

RSVP in orbit: Identification of single and dual targets in motion

Brad Wyble · Mary C Potter · Marcelo Mattar

Published online: 30 December 2011
© Psychonomic Society, Inc. 2011

Abstract Three experiments using rapid serial visual presentation (RSVP) tested participants' ability to detect targets in streams that are in motion. These experiments compared the ability to identify moving versus stationary RSVP targets and examined the attentional blink with pairs of targets that were moving or stationary. One condition presented RSVP streams in the center of the screen; a second condition used an RSVP that was orbiting in a circle, with participants instructed to follow the stream with their eyes; and a third condition had participants fixate in the middle while observing a circling RSVP stream. Relative to performance in stationary RSVP streams, participants were not markedly impaired in detecting single targets in RSVP streams that were moving, either with or without instructions to pursue the motion. In streams with two targets, a normal attentional blink effect was observed when participants were instructed to pursue the moving stream. When participants had to maintain central fixation as the RSVP stream moved, the attentional blink was nearly absent even when a trailing mask was added. We suggest that the reduction of the attentional blink for moving RSVP streams may reflect a reduced ability to perceive the temporal boundaries of the individual items.

Keywords Attentional blink · Motion: Integration · Attention: object-based

Introduction

In rapid serial visual presentation (RSVP), stimuli are presented to participants at rates of about 10/s in order to explore the limits of our ability to detect, identify, and remember visual information (Forster, 1970; Potter, 1976). From these studies, we have learned that even stimuli as complex as natural images can be analyzed to a level of conceptual content at presentation rates as fast as 113 ms per image (Potter, 1976). This rapid conceptual processing allows participants to detect targets in a stream of stimuli that are defined by symbol category (e.g., letter vs. digit; Chun & Potter, 1995) or semantic category (e.g., occupation words; Barnard, Scott, Taylor, May, & Knightley, 2004) or even in scenes defined by the conjunction of multiple components (e.g., *a road with cars*; Potter, 1976).

In addition to revelations about the ability of the visual system to process input quickly, RSVP studies have also revealed dramatic temporal and spatial variation in the attentional state of the viewer in response to targets. For example, when two target items (referred to as T1 and T2) are presented at the same location on the screen and separated in time by 200–500 ms, participants frequently fail to report the second stimulus, a phenomenon known as the *attentional blink* (Raymond, Shapiro, & Arnell, 1992). Paradoxically, when the two target stimuli are presented within about 100 ms of one another, participants can report both of them, producing a so-called lag 1 sparing effect. Sparing is tied to the spatial location of the T1, since numerous studies have found it to be attenuated or even absent when the T2 appears in a different location (Jefferies, Ghorashi, Kawahara,

B. Wyble (✉)
Syracuse University,
430 Huntington Hall,
Syracuse, NY, USA
e-mail: bwyble@gmail.com

M. C. Potter
MIT,
Cambridge, MA, USA

M. Mattar
University of Pennsylvania,
Philadelphia, PA, USA

& Di Lollo, 2007; Shih, 2000; Visser, Zuvic, Bischof, & Di Lollo, 1999). On the other hand, the attentional blink is of similar size whether or not T1 and T2 are presented in the same retinotopic location (Shih, 2000; Visser et al., 1999).¹ These findings are generally taken as evidence that the T1 in an RSVP stream triggers the rapid deployment of attention to its own location (Bowman & Wyble, 2007; Chun & Potter, 1995; Olivers & Meeter, 2008; for a review, see Martens & Wyble, 2010), producing the spatially localized sparing effect. The ensuing attentional blink, which is not spatially localized, is assumed to operate at a central level of processing, as stimuli are encoded into memory (Chun & Potter, 1995; Jolicoeur 1999; Wyble, Bowman, & Nieuwenstein, 2009).

RSVP experiments have used one or more stationary RSVP streams to explore our capability to spot one or more targets. However, stimuli in the natural world are often moving with respect to the observer, leaving a smear of activated representations across the retina. Yet we are nevertheless able to perceive unblurred objects in motion, provided that their velocity falls within an optimal range (Burr, 1980). One way of exploring the ability to identify shape information in a moving stimulus is to display a shifting RSVP stream such that the stimulus is present only for a brief window of time in each location, producing apparent motion. In the following experiments, an RSVP stream orbited around a central point on the screen, changing to a new item about 10 times/s. Participants monitored this stream for letters presented among digit distractors.

Identification of a moving stimulus

One of the questions addressed by this paradigm is how readily participants can identify categorically defined targets within a moving RSVP. There has been a substantial amount of research exploring the degree to which motion perception affects the identification of stimuli in the path of motion. Such studies typically place a to-be-identified stimulus in the path of an apparent motion illusion elicited by two alternating dots (Attenave & Block 1974; Yantis & Nakama, 1998). Other work has looked at how real and apparent motion differ with respect to perception of a Necker cube (Kolers, 1964). A related but distinct question is the degree to which a stimulus that is itself moving can be identified. With regard to this question, there is debate about the degree to which shape information is accumulated over time for objects that move in retinotopic coordinates, while the eyes remain fixed. In a study by Cavanagh, Holcombe, and Chou (2008), it was found that participants were incapable of integrating shape

¹ While eye position was not monitored in either of these studies, the participant had no way of knowing where the T2 would have appeared, and therefore, it would be impossible for participants to move their eyes to ensure that both targets were in the same retinotopic coordinates on more than half of the trials.

information for stimuli that exhibit apparent motion. In their procedure, a circular cue that shifted along sequential positions in a circle acted as an attentional guide that allowed participants to track the spatiotemporal continuity of an alternating pattern. This guide produced apparent motion and greatly improved the ability of participants to report which direction of motion was associated with a particular color. However, this guide had no apparent benefit when participants were attempting to discern the orientation of a reversible character. The authors concluded that motion and color information can be accumulated over multiple frames of an object moving across the retina but shape information cannot. On the other hand, work by Ögmen, Otto, and Herzog (2006) has suggested that shape information is attributed to an object, thus remaining bound to it as it shifts in retinotopic space. In a similar vein, it has been suggested that object substitution masking is the result of a mechanism that can link successive frames of a stimulus together into a coherent representation of an object as it moves (Enns, Lleras, & Moore, 2010). Given such disagreement, it was unclear how readily participants would be able to identify a briefly presented target in a moving RSVP stream. The experiments presented here addressed this question by contrasting the ability to identify a categorically defined target in an RSVP stream that was stationary versus one that was moving either under smooth pursuit by the eyes or with eye position fixed at the center of the display. If shape information cannot be integrated across multiple frames of the moving stimulus, we would expect participants to have substantially lower performance when the eyes were fixed and the stimulus stream was moving than in the other two conditions.

The attentional blink for stimuli in motion

A second question that we address is the degree to which an attentional blink occurs for moving stimuli. The attentional blink is regarded as a nearly ubiquitous effect in visual perception, one that occurs whether stimuli are presented in the same or different locations (Jefferies et al., 2007; Shih, 2000; Visser et al., 1999). However, the attentional blink has never been assessed for objects in motion while the eyes remain fixed, so it is unknown how mechanisms related to motion perception interact with this effect.

In the present experiments, an RSVP stream orbited in a circle around a fixation cross, and participants were asked either to fixate on the cross (while attending to the stream) or to pursue the moving stream with their eyes. In another condition, the stream was stationary at fixation. If participants have difficulty integrating shape information across multiple frames, as was suggested by Cavanagh et al. (2008), the ability to identify individual targets should be impaired when the eyes are fixed and the RSVP stream moves, relative to the stationary and pursuit conditions. With regard to the attentional blink, if the phenomenon is

ubiquitous, there should be a blink in all of the conditions. To anticipate the results, identification of a single target in an RSVP stream was broadly similar in all three conditions. In trials with two targets, there was the expected attentional blink only when the eyes were either fixated on the stationary stream or in pursuit of the moving stream, but not when the RSVP stream was moving in the periphery with central fixation: In that condition, the attentional blink was nearly absent. We discuss possible reasons for this surprising finding in the [General Discussion](#) section.

Experiment 1

The first experiment compared target detection in RSVP while participants were either fixating on a stationary RSVP stream in the center of the display or following a moving stimulus with their eyes. Participants each saw two blocks of trials, one in which the RSVP stream moved in a circular trajectory, and the other in which the stimuli appeared centered at fixation. When the stimulus moved, participants were instructed to follow it with their eyes. Participants reported that this was easy to do.

Method

Participants Eighteen participants from the Syracuse University psychology study pool participated in this experiment and were paid. All had normal or corrected-to-normal vision and were fluent in English.

Stimuli In different blocks, participants viewed a stationary (RSVP) or a moving (MRSVP) stream composed of digit distractors and letter targets in a 70-point Kartika font in uppercase presented on a Windows XP machine using a 19-in. CRT monitor with a refresh rate of 85 Hz, running MATLAB 2007a and Psychtoolbox 3 (Brainard, 1997). The stimuli subtended $1.5^\circ \times 1.0^\circ$ of visual angle, from a viewing distance of 50 cm. Targets were black (40,40,40 RGB values) on a light gray background (150,150,150 RGB values).

The MRSVP stream was composed of an animation sequence in which the stimulus position orbited about a fixation cross at a constant distance of 4° of visual angle. Figure 1a illustrates an example of the sequence of stimuli that would be presented over the course of successive screen updates (frames). Note that each stimulus location was overlapped spatially by the preceding and following stimuli. For the moving display, we use the term *orbital arc* (OA) to refer to distance in degrees around the circumference of the circle that the stimuli traversed. On each refresh cycle (every 12 ms), the moving RSVP stream would advance 2.82 OA degrees clockwise and would complete a circular orbit of the fixation cross every 1.5 s. (The size of the circle was such that 1° of OA corresponded to 0.07° of visual angle.) This velocity was sufficient to create the appearance of a smoothly moving object with little blurring under either smooth pursuit or stationary fixation in the center of the circle. Each stimulus was present for eight successive frames for a total duration of 94 ms. The stationary RSVP stream was presented at a 94-ms stimulus onset asynchrony (SOA) in the middle of the screen with no interstimulus interval.

Procedure The experiment was a 7×8 design with a total of 112 trials in each of two blocks, one with a moving stream and one with a stationary stream. The order of the blocks was counterbalanced between participants. The first factor corresponded to the position of the T1 in the RSVP stream, ranging from 5 to 11. The second factor corresponded to the lag between T1 and T2 and ranged from 1 to 7, with the eighth level representing trials on which there was only a single target. On single-target trials, the lone target appeared in one of the temporal positions that the T2 would have occupied to assess the ability to report a single target from the same time slots the T2 occupied on two-target trials.

Each RSVP stream was 25 items in length. In the moving stream, the first item appeared at a randomly chosen position on the virtual circle. Thus, the T1 could appear with equal probability at any location on the circumference. Once the T1 appeared, the T2 was constrained to appear from one to seven lags later. In the moving stream, T2 appeared from

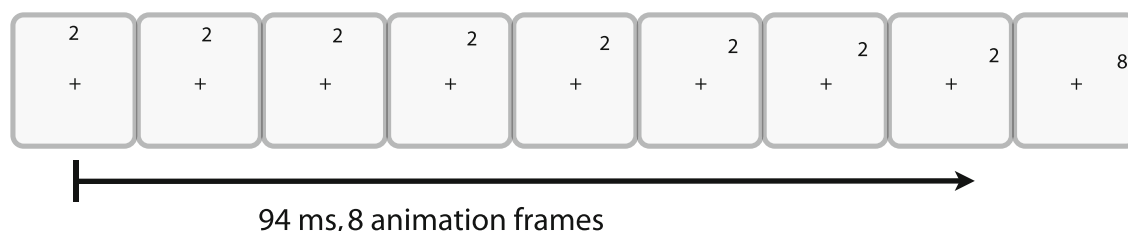


Fig. 1 A moving RSVP stream was created by presenting a sequence of stimuli in orbit around the fixation at 85 Hz. Each item in the RSVP stream was composed of eight animation frames presented at intervals of about 12 ms, for a total duration of 94 ms. Each frame of the presentation was displaced 2.8° around the circle (orbital angle),

producing a displacement of 0.2° of visual angle for the observer. Each RSVP stream contained digit distractors and one or two letters as targets. The apparent motion involved a complete orbit of the fixation cross every 1.5 s. In this diagram, the shift in location per frame has been doubled to make the change of position clearly visible

23 to 158 OA degrees after the T1. At the end of each trial, participants were prompted to enter in order the letters that they had seen, with the option to enter two, one, or zero letters. As is usual in attentional blink studies, responses for T1 and T2 were considered correct in either order. No feedback was given as to accuracy.

Before beginning each block, participants viewed nine practice trials on which the RSVP or MRSVP items changed every 118 ms (10 frames). Each experimental block took approximately 15 min to complete.

Results

The appearance of the MRSVP was of a single item that moved smoothly in a circle, changing identity rapidly according to reports of the participants and the phenomenological experience of the authors. Transitions from one stimulus type to the next in the RSVP did not interfere with this percept, resulting in the appearance of an object that moved in a circle while switching to a new stimulus periodically.

The results of [Experiment 1](#) were clear-cut; a conventional attentional blink pattern was observed whether the RSVP was stationary or moving, with instructions to pursue it with the eyes (Figs. 2a, b). In the analysis of T1, average T1 performance was equivalent between the stationary and pursuit MRSVP conditions ($M = .92$, $SE = .01$, and $M = .93$, $SE = .01$, respectively). A two-factor ANOVA over T1 accuracy, with condition (stationary vs. pursuit MRSVP) and lag (1–7) as factors, demonstrated a significant effect only of lag, $F(6, 102) = 13.6$, $p < .001$, $\eta_p^2 = .45$. As in other attentional blink studies, T1 performance was lower at lag 1 than at other lags. This characteristic pattern suggests that there is competition between the two targets at lag 1 (Potter, Staub, & O'Connor, 2002; Wyble, Bowman, & Nieuwenstein, 2009). When only a single target was presented, performance in the stationary and pursuit conditions was again similar and close to ceiling (.96, $SE = .016$, and .91, $SE = .02$, respectively). This difference did not quite reach significance on a two-tailed t test, $t(17) = 1.87$, $p = .078$, $d = 0.16$.

T2 performance, conditional on a correct T1, exhibited a classic attentional blink pattern with lag 1 sparing and an attentional blink that seemed to recover by lag 5. A two-factor ANOVA over T2|T1 accuracy, with block (stationary vs. pursuit MRSVP) and lag (1–7) as factors, demonstrated a significant effect of lag, $F(6, 102) = 9.9$, $p < .001$, $\eta_p^2 = .37$, and a marginal effect of condition, $F(1, 17) = 4.2$, $p = .055$, $\eta_p^2 = .20$. T2|T1 performance was slightly higher in the pursuit MRSVP condition, with an average T2|T1 accuracy of .82 ($SE = .016$) versus stationary T2|T1 performance of .76 ($SE = .019$). There was no interaction between condition and lag, $F(6, 102) = 0.8$, $p > .57$, $\eta_p^2 = .05$.

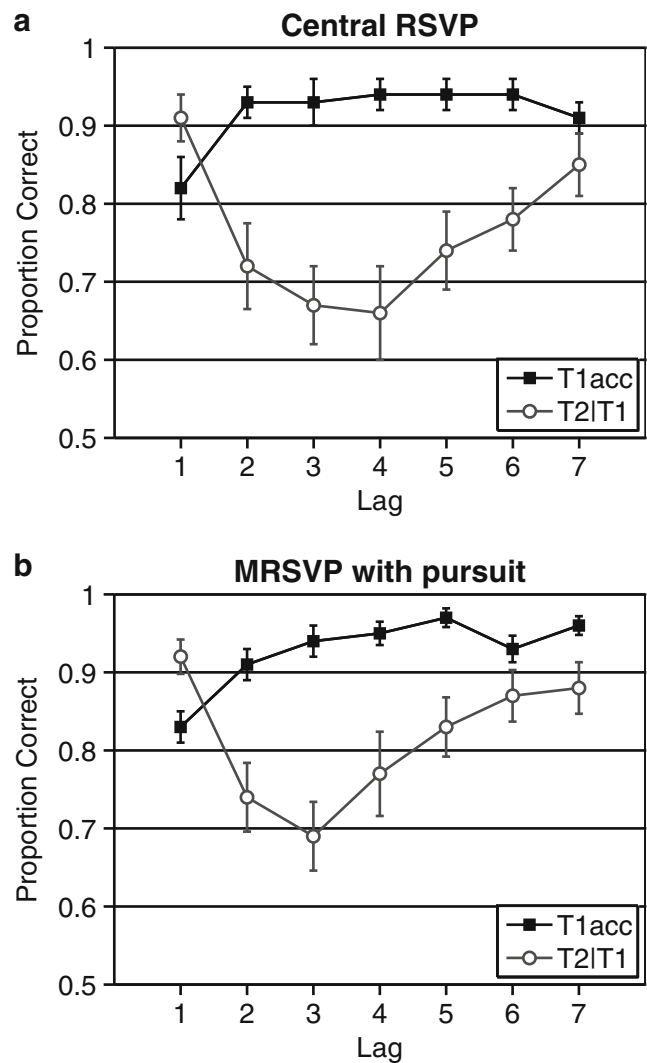


Fig. 2 T1 and T2|T1 accuracy for [Experiment 1](#), in the central RSVP condition (top) and MRSVP with pursuit condition (bottom). Error bars represent 1 standard error in this and all subsequent figures

Discussion

The results of this experiment revealed an excellent ability to detect a target in an MRSVP stream under instructions to pursue the moving stream with the eyes. This suggests that fixing the location of a moving target on the fovea makes target identification broadly similar to performance in a stationary RSVP stream. We next ran an experiment to determine whether this would also be true when participants' gaze was fixated on the middle of the display, rather than following the MRSVP.

Experiment 2

In this experiment, participants viewed the same MRSVP as in [Experiment 1](#), with the addition of instructions to fixate a

cross in the middle of the display. Eye tracking was used to monitor fixation, since participants in pilot experiments had difficulty attending to a moving stimulus without moving their eyes.

Method

Participants Sixteen participants from the MIT community participated in this experiment and were paid. All had normal or corrected-to-normal vision and were fluent in English. Five were eliminated for failure to maintain fixation, as detailed below.

Stimuli The stimuli in [Experiment 2](#) were similar to those in [Experiment 1](#)'s moving condition, but a different computer configured with an eyetracker was used. Participants viewed an MRSVP stream composed of digit distractors and letter targets in the Kartika font presented on a Windows XP machine using a 21-in. CRT monitor with a refresh rate of 70 Hz, running MATLAB 2007a and Psychtoolbox 3. The stimuli subtended $1.5^\circ \times 1.0^\circ$ of visual angle. Targets were black (40,40,40 RGB values) on a light gray background (150,150,150 RGB values). A black fixation cross was presented at the center of the MRSVP's orbit.

On each screen refresh cycle, the MRSVP advanced 3.42 OA degrees in a clockwise direction. Each stimulus in the RSVP stream was repeated for seven screen refresh cycles, for a total presentation duration of 100 ms.

Procedure Before beginning the experiment, participants viewed 9 practice trials on which the MRSVP items changed every 143 ms (10 frames). The experimental block consisted of 224 trials and took approximately 30 min to complete.

Eye tracking Eye movements were monitored at 240 Hz by an ISCAN RK-464 eyetracker. Observers sat 70 cm from a 21-in. CRT monitor with their chin in a headrest. The right eye was tracked, and viewing was binocular. Calibration occurred at the beginning of each experiment and involved fixating five locations. Validation occurred for a set of nine evenly distributed locations and was considered valid if all errors were less than 2° of visual angle. Calibration was repeated until successful.

Participant gaze was monitored throughout each trial. Eye movements that deviated from the fixation cross by more than 2° of visual angle at any point during a window encompassing 400 ms before to 400 ms after the T1 disqualified a trial from consideration. Five participants were eliminated from the sample for failing to maintain fixation on at least one third of the trials. Most participants experienced some difficulty keeping central fixation in the presence of the moving stimulus but learned the skill during the practice session.

Results

As in [Experiment 1](#), the appearance of the MRSVP was of a sequence of items moving smoothly in a circle. Therefore, whether being pursued by eye movements or not, the RSVP stream had the appearance of a single moving object that changed identity.

Performance in identifying a single letter in the MRSVP stream of digit distractors was broadly comparable to that when a stationary, centrally fixated stream was viewed in [Experiment 1](#). The T1, when presented alone in the stream, was identified with a probability of .93 ($SE = .03$), as compared with .96 in the stationary condition in [Experiment 1](#) and .91 in the pursuit condition.

During trials with two targets, average T1 performance across all lags was .86 ($SE = .013$). As in [Experiment 1](#), T1 performance was fairly stable across all lags apart from lag 1, where it dipped by about 10% ([Fig. 3](#)). A single-factor repeated measures ANOVA revealed a significant difference across lags, $F(6, 60) = 4.4$, $p < .001$, $\eta_p^2 = .30$.

In the analysis of T2|T1, average performance across all lags was .83 ($SE = .015$), practically the same as the mean in the pursuit condition in [Experiment 1](#). Unlike in [Experiment 1](#), however, there was no effect of lag, $F(6, 60) = 1.69$, $p = .139$, $\eta_p^2 = .14$. Excluding lag 1 from the analysis failed to reveal a significant lag effect, $F(5, 50) = 1.96$, $p = .1$, $\eta_p^2 = .16$. Numerically, we observed that rather than a robust attentional blink occurring over lags 2, 3, and 4, as in [Experiment 1](#), there was only a modest impairment at lag 2, which recovered entirely by lag 3.

There are two remarkable aspects to these data. The first is that overall performance was relatively good despite the facts that the MRSVP stream was presented 4° of visual

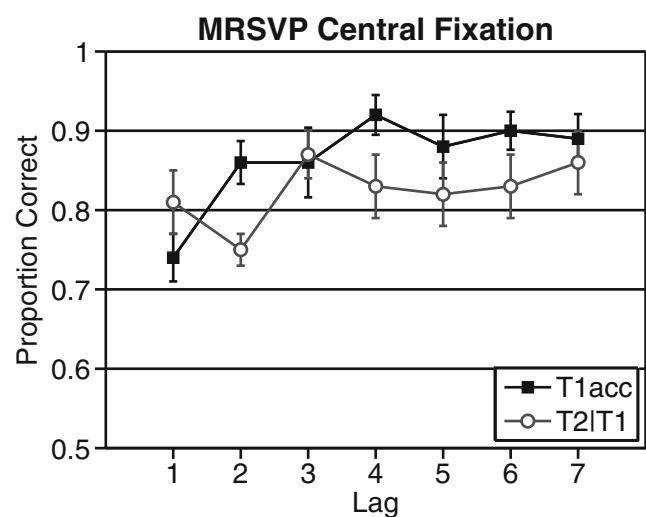


Fig. 3 T1 and T2|T1 accuracy for [Experiment 2](#). Participants monitored a moving RSVP stream for two targets while keeping their eyes fixed in the middle of the display

angle away from fixation, each item was present only for 100 ms before the next item was presented, and the targets were shifted across the retina, with a shift of 0.24° of visual angle every 14.3 ms. Each MRSVP item was composed of seven slightly offset presentations of the same item on the retina that overlapped so much that one might expect strong masking. Nevertheless, the visual system was clearly able to identify these overlapping presentations efficiently enough that target identification was not strongly compromised relative to the previous experiment's conditions, in which the RSVP stream was stationary or moving and pursued. This finding matches well with the percept of moving stimuli as sharply defined images and provides an objective indication of the ability to identify a moving stimulus (Bex, Edgar, & Smith, 1995; Burr, 1980).

The second remarkable aspect is the muted attentional blink, which suggests that while the MRSVP stream is phenomenally similar to a conventional RSVP stream (apart from moving), there may be something fundamentally different about the temporal interaction of the stimuli within the stream.

Discussion

Contrary to our expectations, Experiment 2 indicated only a miniscule attentional blink for targets in an RSVP stream that was moving rapidly across the retina while the eyes remained fixated. Under very similar presentation conditions, altering instructions to the participant to follow the target with their eyes (Experiment 1) produced an attentional blink of normal magnitude and duration. Furthermore, the blink was not absent because the task was easier: T1 performance was, if anything, lower in Experiment 2 than in either condition in Experiment 1.

One possible explanation for the lack of an attentional blink in Experiment 2 is that while MRSVP provides normal backward masking in object-centered coordinates, masking may be reduced in the retinotopic coordinate frame, because the onset of a new item overlaps retinotopically with only part of the preceding item. Backward masking of the T2 is normally necessary for revealing an attentional blink (Giesbrecht & DiLollo, 1998).

To evaluate this possibility, in Experiment 3, we added a trailing mask to the MRSVP on half the trials. Each item was

accompanied by a masking item in the location the current item had occupied when it first appeared. This experiment also allowed us to replicate the results of Experiment 2.

Experiment 3

Method

This experiment was similar to the design of Experiment 2, except where noted.

Participants Nineteen participants from the Syracuse University psychology study pool participated in this experiment for course credit. All had normal or corrected-to-normal vision and were fluent in English. The results of 1 participant were excluded because fewer than 20% of the trials were correct. Four additional participants were eliminated from the sample for failing to maintain fixation on at least one third of the trials.

Stimuli The MRSVP stream was composed of an animation sequence like that in Experiments 1 and 2. On each screen refresh cycle, the MRSVP advanced 3.2° OA degrees in a clockwise direction on a CRT with a refresh rate of 75 Hz. Each item in the MRSVP stream was present for eight frames (106.7 ms).

The masking stimulus appeared as a second stimulus that rotated in lockstep with the MRSVP, lagging spatially by eight positions. The mask was composed of a hash mark (#) superimposed on top of an @ sign and scaled to similar dimensions as the stimuli in the MRSVP. Therefore, when participants were fixating successfully, the mask would provide a strong backward mask that was, in effect, dragged behind the leading stimulus, sweeping over the retinotopic trace of the leading stimulus with a time lag of 107 ms. Figure 4 illustrates a series of frames from this animation.

Each participant saw 8 practice trials on which the MRSVP was slightly slower, 2 trials each of lags 1, 3, and 7, and 2 trials without a T1. Half of these practice trials had no mask; the other half had a mask. The experimental block consisted of a single block of 192 trials evenly distributed

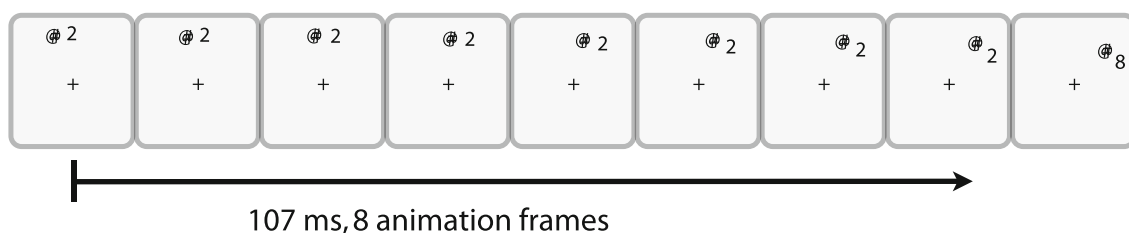


Fig. 4 One item in the MRSVP and its trailing mask. As the MRSVP orbits in clockwise direction, the mask follows behind it at a latency of 107 ms. In this diagram, the shift in location per frame has been doubled to make the change of position clearly visible

across eight conditions (lags 1–7 and T1 only). Half of the trials had a moving mask, and half had no mask; these trial types were intermixed randomly.

Eye tracking Eye movements were monitored at 100 Hz by an Eyelink 1000 eyetracker. Observers sat 70 cm from a 17-in. CRT monitor with their chin in a headrest. The left eye was tracked by default, but tracking was switched to the right eye if necessary (3 participants). The refresh rate of the monitor was 75 hz. Calibration occurred at the beginning of each experiment and involved fixating nine locations. Validation occurred for a set of nine evenly distributed locations and was considered valid if all errors were less than 2° of visual angle. Calibration was repeated until successful. Participant gaze was monitored throughout each trial, and trials were eliminated according to the same criteria as in Experiment 2.

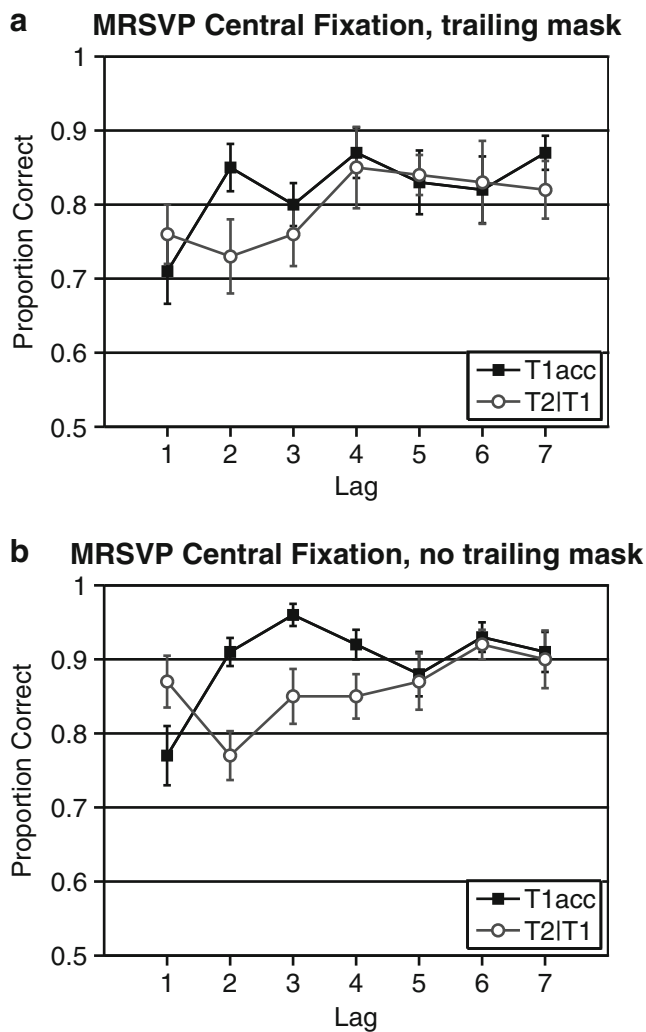


Fig. 5 T1 and T2|T1 accuracy for Experiment 3 in MRSVP conditions, with central fixation in the trailing mask condition (top) and no trailing mask condition (bottom). As in Experiment 2, participants monitored the moving RSVP stream for two targets while keeping their gaze fixed in the middle of the display

Results

The results from the two conditions are graphed separately in Fig. 5. Combining the two conditions together into a single ANOVA with lag (1–7) and condition (maskless vs. mask) as factors, the mask clearly reduced accuracy of reporting targets. There was a main effect of condition for T1 accuracy, $F(1, 13) = 33.4, p < .001, \eta_p^2 = .71$, and average T1 accuracy across all lags was .90 ($SE = .011$) for the maskless condition and .82 ($SE = .014$) for the masked condition. In the conditions for which participants had only one target to report, accuracy was .93 ($SE = .026$) and .87 ($SE = .03$) in the maskless and masked conditions, respectively. There was also a main effect of lag, $F(6, 78) = 8.4, p < .001, \eta_p^2 = .39$, due in large part to the usual dip in T1 accuracy at lag 1.

For T2|T1 accuracy, there was a main effect of condition, $F(1, 13) = 10.6, p < .01, \eta_p^2 = .44$, with performance on the masked trials at .80 ($SE = .018$), as compared with trials in the maskless condition, for which performance was .86 ($SE = .014$). There was also a main effect of lag, $F(6, 78) = 2.3, p < .04, \eta_p^2 = .15$, but no interaction, $F(6, 78) = 0.74, p > .6, \eta_p^2 = .05$, suggesting that the mask was not effective in increasing the attentional blink effect.

The small effect of lag replicated the numerical trends present in Experiment 2 very closely. To facilitate comparison across the three experiments, we present the data from all five conditions compiled into a single graph in Fig. 6. In each condition, raw T2 performance is baselined by subtracting T1 accuracy on those trials when T2 was not

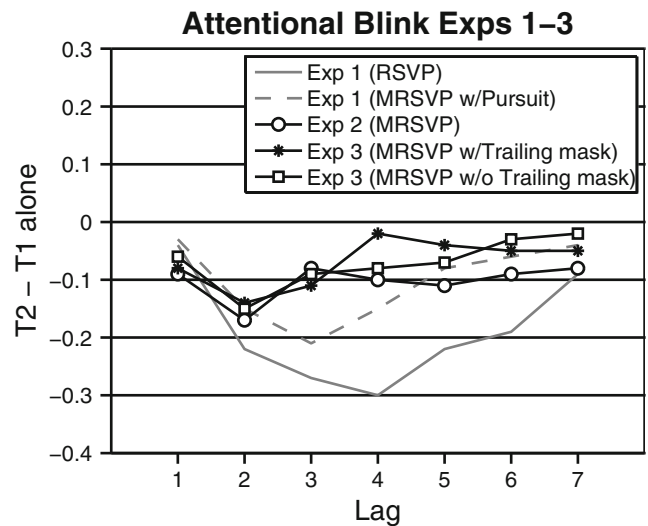


Fig. 6 Comparison of data between all five experimental conditions in Experiments 1, 2, and 3. Shown is the subtraction of accuracy for a single target for T2 accuracy. Negative scores reflect a deeper attentional blink. In Experiment 1, participants were asked to pursue the MRSVP with their eye or to view a central, immobile RSVP stream. Experiments 2 and 3 had participants perceive an MRSVP stream with their eye gaze fixed in the center of the screen

presented. On those T1-only trials, the T1 was presented in any of the temporal positions that T2 would have occupied, so these data provide a good estimate of what T2 accuracy would have been had the T1 not been presented. We observe a clear difference between the small attentional blink in all three conditions in which participants fixated in the center while viewing the MRSVP and the conventional attentional blink in the two conditions in which the participant kept the RSVP near the fovea, either with central fixation or by using pursuit movements.

General discussion

These results are informative in two respects. First, the ability to report targets from an RSVP stream is barely compromised when that stream is shifting at a rate of 17° of visual angle per second. When the RSVP stream was stationary, a solitary target was correctly reported on .96 of the trials; when participants attempted to pursue the MRSVP, the single targets were reported on .91 of the trials. In Experiments 2 and 3, which involved a peripheral MRSVP stream orbiting the fixation cross while the eyes were fixed, trials containing one target had accuracies of .93 (Experiment 2), .93 (Experiment 3 without a trailing mask), and .87 (Experiment 3 with a trailing mask). This is particularly significant because the stimuli were themselves $1.0^\circ \times 1.5^\circ$ of visual angle in size and, therefore, the overlap between sequential frames of the animation sequence would have led to considerable masking had the stimuli not been identical.

Prior behavioral research has found that we fail to integrate shape information over multiple presentations of a stimulus with a moving guide that produced apparent motion (Cavanagh et al., 2008), which suggests that motion should disrupt or at least delay the processing of stimulus identity. However, we found that identification of a target was quite easy in a moving RSVP stream, with or without pursuit, and not much more difficult than identification of a target in a stationary RSVP stream. The discrepancy between the results of Cavanagh et al. and our own work may result from differences in the separation of sequential frames. In the Cavanagh et al. study, the visual system had to integrate over much larger distances (about 4° of visual angle) and at a much higher velocity than in the present study.² In the paradigm they used, the subsequent frames of

the apparent motion did not spatially overlap, and strong masks were added to make the task difficult. Replicating the spatiotemporal parameters of their motion in an RSVP context results in nonoverlapping presentations of individual stimuli; the stimuli would be unmasked and, therefore, quite easy to report. Thus, the difference between the role of motion in the two studies likely stems from the fact that the distance in terms of visual angle from one stimulus presentation to the next was about 4° in the Cavanagh et al. study and about 0.16° in our study. Therefore, we conclude that the results of the Cavanagh et al. study may be particular to situations in which the visual system has to combine visual form information over a considerable retinotopic distance. A second reason for surprise at the relative ease of identifying targets in moving stimuli is that the ability to rapidly identify stimuli, such as in RSVP, is thought to be the result of feedforward processing in the visual system (Thorpe, Fize, & Marlot, 1996; VanRullen, 2007). Consequently, since a visual stimulus that moves retinotopically within our visual field must leave in its wake an overlapping trail of activated neurons in both retinal and early visual cortical areas, one might expect that feedforward processing of retinotopically moving targets would be impaired, strongly reducing RSVP target identification accuracy relative to a stationary stream. This description may, however, underestimate feedforward visual processing. Models that represent object identification as a feedforward process through the ventral stream (Serre, Oliva, & Poggio, 2007) may have little difficulty integrating identity information across multiple frames. In fact, depending on the integration time constant of simulated neurons in the model, computing the identity of a stationary object might be no different from computing the identity of a moving object. The answer to this question awaits the development of feedforward models that process scenes that change over time and do so with temporal characteristics that match single-cell data from monkeys. Our results, which suggest that our participants have a coherent, easily identified percept of an RSVP target, also support theories suggesting that the visual system has mechanisms that update the representation of an object at one location with information from a second object presented closely in space and time (Enns & Di Lollo, 1997; Enns et al., 2010).

The attentional blink

While the attentional blink was of the usual amplitude and duration for stationary RSVP streams and streams under visual pursuit (Experiment 1), the attentional blink was nearly absent for an RSVP stream that was moving while participants held their gaze fixed (Experiments 2 and 3). The results of Experiment 1's pursuit condition are in good agreement with the data in Lunau and Olivers (2010), in

² In the Cavanagh et al. (2008) study, the presentation rate at which the unguided performance (i.e., the condition without guides to assist in tracking the motion) fell below threshold was about 5 Hz, or 100-ms SOA between stimulus and mask. At this rate of presentation, the target stimulus inside the moving guide was moving around the circle in jumps of 4° of visual angle per 100 ms, or a velocity of 40° per second. In our experiment, the stimulus moved around the orbit at a rate of 17° of visual angle per second.

which the T1 and T2 were presented in different spatial positions along a sequence of 27 independent RSVP streams and participants were encouraged to move their eyes along a sequence of these streams, following a moving cue. Their study found a conventional attentional blink effect.

It is unclear why the attentional blink is diminished when the stimulus is moving and the eye position is held fixed. We can rule out some possibilities, however. The obvious explanation to consider when an attentional blink is absent is that processing has become too easy. This is unlikely because the MRSVP was in the periphery for Experiments 2 and 3, rather than in the fovea, and this should have made target detection more difficult rather than easier. Indeed, performance in identifying a single target was numerically worse in all three MRSVP conditions than in the central RSVP condition.

Another possible explanation might be that the attentional blink is simply absent outside of the fovea. However, there are numerous experiments in which peripheral RSVP streams have produced an attentional blink effect (Craston, Wyble, Chennu, & Bowman, 2009; Ho & Cheung, 2010; Jefferies et al., 2007; Shih, 2000; Visser et al., 1999). These findings also eliminate another potential explanation, which would be that separating covert attention from the fovea eliminates the attentional blink.

It is also true that participants were only subjected to eye tracking in Experiments 2 and 3, when the attentional blink was reduced in size. However the attentional blink has been found during eye tracking in other experiments (e.g., Armstrong & Munoz, 2003). Other than the presence of the eyetracker, the experiments were all run on the same type of computer with the same software. Experiments 1 and 3 were run in the same experimental testing room, using participants from the same participation pool. Furthermore, time stamps associated with each stimulus onset were checked to verify that eye tracking did not affect the presentation rate of the MRSVP streams.

It seems likely that the processing required to identify a parafoveal moving stimuli plays a direct role in reducing the attentional blink effect. For example, it is known that a stimulus (e.g., a fingertip) in motion appears to move faster when in the periphery than when under ocular pursuit, an effect called the Aubert–Fleischl effect (Dichgans, Wist, Diener, & Brandt, 1975), so it is true that processing of stimuli differed in perceptually important ways between Experiments 1 and 2 in the present study. It is possible that the integration of multiple frames of motion into a single-object representation obscures the temporal boundaries of targets and distractors in the MRSVP stream. This explanation fits well with a recent explanation of the attentional blink that proposes that the visual system segments an RSVP stream into attentional episodes and the blink reflects these episodic boundaries (Wyble, Bowman, & Nieuwenstein

2009; Wyble, Potter, Bowman, & Nieuwenstein, 2011). According to these models, when RSVP targets are separated by a short temporal gap (e.g., lag 2, or about 200 ms), a competitive interaction between excitatory and inhibitory influences on attention delays encoding the second target to maintain the episodic distinctiveness between the two targets. When the second target is masked, this delay produces an attentional blink effect. This model thus explains why a sequence of continuous targets fails to exhibit the attentional blink at all (Di Lollo et al., 2005; Kawahara, Kumada, & Di Lollo, 2006; Olivers, van der Stigchel, & Hulleman, 2007) and why a blank gap is sufficient to trigger an attentional blink even in the absence of distractors (Nieuwenstein et al., 2009; Wyble et al., 2011). Turning to the present experiments, if the processing of stimuli moving in the parafovea reduces the temporal separability of target representations, the temporal gaps between targets that normally trigger an attentional blink may be obscured. In other words, the fact that an RSVP is moving in the periphery may cause the visual system to treat all of its component stimuli as a single changing object in order to integrate the motion frames into recognizable forms. This explanation is congruent with more direct findings that object continuity reduces the attentional blink, as in the study by Raymond (2003), in which the RSVP stream was composed of a rotating trident symbol and the targets were featural alterations on this rotating stimulus.

Conclusion

The visual system is able to accurately select targets in an RSVP stream that is moving in the retinotopic coordinate frame. We have found, in the present experiments, that untrained participants are able to identify targets very easily even when stimuli are presented in sequential, overlapping video frames that produce a percept of coherent motion, whether they follow the motion with their eyes or with covert attention while their eyes are fixed elsewhere. In an unexpected turn, the latter form of target processing is nearly immune to the attentional blink. This finding suggests that the temporal separability between targets is reduced in RSVP streams that are moving in the periphery.

References

- Armstrong, I. T., & Munoz, D. P. (2003). Attentional blink in adults with attention-deficit hyperactivity disorder: Influence of eye movements. *Experimental Brain Research*, *152*, 243–250.
- Attenave, F., & Block, G. (1974). Absence of masking in the path of apparent movement. *Perception & Psychophysics*, *16*, 205–207.
- Barnard, P. J., Scott, S., Taylor, J., May, H., & Knightley, W. (2004). Paying Attention to Meaning. *Psychological Science*, *15*, 179–186.

- Bex, P. J., Edgar, G. K., & Smith, A. T. (1995). Sharpening of drifting blurred images. *Vision Research*, *35*, 2539–2546.
- Bowman, H., & Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review*, *114*, 38–70.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- Burr, D. (1980). Motion smear. *Nature*, *284*, 164–165.
- Cavanagh, P., Holcombe, A. O., & Chou, W. (2008). Mobile computation: Spatiotemporal integration of the properties of objects in motion. *Journal of Vision*, *8*(12, Art. 1), 1–23.
- Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 109–127.
- Craston, P., Wyble, B., Chennu, S., & Bowman, H. (2009). The attentional blink reveals serial working memory encoding: Evidence from virtual and human event-related potentials. *Journal of Cognitive Neuroscience*, *21*, 550–566.
- Di Lollo, V., Kawahara, J., Gorashi, S., & Enns, J. T. (2005). The attentional blink: Resource depletion or temporary loss of control? *Psychological Research*, *69*, 191–200.
- Dichgans, J., Wist, E., Diener, H. C., & Brandt, T. (1975). The Auber–Fleischl phenomenon: A temporal frequency effect on perceived velocity in afferent motion perception. *Experimental Brain Research*, *23*, 529–533.
- Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of visual masking in unattended visual locations. *Psychological Science*, *8*, 135–139.
- Enns, J. T., Lleras, A., & Moore, C. M. (2010). Object updating: A force for perceptual continuity and scene stability in human vision. In R. Nijhawan (Ed.), *Problems of space and time in perception and action* (pp. 503–520). Cambridge: Cambridge University Press.
- Forster, K. I. (1970). Visual perception of rapidly presented word sequences of varying complexity. *Perception & Psychophysics*, *8*, 215–221.
- Giesbrecht, B., & Di Lollo, V. (1998). Beyond the attentional blink: Visual masking by object substitution. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1454–1466.
- Ho, C., & Cheung, S.-H. (2010). *Temporal resolution of attention in foveal and peripheral vision*. Naples: Poster presented at the Vision Science Society Conference.
- Jefferies, L. N., Ghorashi, S., Kawahara, J., & Di Lollo, V. (2007). Ignorance is bliss: The role of observer expectation in dynamic spatial tuning of the attentional focus. *Perception & Psychophysics*, *69*, 1162–1174.
- Jolicouer, P. (1999). Concurrent response-selection demands modulate the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1097–1111.
- Kawahara, J., Kumada, T., & DiLollo, V. (2006). The attentional blink is governed by a temporary loss of control. *Psychonomic Bulletin & Review*, *13*, 886–890.
- Kolers, P. A. (1964). Apparent movement of a Necker cube. *The American Journal of Psychology*, *77*, 220–230.
- Lunaru, R., & Olivers, C. N. L. (2010). The attentional blink and lag 1 sparing are nonspatial. *Attention, Perception, & Psychophysics*, *72*, 317–325.
- Martens, S., & Wyble, B. (2010). The attentional blink: Past, present, and future of a blind spot in perceptual awareness. *Neuroscience and Biobehavioral Reviews*, *34*, 947–957.
- Nieuwenstein, M. R., Potter, M. C., & Theeuwes, J. (2009). Unmasking the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 159–169.
- Ögmen, H., Otto, T. U., & Herzog, M. H. (2006). Perceptual grouping induces non-retinotopic feature attribution in human vision. *Vision Research*, *46*, 3234–3242.
- Olivers, C. N. L., & Meeter, M. (2008). A boost and bounce theory of temporal attention. *Psychological Review*, *115*, 836–863.
- Olivers, C. N., van der Stigchel, S., & Hulleman, J. (2007). Spreading the sparing: Against a limited-capacity account of the attentional blink. *Psychological Research*, *71*, 126–139.
- Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 509–522.
- Potter, M. C., Staub, A., & O'Connor, D. H. (2002). The time course of competition for attention: Attention is initially labile. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1149–1162.
- Raymond, J. E. (2003). New objects, not new features, trigger the attentional blink. *Psychological Science*, *14*, 54–59.
- Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 849–860.
- Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences*, *104*, 6424–6429.
- Shih, S. (2000). Recall of two visual targets embedded in RSVP streams of distractors depends on their temporal and spatial relationship. *Perception & Psychophysics*, *62*, 1348–1355.
- Thorpe, S. J., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520–522.
- VanRullen, R. (2007). The power of the feed-forward sweep. *Advances in Cognitive Psychology*, *3*, 167–176.
- Visser, T. A., Zuvic, S. M., Bischof, W. F., & Di Lollo, V. (1999). The attentional blink with targets in different spatial locations. *Psychonomic Bulletin & Review*, *6*, 432–443.
- Wyble, B., Bowman, H., & Nieuwenstein, M. (2009). The attentional blink provides episodic distinctiveness: Sparing at a cost. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 324–337.
- Wyble, B., Bowman, H., & Potter, M. C. (2009). Categorically defined targets trigger spatiotemporal attention. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 324–337.
- Wyble, B., Potter, M. C., Bowman, H., & Nieuwenstein, M. (2011). Attentional episodes in visual perception. *Journal of Experimental Psychology: General*, *140*, 488–505.
- Yantis, S., & Nakama, T. (1998). Visual interactions in the path of apparent motion. *Nature Neuroscience*, *1*, 508–512.