An unattended mask makes an attended target disappear

Arielle A. Veenemans

Laboratoire Psychologie de la Perception,
Université Paris Descartes, Paris, France ✉

Patrick Cavanagh

Laboratoire Psychologie de la Perception,
Université Paris Descartes, Paris, France
Department of Psychological and Brain Sciences,
Dartmouth College, Hanover, NH, USA 🏠 ✉

In pattern masking, the target and mask are presented at the same location and follow one another very closely in time. When the observer attends to the target, he or she must also attend to the mask, as the switching time for attention is quite slow. In a series of experiments, we present mask–target–mask sequences staggered in time and location (Cavanagh, Holcombe, & Chou, 2008) that allow participants to attentively track the target location without attending to the masks. The results show that the strength of masking is on average unaffected by the removal of attention from the masks. Moreover, after isolating the target location perceptually with moving attention, it is clear that the target, when at threshold, has not been degraded or integrated with a persisting mask but it has vanished. We also show that the strength of masking is unaffected by the lateral spacing between adjacent target and mask sequences until the spacing is so large that the apparent motion driving the attentive tracking breaks down. Finally, we compare the effect of the pre- and postmask and find that the premask is responsible for the larger part of the masking.

Introduction

In this article, we address the relation between masking and attention. Masking is widely used in visual and cognitive research to control or limit target processing and generally produces a reduction in visibility of a target due to close temporal or spatial proximity of a mask (Werner, 1935; Alpern, 1953). Although there are several distinct types of masking (pattern masking, lateral masking, metacontrast masking; for review, see Breitmeyer, 1984; Enns & Di Lollo, 2000), we focus here on forward and backward pattern masking (e.g., Smith & Schiller, 1966; Nachmias & Rogowitz, 1983; Foley & Boynton, 1993) where

the target and the mask are superimposed spatially, but separated temporally.

In these experiments we examine whether the strength of masking is influenced by whether the mask itself is attended or not using a moving attention technique that lets observers see and attend to the target without attending to the masks (Cavanagh, Holcombe, & Chou, 2008). This technique also lets observers inspect what the target looks like in isolation at near threshold levels so we can ask whether the masked target becomes an unrecognizable mixture of target and mask elements or if it is simply suppressed to invisibility. In standard temporal masking, the target and mask follow each other in quick succession at the same location, so directing attention to the target also unavoidably directs attention to the mask (see Figure 1 for the version of this stationary attention condition used in the current study).

The amount of time necessary to redirect attention from one stimulus to another, also called attentional dwell time, is thought to be in the order of 200–500 ms (Moore, Egeth, Berglan, & Luck, 1996; Ward, Duncan, & Shapiro, 1997; Theeuwes, Godijn, & Pratt, 2004). Generally, the strength of masking is determined by the time separating the mask and the target (cf. Breitmeyer, 1984). Since masking is strongest within 100 ms before or after the target appears (Breitmeyer, 1984; Breitmeyer & Öğmen, 2000; Polat & Sagi, 2006; Sterkin, Yehezkel, Bonneh, Norcia, & Polat, 2009; Polat, Sterkin, & Yehezkel, 2010), attention to the target for this timing will necessarily also pick up the mask. In other words, with standard forward and backward masking, it is not possible to present a mask and target in sequence fast enough to create masking but slow enough to attend to the target but not the mask, and so it is not possible to determine if the masking depends on the attention allocated to the mask.
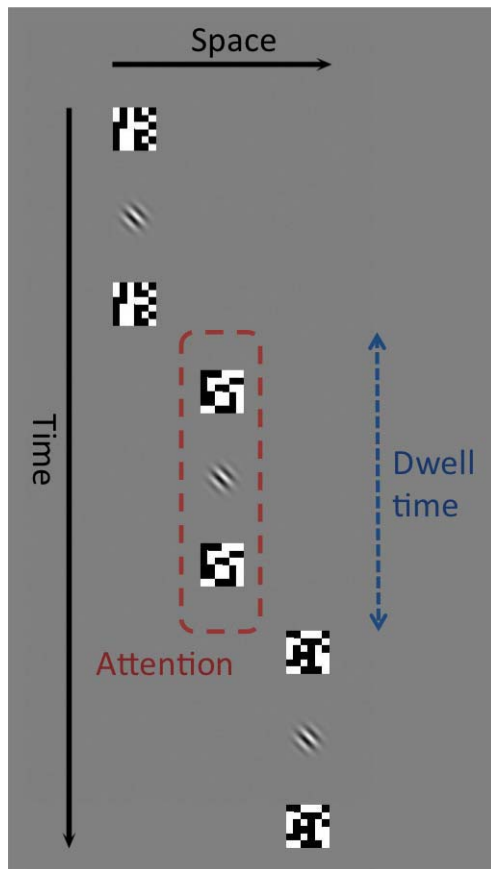
Figure 1. Stationary attention. The mask–target–mask sequence is presented in one location before shifting to an adjacent location after the three full frames are completed. Because of its minimum "dwell time," attention directed to the target cannot avoid also selecting one or both of the masks.

These characteristics of standard forward and backward masking were obtained using stationary attention, but in the case of moving attention, the dwell time of attention at a given location can be significantly reduced, allowing attention and stimulus timing to be manipulated independently (Figure 1). In particular, Cavanagh et al. (2008) showed that attention can move to follow a target and single out a particular location in as little as 50 ms without selecting information that was at that same location just before and after the target. Moving attention is thus fast enough to extract the target information from a rapidly alternating stream of masks and targets and it allows observers to process information that would otherwise have remained inaccessible.

This article extends upon this earlier study on the effectiveness of unattended masks on recognition (Cavanagh et al., 2008) to investigate the effect that attention to the mask has on the strength of masking and to determine the nature of the target representation near threshold. In the moving target technique developed by Cavanagh et al., attention can avoid the masks, allowing the target location to be inspected in isolation (Figure 2 shows the version of this stepping
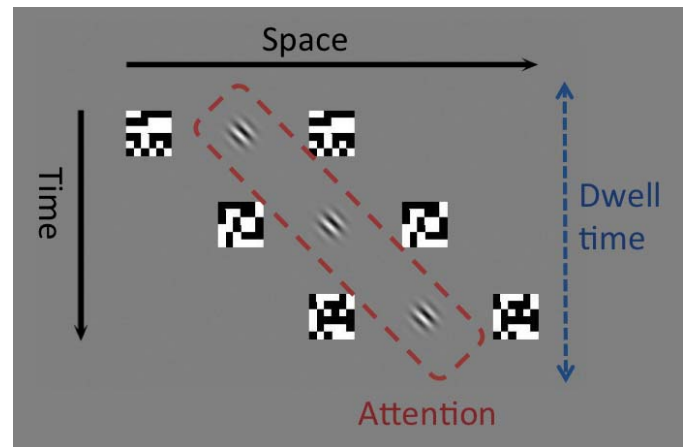

Figure 2. Moving attention. The mask–target–mask sequence remains the same at each location, but the next sequence appears after the first frame. In this case, the target appears between the two masks and attention can track and isolate the target while avoiding the masks that still precede and follow the target at each location.

attention method as used in the current study). Specifically, as can be seen in demonstration Movie 1, when a temporal sequence of mask–target–mask is presented at each adjacent location but with each new sequence delayed by one frame, the perceptual organization of the masks and targets is radically altered. Instead of seeing each sequence in place, followed by the next (as is the case in the static condition, see Figure 1), the target now appears flanked by two masks in each frame. This spatial group of mask–target–mask then steps to the next location and by following the target from frame to frame, it is possible to get a clear view of the target location that is unobstructed by the masks.

The first condition in the Cavanagh et al. (2008) study used a replacement masking task where two similar patterns alternated continuously at each location. In this case, it was far easier to report the properties of a pattern when stepping attention allowed it to be selected while the other remained unattended. The stimuli were random dots that alternated in color and direction (e.g., target of red dots moving upward alternating with mask of green dots moving downward). Moving attention provided a 250% improvement over static attention. This suggested that masking in this task was occurring between representations in the stream of attended input so that if the competing pattern was not selected (in the moving attention case), it had much less effect. This effect of attention on higher level, *interruption* masking (Spencer & Shuntich, 1970; Breitmeyer, 1984; Enns & Di Lollo, 1997; Lleras & Moore, 2003) shows that the stepping attention technique can reveal attentional effects on masking when they are present.
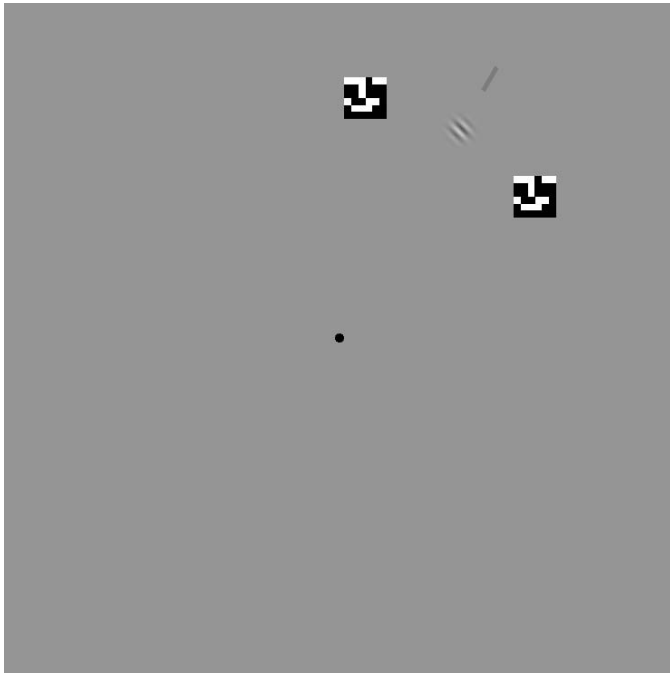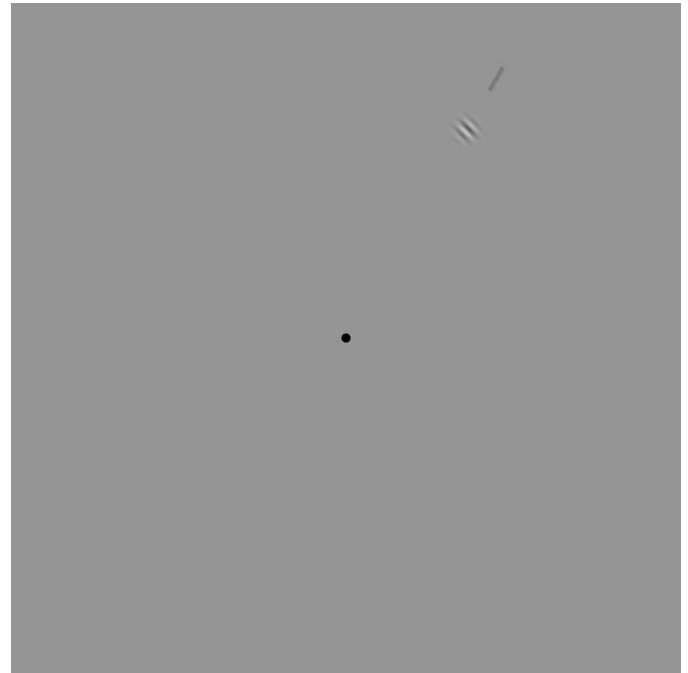
Movie 1. Dynamic attention condition. Target and masks rotate together around fixation while target contrast ramps up.

Movie 2. Stationary attention condition. Alternation of mask–target–mask sequence at each location while target contrast ramps up.

In contrast, in the second pattern-masking task of the Cavanagh et al. (2008) study, participants had to report the identity of a target letter that alternated with a random-dot mask. In this case, moving attention provided no advantage, suggesting that the target letter was already degraded by the mask when the letter's location was selected by attention. This simple pattern masking corresponded to low-level *integration* masking, where stimulus and mask features are integrated at an early level prior to selection by attention (Schiller, 1966; Breitmeyer, 1984). As a consequence, masking was strong whether or not the mask was selected by attention in the moving attention procedure.

However, Cavanagh et al. (2008) did not address the question of what was seen at the target location when the pattern masking was effective. At identification threshold, the target may still have been above its detection threshold—something may have been visible—even though it could not be identified. We were interested to see whether any target information was still accessible when the masks were unattended, so we used the stepping attention technique with a detection task to find out.

We first compared thresholds for identification and detection using procedures that were identical except for the response. In both cases we tested a static condition where the target was seen and attended *together* with the mask (Figure 1) and a moving condition where the target could be seen and attended *separately* from the mask (Figure 2). If the effect of the mask is to produce a degraded mixture of target and

mask that remains visible but unrecognizable at the target location, then the detection threshold should be lower than the identification threshold in the moving attention condition. Additionally, the detection threshold itself should be lower in the moving than in the static condition. Specifically, in the moving case, the target location can be isolated and any residual target–mask mixture at that location has only to be distinguished from a blank location, whereas in the static case, the residual target–mask mixture must be distinguished from the mask itself since both mask and target are necessarily selected together. So if masking obscures the target by integrating with it, leaving an unrecognizable mixture at the target location, then the detection threshold should be lower than the identification threshold in the moving condition, but additionally the detection threshold in the moving condition should also be lower than the detection threshold in the static condition. Conversely, if these three thresholds do not differ, it would suggest that the target is simply suppressed to invisibility by the mask, leaving no residual target–mask mixture at the target location.

Additionally, as can be seen in the demonstration movies (Movies 1 and 2), the stepping stimulus reorganizes the spatial layout of the adjacent mask–target–mask sequences so that the target appears flanked by two masks. This spatial adjacency could introduce a component of lateral masking (Werner, 1935; Alpern, 1953; Growney, Weisstein, & Cox, 1977) or, with a spacing of about 50% of eccentricity

or less, it could generate crowding (Bouma, 1970; Toet & Levi, 1992). To determine whether the spatial layout introduced any lateral masking or crowding, we measured the effect of the spacing between the target and the masks in each frame on detection thresholds (Werner, 1935; for review, Herzog, 2006).

Finally, we compared the contribution of onset and offset transients of the target to the effectiveness of the masking. There is still some debate about the extent to which luminance transients capture attention (Yantis & Jonides, 1984; Theeuwes, 1995; Franconeri, Hollingworth, & Simons, 2005). Nevertheless, there is a general consensus that onset transients play a larger role in the allocation of attention than offset transients (Yantis & Jonides, 1984; Jonides & Yantis, 1988; Miller, 1989). A recent study by Motoyoshi and Hayakawa (2010) showed that when adaptation suppressed the onset transient of a stimulus, it could be blocked from awareness even at 100% contrast. To compare the effectiveness of onset and offset transients of the target, we masked the target separately with a premask or a postmask, or both.

Experiment 1: Detection versus identification

In the first experiment we compared masking thresholds for identification and detection tasks. The effect is tested both for dynamic targets that can be attended without attending to the masks (Figures 2 and 3) and for static targets that are attended together with the masks (Figures 1 and 4).

To measure the thresholds in the following experiments, we used the anticipated threshold technique developed by Brussell and Cavanagh (1984). Similar to an ascending staircase, the contrast started at zero and increased until the subject detected the target. In this method, however, the increase was continuous within a trial, not across trials. To reduce overshoot effects caused by the response delay, the contrast rose rapidly until just before the threshold for that particular condition was reached (10% contrast before threshold), after which the contrast continued to rise more slowly until detection.

Methods

Participants

Six subjects participated in the experiment, including author AV. They ranged in age from 25 to 35 years and had normal or corrected-to-normal visual acuity. All
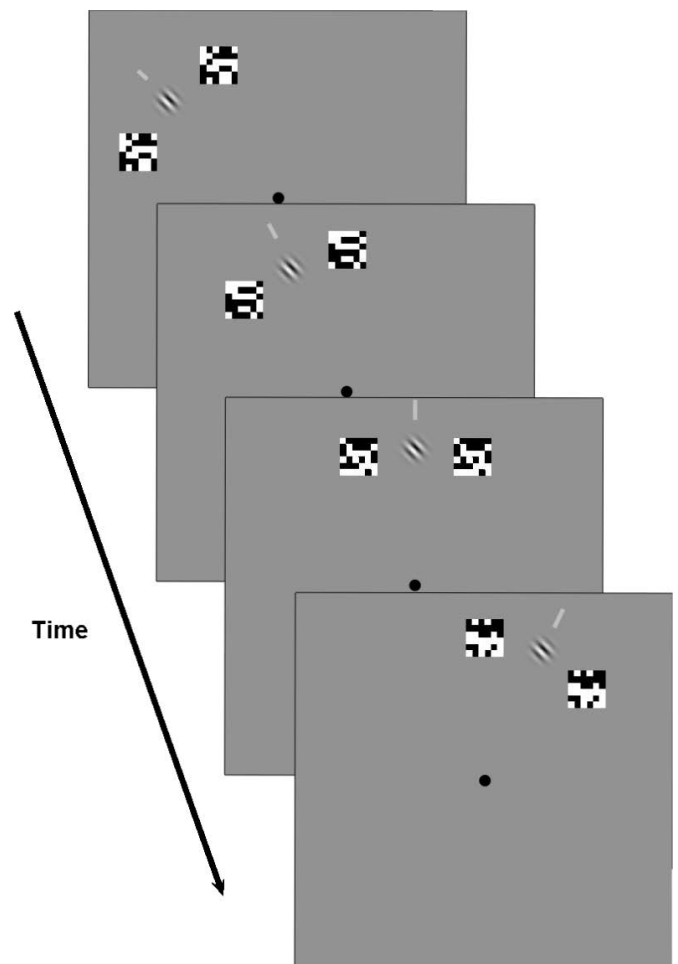

Figure 3. Time course of the dynamic condition of Experiment 1. The stimuli rotate around fixation with every element taking one step forward on each new frame. Nevertheless, note that the standard mask–target–mask sequence is presented at each location across three successive frames (compare, for example, the topmost element in the second, third, and fourth panels here). A tick mark is always present adjacent to the target location to help guide tracking.

participants gave informed consent in writing prior to participation and the protocols for the study were approved by the Université Paris Descartes Review Board, CERES, in accordance with French regulations and the Declaration of Helsinki. Participants were compensated 10€ per hour for their time.

Apparatus

Stimuli were displayed on a 22-in. LaCie Electron22 Blue IV CRT monitor (LaCie, Paris, France) at a vertical refresh rate of 100 Hz, resulting in 10 ms per refresh, with a resolution of 1024 × 768 pixels and stimuli were generated on a Macintosh G4 computer (Apple, Inc., Cupertino, CA) using MATLAB (The MathWorks, Inc., Natick, MA) and the Psychophysics
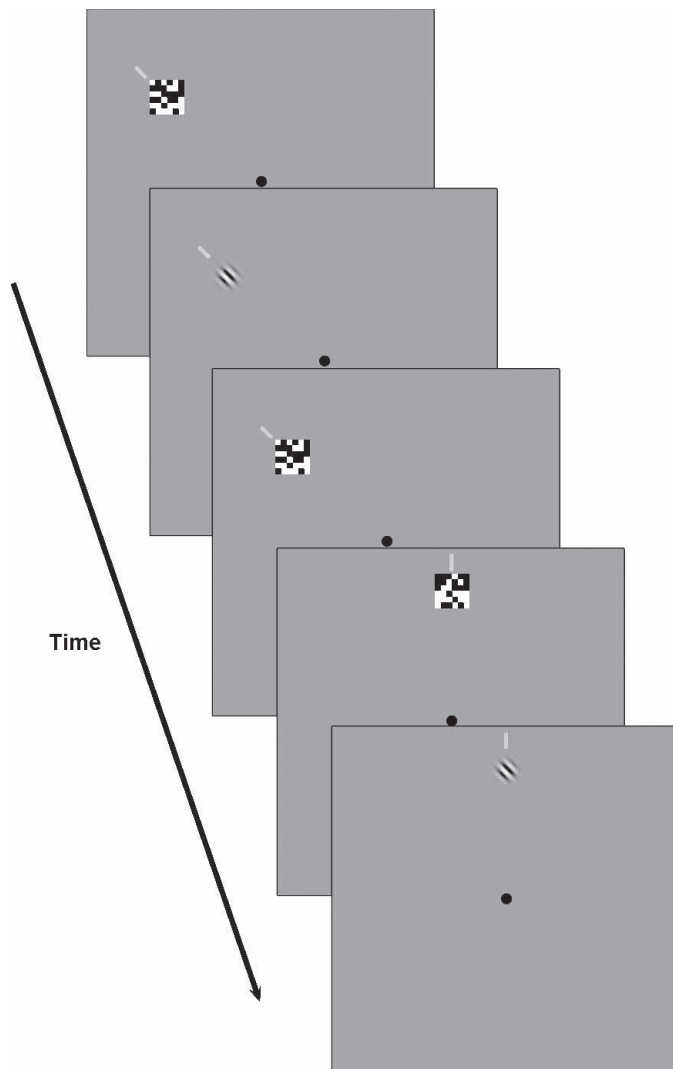
Time

Figure 4. Time course of the static condition of Experiment 1. The stimuli rotate around fixation with one single element displayed on each frame. The mask–target–mask sequence is presented at each location before jumping to the next location. A tick mark is always present next to the target location to help guide tracking.

Toolbox functions (Brainard, 1997). Subjects viewed the screen from 57-cm distance with their head stabilized by a chin rest.

Stimuli

A fixation mark (a black dot) was presented in the middle of a midlevel gray background on the screen. *Dynamic condition*: Two masks were presented at 13° from fixation and a target increasing slowly in contrast from 0% to 100% was presented between the two masks. A tick mark was placed 1° away from the outer edge of the position of the target to help guide attention to the target location. Pilots with and without the tick mark

did not result in a difference in threshold, indicating that it does not influence the detection of the stimulus.

Masks were composed of 18 black and 18 white blocks arranged in a randomized six by six pattern. This pattern was rerandomized for each sequence. The width of the masks was 2.8°, and it was always shown at full contrast. The target was a Gabor patch, a sinewave grating of 1.79 cpd, slightly smaller than the mask, that was tilted 45° to the right or to the left. Target and masks were arranged and spaced so that they never had any simultaneously overlapping locations with targets and masks in the adjacent sequences (they always overlapped within each sequence at a given location).

There were 15 possible positions around fixation where masks or target could appear, resulting in a 24°-of-rotation displacement of target and mask location with each step. Each frame had two masks and a central target mask, each separated by 24° of rotation (Figure 3; Movie 1). The target moved by one step each frame and there were six frames per second, giving an alternation rate of 3 Hz. The target and mask durations were a one frame duration (160 ms, 0 ms interstimulus interval (ISI); the stimulus onset asynchrony (SOA) is therefore the same as duration, 160 ms), so one complete revolution around the circle took 2.4 s.

On each trial the masks and target stepped around fixation, increasing contrast on each step until the participant responded or until 100% contrast was reached. The start location on each trial was randomized. Contrast increased quickly until 10% before the average threshold for that condition was reached (thresholds updated on each trial) and much more slowly after that. If 100% contrast was reached and no response had yet been made, the stimulus would appear a fixed amount of additional frames before the trial ended and a new one begun.

Static condition: To compare detection thresholds between tracked and untracked target conditions, we made a jumping, static version of the stimuli. We presented the sequence of masks and targets at successive locations as in the dynamic, tracked target case, but now each mask–target–mask sequence was completed at the same location before the next sequence began at the next adjacent location. This eliminated any impression of target motion that would allow the attention to focus on the target and avoid the masks (see Figure 4; Movie 2).

In the static condition, each frame had only a single location occupied by a mask or a target and that location stepped by 24° of rotation after the three frames required to complete the mask–target–mask sequence (Movie 2). Again there were six frames per second giving an alternation rate of 3 Hz. Because the target only changed location on every third frame in this condition, the revolution rate was one third that of the dynamic condition, and so one complete revolution
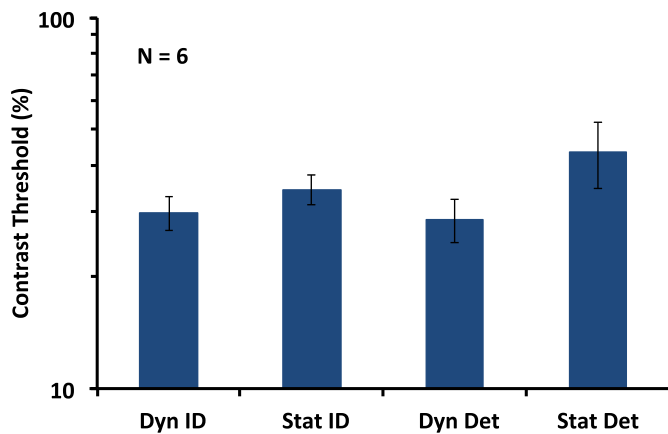
Figure 5. Experiment 1: Effect of task. Average detection threshold (log scale) at which participants were able to detect or identify the target for the dynamic or static condition. Error bars indicate the standard error of the mean.

of the circle took 7.2 s. All other parameters were similar to the dynamic condition.

Procedure

Participants were given practice trials in order to familiarize them with the task and the stimuli. They were instructed to maintain fixation, indicated by a black dot in the middle of the screen, at all times and attentively track the location, indexed by a tick mark just outside the target trajectory, as the masks traveled on a circular path around fixation. The contrast of the target would increase on each successive presentation. In the dynamic condition, the tick mark indexed the blank location between two adjacent masks and the target would slowly ramp up in contrast in this empty space. In the static condition, the tick mark again indexed the target location that in this case was perceived at the same location as the masks. In both cases, the target would slowly ramp up in contrast from one position to the next. In the detection block, observers were instructed to detect the target and respond with a key press as soon as they saw anything at the blank indexed location in the dynamic condition or anything in addition to the masks in the static condition. In the identification block, observers were instructed to report the target's orientation as soon as they were able, by pressing one of two arrow keys. Half of the subjects performed the detection block first, followed by the identification block, while the other half were given the identification block first, followed by the detection block.

Each participant performed a total of 200 trials for this experiment, 100 trials per block. Some trials were discarded due to accidental button presses right after the start of a trial when contrast was still 0% or because the participant paused that trial to take a break.

Results

No trials were discarded because the observer didn't respond before the trial ended at 100% contrast; 0.58% of trials were discarded because the participant responded while the target was still at 0% contrast due to accidental button press or a break. The error rate for the identification conditions, the percentage of incorrect orientation responses, was 2.69% in the dynamic condition and 4.77% in the static condition. This low error rate suggests that the rate of false alarms in the detection condition was similarly low.

The detection threshold (Figure 5) does not differ significantly from the identification threshold, ANOVA $F(1, 1) = 1.01$, $p = 0.361$, and there is no significant main effect of static versus dynamic condition, ANOVA $F(1, 1) = 6.201$, $p = 0.055$, or any interaction effect. The absence of an effect of attention to the masks for the identification task replicates the finding of Cavanagh et al. (2008), indicating again that the effect of the mask is preattentive—an early influence on the target that precedes attention's access to the target or the masks.

In addition, the absence of a difference in masking threshold for identification and detection suggests that the effect of masking is not to obscure the target in an integrated mixture of target and mask, but simply to suppress it. It does not become an unrecognizable mixture at identification threshold; it just vanishes so that detection is lost at around the same contrast as identification. Earlier studies (e.g., Kulikowski & Tolhurst, 1973; Thompson, 1983) have suggested that identification and detection thresholds may be the same when both tasks rely on the same signal. In this case, as soon as the target can be detected, it can also be identified, and conversely, if the target cannot be identified, it is simply not visible.

In addition to the lack of difference between detection and identification threshold, there was also a lack of significant difference between the detection thresholds for static and dynamic conditions. Again, if the effect of the mask at threshold was to generate a mixture of target and mask components, this should have been easier to detect in the dynamic condition where the target location is seen in isolation and the mask–target residual has only to be distinguished from a blank location rather than from the mask itself in the static condition. The absence of advantage for the dynamic condition suggests that at threshold the target is suppressed, leaving no residual trace.

Experiment 2: Timing

In our next experiment, we examined the influence of the frame alternation rate on detection thresholds.

Faster frame alternation decreases the duration of the target presentation and decreases the time between mask and target onset, both of which should make target detection more difficult. The increased alternation rate also increases the rate at which the targets and masks orbit fixation. The effects of the different rates are again tested for both the tracked dynamic targets (Figures 2 and 3) and the nontracked, static targets (Figures 1 and 4).

Methods

Participants

Five subjects participated in the experiment, all of whom participated in Experiment 1. They ranged in age from 24 to 35 years and had normal or corrected-to-normal visual acuity. All participants gave informed consent in writing prior to participation and the protocols for the study were approved by the Université Paris Descartes Review Board, CERES, in accordance with French regulations and the Declaration of Helsinki. Participants were compensated 10€ per hour for their time.

Apparatus

The apparatus was unchanged from Experiment 1.

Stimuli

The same stimuli were presented as in Experiment 1 except that the Gabor patch was always tilted 45° to the left. Also, the frame alternation rate was varied from 1 to 5 Hz. In the dynamic condition, the target moved by one step each frame thus completing one revolution of the circle in 1.5 s (5 Hz) to 7.55 s (1 Hz). In the static case, because the target only changed location on every third frame, the revolution rate was one third that of the dynamic condition, so 4.5 s (5 Hz) to 22.5 s (1 Hz).

Procedure

The procedure was unchanged from Experiment 1 except that each participant performed 510 trials per experiment and they were only instructed to detect the target and respond as soon as they saw something in the blank space between the masks in the dynamic condition or something in addition to the masks in the static condition.

Results

For this experiment 2.34% of trials were discarded due to premature responses from the observers when
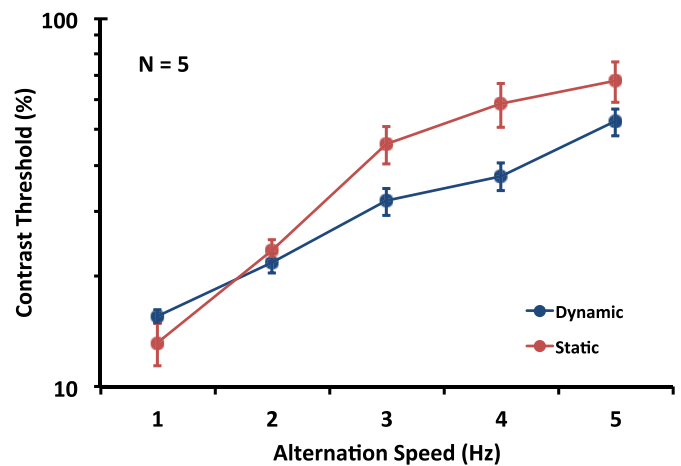

Figure 6. Experiment 2: Effect of alternation speed. Average detection threshold (log scale) at which participants were able to detect the target as a function of speed for both the dynamic and static condition. Error bars represent the standard error of the mean.

the target was still at 0% contrast. Additionally, the percentage of trials that were rejected due to the participant not responding before the trial ended at 100% contrast was 1.69%. The lowest thresholds we measured, at 1 Hz alternation rate, were in the range of 12% to 15% contrast. At this rate, the target is present for 500 ms and it is unlikely that there is much effective masking at this duration. This is similar to or higher than unmasked contrast detection thresholds in other studies although none matched our conditions. For example, Wright and Johnston (1983) reported about 3% threshold for continuous presentation, 3.5°-wide, 2 cpd test at 12° eccentricity; Petrov and McKee (2009) reported up to 15% for brief presentation at 8° eccentricity of a 1.5 cpd Gabor, 1° in diameter (Petrov & McKee, 2009).

Participants' detection thresholds increased as the rate of alternation increased (Figure 6). The effect of alternation speed on the detection threshold was highly significant, ANOVA $F(4, 16) = 70.19$, $p < 0.001$.

These results held true for both the static and the dynamic conditions. The thresholds for the dynamic condition are also lower than those for the static condition (mean difference = 9.83%), however an ANOVA revealed that there was no significant main effect of static versus dynamic condition, $F(1, 16) = 6.52$, $p = 0.06$. However, there was a significant interaction effect, ANOVA $F(4, 16) = 8.64$, $p = 0.001$, that is mostly explained by the difference of the two linear components, the slope for the static case being steeper than for the moving case (paired t test: $t[4] = 2.722$, $p = 0.053$). Further testing at each of the five individual speeds revealed that none of the individual differences between static and dynamic conditions reached the Bonferroni corrected significance level of p

< 0.01 although the difference at 3 Hz was closest. The noncorrected significance levels for 1–5 Hz were $p = 0.096$, $p = 0.154$, $p = 0.026$, $p = 0.040$, and $p = 0.105$, respectively.

The results of Experiment 2 show that the detection threshold increases as the alternation rate increases and the target duration decreases. There is again very strong masking in the dynamic condition where the target location can be attended to in isolation and where attention can avoid focusing on the masks. The strength of this masking did not differ overall from the strength of masking in the static condition where the masks were attended. We ran a pilot experiment to determine the range of speeds over which participants could correctly track the target. They failed at rates of 6 Hz or beyond, a limit similar to that found by Verstraten, Cavanagh, and Labianca (2000). Accurate tracking was possible up to 5 Hz, so the thresholds over the tested range of 1 to 5 Hz appropriately represent detection for the dynamic condition where the target is isolated by attention and the masks are unattended.

Experiment 3: Spacing

The masking found in Experiments 1 and 2 in the dynamic tracking condition may have been caused by the temporally preceding and following masks that are spatially superimposed on the target and thus occupy the same position as the target at different times, Alternatively, it may be the result of lateral masking (Werner, 1935; Alpern, 1953; Growney et al., 1977) or crowding (Bouma, 1970) from the spatially adjacent masks that appear at the same time as the target in this condition. Since the lateral flankers are masks in our stimuli rather than other Gabor patterns, we do not expect any collinear facilitation (e.g., Lev & Polat, 2011). It is generally thought that crowding degrades target identification but not detection (Pelli, Palomares, & Majaj, 2004); nevertheless, recent studies have shown that in some conditions, detection can indeed be affected by crowding (Põder, 2008; Allard & Cavanagh, 2011). In Põder's (2008) study, for example, multiple flankers yielded crowding on a detection task, whereas two flankers alone did not. Crowding can also be sensitive to presentation duration where a decrease in presentation time results in an increase in the size of the crowding zone (Tripathy & Cavanagh, 2002; Chung & Mansfield, 2009; Tripathy, Cavanagh, & Bedell, 2014). In our tests here, presentation duration was fixed at 167 ms where the crowding zones should be close to about 0.2 times the eccentricity for tangentially arranged flankers (Toet & Levi, 1992).

We addressed this possibility of lateral interactions by varying the spacing of the targets and masks in the

moving condition alone since crowding and lateral masking would not be a factor in the static condition where the target appears by itself.

Methods

Participants

Six subjects participated in this experiment, five of whom participated in Experiments 1 and 2 and including one of the authors (AV). They ranged in age from 24 to 35 years and had normal or corrected-to-normal visual. All participants gave informed consent in writing prior to participation and the protocols for the study were approved by the Université Paris Descartes Review Board, CERES, in accordance with French regulations and the Declaration of Helsinki. Participants were compensated 10€ per hour for their time.

Apparatus

The apparatus was unchanged from the previous experiments.

Stimuli

The same stimuli were presented as in Experiment 2's moving condition except that the frame alternation rate was fixed at 3 Hz but the spacing of the targets was varied. There were 5, 10, 15, 20, or 25 possible positions around fixation where masks and targets appeared. This translates to target-to-mask distances of 72°, 36°, 24°, 18°, and 14.4° of rotation around the circle.

Procedure

The procedure was unchanged from Experiment 2 except that each session consisted of 255 trials and only moving attention trials were included.

Results

The percentage of trials that were discarded because the subjects responded when the target contrast was still 0% was 0.07%, while the amount of trials that were rejected because the participant did not reply before the end of the trial, while the target had already reached 100% contrast, was 4.78%.

Spacing had little effect on the detection threshold (Figure 7) except at the widest spacing between successive locations (72° of rotation) where thresholds increased compared to the other spacings, ANOVA $F(4, 20) = 16.08$, $p < 0.001$.

This increase in threshold at the largest spacing most likely corresponds to the loss of clear apparent motion
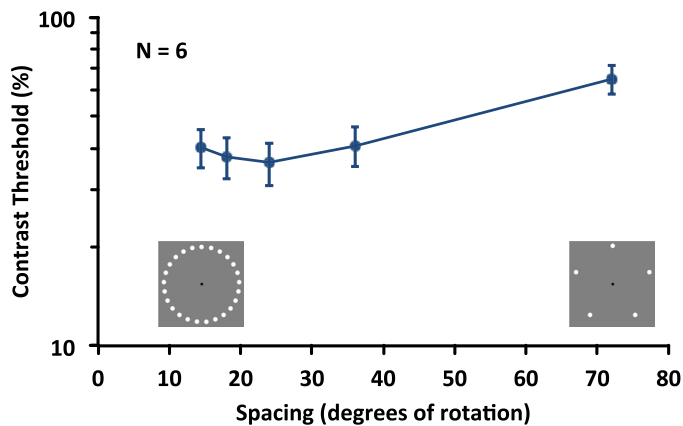
Figure 7. Experiment 3: Effect of spacing. Average contrast threshold (log scale) at which participants were able to detect the target for each spacing. Error bars indicate the standard error of the mean. The spacing of the target and masks for the closest and farthest condition is represented here by white dots in the diagrams below the corresponding data points (where each frame has only one target and two masks covering an adjacent set of three of these locations).

for these large steps. Without the strong apparent motion, attentive tracking would become harder to maintain. A loss of performance for large spacings was reported in an earlier study on attentive tracking already when step size reached 45° of rotation, the largest step tested (Verstraten et al., 2000).

More importantly, we see no increase in threshold at closer spacings that could indicate any effect of crowding. This result shows that the masking effect seen at the spacings used in Experiment 1 must be due to the backward and/or forward masking at the target location, and not from lateral masking or crowding.

Experiment 4: Pre- versus postmasks

In this experiment we compared the contribution of pre- and postmasks to the effectiveness of the masking on detection thresholds to determine whether onset and offset transients of the target play different roles. To do so, we removed either the leading mask or the trailing mask. In order to compare these two conditions, we also tested with both masks present and without any masks present. We included both dynamic and static conditions.

Methods

Participants

Six subjects participated in the experiment, five of whom participated in the previous experiments. They

ranged in age from 24 to 35 years and had normal or corrected-to-normal visual acuity. All participants gave informed consent in writing prior to participation and the protocols for the study were approved by the Université Paris Descartes Review Board, CERES, in accordance with French regulations and the Declaration of Helsinki. Participants were compensated 10€ per hour for their time.

Apparatus

Apparatus was unchanged from previous experiments.

Stimuli

The same stimuli were presented as in previous experiments with the following exceptions. The alternation rate was fixed at 3 Hz and the number of positions at 15. Depending on the condition either both masks were present, only the premask was present, only the postmask was present or no masks were present. The masks that were dropped were replaced by a blank of equal duration to keep the timing across conditions consistent.

Procedure

Unchanged from Experiments 2 and 3.

Results

The detection contrast thresholds are shown in Figure 8. There was a highly significant effect of masking condition on the detection threshold, ANOVA $F(3, 15) = 83.37$, $p < 0.0001$. Post hoc paired t tests
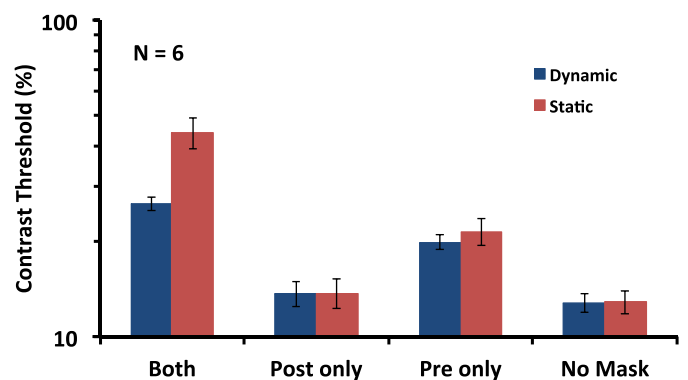

Figure 8. Experiment 4: Effect of pre- versus postmask. Detection thresholds (log scale) for both masks present, postmask only, premask only, and no-mask present in both the dynamic and the static condition. Error bars denote the standard error of the mean.

showed that there was no significant difference between the postmask-only and no-mask conditions (dynamic condition: $t[5] = 1.62$, $p = 0.166$; static: $t[5] = 1.07$, $p = 0.336$). However, thresholds for premask-only were significantly higher than for no-mask, $t(11) = 9.240$, $p < 0.001$, and postmask-only, $t(11) = 13.03$, $p < 0.001$. The threshold when both masks were present is again significantly higher than all the other conditions ($p < 0.001$).

There was also a main effect of dynamic versus static condition on the detection threshold across all conditions, ANOVA $F(1, 15) = 11.40$, $p = 0.020$, which was driven by the dynamic versus static difference in the both-masks-present condition, paired t test: $t(5) = 4.27$, $p = 0.008$. This difference with both masks present is comparable to the difference seen in Experiment 2 at this alternation rate (3 Hz).

Although the pattern of results was similar for the dynamic and static conditions, there was a significant interaction effect between the masking and dynamic-versus-static conditions, ANOVA $F(3, 15) = 20.41$, $p < 0.001$. This was driven by the significant difference between static and dynamic found only in the both condition.

These results indicate that the failure to see the attended target is primarily caused by the premask, whereas the postmask alone has little effect. However, the postmask is not completely ineffective; it does seem to have an interactive effect when the premask is present.

General discussion

We evaluated the effectiveness of masks when they were attended (our static conditions) versus unattended (dynamic conditions) and the nature of the target representation when masked. We found that the strength of masking was on average unaffected by whether or not the masks were attended, although there was a trend for dynamic thresholds to be lower at higher speeds. This result suggests that the effect of pattern masking is not the integration of mask and target (Breitmeyer, 1984), but instead the complete suppression of the target at threshold. Indeed, the detection thresholds were not different from identification thresholds, a result that indicates that the masking did not leave some unrecognizable mixture of target and mask at identification threshold but instead rendered the target invisible, causing the detection threshold to match the identification threshold. This suggests that low-level masking that is often referred to as integration masking (e.g., Breitmeyer, 1984) may not obscure the target by integrating its features with those of the mask but simply suppress the target from awareness.

Here we used a moving attention technique that allowed participants to see the target location without seeing or attending to the mask. This was the case even at rates where the target and mask would normally be selected and perceived together in ordinary forward and backward masking. Note that in the moving attention condition, the mask–target–mask sequence was always presented in order at each location, just as it was in the static condition. However, the stepping motion, when followed, organized the stimuli laterally in space with a mask on the left, target in the center, and a mask on the right. This set of three stimuli then shifted one location with each step. This organization allowed participants to focus on the target location and detect whatever was there without also picking up the mask.

Of course, selecting only the target location does not guarantee that masking itself is avoided as the mask could affect the target before the selection stage. In that case, the pattern seen at the target location should reveal the outcome of that preattentive process. The effect of the masking may be to generate some degraded target pattern or mixture of target and mask (integration masking; Breitmeyer, 1984). Given that this degraded target is to be detected against a blank background, its detection threshold in the dynamic condition should be lower than its identification threshold, and lower as well than the detection threshold in the static condition when this degraded pattern would be detected against the context of the mask itself.

This did not happen. The detection threshold was the same as the identification threshold and, on average, the same for the dynamic and static conditions. This result is a clear departure from many other masking studies that showed a positive effect of attention on detection (i.e., in metacontrast masking; Werner, 1935; Johnston & Dark, 1986; Ramachandran & Cobb, 1995). These studies showed that performance is better with more rather than with less attention. What we have done in our study is keep the amount of attention constant but vary whether it can be distributed to targets alone (dynamic condition) or, unavoidably, to the targets and masks together (static condition). Here we find no significant advantage in the case of forward and backward masking when attention can avoid the masks, replicating the earlier study of Cavanagh et al. (2008). Additionally, we find no evidence that the masking acts by integrating mask and target elements but rather that it acts by suppressing the target altogether.

Previous experiments have demonstrated the integration of information along a motion trajectory for properties such as such as luminance (Burr, 1981; Burr & Ross, 1986), color and motion (Cavanagh et al., 2008), and color alone (Nishida, Watanabe, Kuriki, & Tokimoto, 2007; Watanabe & Nishida, 2007), which could enhance or mask detection (Hidaka, Nagai, Sekuler, Bennett, & Gyoba, 2011) depending on

whether the successive instances were similar or different, respectively. In contrast, the results in our dynamic conditions here do not support any motion-based integration across the sequential locations of the target. Specifically, the thresholds in the dynamic case were not significantly lower from those in the static case where there is no motion impression and integration would be unlikely. Cavanagh et al. (2008) also found no evidence for integration across locations in the moving target conditions for identification thresholds.

By using the sequential presentation, we change the perceived organization of the target–mask–target sequence that repeats at each location. Instead of seeing the sequence in its actual temporal order, the synchronized stepping generates a spatial array of the target flanked in space by the last mask at the previous location and the first mask at the next. This spatial mask–target–mask array then steps one position ahead on each new frame. This spatial adjacency raises the question of whether there are any lateral interactions contributing to the masking in the dynamic case. Typically, in lateral masking, the further the mask is from the target, the weaker the masking becomes (Werner, 1935; Alpern, 1953; Growney et al., 1977). In Experiment 2, however, the strength of masking did not change consistently with spacing, showing that the effect of the masks was not due to the lateral spacing between the adjacent sequences of mask–target–mask. The effect of spacing only became important when the steps between sequences were large, perhaps too large to support the apparent motion of the target from location to location, as was reported by most subjects.

We were also able to rule out crowding as a factor. Although crowding is frequently described not to affect detection (Pelli et al., 2004), recent studies do show crowding effects for detection in some situations (Allard & Cavanagh, 2011). In either case, there was no evidence of crowding over the range of spacings that we tested and no significant increase in threshold as the spacings got closer. These results suggest that the masking that we observed was due to the temporal sequence at each location, the forward and backward pattern masking. This is the case even though, in the moving attention condition, the pre- and postmasks were not perceived as preceding and following the target. Instead they were perceived as spatially adjacent to the previous and the following targets. This perceptual effect, grouping the masks with the targets over space instead of time, had little effect on the measured thresholds.

In the last experiment we tested the effectiveness of the premask and postmask individually to see whether the onset and offset transients of the target might contribute differently to the masking. The data from the last experiment show that the premask in each mask–target–mask sequence is much more important

than the postmask, indicating that forward masking is more effective here. This would reiterate the importance of onset transients of the target in order for it to enter awareness (Yantis & Jonides, 1984; Öğmen, Breitmeyer, & Melvin, 2003; Cavanagh et al., 2008; Motoyoshi & Hayakawa, 2010). The masking in our conditions appears to operate prior to attentional selection but its effect is to suppress the target by masking its onset transient, rendering it invisible.

Although there was no significant main effect of dynamic versus static thresholds in Experiment 1, there was an interaction showing that the dynamic thresholds did get increasingly lower than the static thresholds at higher rates, although the difference was not significant at any individual rate. This difference was also seen in the fourth experiment for the condition with both pre- and postmasks, which is essentially a replication of the 3 Hz condition from Experiment 2. There are (at least) two possible explanations for these differences. First, the absence of attention to the masks may in fact decrease their effectiveness, lowering detection thresholds. Second, the context of the detection is quite different in the two conditions: detecting the target against the blank field in the dynamic case versus detecting the target against the mask. These factors as well as other differences between the task demands in the dynamic and static conditions could produce the threshold differences that were seen in some conditions.

Nevertheless, the lack of difference between identification and detection thresholds (in Experiment 1) and the absence of a main effect of dynamic versus static conditions for detection thresholds in Experiment 2 suggests that the effect of the mask in backward and forward masking is not to integrate the mask and target patterns but to block the target from awareness. In particular, if the target became, through integration, a combination of target and mask patterns, it would be easily lost within the mask and target stream in the static condition. In the dynamic condition, however, the target location is visible in isolation. If the effect of the masking were to generate a combined mask and target mix, it would still be visible as such, lowering the threshold until nothing was visible. The finding that the two thresholds are so similar supports the hypothesis that the target representation at threshold has vanished in both cases. The mask suppresses the visibility of the target rather than mixing with it.

Keywords: visual attention, masking, onset transient

Commercial relationships: none.
Corresponding author: Arielle A. Veenemans.
Email: aveenemans@gmail.com.
Address: Université Paris Descartes-CNRS, Laboratoire Psychologie de la Perception, Paris, France.

References

Allard, R., & Cavanagh, P. (2011). Crowding in a detection task: External noise triggers change in processing strategy. *Vision Research*, *51*(4), 408–416.

Alpern, M. (1953). Metacontrast. *Journal of the Optical Society of America*, *43*(8), 648–657.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177–178.

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

Breitmeyer, B. G. (1984). *Visual masking: An integrative approach.* New York: Oxford University Press.

Breitmeyer, B. G., & Öğmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics*, *62*(8), 1572–1595.

Brussell, E. M., & Cavanagh, P. (1984). An anticipated threshold technique for measuring contrast sensitivity. *American Journal of Optometry & Physiological Optics*, *61*(2), 125–126.

Burr, D. C. (1981). Temporal summation of moving images by the human visual system. *Proceedings of the Royal Society of London Series B: Biological Sciences*, *211*, 321–339.

Burr, D. C., & Ross, J. (1986). Visual processing of motion. *Trends in Neuroscience*, *9*, 304–307.

Cavanagh, P., Holcombe, A. O., & Chou, W. (2008). Mobile computation: Spatiotemporal integration of the properties of objects in motion. *Journal of Vision*, 8(12):1, 1–23, doi:10.1167/8.12.1. [PubMed] [Article]

Chung, S. T. L., & Mansfield, J. S. (2009). Contrast polarity differences reduce crowding but do not benefit reading performance in peripheral vision. *Vision Research*, *49*, 2782–2789.

Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science*, *8*, 135–139.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*(9), 345–352.

Foley, J. M., & Boynton, G. M. (1993). Forward pattern masking and adaptation: Effects of duration, interstimulus interval, contrast, and spatial and temporal frequency. *Vision Research*, *33*, 959–980.

Franconeri, S. L., Hollingworth, A., & Simons, D. J. (2005). Do new objects capture attention? *Psychological Science*, *16*(4), 275–281.

Growney, R., Weisstein, N., & Cox, S. I. (1977). Metacontrast as a function of spatial seperation with narrow line targets and masks. *Vision Research*, *17*, 1205–1210.

Herzog, M. H. (2006). Spatial processing and visual backward masking. *Advances in Cognitive Psychology*, *3*(1–2), 85–92.

Hidaka, S., Nagai, M., Sekuler, A. B., Bennett, P. J., & Gyoba, J. (2011). Inhibition of target detection in apparent motion trajectory. *Journal of Vision*, 11(10):2, 1–12, doi:10.1167/11.10.2. [PubMed] [Article]

Johnston, W. A., & Dark, V. J. (1986). Selective attention. *Annual Review of Psychology*, *37*, 43–75.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, *43*, 346–354.

Kulikowski, J. J., & Tolhurst, D. J. (1973). Psychophysical evidence for sustained and transient detectors in human vision. *Journal of Physiology*, *232*(1), 149–162.

Lev, M., & Polat, U. (2011). Collinear facilitation and suppression at the periphery. *Vision Research, 51,* 2488–2498.

Lleras, A., & Moore, C. M. (2003). When the target becomes the mask: Using apparent motion to isolate the object-level component of object substitution masking. *Journal of Experimental Psychology: Human Perception & Performance*, *29*(1), 106–120.

Miller, J. (1989). The control of attention by abrupt visual onsets and offsets. *Perception & Psychophysics*, *45*(6), 567–571.

Moore, C. M., Egeth, H., Berglan, L. R., & Luck, S. J. (1996). Are attentional dwell times inconsistent with serial visual search? *Psychonomic Bulletin & Review*, *3*, 360–365.

Motoyoshi, I., & Hayakawa, S. (2010). Adaptation-induced blindness to sluggish stimuli. *Journal of Vision*, 10(2):16, 1–8, doi:10.1167/10.2.16. [PubMed] [Article]

Nachmias, J., & Rogowitz, B. E. (1983). Masking by spatially-modulated gratings. *Vision Research, 23*(12), 1621–1629.

Nishida, S., Watanabe, J., Kuriki, I., & Tokimoto, T. (2007). Human visual system integrates color signals along a motion trajectory. *Current Biology, 17*, 366–372.

Öğmen, H., Breitmeyer, B. G., & Melvin, R. (2003). The what and where in visual masking. *Vision Research, 43*(12), 1337–1350.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision, 4*(12):12, 1136–1169, doi:10.1167/4.12.12. [PubMed] [Article]

Petrov, Y., & McKee, S. P. (2009). The time course of contrast masking reveals two distinct mechanisms of human surround suppression. *Journal of Vision, 9*(1):21, 1–11, doi:10.1167/9.1.21. [PubMed] [Article]

Põder, E. (2008). Crowding with detection and coarse discrimination of simple visual features. *Journal of Vision, 8*(4):24, 1–6, doi:10.1167/8.4.24. [PubMed] [Article]

Polat, U., & Sagi, D. (2006). Temporal asymmetry of collinear lateral interactions. *Vision Research, 46*(6–7), 953–960.

Polat, U., Sterkin, A., & Yehezkel, O. (2010). Spatio-temporal low-level neural networks account for visual masking. *Advances in Cognitive Psychology, 3*(1–2), 153–165.

Ramachandran, V. S., & Cobb, S. (1995). Visual attention modulates metacontrast masking. *Nature, 373*, 66–68.

Schiller, P. H. (1966). Forward and backward masking as a function of relative overlap and intensity of test and masking stimuli. *Perception & Psychophysics, 1*(3), 161–164.

Smith, M. C., & Schiller, P. H. (1966). Forward and backward masking: A comparison. *Canadian Journal of Psycholology, 20*(2), 191–197.

Spencer, T. J., & Shuntich, R. (1970). Evidence for an interruption theory of backward masking. *Journal of Experimental Psychology, 85*, 198–203.

Sterkin, A., Yehezkel, O., Bonneh, Y. S., Norcia, A., &

Polat, U. (2009). Backward masking suppresses collinear facilitation in the visual cortex. *Vision Research, 49*, 1784–1794.

Theeuwes, J. (1995). Abrupt luminance change pops out; abrupt color change does not. *Perception & Psychophysics, 57*(5), 637–644.

Theeuwes, J., Godijn, R., & Pratt, J. (2004). A new estimation of the duration of attentional dwell time. *Psychonomic Bulletin & Review, 11*, 60–64.

Thompson, P. (1983). Discrimination of moving gratings at and above detection threshold. *Vision Research, 23*(12), 1533–1538.

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

Tripathy, S., & Cavanagh, P. (2002). The extent of crowding in peripheral vision does not scale with target size. *Vision Research, 42*, 2357–2369.

Tripathy, S. P., Cavanagh, P., & Bedell, H. E. (2014). Large crowding zones on peripheral vision for briefly presented stimuli. *Journal of Vision, 14*(6): 11, 1–11, doi:10.1167/14.6.11. [PubMed] [Article]

Verstraten, F. A. J., Cavanagh, P., & Labianca, A. T. (2000). Limits of attentive tracking reveal temporal properties of attention. *Vision Research, 40*, 3651–3664.

Ward, R., Duncan, J., & Shapiro, K. (1997). Effects of similarity, difficulty, and nontarget presentation on the time course of visual attention. *Perception & Psychophysics, 59*, 593–600.

Watanabe, J., & Nishida, S. (2007). Veridical perception of moving colors by trajectory integration of input signals. *Journal of Vision, 7*(11):3, 1–16, doi:10.1167/7.11.3. [PubMed] [Article]

Werner, H. (1935). Studies on contour: I. Qualitative analyses. *American Journal of Psychology, 47*(1), 40–64.

Wright, M. J., & Johnston, A. (1983). Spatiotemporal contrast sensitivity and visual field locus. *Vision Research, 23*(10), 983–989.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective visual attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception & Performance, 10*(5), 601–621.