

NIH Public Access

Author Manuscript

Vision Res. Author manuscript; available in PMC 2008 October 1.

Published in final edited form as: *Vision Res*. 2007 October ; 47(23): 2973–2981.

Attentional blinks as errors in temporal binding

Ariella V. Popple* and **Dennis M. Levi**

Abstract

In the attentional blink, the second of two targets in a Rapid Serial Visual Presentation (RSVP) stream is difficult to detect and identify when it is presented soon but not immediately after the first target. We varied the Stimulus Onset Asynchrony (SOA) of the items in the stream and the color of the targets (red from gray or vice versa), and looked at the responses to the second target. Exact responses to the second target (zero positional error) showed a typical attentional blink profile, with a drop in performance for an interval of 200–500 ms after the first target. Approximate responses (positional error no greater than 3 frames) showed no such drop in performance, although results were still dependent on color (better for red) and increased with increasing SOA. These findings are consistent with a 2-stage model of visual working memory, where encoding of the first target disrupts attention to (and temporal binding of) the second target. We suggest that this disruption occurs within a certain time (≈ 0.5 s) after the first target, during which period salient distractors are as likely as the second target to enter working memory.

Keywords

temporal attention; attentional blink; RSVP; stimulus onset-asynchrony; temporal binding

Introduction

When we look at the world, we make saccadic eye-movements to inspect our surroundings. These eye-movements are largely unconscious, and typically we make one every 150–300 ms. This is because the visual system selectively processes only the central 1–2 deg of visual angle in fine detail and with full color vision, using the dense cone receptors in the retinal fovea whose input dominates the higher visual areas of the brain responsible for abstract representations. We need to make eye-movements to salient and behaviorally relevant objects such as another person's eyes or mouth, or the words on a page, in order to see and read them accurately. Our image of the environment is patched together from a series of such fixations, elaborating on the more vague information that we get from vision outside the fovea which tells us that we are looking at a crowd of people, or a page in a journal. The processes that lead to the binding of information from multiple fixations over time and space can be partially elucidated by examining the limits of human performance when items are presented rapidly one after the other in the central foveal region, using Rapid Serial Visual Presentation (RSVP).

Bowman and Wyble (2007) recently proposed a new and comprehensive theory of encoding RSVP targets in visual working memory. Their theory brings together several previous models and ideas, particularly those relating to the attentional blink. The theory has two stages: 1) A perceptual stage, in which many 'types' or object representations become activated in a parallel

^{*}Corresponding author: ariellap@berkeley.edu; School of Optometry & Helen Wills Neuroscience Institute, University of California Berkeley, Berkeley CA 94720-2020, Phone: (510) 643-8685, Fax: (510) 643-8733

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

fashion, depending on their presence in the environment, and the task-dependent operations of a 'salience filter'; 2) An encoding stage, when a series of time-labeled 'tokens' are bound to the activation of the salience-filtered 'types'. This 'tokenization' requires the action of a 'temporary attentional enhancement' (TAE) mechanism that selectively enhances the activation of salient 'types', so that they can be bound to 'tokens' in the 'binding pool'. According to this theory, in the attentional blink the TAE is still suppressed by the processing of the first target when the second one arrives.

We predicted that under these circumstances the binding of 'types' and 'tokens' will become momentarily more random, while still depending on the available salience-filtered types. This is consistent with the finding that attentional blink errors typically involve the misreporting of neighboring non-targets (Chun, 1997). Specifically, we predicted that these positional errors will depend on the Stimulus Onset Asynchrony (SOA) of the items in the RSVP stream and on the salience of the distractors, as well as the number of frames after the first target.

The Discussion section considers the implications of other theories of the attentional blink, including the Interference Theory (Shapiro, Arnell & Raymond, 1997), Chun & Potter's (1995) Two Stage Model, Temporary Loss of Control (DiLollo et al., 2005), and the Locus Coeruleus Model (Niewenhuis et al., 2005).

Methods

Subjects

Twelve observers with normal or corrected to normal visual acuity participated in this study. Their ages ranged from 18 to 36, and five of the participants were male. Six observers were tested at all the SOA conditions, and six further observers were only tested at the 100 ms SOA typical of attentional blink experiments.

Apparatus and Stimuli

The stimuli were 40-point Times font letters displayed using Presentation on a Dell Inspiron 9100 laptop with LCD screen set at 1024 x 768 pixel (32 bit color) resolution, using a frame rate of 60 Hz. The screen was viewed from about 50 cm, such that letters subtended about 1.5° of visual angle, and were highly visible. The recorded stimulus duration on the computer was 33–34 ms, although LCD persistence may have in fact lengthened the stimuli somewhat.

Stimuli were upper-case letters in Times New Roman font presented in 'Red' (CIE1931 $x =$ 0.60, y = 0.33, lum = 16.75 cd m⁻²) or 'Gray' (CIE1931 x = 0.27, y = 0.32, lum = 16.15 cd m⁻²) on a black (near zero luminance) background. Note that both colors have almost identical luminance, and vary only in their chromaticity.

Design

We used a 3-way $5 \times 2 \times 7$ factorial design, and varied: 1) The Stimulus Onset Asynchrony (SOA) of letters presented in an RSVP stream (67, 83, 100, 117, or 133 ms); 2) The color of the target and non-target letters ('Red' from 'Gray' or vice versa); 3) The number of frames between the two targets $(1, 2, 3, 4, 5, 6)$ or 7 frames).

We measured performance using 4 different measures: 1) Positional error < 1 (i.e. correct responses); 2) Positional error < 2; 3) Positional error < 3; 4) Positional error < 4 (i.e. approximately correct responses).

Procedure

Letters were selected randomly from the full 26-letter English alphabet, without replacement, so that we could look at the positional errors within each trial. Each trial was preceded by an 'X' fixation marker, and was initiated by pressing the space bar. The task was to identify and type in the two target letters at the end of the trial. Participants had an option to repeat a trial at will if they made a typing error, in which case that trial was not recorded and a replacement trial was generated. Figure 1 shows a schematic illustration of a trial.

Each trial consisted of 14 frames. The first target could appear on frame 2–5 at random, and the second target arbitrarily 1–7 frames later. Trials were blocked according to target color and SOA, in blocks of 70 trials in which each of the 7 possible frame delays between the two targets was presented 10 times. Subjects completed 3 blocks (i.e. 30 trials) for each datum point in the results.

Results

The results (Figure 2) clearly indicate that most errors caused by the first target are positional errors in time. The figure shows all the responses made by the 12 observers (100 ms SOA) or 6 observers (all other SOA's) as follows: blue bars indicate first responses (R1), and red bars indicate second responses (R2). The number of responses to each target or distractor frame is plotted against distance in frames from T1. Thus, at position 0 (on the x-axis) there are two bars, a blue bar indicating the number of R1 responses, and a red bar indicating the number of R2 responses. Each histogram, going along the x-axis of the table, shows a different T1–T2 lag. Figure 2a shows the results for Red and Gray targets on two separate rows. Figure 2b shows the results (summed across color) for the different SOA's, on five successive rows. In Figure 2c, these data have been conditioned by the correct report of T1.

Two different kinds of errors are apparent. The first kind of error results in the report of immediate T2 neighbors in place of T2, particularly the neighbor immediately after T2 but also sometimes the one before and up to two or three frames away from T2. The second kind of error results in the report of T1 or its neighbors. When the interval between T1 and T2 is only two or three frames, the distractors between them (neighbors of both) are frequently reported in place of one target or the other. This is particularly true for gray targets (shown separately in Figure 2a), where the second response distribution blurs T1 and T2 positions together. Responses to more distant neighbors happen more frequently at the briefer SOA's (Figure 2b). It is notable that the errors involving the incorrect report of T1 neighbors in place of T2 occur even on those trials conditioned by correct report of T1 (Figure 2c).

The raw data that went into compiling Figure 2 are available as Supplementary Materials on the Vision Research website, for use in future models and theories.

In order to make these observations more concrete, we entered the data into an ANOVA analysis, comparing correct report of T2 with responses to T2 $+/-$ 3 frames, to determine whether such responses accounted for the majority of errors under the attentional blink rubric. This analysis is somewhat problematic as there were also positional errors resulting from the report of T1 and its neighbors further than 3 positions prior to T2, but such responses were relatively infrequent where compared with responses to closer neighbors of T2. When considering the different SOA's shown in Figure 3, it becomes clear that the effect of the first target lasted 7 frames (about 470 ms) at the briefest SOA (67 ms), but only about 400 ms (3 frames) at the longest SOA (133 ms). A comparison between Correct Responses (Positional Error = 0) and Approximate Responses (Positional Error \leq 3) shows that the effect of the first target goes away when considering the Approximate Responses, but the effects of SOA and

color remain, suggesting that these effects may be due at least in part to the masking effects of the distractors directly on the second target itself (Figure 4).

These data were entered into a 4-way Repeated Measures Analysis of Variance (ANOVA) design, using the General Linear Model tool in SPSS. We considered 4 measures of Positional Error (PosErr), 5 levels of SOA, 2 Colors, and 7 levels of Frame interval (as outlined in the Design part of the Methods section). F-values and degrees of freedom were corrected in order to obtain p-values that indicate the 'significance' or chance probability of the results obtained. We used the Huynh-Feldt correction for non-sphericity, because of the insufficient residual degrees of freedom of the chosen design (too many different conditions for the number of subjects tested). According to the SPSS website, this correction provides appropriate estimates given the available data.

The results (summarized in Table 1) highlight the effects of Positional Error on performance, and its interactions with the effects of SOA, Color and Frame. The effect of the first target can be considered as an interaction between SOA and Frame, assuming that it depends on the time after T1. This measure interacted significantly with Positional Error (PosErr*SOA*Frame interaction, $p < .001$, details in Table 1), as further confirmed by the contrast of the level 4 vs. level 1 PosErr effect on the interaction between the linear component of SOA and the level 3 vs. level 1 Frame contrast ($p < .001$; F(1,5) = 142.90).

Further analysis of the subset of Correct (PosErr $= 0$) responses using a 3-way ANOVA confirmed the significance of time after T1, as indicated by the interaction between SOA and Frame $(p < .001$, details in Table 2), and specifically the interaction between the linear component of SOA and the quadratic component (peak/trough) of the Frame effect ($p = 0.003$; $F(1,5) = 27.24$. Additionally, there was a significant 3-way interaction between Color, SOA and Frame ($p = .001$, details in Table 2).

Analysis of the subset of Approximately Correct (PosErr \leq 3) responses using 3-way ANOVA revealed that there was no longer a significant effect of Frame ($p = .34$, details in Table 3), suggesting that there was little or no effect of T1 on these results. There was a marginally significant interaction between SOA and Frame ($p = .034$, details in Table 3). Analysis of the results for all 12 subjects at 100 ms SOA confirmed that there was still a small effect of Frame at PosErr ≤ 3 (p = 0.033, F(2.8, 30.9) = 3.37). However, this was due to a contrast between level 1 and level 6 of Frame ($p = 0.036$, $F(1,11) = 5.69$). There was no contrast between level 1 and level 2 of Frame ($p = 0.30$, $F(1,11) = 1.19$) or between level 1 and level 3 of Frame ($p = 1.19$) 1.000, $F(1,11) = .000$. The absence of any significant drop in performance at 2 or 3 frames after T1 strongly suggests that there was no attentional blink when considering responses to T2 and its 3 nearest neighbors as approximately correct (see Figure 4).

Discussion

In summary, we found that the attentional blink, as measured by the effect of first-target timing on second-target performance, was absent when performance was considered in terms of responses to the second-target position +/− 3 frames. This main finding is consistent with the theory that during the attentional blink, the temporal binding of perceptual-stage 'types' with memory-stage 'tokens' is disrupted by the lack of available resources at the TAE ('temporary attentional enhancement'), which is still occupied with processing the first target. As a result, other salient 'types' still available in the binding pool because they have not yet become extinguished by the passage of time erroneously bind to the second-target 'token'. This explanation, in terms of Bowman and Wyble's (2007) theory, adequately accounts for our results. The results can also be explained, however, within the framework of earlier theories

of the attentional blink. In the following sections, we consider each theory in turn, concluding with a section containing reservations and caveats.

Interference Theory

Interference Theory (IT; Shapiro, Arnell & Raymond, 1997) considers four relevant items, the first and second targets (T1 and T2), and the items immediately succeeding them $(T1+1)$ and T2+1). Because of its simplicity, this model makes no predictions concerning the effects of SOA, or items before and after these four.

Two kinds of interference have been considered in the past, low level masking (Seiffert & Di Lollo, 1997) and category-level interference (Isaak, Shapiro & Martin, 1999). In the present experiment, we did not vary the categories of the targets and non-targets; however the differences between 'Red' and 'Gray' targets address low-level masking. According to lowlevel masking, differences in T2 performance might arise as a result of differences in the effectiveness of the T2+1 non-target or post-stimulus mask. When targets were 'Red' the poststimulus mask was 'Gray', and vice versa for 'Gray' targets. Low-level masking provides a possible explanation for the difference in performance between 'Gray' and 'Red' targets, by assuming that 'Red' non-targets are more effective masks, either because of their cultural signification (red for danger) or because of their longer persistence in the visual system (Burr & Morrone, 1996). The interference model would predict frequent report of $T1+1$ and $T2+1$ non-targets, consistent with what we found (Figure 2).

Two Stage Model

According to the Two Stage Model (TSM; Chun & Potter, 1995), items in Stage 1 are rapidly identified and available in a post-categorical short-term buffer, but need to be consolidated by Stage 2 for the purpose of conscious perception and report. Stage 2 is limited in capacity, and cannot process other targets while it is suppressed by processing T1. As a result, and also because representations in Stage 1 are ephemeral, by the time Stage 2 is available, the codestrength of post-T2 items in stage 1 may be stronger than the code-strength of T2 itself. This can result in the preponderance of post-target positional errors reported by Chun (1997), and consistent with our own findings in the present study. Although this model does not explicitly address the effects of SOA, by adding a temporal component to the Stage 2 bottleneck it is easily translated into the more elaborate model of Bowman and Wyble (2007). The main distinction between the two models, with regard to the present study, is that the latter more easily explains the prevalence of pre-target as well as post-target positional errors through the erroneous association of 'types' and 'tokens' in the binding pool without needing to wait for an unavailable 'TAE' stage. In contrast, according to TSM, the operation of the serial Stage 2 is quintessential for conscious perception and subsequent report.

Temporary Loss of Control

Temporary Loss of Control (TLC; Di Lollo et al., 2005) is an innovative theory of the attentional blink that posits a central processor whose engagement with T1 processing and response planning makes it unavailable for filtering out unwanted non-target items. These unwanted items influence the system so that it is no longer optimally tuned for the detection of T2. Evidence for this theory comes from the elegant experiments of Olivers et al. (2005), who showed that up to 4 targets can be reported more-or-less correctly, providing that no nontargets are interspersed between them. TLC is consistent with many of our present findings. The positional errors that include pre-target as well as post-target items, the effects of color, and the effects of SOA are all compatible with TLC. However, TLC makes no predictions concerning those trials in which T2 appeared in the frame right after T1, where we found successive improvements in performance when considering Positional Errors < 2 (i.e. including swaps with T1), and even Positional Errors < 3, especially at the briefer SOA's. It is unclear

how TLC is compatible with our data, if it is taken as a comprehensive theory of all attentional blink phenomena, but our data do not address its validity with respect to those experiments it was designed to explain. A possible alternative hypothesis is outlined in the 'Speculation' section below.

Locus Coeruleus Model

The Locus Coeruleus (LC) is a small structure in the brain that responds to salient stimuli, and has a refractory period of about 200 ms, consistent with the peak timing of the attentional blink (Niewenhuis et al., 2005). If it is involved in the processing of RSVP targets, this would explain why attention blinks when it does – because the LC is still responding to T1, and is unavailable for T2. This model accounts for the temporal tuning featured in the results of the present study. The refractory period of the LC varies somewhat with the amplitude of its activation, and hence with stimulus saliency. Assuming that the differences we found between Red and Gray targets are due to saliency, one might expect a change in the temporal tuning of T2 performance depending on color. This is not what we found, but it remains a possibility, given the sensitivity limitations imposed by the crudeness of our methods. For the purpose of the present study, the LC may be the implementation of the TAE, however see Bowman & Wyble (2007) for further qualifications on this option. The LC model would seem to predict a preponderance of $T1+1$ intrusion errors, and indeed the large part of T1 errors fall in this category (Figure 2).

Delayed Attention Model

The delayed attention model was proposed by Nieuwenstein (Nieuwenstein et al., 2005). This model proposes that it is not the memory consolidation of T1 which is responsible for the attentional blink, but instead the problem is with the delayed allocation of attention to stimuli following T1. In support of this hypothesis, Nieuwenstein and Potter (2006) found that performance on 'whole report' (where all stimuli after the first target were reported) was superior to 'partial report' (only T1 and T2 reported). This finding suggests that the problem lies with temporal attention to T2. The delayed attention model is entirely consistent with Bowman and Wyble's (2007) theory, with the occupation of the TAE by T1 causing the delay in attention. The results of Nieuwenstein and Potter (2006) are also consistent with our results, as they did not require stimuli to be reported in the correct order. Our results show that stimuli before and after T2 are frequently recalled in place of T2, and Niewenstein and Potter (2006) showed that when required to do so, subjects can recall up to 4 stimuli before, after and including T2. They did not require subjects to report the stimuli in the correct order, and we predict that once order is taken into account, the difference between 'whole report' and 'partial report' will vanish, as the problem lies with the temporal binding of T2 or temporal attention to T2, and not with its encoding in memory. The delayed attention model would predict frequent T2+1 intrusion errors, and in some conditions this was what we found (see Figure 2).

Reservations and Caveats

Our experimental paradigm does not distinguish errors in temporal binding between 'types' and 'tokens' on the one hand, and errors in temporal binding between color and letter-identity on the other. In other words, we don't know if errors were the result of erroneously binding the target color ('Red' or 'Gray') with a letter of the non-target color, or erroneously binding the 'targetness' of the target with a letter that appeared before or after it. Either way, what can be said for certain is that the attentional blink disrupts temporal binding. What is not clear is – the temporal binding of what with what? Subjectively, some participants reported that they always saw a second red flash, when the targets were red. When the targets were white, all the observers occasionally missed seeing the second target altogether. Nevertheless, their guesses reflected a consistent pattern of temporal binding errors.

Of the theories presented, only that of Bowman and Wyble (2007) accounts for all the data, although none of the other theories can be disproved on the basis of these results. The neuralnetwork implementation of their model would have to be changed to make quantitative predictions concerning our experimental paradigm. Until these quantitative predictions have been calculated, these results can only be said to qualitatively support Bowman and Wyble's (2007) theory.

The attentional blink is frequently evoked in paradigms where the targets and distractors are not confusable (e.g. letters from numbers, or vice versa). In such paradigms, positional errors cannot occur. Performance under these conditions is separately accounted for by Bowman and Wyble's (2007) model.

Speculation

Several features of the raw data (as shown in Figure 2) differ from the expectations that might arise out of inspecting the existing literature on the Attentional Blink, including Bowman and Wyble's (2007) theory, and Chun's (1997) paper on intrusion errors, which would seem to be the closest in scope to the current experiment. One such feature is the symmetrical pattern of pre- and post-target intrusion errors observed at Lag 2 in all but the longest SOA conditions. We speculate that this is due to the merging together of two response distributions, one centered in time around T1 and the other centered in time around T2. Figure 2c suggests that these two distributions remain present as a feature of the data, even when conditioned on the correct report of T1, as is customary in the Attentional Blink literature. This observation is consistent with an hypothesis that only the most salient features of a temporal stream are encoded in visual working memory, and that they are assigned temporal order and timing tags by a subsequent process relying on top-down knowledge or expectations regarding the likely sequence of events. Such speculation would need to be grounded in further empirical and computational investigations, which we hope may be motivated by the present report.

Summary and Conclusions

In conclusion, when visual items are presented in a rapid sequence at the fovea, there is a limit to the speed and accuracy with which targets after the first one presented can be correctly stored in visual working memory. Beyond this limit, although the second target is perceived, it is frequently confused with items appearing closely before or after it in time. Perhaps this limit is one of the factors restricting the rate at which saccadic eye-movements can be used to scan the visual environment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Supported by grant RO1EY01728 from the National Eye Institute, NIH, Bethesda, MD, to Dennis Levi. With thanks to Christopher Cantor for suggesting the color reversal conditions, and Kevin Yuen for comments and help with data analysis.

References

- Bowman H, Wyble B. The simultaneous type, serial token model of temporal attention and working memory. Psychological Review 2007;114:38–70. [PubMed: 17227181]
- Burr DC, Morrone MC. Impulse-Response Functions for Chromatic and Achromatic Stimuli. Journal of the Optical Society of America a-Optics Image Science and Vision 1996;10(8):1706–1713.
- Chun MM. Temporal binding errors are redistributed by the attentional blink. Perception & Psychophysics 1997;59(8):1191–1199. [PubMed: 9401454]

- Chun MM, Potter MC. A two-stage model for multiple target detection in rapid serial visual presentation. Journal of Experimental Psychology: Human Perception & Performance 1995;21(1):109–127. [PubMed: 7707027]
- Di Lollo V, Kawahara J, Shahab Ghorashi SM, Enns JT. The attentional blink: resource depletion or temporary loss of control? Psychological Research 2005;69(3):191–200. [PubMed: 15597184]
- Isaak MI, Shapiro KL, Martin J. The attentional blink reflects retrieval competition among multiple rapid serial visual presentation items: tests of an interference model. Journal of Experimental Psychology: Human Perception & Performance 1999;25(6):1774–1792. [PubMed: 10641317]
- Nieuwenhuis S, Gilzenrat MS, Holmes BD, Cohen JD. The role of the locus coeruleus in mediating the attentional blink: a neurocomputational theory. Journal of Experimental Psychology, General 2005;134(3):291–307. [PubMed: 16131265]
- Nieuwenstein MR, Chun MM, van der Lubbe RHJ, Hooge ITC. Delayed attentional engagement in the attentional blink. Journal of Experimental Psychology: Human Perception and Performance 2005;31 (6):1463–1475. [PubMed: 16366802]
- Nieuwenstein MR, Potter MC. Temporal limits of selection and memory encoding. Psychological Science 2006;17(6):471–475. [PubMed: 16771795]
- Olivers CN, van der Stigchel S, Hulleman J. Spreading the sparing: against a limited-capacity account of the attentional blink. Psychological Research 2005;8:1–14.
- Potter MC, Staub A, O'Connor DH. The time course of competition for attention: attention is initially labile. Journal of Experimental Psychology: Human Perception and Performance 2002;28(5):1149– 1162. [PubMed: 12421061]
- Raymond JE, Shapiro KL, Arnell KM. Temporary suppression of visual processing in an RSVP task: an attentional blink? Journal of Experimental Psychology: Human Perception and Performance 1992;18 (3):849–860. [PubMed: 1500880]
- Seiffert AE, Di Lollo V. Low-Level Masking in the Attentional Blink. Journal of Experimental Psychology: Human Perception & Performance 1997;23(4):1061–1073.
- Shapiro KL, Arnell KM, Raymond JE. The Attentional Blink. Trends in Cognitive Sciences 1997;1(8): 291–297.

Figure 1. Schematic illustration of a trial sequence.

Popple and Levi Page 10

Figure 2.

Response and error distributions summed across observers – raw counts out of 360 trials in each condition. The maximum value shown on the y-axis is 100/360 but in most cases there were >100 correct responses, and these bars are cut off on the graphs. (a) for Red targets and Gray targets (12 observers). (b) For the different SOA's (6 observers, summed across Red and

Gray targets). (c) As in (b), but conditioned on correct T1 report. Note that the temporal smear of responses around the target frame varies smoothly with SOA, and extends over more frames at the briefer SOA's suggesting that it is bounded by time as well as the number of frames.

Figure 3.

Results for the different SOA conditions. At SOA = 100 ms these are means across 12 observers. At other SOA's, results were averaged across 6 observers and Red and Gray targets. The figure shows T2 Performance as a function of Frames After T1 for the 4 measures of Positional Error.

Figure 4.

Results for 'Correct' responses (Positional Error < 1) and 'Approximate' responses (Positional Error < 4), for Red Targets and Gray Targets (means across 6 observers). The figure shows T2 Performance as a function of Frames After T1, for the 5 different SOA conditions.

PosErr ≤ 3 (Approx. Correct)

Table 3

