Title
Expectations About the Temporal Structure of the World Result in the Attentional Blink and Repetition Blindnsess

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 34(34)

ISSN
1069-7977

Authors
Rieth, Cory
Vul, Edward

Publication Date
2012

Peer reviewed

Expectations About the Temporal Structure of the World Result in the Attentional Blink and Repetition Blindness

Cory A. Rieth & Edward Vul (crieth, evul@ucsd.edu)
Department of Psychology, 9500 Gilman Dr. # 109
La Jolla, CA 92093-109 USA

Abstract

We consider how repetition blindness and the attentional blink might arise from prior assumptions about the occurrence of task-relevant states of the world. Repetition blindness and the attentional blink are behavioral deficits in the identification of items during rapid serial visual presentation at varying delays after identifying the first "target." Here we propose that both of these effects are explained by rational inference given prior expectations about the timing of task-relevant world transitions. While such expectations would be helpful in the natural world, they may result in unanticipated biases in laboratory settings. We show that a rational model using prior expectations of the timing of task-relevant information captures the basic repetition blindness and the attentional blink effects, and also the specific distributions of errors made during the attentional blink.

Keywords: Attentional blink; Repetition blindness; Attention; Computational Modeling; Bayesian Models

Introduction

Repetition blindness and the attentional blink refer to common failures in identifying stimuli shortly after an important event. Both phenomena are typically studied in rapid serial visual presentation (RSVP) tasks, in which many stimuli (often letters or digits) are displayed in quick succession. In the task people are asked to pick out some specific "targets" from the RSVP stream (for instance, letters among digits, or letters cued by a ring). After the first target ($\mathbf{T}_1$), a second target ($\mathbf{T}_2$) appearing within a few stimuli of the first tends to be misreported – this is called the "attentional blink" (AB; Figure 1, top; Raymond, Shapiro, & Arnell, 1992; Weichselgartner, 1987). If an identical target is presented twice without much time intervening, people tend to not notice the repetition, yielding "repetition blindness" (RB; e.g., the 'E' in the bottom section of Figure 1; Kanwisher, 1987).

Although both RB and AB have many surface similarities, they follow a qualitatively different time-course as a function of the "stimulus onset asynchrony" (or lag) – the time between the onset of $\mathbf{T}_1$ and the onset of $\mathbf{T}_2$. The attentional blink exhibits lag-one sparing – if $\mathbf{T}_2$ appears immediately after $\mathbf{T}_1$, detection is often not impaired; in contrast, the RB deficit is most pronounced at such short delays. Moreover, when manipulated within one experiment, these effects have independent, dissociable effects (Chun, 1997). Several accounts of RB and AB postulate resource limitations (Shapiro, Raymond, & Arnell, 1997) in individuating items into unique tokens and binding them to type identities (Kanwisher, 1987; Bowman & Wyble, 2007). Moreover, the time-course of these effects has been the target of several computational process models that utilize specific mechanistic dynamics of attention and memory (Bowman & Wyble, 2007; Olivers & Meeter, 2008).

Here we do not aim to supplant these models, but only to provide a rational account of these dynamics, to explain why the visual system may exhibit AB and RB. We remain agnostic to the particular processes and implementations, and instead seek an understanding of the computational principles that may be in play. We develop a probabilistic model of target identification by *co-occurrence detection*, which aims to infer which item occurred simultaneously with the cue, given temporal uncertainty and expectations about world dynamics. First we describe the derivation of the model based on assumed temporal statistics, then illustrate its application to the basic RB and AB paradigms. Finally, we will apply this model to AB response distributions.

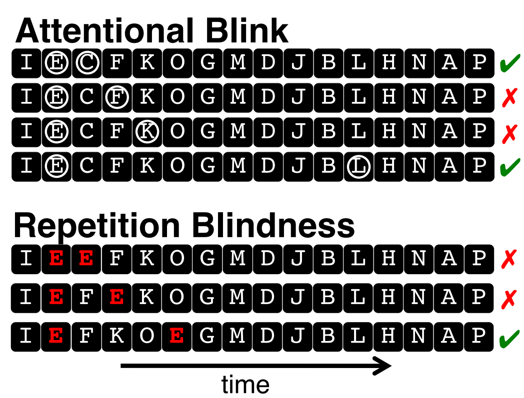

Figure 1: Repetition blindness and attentional blink paradigms. Each row of letters represents a particular type of trial where the letters appear in rapid succession. Repetition blindness arises when a target repeats, and participants indicate whether it appeared once or twice: performance is worse at short inter-target intervals when the two targets are identical compared to when they are different. The attentional blink occurs when the task requires identifying two designated targets: Participants are can report items in quick succession, as in the top attentional blink row ("lag-one sparing"); however, when the second target appears a few items later, participants are less accurate.

RSVP Identification Model

We assume that people identify targets in an RSVP task via the probability that a given item co-occurred with the cue. Under this assumption, the observer's task is to infer the letter identity ($\mathcal{L}$) of the second target ($\mathbf{T}_2$), by marginalizing over all possible stimuli used in the experiment (x_i) that may have coincided with the cue. The probability that the cue co-occurred with item i at time t is the product of the probability that the item ($\mathbb{X}_i(t)$) and cue ($\mathbb{C}(t)$) were present at that time. This quantity is integrated over time to obtain an unnormalized probability of item i co-occurring with the cue. The core of our model is an assumption that participants use prior expectations about the times when salient or task-relevant events are likely to occur ($P(\text{any}|t)$). The probability of co-occurrence at a given time is thus weighted by the probability the stimulus at time t will be relevant for the task. Thus, the participant's subjective probability distribution over possible identities of the second target is given by:

$$P(\Upsilon(\mathbf{T}_2) = \mathcal{L}|\mathbf{S},\Omega) \propto \sum_i P(x_i = \mathcal{L}) \int_{t=0}^{\infty} \mathbb{C}(t)\mathbb{X}_i(t)P(\text{any}|t), \quad (1)$$

where $P(x_i = \mathcal{L})$ is 1 when x_i is $\mathcal{L}$, and 0 otherwise. To distinguish between the letter identity, or *type*, and a particular instance of a letter, or the *token*, we use $\mathbf{T}_2$ to denote the token of the second target, and $\Upsilon(\mathbf{T}_2)$ to denote it's type.[1] $\mathbf{S}$ refers to the stimulus sequence of the trial, including the particulars of which items occurred at which points in time, and when the cue occurred. Ω includes the set of parameters which our model uses to describe subjective uncertainty and expectations about the task. The following sections describe the parameters used to calculate $\mathbb{C}(t)$, $\mathbb{X}_i(t)$, and $P(\text{any}|t)$.

Perceptual Likelihood

In RSVP tasks, participants often misreport item identities (Vul & Rich, 2010; Vul, Hanus, & Kanwisher, 2009) and their presentation order (Wyble, Potter, Bowman, & Nieuwenstein, 2011), indicating that even without RB and AB, there is considerable uncertainty about the relative timing of cues and items. The cue-function ($\mathbb{C}(t)$) and the item-function ($\mathbb{X}_i(t)$), account for this temporal uncertainty: they represent the subjective probability that these stimuli were presented at a given point in time. Both of these functions are obtained by convolving some temporal uncertainty kernel with the physical time series of stimulus presentation. We use the Student's t-distribution as the uncertainty kernel.[2]

Thus, the cue-function ($\mathbb{C}(t)$) and the item-function ($\mathbb{X}_i(t)$), are defined as a boxcar function – representing the physical presence of the stimulus – convolved with a t-distribution – representing perceptual uncertainty about the precise time of stimulus onset and offset:

$$\mathbb{C}(t) = P(\text{cue on}|t) = \text{boxcar}(t|\text{onset}), \text{offset}) \otimes \text{Student}(t|\sigma_c, \nu_c),$$

where ($\otimes$) is the convolution operator, the boxcar function is 1 between the onset and offset of the cue, and 0 otherwise, and Student$(t|\sigma_c, \nu_c)$ is a scaled Student's t-distribution with mean 0, ν_x degrees of freedom, and standard deviation σ_c. This convolution operation effectively captures uncertainty about the exact onset and offset of the cue. Similarly, the item function can be written as:

$$\mathbb{X}_i(t) = P(x_i|t) = \text{boxcar}(t|\text{onset}, \text{offset}) \otimes \text{Student}(t|\sigma_x, \nu_x).$$

In the simulations presented later $\nu_x = \nu_c = 2$, $\sigma_x = 30$, and $\sigma_c = 75$. The larger uncertainty for the cue reflects the fact that the cue in a RSVP stream is a rare and unpredictable occurrence relative to the steady stream of items.

Prior for Task-Relevant Events

Both RB and AB are conditioned on correct detection/identification of the first target ($\mathbf{T}_1$) in an RSVP stream. Therefore, we consider expectations about whether any task-relevant stimulus is likely to be visible at time t after the onset of the $\mathbf{T}_1$. We write this as $P(\text{any}|t)$. Since $\mathbf{T}_1$ is, by definition, task-relevant, $P(\text{any}|t)$ decomposes into the mutually exclusive probabilities that old task-relevant information is still visible ($P(old|t)$) and that new task-relevant information has appeared ($P(new|t)$):

$$P(\text{any}|t) = P(old \cup new|t) = P(\text{old}|t) + P(\text{new}|t). \quad (2)$$

To obtain these time-varying probability functions, we must assume a transition distribution for the world, that is: what is the probability distribution of the interval between changes in the world? We write this distribution as $P(R_1 = t)$, the probability that the first transition (R_1) occurred at time t. This distribution reflects implicit beliefs about the temporal structure of the world (we will later discuss different choices of this transition distribution). Figure 2 illustrates the derivation of the different elements of $P(\text{any}|t)$, from $P(R_1 = t)$. Based on this transition distribution, the probability that an old task-relevant stimulus is still visible at time t is given by the probability that the first transition has not yet happened – that is, the probability that the first transition will happen at a time later than t:

$$P(\text{old}|t) = P(R_1 > t) = 1 - \int_{t'=0}^{t} P(R_1 = t') \quad (3)$$

To find the probability that a new task-relevant stimulus has appeared by a given point, we must take into account that not every transition in the world produces a task-relevant stimulus. We define the parameter θ as the probability that a transition yields a task-relevant stimulus, and define $P(\text{new}|t)$ as the sum of a convolution series of transitions. The probability that the nth transition will occur at time t ($P(R_n = t)$) is computed as the nth convolution power of the transition dis-

[1] In general, $P(\Upsilon(x_i) = \mathcal{L})$ could capture perceptual uncertainty about stimulus identity (to reflect that some letters are more confusable than others), but here we ignore this complication, and assume no pairwise letter confusion, by assigning probability 1 to the actual identity of x_i, and 0 to all other alternatives.

[2] Our qualitative results do not much depend on the specific functional form of the temporal uncertainty kernel – simple Gaussian distributions are sufficient – but our numerical results are more consistent with human behavior when using a heavy-tailed distribution, like the t-distribution.

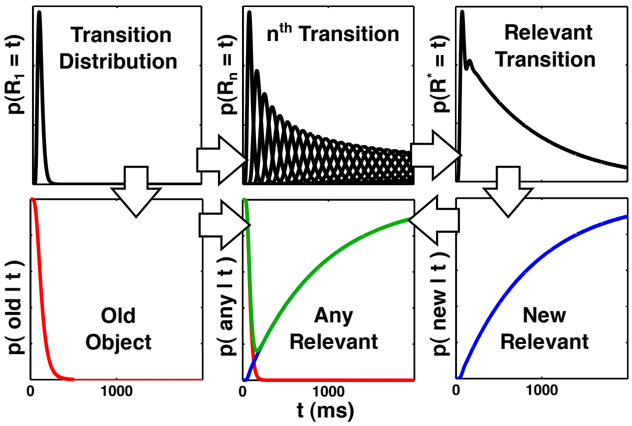

Figure 2: Construction of the prior probability of any task-relevant information at time t from an assumed transition distribution between world states (see text).

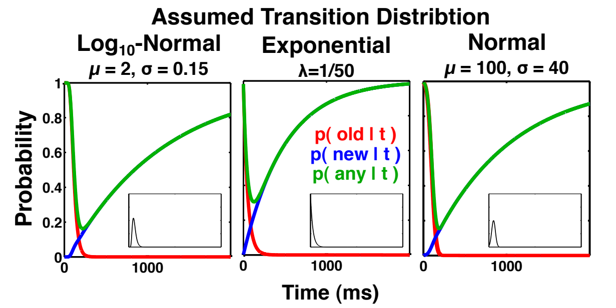

Figure 3: The resulting $P(old|t)$, $P(new|t)$, and $P(any|t)$ from several theoretically-motivated possible transition distributions. The probability density of each assumed transition distribution is illustrated in the insets.

tribution (where $\otimes$ is the convolution operator):

$$P(R_n = t) = P(R_1 = t)^{\otimes n} = \underbrace{P(R_1 = t) \otimes P(R_1 = t) \otimes \ldots \otimes P(R_1 = t)}_{n}. \tag{4}$$

Since each transition yields a task-relevant stimulus with probability θ, the probability that the first task-relevant stimulus (R^*) appears at time t is given by the probability of any number of non-task-relevant transitions occurring followed by a relevant observation (in other words, the sum of the geometrically weighted convolution series of transitions):[3]

$$P(R^* = t) = \sum_{n=1}^{\infty} \theta(1-\theta)^{n-1} P(R_n = t) \tag{5}$$

The probability that a new item has appeared at some point before time t – that is, $P(\text{new}|t)$ – can be calculated by evaluating the cumulative distribution of $P(R^* = t)$:

$$P(\text{new}|t) = P(R^* \leq t) = \int_{t'=0}^{t} P(R^* = t') \tag{6}$$

Choice of Transition Distribution

What might be a reasonable form of $P(R_1 = t)$? If we consider world changes to be the result of saccades, $P(R_1 = t)$ would be a distribution of inter-saccadic intervals, which is well-approximated by a log-normal distribution (Wang, Freeman, Merriam, Hasson, & Heeger, 2012). Alternatively, assuming transitions in the world follow a Poisson process will yield an exponential distribution of transition times. Another approach is to consider transition times to be normally distributed around the rate of the RSVP task itself. This would be the case if subjects were noisily calibrated to the RSVP stream used, which has a regular structure. The insets of Figure 3 illustrate these three possibilities. The larger graphs show, in green, $P(\text{any}|t)$ from each of these options. These resulting priors on a task-relevant stimulus being visible (in green) are similar for all three possible transition distribution functions. At short delays from the first target, the probability of a task-relevant object is initially high, since the first transition is unlikely to have yet occurred. However, after the first transition, with $\theta < 1$, a new object is not certain to be task-relevant; thus the probability that a new task-relevant item has appeared remains low. With more time, the probability of observing a novel task-relevant item increases. This combination results in the initial dip and later recovery of the prior on task-relevance – this shape of $p(\text{any}|t)$ is the backbone of our model.

Throughout the results sections, we use the $\log_{10}$-normal distribution with $\mu = 100$ and $\sigma = .15$, and $\theta = .1$. A higher θ results in a smaller dip in the resulting prior, and a smaller RB or AB effect; decreasing μ results in a faster time-course, while increasing μ results in a slower time-course.[4]

Modeling Repetition Blindness

In an RB paradigm, participants must decide whether a given stimulus was presented once or twice. In other words, they must decide if there were two tokens of the same type, or just one token (e.g., one or two "E"s in Figure 1). In our notation, the participants challenge is to determine if the two identified targets correspond to only one token ($\mathbf{T}_1 = \mathbf{T}_2$). This posterior belief is given by:

$$P(\mathbf{T}_1 = \mathbf{T}_2|\mathbf{S},\Omega) \propto \begin{cases} \Upsilon(\mathbf{T}_1) = \Upsilon(\mathbf{T}_2) & \int_{t=0}^{\infty} \mathbb{C}(t) \frac{P(old|t)}{P(old|t)+P(new|t)} \\ \Upsilon(\mathbf{T}_1) \neq \Upsilon(\mathbf{T}_2) & 0. \end{cases} \tag{7}$$

In words: if the types of $\mathbf{T}_1$ and $\mathbf{T}_2$ are different, then they must be different tokens; if their types are the same, then the

[3]In practice, we evaluate this infinite sum numerically by truncating higher order elements of the series (dropping ns for which the probability of having not yet transitioned to a task-relevant state ($(1-\theta)^{n-1}$) is less than .001).

[4]To the best of our knowledge, any transition distribution defined on the support of $(0,\infty)$ will produce qualitative results similar to RB and AB. This is because $P(\text{new}|t)$ is defined as the cumulative distribution of a series of convolutions of $P(R_1 = t)$, making the prior robust to the specific choice. The time-course of the RB/AB effects will be determined by the median of the transition distribution. In our case, anything with a median of about 100 seems to yield a qualitatively appropriate time course.

probability of them being the same token is given by the probability that the participant perceived the cue before a transition from $\mathbf{T}_1$ was expected.

We assume that participants adopt a soft-max response strategy (effectively sampling from this distribution); thus the expected frequency of correct detections of a repetition is $P(\mathbf{T}_1 \neq \mathbf{T}_2|\mathbf{S},\Omega) = 1 - P(\mathbf{T}_1 = \mathbf{T}_2|\mathbf{S},\Omega)$. The left graph of Figure 4 shows simulation results for a generic RB task with 40 ms targets and 40 ms gaps between items. Overall the qualitative pattern of data is correct. There is no "lag-one sparing", and the probability of a correct response increases with greater delays.

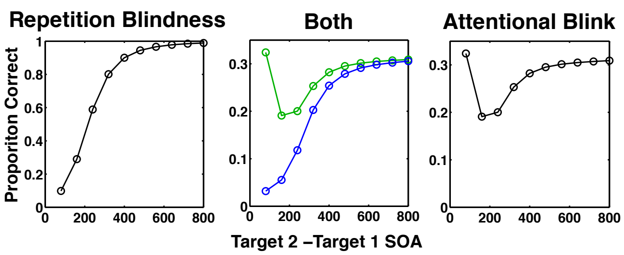

Figure 4: Model results for the RB, AB, and combination tasks. RB accuracy corresponds to the probability that the target item is a new token. AB accuracy is the probability that the cue is inferred to coincide with the second target item. The middle plot shows the combination of both effects, as in Chun, (1997).

Modeling the Attentional Blink

In the attentional blink task, the participant needs to determine the identity of the stimulus co-occurring with the cue. Therefore the posterior distribution that the correct $\mathcal{L}$ was $P(\Upsilon(\mathbf{T}_2))$ is given directly by Equation 1. Responses from the posterior are again considered to follow a soft-max rule. That is, the relative frequency of reporting the item $\mathcal{L}$ as the second target is $P(\Upsilon(\mathbf{T}_2) = \mathcal{L}|\mathbf{S},\Omega)$. The right panel in Figure 4 shows that the time-course of $\mathbf{T}_2$ accuracy matches the typical AB effect: Accuracy for target detection just after $\mathbf{T}_1$ is high (lag-one sparing), and then drops off, recovering after a few hundred milliseconds. The center panel illustrates simultaneous AB and RB effects using the same task and model structure, as has been observed behaviorally (Chun, 1997). Here we can account for both effects as a natural consequence of temporal uncertainty and expectations. While capturing of the qualitative patterns of the AB and RB results is interesting, there are many models that can fit this data (Bowman & Wyble, 2007; Olivers & Meeter, 2008). We now turn to the specific error distributions in the attentional blink to test finer-grained predictions of our model.

Error Distributions in the Attentional Blink

Our model has the fidelity required to make not only predictions of changes in accuracy as a function of inter-stimulus interval in the attentional blink, but also the error distributions – when $\mathbf{T}_2$ is misreported, which items are reported instead? Vul, Nieuwenstein, and Kanwisher (2008) presented subjects with an RSVP stream containing all 26 English letters at 12 items/sec. Two letters were cued as targets by flashes of an annulus around the stream (as illustrated in Fig 1). Participants were asked to report the identities of the letters that appeared simultaneously with the cues, in order. Since each english letter appeared only once in the RSVP stream, the authors could identify the exact serial position where the reported letter occurred on each trial; thus they could determine exactly which items, relative to $\mathbf{T}_2$, tended to be reported in its place (Figure 5). The attentional blink can be decomposed into three changes in error distributions that follow different time-courses (Figure 6): (1) At $\mathbf{T}_1 - \mathbf{T}_2$ lags of 100 to 500 ms, participants tend to make more random guesses: their responses are less likely to come from a window around $\mathbf{T}_2$; (2) At lags between 50 and 400 ms, responses tend to be more highly dispersed around $\mathbf{T}_2$: participants are more likely to report letters several items away from $\mathbf{T}_2$; (3) At lags shorter than 300 ms, items *preceding* $\mathbf{T}_2$ tend to be misreported in its place, but after 400 ms (and lasting to lags as long as several seconds) errors instead come from items *following* $\mathbf{T}_2$. We now test whether these three effects and their unique time-courses can also be accounted for by our model based on expectations about temporal structure of the environment.

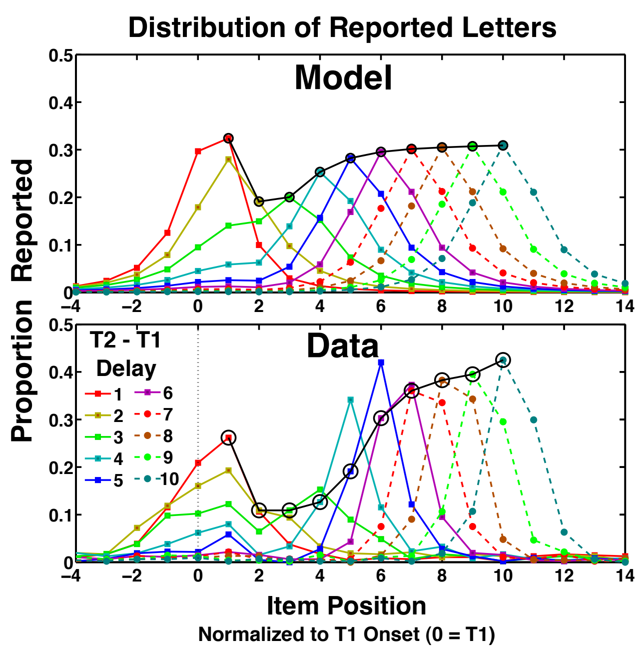

Figure 5: Response distributions for cues appearing at various positions. Behavioral data are replotted from (Vul et al., 2008). Each curve is the distribution of responses from a second cue appearing at a particular SOA relative to the first target. The dark curves show the correct responses that make up the basic AB effect.

Error Distributions in the Model

To model error responses, we calculate the model's prediction about the probability of each possible letter identity $\mathcal{L}$ being reported (following Equation 1). These error distributions are presented in Figure 5. Each colored curve plots the distribution of responses for a particular delay between $\mathbf{T}_1$ and $\mathbf{T}_2$ as a function of the reported item position relative to $\mathbf{T}_1$. During the attentional blink, the distribution of responses is distorted by the prior about when task-relevant items occur.

The most notable effect is the reduction in accuracy itself: the items corresponding to $\mathbf{T}_2$ itself are less likely to be reported at $\mathbf{T}_1 - \mathbf{T}_2$ lags between 150 to 500 ms. Critically, not only does the model match the accuracy time-course of human data, but it also captures the changes in error distributions. At short delays between $\mathbf{T}_1$ and $\mathbf{T}_2$, $\mathbf{T}_2$ errors tend to come from items preceding $\mathbf{T}_2$, and this pattern reverses at longer delays such that items following the second target are more common intrusions instead. Finally, our model also demonstrates increased variability in which items are reported during the attentional blink. These changes in error distributions are quantified identically to the behavioral data (Vul et al., 2008), and presented in Figure 6.

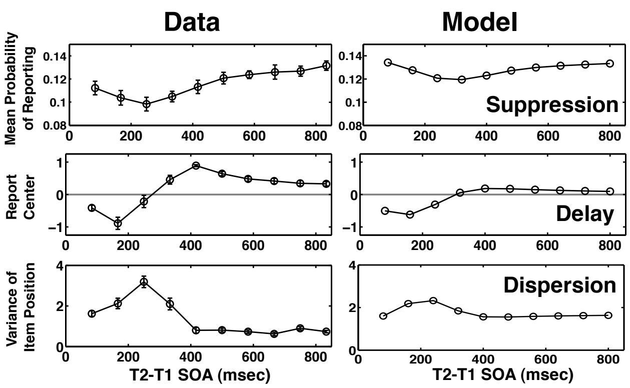

Figure 6: Summary statistics of $\mathbf{T}_2$ response distributions during the attentional blink, as a function of $\mathbf{T}_1$-$\mathbf{T}_2$ delay (stimulus onset asynchrony; SOA). Each statistic is computed within a seven item window centered on the position of the second target. Top row: "Suppression" refers to a decrease in the average probability of reports (y-axis) of items within the seven item window around $\mathbf{T}_2$. Middle row: "Delay" refers to the tendency to report items preceding $\mathbf{T}_2$ at short SOAs, and items following $\mathbf{T}_2$ at long SOAs; the y-axis is the mean position (relative to $\mathbf{T}_2$) of reported items within the window. Bottom row: "Dispersion" refers to the increased variance of reported positions within the window around $\mathbf{T}_2$ at SOAs around 150 to 300 msec.

Discussion

We proposed an account of perception in RSVP paradigms based on probabilistic inference about the co-occurrence of cues and items under expectations about the transitions of interesting items in the world. Our model captures patterns of accuracy in RB and AB, as well as the specific error distributions in AB. The core of this account is the premise that not every world transition brings an important item into view, and a prior constructed based on this premise produces qualitatively similar results regardless of the specific choice of transition distribution.

Why would participants use this prior, since it seems to result in performance deficits? We can only speculate that performance deficits in laboratory RSVP paradigms do not reflect the benefits that such a prior could confer to people in a more natural environment. Recent work has suggested a similar account of inhibition of return (Posner & Cohen, 1984) and immediate priming (Huber, 2008), by demonstrating that the time-courses of these effects are sensitive to learned temporal properties of the environment (Rieth, 2012).

Our account captures a number of other RSVP phenomena. First, because the prior is defined with respect to time, not serial position, the dynamics of lag-one sparing and the detriment to $\mathbf{T}_2$ reports should be constant in time, as has been experimentally shown (Wyble & Bowman, 2005; Vul et al., 2008). Second, because the relative delay of $\mathbf{T}_2$ reports modulates $\mathbf{T}_2$ accuracy, pre-cueing should improve $\mathbf{T}_2$ accuracy, as is the case (Nieuwenstein, Chun, van der Lubbe, Hooge, 2005; Vul et al, 2008). Third, because we assume that perception of particular targets are driven by a *sampling* process (Vul, Hanus, & Kanwisher, 2009), perception of $\mathbf{T}_2$ accuracy will be all-or-none, which appears to be the case (Sergent & Dehaene, 2004). Finally, since we explicitly consider the transition probability from one event to another, our account is consistent with results that the attentional blink facilitates item individuation (Wyble, Bowman, & Nieuwenstein, 2009).

A number of other findings that are not adequately captured by our account provide future challenges. Because we only consider perception of $\mathbf{T}_2$ contingent on $\mathbf{T}_1$ identification, we cannot account for "spreading the sparing" – the finding that multiple cues presented during the "lag-one sparing" interval extend the sparing window to later items (Olivers, van der Stigchel, & Hulleman, 2007) – or whole report errors and order confusion more generally (Wyble, et al., 2009; Nieuwenstein & Potter, 2006). Another challenge is the independence of the *types* of items and the assumed dynamics of the world; because of this independence, our account cannot capture the fact that sometimes words, and sometimes letters, may be treated as the relevant items of interest (Kanwisher & Potter, 1990). Furthermore, additional assumptions would be needed for our account to be consistent with findings that missed second targets in the attentional blink produce semantic priming (Shapiro, Driver, Ward, & Sorensen, 1997). Additionally, there are some aspects of human performance for which the model is quantitatively off. Overall, suppression is smaller in the model (compare the graphs in the top row of Figure 6). For longer SOAs there is less delay in the model, slightly more dispersion, and a less extreme peak in dispersion. Furthermore, in the data – but not the model – the modal response for $\mathbf{T}_2$ at SOAs of four to six items is actually an incorrect answer. Finally, in the AB effect, accuracy for the item just after the first cue is too high relative to other positions.

Despite these challenges, we are encouraged that our rational inference account is consistent with classic type-token theories (Kanwisher, 1987; Chun, 1997): the discrepancy between identifying, and individuating, items in an RSVP stream seems to be of central importance to combining and distinguishing between repetition blindness and the attentional blink, and our account formalizes this intuition in probabilistic inference. Furthermore, the mechanistic dynamics of attention postulated in process models of these phenomena match our prior about task transitions (Bowman & Wyble, 2007; Olivers & Meeter, 2008). We hope that future work might provide a synthesis of these accounts.

One promising direction for future work is to extend our model to include the perception of $\mathbf{T}_1$. Because our objective was to demonstrate that RB and AB effects can result from rational use of expectations about world transitions, our current account is conditioned on the identification of the first target; however, a complete model of RSVP perception must include $\mathbf{T}_1$. This extension might be achieved via an online inference process to account for changing beliefs over the time course of a trial, capitalizing on particle filter approximate inference algorithms (Doucet, Freitas, & Murphy, 2000) and expected duration hidden Markov models (Rabiner, 1989). With an online inference process the model could capture whole report paradigms (Nieuwenstein & Potter, 2006) and might explain the effect of "spreading the sparing" (Olivers, et al., 2007). Moreover, such an extension might help connect our computational level (Marr, 1982) account to to previous models of the attentional blink (Bowman & Wyble, 2007; Olivers & Meeter, 2008) by postulating approximate inference algorithms, and their possible neural implementations (Fiser, Berkes, Orbn, & Lengyel, 2010; Vul & Pashler, 2008; Moreno-Bote, Knill, & Pouget, 2011).

Conclusions

Our results suggest that the attentional blink and repetition blindness are both consequences of rational inference about cue-item co-occurrence given a prior about the rate of transitions to task-relevant items. This framework accounts for the attentional blink and repetition blindness effects as well as the error distributions of in the attentional blink paradigm. The difference between these tasks under our framework supports the intuition that repetition blindness and the attentional blink are two sides of the same coin: both are consequence of rational inference under identical assumptions, but with the observer asked to identify types, or distinguish tokens.

Acknowledgments: EV and CR were supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of the Interior (DOI) contract D10PC20023. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DOI, or the U.S. Government.

References

Bowman, H., & Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review, 114(1)*, 38-70.

Chun, M. M. (1997). Types and tokens in visual processing: A double dissociation between the attentional blink and repetition blindness. *Journal of Experimental Psychology: Human Perception and Performance, 23(3)*, 738-755.

Doucet, A., Freitas, N. D., & Murphy, K. (2000). Rao-Blackwellised particle filtering for dynamic Bayesian networks. *Proceedings of the 16th Annual Conference on Uncertainty in Artificial Intelligence*, 176-183.

Fiser, J., Berkes, P., Orbn, G., & Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. *Trends in Cognitive Sciences, 14(3)*, 119-130.

Huber, D. E. (2008). Immediate priming and cognitive aftereffects. *Journal of Experimental Psychology: General, 137(2)*, 324-347.

Kanwisher, N. (1987). Repetition Blindness: Type recognition without token individuation. *Cognition, 27(2)*, 117-143.

Kanwisher, N. G., & Potter, M. C. (1990). Repetition blindness: levels of processing. *Journal of experimental psychology. Human Perception and Performance, 16(1)*, 30-47.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman.

Moreno-Bote, R., Knill, D. C., & Pouget, A. (2011). Bayesian sampling in visual perception. *Proceedings of the National Academy of Sciences, 108(30)*, 12491-12496.

Nieuwenstein, M. R., Chun, M. M., van der Lubbe, R. H. J., & Hooge, I. T. C. (2005). Delayed Attentional Engagement in the Attentional Blink. *Journal of Experimental Psychology: Human Perception and Performance, 31(6)*, 1463-1475.

Nieuwenstein, M. R., & Potter, M. C. (2006). Temporal Limits of Selection and Memory Encoding A Comparison of Whole Versus Partial Report in Rapid Serial Visual Presentation. *Psychological Science, 17(6)*, 471-475.

Olivers, C. N. L., & Meeter, M. (2008). A boost and bounce theory of temporal attention. *Psychological Review, 115(4)*, 836-863.

Olivers, C. N. L., van der Stigchel, S., & Hulleman, J. (2007). Spreading the sparing: Against a limited-capacity account of the attentional blink. *Psychological Research, 71(2)*, 126-139.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and performance X: Control of language processes, 32*, 531-556.

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE, 77(2)*, 257-286.

Raymond, J., Shapiro, K. L., & Arnell, K. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink?. *Journal of Experimental Psychology. Human Perception and Performance, 18(3)*, 849-860.

Rieth, C. A. (2012). Adaptations of temporal dynamics: Faces, places, and words. Doctoral dissertation, UCSD

Shapiro, K. L., Raymond, J., & Arnell, K. (1997). The attentional blink. *Trends in Cognitive Sciences, 1(8)*, 291-296.

Sergent, C., & Dehaene, S. (2004). Is Consciousness a Gradual Phenomenon? Evidence for an All-or-None Bifurcation During the Attentional Blink. *Psychological Science, 15(11)*, 720-728.

Shapiro, K., Driver, J., Ward, R., & Sorensen, R. E. (1997). Priming from the Attentional Blink: A Failure to Extract Visual Tokens but Not Visual Types. *Psychological Science, 8(2)*, 95-100.

Vul, E., & Pashler, H. (2008). Measuring the Crowd Within. *Psychological Science, 19(7)*, 645-647.

Vul, E., & Rich, A. N. (2010). Independent sampling of features enables conscious perception of bound objects. *Psychological Science, 21(8)*, 1168-1175.

Vul, E., Hanus, D., & Kanwisher, N. (2009). Attention as inference: Selection is probabilistic; responses are all-or-none samples. *Journal of Experimental Psychology: General, 138(4)*, 546-560.

Vul, E., Nieuwenstein, M., & Kanwisher, N. (2008). Temporal selection is suppressed, delayed, and diffused during the attentional blink. *Psychological Science, 19(1)*, 55-61.

Wang, H., Freeman, J., Merriam, E., Hasson, U., & Heeger, D. J. (2012). Temporal eye movement strategies during naturalistic viewing. *Journal of Vision, 12*, 1-27.

Weichselgartner, E. (1987). Dynamics of automatic and controlled visual attention. *Science, 238(4828)*, 778-780.

Wyble, B., & Bowman, H. (2005). Computational and experimental evaluation of the attentional blink: Testing the simultaneous type serial token model. *Proceedings of the 27th Annual Conference of the Cognitive Science Society*, 2371-2376.

Wyble, B., Bowman, H., & Nieuwenstein, M. R. (2009). The attentional blink provides episodic distinctiveness: Sparing at a cost. *Journal of Experimental Psychology. Human Perception and Performance, 35(3)*, 787-807.

Wyble, B., Potter, M. C., Bowman, H., & Nieuwenstein, M. (2011). Attentional episodes in visual perception. *Journal of Experimental Psychology. General, 140(3)*, 488-505.