, as depicted in Figure 7.

Figure 6. The speed limit for tracking two targets is so far below the speed limit for tracking one that it falls below that predicted by a one-target capacity limit (dashed bars). Participants would have done better by ignoring one target, as they then should have matched the dashed bars. Error bars show one standard error across participants.
In the experiments of Holcombe and Chen (2012), we also found performance to be lower than the capacity-one model, but the difference was not statistically significant in those experiments.

At very low speeds/temporal frequencies, participants have no problem tracking two or even three targets, consistently achieving performance levels above 95% (as seen in Figure 3). This implies that the performance versus resource function is different (shallower) for low speeds. Going from 100% resource (one target) to 50% (two targets) or 33% (three targets) yields little impairment. The reason may be that in this regime of low speeds, even if participants fail to move their tracking focus with the targets for an extended interval (a few hundred milliseconds), they can still recover the targets by finding the objects closest to the positions they last registered. Only at limiting temporal frequencies or speeds is there a consequence to having noisy (parallel tracking models, see General discussion) or intermittent (serial tracking models) position updating.

A role for spatial interference?

If the number of objects in a ring were increased enough, a spatial frequency limit would be reached rather than a temporal frequency limit. Although acuity limits were certainly not exceeded here, of possible concern is the crowding limit, at which one can still resolve a pattern but cannot attentionally isolate a target among nearby distracters (Intriligator & Cavanagh, 2001). The spatial characteristics of crowding with a single target have been studied extensively (Toet & Levi, 1992; Pelli & Tillman, 2008). The results have supported “Bouma’s Law” (Bouma, 1970)—for objects arrayed radially from fixation outside the fovea, crowding only occurs when distracters are separated from the target by approximately less than half the target’s eccentricity (Pelli & Tillman, 2008). For objects arrayed at a common eccentricity (isoeccentrically), as are the objects within each of our rings, the zone of spatial interference is substantially smaller than the Bouma figure of half the eccentricity (Toet & Levi, 1992). But to be conservative, we will calculate the implications of the half-the-eccentricity figure. It implies that crowding should not occur as long as the number of objects \( n \) is less than 13. The reason is that when \( n \) objects are equally spaced about a circle centered on fixation, they will be separated by greater than half their eccentricity \( e \) as long as \( n \) is less than 13. Consider that the separation between the objects is \( 2 * \pi * e / n \), which is less than 0.5 \( * e \) as long as \( n \) is less than 13. As crowding within our isoeccentric rings should not occur until the separation is substantially smaller than this, crowding should not occur when \( n \leq 12 \), so we restricted our trajectories to 12 objects or less.

To avoid crowding between rings, the distances between them were chosen to be large enough to be significantly greater than half each ring’s eccentricity. In the experiments by Franconeri et al. (Franconeri et al., 2008; Franconeri et al., 2010; Franconeri, in press) that led them to propose spatial interference is the only reason for decreased performance with additional targets, they used displays in which collisions were allowed between objects and therefore minimum separations were much smaller than in our displays. This suggests that spatial interference was largely or entirely avoided in our experiments, but nevertheless we below report data analyses to address the possibility that tracking yields spatial interference extending over a larger range than that documented in traditional crowding research.

Spatial interference should manifest as a reduction of the maximum performance level for higher numbers of objects in a ring. That is, a spatial limitation would reduce performance at all speeds, which corresponds to lowering the ceiling on proportion correct or increasing the lapse term in the psychometric fit. Franconeri et al. have emphasized that spatial interference may become greater with speed because in a typical MOT display with random trajectories, at higher speeds targets will travel farther and be involved in more close passes. But
separation was controlled here and in any case spatial interference should occur at all speeds, including slow ones. If it did not, it would not be truly spatial. The spatial interference account therefore predicts that performance should be worse for more targets and/or more distracters, even at very slow speeds. Therefore we tested for spatial interference in two ways:

1. Examining the effect of number of targets and objects on performance at only the slowest speed
2. Examining the effect of number of targets and objects on the lapse rate

**Examining the effect of number of targets and objects on performance at only the slowest speed**

The slowest speed varied somewhat across conditions and participants, as it was set on a per-participant basis to be slow enough to allow near-perfect performance. According to the spatial interference hypothesis of Franconeri et al. (Franconeri et al., 2008; Franconeri et al., 2010; Franconeri, in press), as the number of targets increases, the maximum performance level should decline with numbers of targets and/or number of objects. However the data do not show a significant effect in that direction.

A repeated-measures ANOVA with number of objects and number of targets as factors yielded no significant effect contrasting one versus two targets (the two-ring experiment) or two versus three targets (the three-ring experiment). Specifically, there was no significant effect of target number for one versus two targets, $F(1, 5) = 4.16\), $p = 0.097$, or for two versus three targets, $F(1, 6) = 5.76\), $p = 0.053$, a nonsignificant object number effect for one versus two targets, $F(3, 15) = 2.88\), $p = 0.071$, and for two versus three targets, $F(3, 18) = 0.972\), $p = 0.428$, and no interactions between target and number of objects for one versus two targets, $F(3, 15) = 3.048\), $p = 0.61$, and for two versus three targets, $F(3, 18) = 0.385\), $p = 0.765$.

Post-hoc tests revealed that the near-significant effect with a second target was caused by a deficit for the six-object condition. For example, the accuracy for the lowest speed for six objects (91%) is significantly worse than for the nine-object condition (100% correct), $p = 0.018$. The data (Figure 3) suggest that this was a result of not using slow enough speeds for the six-object condition for participants SM and LH (purple lines in the fifth column of the top of Figure 3). It certainly does not fit with the spatial interference prediction of worst performance for the 12-object condition. Regarding the three-target condition, from the graphs in Figure 3 it is apparent that the speed thresholds were so slow that for some conditions the speeds included were not slow enough to capture the asymptote of the psychometric function, which may explain the near-significant effect of number of target ($p = 0.053$) and number of objects ($p = 0.071$) reported in the previous paragraph. Slower speeds would be needed to assess spatial interference at speeds where the hypothesized low temporal frequency limit has not yet begun to affect performance.

**Absence of differences in lapse rate for different numbers of objects and targets**

The lapse rate parameter in our psychometric function fit is a measure of the errors that do not depend on speed. It is conventionally called the lapse rate in psychophysics because it includes complete lapses on the part of the participant, such as hitting the wrong key, which presumably occurs as often for trials with fast speeds as for trials with slower speeds. If spatial crowding (which by definition is also independent of speed) were impairing tracking it should occur more for conditions with higher object numbers and inflate the lapse rate. To the contrary, we found little to no change in lapse rate as the number of objects and targets increases. The fitted lapse rates are shown in Table 1. There is little to no evidence of crowding or other spatial interference, as the 12-object condition is very similar to the others. In fact the highest lapse rate (.05) occurred for nine objects in the two-target condition. That this high figure occurred in one of the two experiments only and did not extend to the 12-object conditions suggests it is not due to crowding and may be due to random variation.

Holcombe and Chen (2012) also did not find an elevation in lapse rate for two targets compared to one (they did not vary the number of objects in their rings, so they did not test the effect of that factor).

**Effect of eccentricity**

If the 7 Hz temporal frequency limit on tracking a target reflects the same processing that limits relative phase judgments of two flickering lights (more on that in the discussion), then there should be little effect of eccentricity. This presumes that Aghdaee and Cavanagh (2007) were correct to conclude that there is little effect of eccentricity on relative phase judgments. They tested long-range phase discrimination at 4° and 14° and found that 75% thresholds were 11.4 Hz for 4° and 8.9 Hz for 14°. This 28% decline in temporal resolution is small relative to the 288% decline in spatial resolution of attention for the same eccentricities (Intriligator & Cavanagh, 2001).

In addition to a temporal frequency limit change with eccentricity, the speed limit that appears to constrain three-object performance also might change with eccentricity. Holcombe and Chen (2012) reported
The revolutions per second limit merely proportional, requires that the scaling constant second to be the same across eccentricities, rather than eccentricity. However for speed limit in revolutions per cortex, as receptive field size may scale linearly with tracking object movement per unit area of retinotopic This could conceivably be explained by a limit on speed, which was much greater for larger eccentricities. The speed limit, this suggests it is imposed not by linear significant effects, as detailed below. In the case of speed limits and temporal frequency limits, we report statistics separately for the three-object condition as it therefore remains mysterious. Verstraten et al. (2000) suggested it may correspond to a limit on mental rotation (Shepard & Metzler, 1971; Cooper, 1976).

For the two-ring experiment, a three-way repeated-measures ANOVA with threshold speed as the dependent variable was conducted with target number (tracking one or tracking two targets), object number (three, six, nine, and 12 object conditions), and eccentricity (2.5° and 5.5°), and their interactions as factors. For the three-objects condition, the eccentricity effect, \( F(1, 5) = 2.142, p = 0.203 \), was not significant according to a repeated-measures ANOVA with eccentricity and number of targets as factors and neither were its interactions. For the remaining conditions (six to 12 objects), the ANOVA included number of objects as well as eccentricity and target number and eccentricity was also nonsignificant, \( F(1, 5) = 0.001, p = 0.974 \), as were the interactions. The average speed limit across all conditions for a single target was 1.1 rps at 2.5° and 1.0 rps at 5.5° and for two targets 0.7 rps at 2.5° and 5.5°.

For the three-ring experiment, three-objects condition, in a repeated-measures ANOVA with eccentricity and number of targets as factors, eccentricity was not significant, \( F(2, 12) = 0.409, p = 0.673 \), and neither was the interaction. For the remaining conditions (six to 12 objects), the interactions were not significant but the eccentricity effect was marginally significant, \( F(2, 12) = 3.503, p = 0.063 \). Post-hoc tests revealed that the middle ring (0.45 rps) threshold speed was significantly faster than the outer ring (0.38 rps). The average threshold speed across four object conditions for two targets was 0.7 rps at 1.5° and 4.5° and 0.6 rps at 12° and for three targets 0.4 rps at 1.5°, 0.5 rps at 4.5°, and 0.6 rps at 12°. Possible differences like these with more than one target must be interpreted with caution, because with more than one target present, participants may pay more (or less) attention to rings that seem to them to be harder to track, distorting any intrinsic difference in threshold speeds across the rings (which were nonsignificant in the one-target condition, as reported above).

### General discussion

In agreement with Verstraten et al. (2000), our data indicate the presence of both a speed limit and a temporal frequency limit on tracking a single target. Tracking additional targets lowers the temporal frequency limit from approximately 7 Hz for one target to about 4 Hz for two targets and 2.6 Hz for three targets. This reduction of temporal resolution was sufficiently large that the speed limit could not be reliably assessed when more than one target was tracked. Tentatively, the speed limit appears to have decreased from 1.7 rps to 1.3 rps for two targets. Decisive evidence must await future studies testing fewer than three objects in an array.

We know of no other work testing the effect of load on temporal resolution of any visual task. This issue of the temporal resolution of vision generally is discussed below after consideration of current theories of tracking.

Published theories of tracking have been designed to explain effects of speed, target number, and spacing. None have explicitly addressed temporal frequency, yet here we varied spacing and speed widely and found temporal frequency to be the primary constraint on performance.

Theories positing that spatial interference or resolution is the primary constraint on tracking performance do not explain these results (Franconeri et al., 2008, 2010). Spatial interference is also unable to explain the load-dependent speed limit on tracking, particularly when widely-spaced objects are used, as here, or travel distance is equated across speeds, as in Holcombe and Chen, 2012. Spatial interference certainly occurs in situations where the objects come closer to each other than they did in the present experiments. Indeed, the papers supporting a role for spatial interference used displays with objects that approached closer than the

<table>
<thead>
<tr>
<th>Objects</th>
<th>Two-ring experiment</th>
<th>Three-ring experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>One target</td>
<td>Two targets</td>
</tr>
<tr>
<td>3</td>
<td>0.01 ± 0.00</td>
<td>0.02 ± 0.01</td>
</tr>
<tr>
<td>6</td>
<td>0.02 ± 0.01</td>
<td>0.01 ± 0.00</td>
</tr>
<tr>
<td>9</td>
<td>0.03 ± 0.02</td>
<td>0.05 ± 0.05</td>
</tr>
<tr>
<td>12</td>
<td>0.02 ± 0.01</td>
<td>0.02 ± 0.02</td>
</tr>
</tbody>
</table>

Table 1. Estimated lapse rates (SE across participants).
Serial switching theory of tracking

Given the massively parallel architecture of the visual system, serial theories may seem implausible. However, the recent evidence that neural oscillations may coordinate serial sampling of visual stimuli (VanRullen, Carlson, & Cavanagh, 2007; Landau & Fries, 2012) suggests they should be considered.

According to the serial switching idea (Pylyshyn & Storm, 1988), target positions are updated one at a time. When it comes time to update a target’s position, the only information available is its last-sampled position, so the best that can be done is to go to the object that is nearest the target’s last-sampled position and assume it is the target. The critical stimulus variable determining whether tracking succeeds is that which determines whether, after a given duration (the intersampling interval), the object speeds and their spacing are such that a distracter is closest to the previous position of the target. This is half the limiting cycle duration, so the limiting temporal frequency is the inverse of twice that duration.

For a particular temporal frequency (e.g., 7 Hz or 143 ms in a cycle), a distracter will always be closer to the target’s sampled position after half a cycle (71.5 ms). A particular duration between position updates therefore naturally predicts a corresponding temporal frequency limit.

The serial theory predicts that the maximum temporal frequency will decline dramatically with the number of targets to track. With two targets, each target will be visited only half as often, and therefore the required cycle duration should double.

Serial switching theory of tracking

Bouma crowding distance (Franconeri et al., 2008, 2010; Tombu & Seiffert, 2010).

What are the consequences of these temporal limits for tracking with the random trajectories of typical MOT experiments? Other researchers rarely or never use speeds that exceed the speed limit documented here, but the targets in their displays likely do run afoul of the temporal frequency limits. For example, a 2.6 Hz limit (found here with three targets) suggests that when a distracter occupies a region within 386 ms of a target occupying it, tracking performance will be poor. This will occur frequently with the unconstrained trajectories of typical MOT tasks. Many laboratories routinely require participants to track four targets, for which the temporal frequency limit may be even lower.

With the exception of Franconeri’s spatial interference theory, most tracking theories have some scope to predict temporal limits that worsen with load. But while some theories can accommodate the phenomenon because they are vague, there is one theory for which the temporal frequency effects are a consequence of its core assumptions. This is the serial sampling theory of tracking multiple objects.

Serial switching theory of tracking

The serial theory predicts that the maximum temporal frequency will decline dramatically with the number of targets to track. With two targets, each target will be visited only half as often, and therefore the required cycle duration should double.

Strange, this fundamental prediction does not seem to have been stated in any published discussion of serial sampling theory (e.g., Oksama & Hyönä, 2008; Tripathy, Ogmen, & Narasimhan, 2011). According to its logic, the 4.15 Hz limit observed here for two targets indicates that 60 ms are required to sample one target position and switch to the other. The predicted results for three targets are then 2.78 Hz, not far from the mean 2.6 Hz observed. The one-target case may be special and is discussed in the Temporal resolution of high-level vision section below.

The distinct speed limit that appears to constrain performance with three objects before the temporal frequency limit is reached requires a separate assumption—unlike the temporal frequency limit, it cannot be a direct consequence of serial switching.

The speed limit and its decrease with load might be accommodated in this framework by assuming that the farther an object is from its last recorded location the more likely the spotlight will not reacquire it, even without any ambiguity regarding the nearest object. With higher loads, the target will have traveled further since the last update, reducing the speed limit.

To explain the hemispheric independence of tracking (Alvarez & Cavanagh, 2005; Holcombe & Chen, 2012), more than one sampler must be proposed—at least one in each hemisphere operating concurrently.

Howe, Cohen, Pinto, and Horowitz (2010) found evidence against serial theories using a sequential versus simultaneous comparison of object movement conditions. In the simultaneous condition with two targets, the targets in the display all moved simultaneously. In the sequential condition when one target was moving, the other was always stationary. Under a serial model, in the sequential condition the tracking might be able to ignore each target when it was stationary and thus sample the moving target more frequently, improving performance relative to the simultaneous condition. But Howe et al. (2010) found that performance was no better in the sequential condition which supports parallel models. This conclusion may however only rule out a certain class of serial models—those in which the serial process can rapidly (more often than every 500 ms) vary which targets it visits without any cost.

The serial theory predicts that represented target positions should lag their actual positions and that this lag should increase with the number of targets tracked. Evidence for this was reported by Howard and Holcombe (2008). Iordanescu, Grabowecky, and Suzuki (2009) found extrapolation rather than lag, but when Howard, Masom, and Holcombe (2011) did a very similar experiment with several variants, in every case they found lags. The discrepant results may reflect
the presence of an extrapolation mechanism (Roach, McGraw, & Johnston, 2011; Howe & Holcombe, 2012) that can counteract the lag to varying extents, depending on stimulus parameters.

Parallel theories of tracking: Flexible resource theory

While the load-dependent temporal frequency limit effect falls directly out of serial theories, for parallel theories to explain it, modification with suitable post-hoc assumptions is required.

According to flexible resource theory (Alvarez & Franconeri, 2007) tracking multiple objects occurs in parallel and is mediated by a finite attentional resource that is distributed among the targets. Targets are processed independently, and the more resource allocated to a target, the better the tracking. The theory does not specify however in what way more resource is beneficial. It might reduce spatial interference by narrowing the tracking foci. It might improve temporal resolution by reducing the temporal selection window or decreasing sampling variability in time, both of which could explain the present results. More resource might also somehow increase the speed at which the tracking foci can move, which would accommodate the speed limit and our tentative evidence for a decline from one to two targets.

To explain the present data, flexible resource theory must be fleshed out in a particular way. It should predict that tracking is limited by temporal frequency (rather than speed or spatial interference) over the wide range of display parameters used here and predict a large effect of available resource on temporal uncertainty or temporal resolution.

Parallel theories of tracking: Oscillator network model of Kazanovich and Borisuyk (2006)

Kazanovich and Borisuyk (2006) proposed a model with similarities to resource theory that should be favored because it is a working computational model rather than a vague theory. Moreover, it predicts speed and temporal frequency limits and may correctly predict that they worsen with load.

Targets are tracked by maintaining synchronized assemblies of neurons in a retinotopic array with each target represented by an assembly. As a target moves, it stimulates a new population of neurons in the retinotopic array and if the target is to remain tracked, this new population of neurons must join the synchronized assembly and the neurons representing the old location must leave the assembly. The synchrony and desynchrony have slow dynamics, imposing a speed limit that was explored in computational experiments by Kazanovich and Borisuyk (2006).

Kazanovich and Borisuyk (2006) did not consider the issue of temporal frequency limits, but it appears their model predicts a limit and may also predict that the limit decreases with load. The synchrony and desynchrony needed when an object moves takes time, and if a distracter arrives at the former target location when those neurons are still partially synchronized, then tracking may be prevented.

Performance decreases with load in their model because of the limited availability of phase space. Each target must be represented as a synchronous assembly that is distinct from the others by the phase of its firing. When only two targets are present the assemblies can fire in antiphase, but as the number of targets increases their firing times become more similar, increasing the probability that the phase separation cannot be maintained by the synchronizing interactions. This effect of load may have negative consequences for temporal resolution, although it is not clear whether the parameters of the model can be set to yield the present findings, in particular a large effect of load on temporal resolution and speed limit before a spatial interference regime is encountered.

Temporal resolution of high-level vision

The approximately 7 Hz tracking limit with one target (this article and Verstraten et al., 2000) may reflect a general limitation on high-level visual and possibly multimodal processing. It is often found for synchrony/asynchrony discriminations that do not require identifying constituent features (Rogers-Ramachandran & Ramachandran, 1998; Battelli et al., 2001; Holcombe, 2009; Fujisaki & Nishida, 2010). Aghdaee and Cavanagh (2007) found little effect of eccentricity on the temporal limit in such a task, consistent with the absence of an effect of eccentricity on temporal limits as found here.

As opposed to synchrony/asynchrony discriminations and the present tracking task, tasks that require labeling the constituent features of the stimuli have yielded slower limits of approximately 3 Hz (Holcombe & Cavanagh, 2001; Fujisaki & Nishida, 2010). We suggest that identifying two arbitrary features such as color and orientation requires an additional time-consuming step not required for synchrony discrimination nor attentional tracking.

Many researchers describe the attentional blink as a manifestation of the temporal limits of attention. However, almost all attentional blink studies are focused on how the successful selection of a first target disrupts the processing of a second target, which as
explained in the Introduction, may reflect limitations not relevant here.

Processing of the first target in attentional blink tasks may be limited by the temporal imprecision of attentional selection, but unfortunately it is rarely studied. The evidence available is that the precision of selection of the first target in a rapid serial visual presentation (RSVP) stream (Martini, 2012; Vul, Nieuwenstein, & Kanwisher, 2008) has a standard deviation of approximately 70 ms. Approximately the same figure is found for judging the instantaneous position of a moving object at the time of a cue (Murakami, 2001; Linares, Holcombe, & White, 2009). If the only factor affecting selection of a target from a stream were the uncertainty about its time, then the 75% threshold would be approximately 160 ms/item or 6.2 Hz (if the standard deviation of sampling about a cue is 70 ms, the area under the central ±80 ms of the corresponding Gaussian is 75%). But modeling position tracking rather than individual item selection in terms of this sampling dispersion requires many assumptions such as the rate of sampling, independence of successive samples, and how distracters are sampled, so the meaning of the similarity of these temporal limits is uncertain.

The studies of the attentional blink and of the position of moving objects both required determining which stimulus value (e.g., letter or object position) occurs at the same time as the cue. Some of the temporal imprecision in these cases may therefore reflect a temporal pairing process rather than the narrowest interval that processes like selection can access.

Unlike such tasks, attentional tracking does not require temporal binding of arbitrary events. It does require the timing of successive positions of an object, but this is an integral part of motion perception for which the brain has specialized processes with high temporal resolution (for a review see Holcombe, 2009). The limiting factor may therefore be the selection process rather than its inputs. If ever an attentional process was to have high temporal resolution, tracking might have been it.

One possible explanation for the 7 Hz one-target limit is that attention acts like a “flickering spotlight” (VanRullen, Carlson, & Cavanagh, 2007) that takes samples periodically, even when not switching among multiple targets. Note that a 7 Hz limit implies a sampling rate of at least 14 Hz in order to avoid direction ambiguity and track a moving object. An alternative possibility is that the tracking system capitalizes on a lower-level, less limited-capacity motion system that helps solve the correspondence problem. Under this scenario, if the attentional system can successfully select an individual disc to isolate it, the selection focus may then be carried along by the motion system so that tracking succeeds. In that case, the observed limit of 7 Hz may be the sampling rate rather than 14 Hz.

Verstraten et al. (2000) assessed the tracking temporal limit both for continuous motion like we used here and also tested ambiguous apparent motion (counterphasing dots) for which attention would have to solve the correspondence problem as well as track. The temporal limit was lower for ambiguous (counterphasing) apparent motion, but only by a few hertz. Yet if 7 Hz were the sampling rate, with counterphasing stimuli the tracking limit could not have exceeded 3.5 Hz. That they observed a limit higher than that suggests that a sampling-theory explanation of the 7 Hz continuous motion limit requires a sampling rate of at least 14 Hz. Using counterphasing stimuli, Kanaya and Sato (2012) found that cross-attribute attentional tracking with counterphasing stimuli had a lower temporal frequency limit (2–3 Hz) than within-attribute tracking (4–5 Hz), suggesting that the lower-level motion system contributes to the attentional tracking limit. This suggests that with unambiguous motion as used here, the one-target tracking limit is not set entirely by the temporal properties of attention (such as an attentional sampling rate).

Due to the contribution of low-level motion in the tracking case, the relationship of its temporal limit to that of other high-level tasks may be complex. It remains unclear whether the limiting high-level factor in both cases is a sampling rate, poor temporal precision, or an extended interval of temporal integration.

Multiple object tracking may be an elementary act of selection necessary for all attention-mediated processing of stimuli at the corresponding locations (Pylyshyn, 2001; Cavanagh & Alvarez, 2005; Holcombe, Linares, & Vaziri-Pashkam, 2011; Cavanagh, Battelli, & Holcombe, in press). Future work should investigate, then, whether dividing selection among multiple locations diminishes temporal resolution markedly for other attentional tasks as it did here for tracking.

Acknowledgments

We thank Piers Howe for extensive comments on the manuscript and Patrick Cavanagh for digging up lab notebooks from the Verstraten, Cavanagh, and LaBianca (2000) project. This project was supported by Australian Research Council (ARV) grants DP110100432 and FT0990767 to A. O. H.

Commercial relationships: none.
Corresponding author: Alex O. Holcombe
Email: alex.holcombe@sydney.edu.eu.
References


of English and Chinese word parts is limited to low temporal frequencies. *Perception*, 36(1), 49–74.


Pylyshyn, Z. (1989). The role of location indexes in...


