Cognition 165 (2017) 73–81

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Original Articles

The effect of phasic auditory alerting on visual perception

Anders Petersen [a,*], Annemarie Hilkjær Petersen [b], Claus Bundesen [a], Signe Vangkilde [a], Thomas Habekost [a]

[a] Department of Psychology, University of Copenhagen, Denmark
[b] Center for Rehabilitation of Brain Injury, Copenhagen, Denmark

ARTICLE INFO

Article history:
Received 1 December 2016
Revised 4 April 2017
Accepted 12 April 2017
Available online 11 May 2017

Keywords:
Phasic alerting
Theory of visual attention
Mathematical modeling
Pupillometry

ABSTRACT

Phasic alertness refers to a short-lived change in the preparatory state of the cognitive system following an alerting signal. In the present study, we examined the effect of phasic auditory alerting on distinct perceptual processes, unconfounded by motor components. We combined an alerting/no-alerting design with a pure accuracy-based single-letter recognition task. Computational modeling based on Bundesen's Theory of Visual Attention was used to examine the effect of phasic alertness on visual processing speed and threshold of conscious perception. Results show that phasic auditory alertness affects visual perception by increasing the visual processing speed and lowering the threshold of conscious perception (Experiment 1). By manipulating the intensity of the alerting cue, we further observed a positive relationship between alerting intensity and processing speed, which was not seen for the threshold of conscious perception (Experiment 2). This was replicated in a third experiment, in which pupil size was measured as a physiological marker of alertness. Results revealed that the increase in processing speed was accompanied by an increase in pupil size, substantiating the link between alertness and processing speed (Experiment 3). The implications of these results are discussed in relation to a newly developed mathematical model of the relationship between levels of alertness and the speed with which humans process visual information.

1. Introduction

Alertness refers to the brain's general readiness to respond to an upcoming event (Posner & Petersen, 1990). The ability to prepare and sustain alertness is an important attentional function, which, if impaired, can result in severe attentional problems. A broad distinction is made between tonic and phasic alertness (Sturm & Willmes, 2001): Tonic (intrinsic) alertness denotes the sustained long-term intensity level of attention, whereas phasic (extrinsic) alertness refers to a short-lived increase of attention elicited, for example, by a warning signal preceding an upcoming event. Phasic alertness has been found to reduce reaction times in response to various stimuli (Fan, McCandliss, Sommer, Raz, & Posner, 2002; Posner & Boies, 1971), which was originally attributed to faster preparation and/or execution of the motor response (Posner, 1978; Sanders, 1980). Evidence, however, is accumulating showing that phasic alertness also affects earlier perceptual processes. For instance, Matthias et al. (2010) presented a visual alerting cue prior to a whole report letter display (Sperling, 1960) and used mathematical modeling based on the Theory of Visual Attention (TVA; Bundesen, 1990) to show that phasic visual alerting increased visual processing speed and changed the spatial distribution of attentional resources. These findings are particularly interesting because they were obtained using pure accuracy-based measures, unconfounded by motor components. Phasic auditory alertness has also been reported to affect early perception (Jepma, Wagenmakers, Band, & Nieuwenhuis, 2009; Kusnir, Chica, Mitsumasu, & Bartolomeo, 2011; Robertson, Mattingley, Rorden, & Driver, 1998; Weinbach & Henik, 2011). However, most of these studies rely on reaction time-based measures, by which it is difficult to disentangle effects on early perception from effects on later, response-based processes (although see Finke et al., 2012; Robertson et al., 1998; and Brown et al., 2015). Thus, in this article we use pure accuracy-based measures and TVA-based modeling to examine how phasic auditory alertness influences early visual perception. In contrast to most previous research on phasic alerting, we do not rely on uniform foreperiod distributions (i.e., distributions of the time interval between the cue and the target) because such distributions confound the alerting effect with the build-up of temporal expectancy as time elapses (Weinbach & Henik, 2012). Instead, we use a non-aging foreperiod distribution (Niemi & Näätänen, 1981) to reduce the effect of temporal expectancy (but see Lawrence & Klein, 2013 for a more sophisticated, yet less

* Corresponding author.
E-mail address: anders.petersen@psy.ku.dk (A. Petersen).

http://dx.doi.org/10.1016/j.cognition.2017.04.004
0010-0277/

simple, approach). A non-aging distribution has a constant hazard rate (i.e., a constant probability that the target appears during the next little period of time, given that the target has not yet appeared), resulting in a distribution with many short and only a few long foreperiods. Experiment 1 was a single-letter recognition task in which a high-intensity alerting cue (85 dB) was presented prior to a backward-masked target letter in 1/3 of the trials, and TVA-estimates of visual processing speed and perceptual threshold were compared between the alerting and no-alerting conditions. In Experiment 2, we investigated the effect of alerting intensity on TVA-estimates by including both high (85 dB) and low (40 dB) intensity cues in an experimental design similar to the design used in Experiment 1. Finally, in Experiment 3 we replicated Experiment 2 but included measures of pupil size to examine physiological effects of auditory alerting (see, e.g., Kahneman, 1973; Tona, Murphy, Brown, & Nieuwenhuis, 2016).

Based on previous findings by Matthias et al. (2010), we hypothesized that presentation of an auditory alerting cue, similar to the presentation of a visual alerting cue, increases visual processing speed. Further, theoretical considerations by Bundesen, Vangkilde, and Habekost (2015; see discussion) suggest that the intensity of alerting should correlate with the observed estimates of processing speed. It has previously been reported that increasing levels of auditory alerting prior to an imperative stimulus lead to faster response times (Behar & Adams, 1966; Keuss, 1972), but to the best of our knowledge this is the first time effects of alerting intensity on early perceptual processes have been investigated using pure accuracy-based measures, unconfounded by motor processes.

2. Theory of Visual Attention

The behavioral data in this article were analyzed by use of Bundesen's (1990) Theory of Visual Attention (TVA). In this section, we introduce TVA and the way it was used to analyze data from the single-letter recognition task. In general, TVA proposes that an object x in the visual field is encoded into visual short-term memory (VSTM) by encoding one or more categorizations of the object into VSTM. A categorization has the form "object x belongs to category i" (or equivalently "object x has feature i"), where i is a perceptual category (e.g., a certain letter shape, color, size, or orientation). Consider the hazard rate, $v(x,i)$, of the event that the categorization "x belongs to i" becomes encoded into VSTM at a given time t. If $f(t)$ and $F(t)$ are the probability density and distribution functions of the event, the hazard rate is $f(t)/[1-F(t)]$, which is a measure of the speed of processing at time t. By the rate equation of TVA,

$$v(x,i) = \eta(x,i)\beta_i \frac{w_x}{\sum_{z\in S} w_z}, \qquad (1)$$

where $\eta(x,i)$ is the strength of the sensory evidence that object x belongs to category i, β_i is the perceptual bias associated with category i, and w_x is the attentional weight of object x, which is divided by the sum of attentional weights across all objects in the visual field, S. When several objects are presented simultaneously in the visual field, they compete for access to the limited storage space of VSTM. Objects that are likely to win the competition are those that differ from their local surroundings (feature contrast) and those that are relevant for the task (feature relevance). Such objects get high attentional weights (see Nordfang, Dyrholm, & Bundesen, 2012), and task-relevant categorizations of objects with high attentional weights are likely to become encoded into VSTM (see Eq. (1)).

As an example, consider the task of reporting the identity of red letters (targets) among equally salient blue letters (distractors). According to TVA, the visual system solves this task by setting the attentional weights high for red objects and low for blue objects. This makes categorizations of red objects faster than categorizations of blue objects. Furthermore, to facilitate that only task-relevant categorizations of letter shapes are encoded into VSTM, the perceptual biases associated with letter shapes (i.e., $\beta_A, \beta_B, \ldots, \beta_Z$) are set high, whereas perceptual biases for task-irrelevant categories are set at values near zero.

TVA is a general model of visual perception (see Bundesen, 1990; Bundesen, Habekost, & Kyllingsbæk, 2005). In this article, we use TVA to model data from a simple, single-letter recognition task, which greatly reduces the complexity of the model. If we neglect the storage limitation of VSTM and neglect potential perceptual confusions by setting $\eta(x,i) = 0$ for all incorrect categorizations of object x, the probability of encoding object x into VSTM can be expressed as,

$$p = \begin{cases} 1 - e^{-v_x(\tau - t_0)} & \tau > t_0 \\ 0 & \tau \leqslant t_0 \end{cases}, \qquad (2)$$

where v_x is the processing speed of object x (i.e., the speed of the process of correctly categorizing object x), τ is the exposure duration of object x, and t_0 is the longest ineffective exposure duration (a.k.a. the threshold of perception). That is, if the exposure duration of an object is shorter than or equal to t_0, the probability of encoding the object into VSTM is zero. However, if the exposure duration of the object is longer than t_0, the probability of encoding the object into VSTM increases exponentially as a function of $\tau - t_0$. An instance of the model (corrected for guessing, see method section of Experiment 1) is provided in Fig. 2 with t_0 interpreted visually as the exposure duration at which the curve rises from pure guessing performance and v as the slope of the curve at t_0.

3. Experiment 1

To investigate the effect of phasic auditory alerting on visual perception we conducted an experiment in which participants were to report the identity of a post-masked letter presented for varying exposure durations. This made it possible to perform a TVA-based modeling of the data for each individual subject. In one third of the trials, a loud 85 dB auditory alerting cue preceded the presentation of the letter. In the remaining two thirds of the trials, the letter was presented without a preceding cue. The TVA-based modeling of the data was performed independently for the two conditions.

3.1. Method

3.1.1. Participants

28 Danish students (23 females, 5 males, mean age = 22.8 years, *SD* = 1.8 years) were paid a standard fee by the hour (18 students) or received course credits (10 students) for participating in the experiment. All had normal or corrected-to-normal vision and performed within the normal hearing range on a screening audiometer test (Oscilla® USB-310). The study was approved by the departmental board of ethics (No. 2012/2).

3.1.2. Design

In one third of the trials, an 85 dB auditory cue was played prior to the presentation of a target letter (85 dB cue condition), whereas in the remaining two thirds of the trials no auditory cue was given (no cue condition; see Fig. 1). This distribution of trials was chosen to allow for longer periods with low alertness while leaving enough trials in the cue condition for TVA-based modeling. To avoid habituation to a specific auditory stimulus, the cue was played equally often at either a high pitch of 900 Hz or at a low