A model of visible persistence and temporal integration

[1]*MARINA T. GRONER, [2]WALTER F. BISCHOF and [2]VINCENT DI LOLLO
[1]*Department of Psychology, University of Bern, Laupenstrasse 4, 3008 Bern, Switzerland and*
[2]*Department of Psychology, University of Alberta, Edmonton, Alberta T6G 2E9, Canada*

Received 5 August 1988; revised 20 November 1988; accepted 23 November 1988

Abstract—A quantitative model of temporal integration and visible persistence is described and tested. The model treats visible persistence as resulting from processing activity within sustained visual channels whose temporal response is modelled using a second-order control system. Temporal integration of two successive stimuli is assumed to be enabled by the overlap between the two periods of activity and to be impaired by the non-overlapping activity. The model successfully predicts the effects of inter-stimulus interval and of stimulus duration on goodness of temporal integration.

INTRODUCTION

Visible persistence may be described as the continued visibility of a stimulus for a brief period after its physical termination. At first, visible persistence—also known as *iconic memory*—was likened to a sensory store whose contents decayed rapidly after stimulus termination (e.g., Sperling, 1960; Neisser, 1967). However, later findings showed a simple storage hypothesis to be inadequate (see Coltheart, 1980). An alternative hypothesis regards visible persistence as the outcome of processing activity time-locked to the onset of the inducing stimulus. This hypothesis, first suggested by Di Lollo (1977, 1980), is elaborated further in the present work.

Visible persistence has been the subject of a large number of studies that have identified some of its fundamental characteristics (see review by Coltheart, 1980). Although some quantitative models of temporal integration have been proposed (e.g., Kelly, 1971), theoretical interpretations of visible persistence have been qualitative rather than quantitative. It is the purpose of the present work to examine the possibility of quantifying the main temporal factors underlying visible persistence. The plan is to specify the known temporal characteristics of the phenomenon, and to formulate a quantitative model capable of explaining the data.

Amongst the temporal variables known to affect the strength of visible persistence, two have been found to yield particularly reliable effects: inter-stimulus interval (ISI), and stimulus duration. Inter-stimulus interval is the period that elapses between the termination of a stimulus whose persistence is to be estimated and the onset of a trailing stimulus that acts as a probe. As might be expected, strength of persistence diminishes as the ISI is increased. The effect of duration, on the other hand, is somewhat counterintuitive: as stimulus duration is increased, there is a corresponding *decrement* in duration of persistence (e.g., Efron, 1970; Bowen *et al.*, 1974; Di Lollo and Bourassa, 1983).

*To whom all correspondence should be addressed

A quantitative model of the effects of ISI and stimulus duration was formulated and tested in the present work. Experiment 1 replicated the well-known effect of ISI, and provided the data on which the model was developed. The model's predictions were then tested in Experiment 2, in which ISI was held constant and stimulus duration was systematically varied.

EXPERIMENT 1

Experiment 1 had two related aims. While replicating the known effect of ISI, the study was designed to provide the data on which a quantitative model of visible persistence could be based.

To study visible persistence, we employed a paradigm first developed by Eriksen and Collins (1967), and later named by Coltheart (1980) as "Temporal integration of form parts". In its present application, the paradigm comprised two brief stimuli, displayed successively, separated by an ISI. Viewed in isolation, each stimulus was a meaningless configuration of dots. However, when seen in combination, the two stimuli formed a 5×5 square matrix with one dot missing from a randomly chosen location. The observer was required to report the matrix location of the missing dot.

The task is easy and engaging, provided that the two portions of the display are perceptually integrated and are, therefore, seen as one. Performance is virtually errorless at brief ISIs (at which a unitary pattern is perceived), but deteriorates rapidly at longer ISIs (at which two sequential stimuli are seen and identification of the missing element becomes impossible). Perception of a unitary configuration clearly requires that some form of visible persistence of the leading display be available to bridge the temporal gap (ISI) between the two successive patterns. Duration of visible persistence is estimated by the longest ISI at which temporal integration remains possible, and beyond which errors are made.

In Experiment 1 the intensity and duration of the two successive portions of the display were held constant, while the duration of the ISI was systematically varied.

Method

Observers. Two authors (MTG and VDL) and one paid student unaware of the purpose of the research served in both experiments. All had normal or corrected-to-normal vision.

Visual displays. The display consisted of 24 of the 25 dots defining a 5×5 matrix that formed a square having a side of 20 mm on the face of a Hewlett-Packard 1333A oscilloscope equipped with P15 phosphor. The diameter of each dot was approximately 0.2 mm. At the viewing distance of 57 cm, set by a head-rest, the dot-matrix subtended a visual angle of 2.0 deg.

The dot-matrix was displayed in two successive flashes of 12 dots each. The 12 dots in each flash were chosen randomly on each trial from the 25-dot pool. Therefore, the location of the missing dot varied randomly from trial to trial. Each flash was displayed for 40 ms at a dot intensity of 153 cd/m^2 on a background of approximately 0.5 cd/m^2.

Design and procedure. The observers sat in a dimly-illuminated test chamber, and viewed the displays binocularly with natural pupils. Four fixation dots defined a 3-deg square area in the centre of the screen. All stimuli were displayed within this square area. The observer initiated each display by pushing a hand-held button. Upon a button-press, the first 12 dots were displayed for 40 ms. Next, there was an ISI of either 0, 20, 40, 60, 80, or 100 ms (the nominal value of an ISI of 0 was less than 20 μs). Finally, the remaining 12 dots were displayed for 40 ms. The observer then identified the matrix location of the missing dot (guessing if unsure) by encoding its matrix coordinates in a 5-button response box. No feedback was provided to the observer.

An experimental session consisted of 150 randomly ordered trials comprising 25 presentations of the dot matrix at each of the six ISIs. The 150 trials occurred in a different random order in each session, and were completed within about 10 min. Each observer served in four sessions, yielding a total of 100 observations at each ISI.

Results

Figure 1 shows the percentage of correct responses made by each observer at each of the six durations of the ISI. Within each graph, the individual symbols represent the empirical results, and the continuous curve shows the prediction based on the model described below.

For each observer, accuracy of performance was high at brief ISIs, but deteriorated rapidly as the duration of the ISI was increased. The phenomenal appearance of the

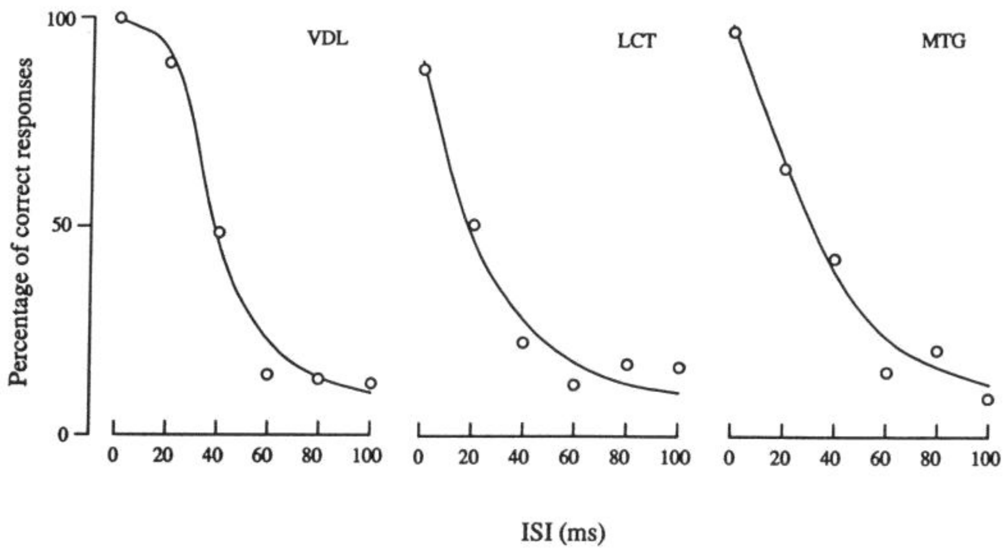

Figure 1. Percentage of correct responses as a function of ISI for each of the three observers in Experiment 1. The continuous lines show the model's predicitons. Standard deviations ranged from 0 to 5%.

displays is worth noting. At brief ISIs, the two sets of 12 dots were seen as a single integrated matrix form, and the location of the missing dot could be identified with ease. At longer ISIs, the two sets of dots were seen as separate configurations, and the observer was forced to guess the answer from what appeared to be a large number of empty matrix locations. No coherent motion was seen in the displays, although evidence of local motion was noted at the longer ISIs. Such motion could not have facilitated identification of the empty matrix location; indeed, it might have interfered with correct responding by suppressing the visible persistence of the temporally leading dots through a process akin to metacontrast masking (e.g., Di Lollo and Hogben, 1987). These results are in good agreement with the outcomes of other investigations that employed similar paradigms (Eriksen and Collins, 1967; Hogben and Di Lollo, 1974).

A model of temporal integration

In designing a model of temporal integration, there is a potentially large range of possibilities, both in terms of types of models to be considered and in terms of the empirical findings to be used in constraining the model space. With respect to the former, temporal integration could be modelled using linear or nonlinear systems (Ogata, 1970). Since there have been only few attempts at modelling temporal integration (e.g., Kelly, 1971), we consider it important to begin by exploring the power and limitations of simple linear models, despite the fact that some experimental evidence seems to suggest that nonlinear processes may be involved in the early stages of vision (Watson, 1987).

Although it is obvious that performance in temporal integration is determined by both temporal and spatial factors (Farrell, 1984; Di Lollo and Hogben, 1985, 1987), we restrict our model to explaining the former only. Further, we require that any model of temporal integration should generate predictions for arbitrary temporal patterns, thus excluding any approach based on modelling the response to a fixed temporal pattern. In our simulations, we have obtained good results using a model based on a second-order control system proposed by Caelli *et al.* (1985) which was applied to data on flicker frequency, subthreshold summation, and masking.

We describe the model in three steps: first, we outline the temporal properties of the system's response, characterized by its impulse response function. Second, we describe the temporal integration process, and, third, we describe the model's linkage between temporal integration and response generation.

The response $r(t)$ of a linear system to an arbitrary signal $s(t)$ can be characterized by the impulse response $g(t)$; more precisely, the response of a linear system is given by the convolution integral

$$r(t) = \int_0^t s(\tau)g(t-\tau)\mathrm{d}\tau. \quad (1)$$

There is ample empirical evidence that the response of the visual system to temporal patterns must take into account both transient and sustained channels (Tolhurst, 1973; Breitmeyer and Ganz, 1976; Di Lollo and Woods, 1981; Breitmeyer, 1984). However, we assume that temporal integration of form parts (the paradigm employed in the present work) involves mainly sustained mechanisms. Thus we use a single sustained mechanism (Caelli *et al.*, 1985) characterized by the impulse response function $g(t)$

$$g(t) = \frac{ab}{b-a}(\mathrm{e}^{-at} - \mathrm{e}^{-bt}) \quad a \neq b, a > 0, b > 0 \quad (2)$$

shown in Fig. 2a with parameters $a = 0.0167\ \mathrm{ms}^{-1}$, and $b = 0.0178\ \mathrm{ms}^{-1}$. The response of the system to a typical stimulus of Experiment 1, i.e. two pulses separated by an ISI, is shown in Fig. 2b.

Given the temporal integration paradigm employed in the present work, there are three stimulus configurations that could possibly be seen: the configuration formed by the leading stimulus, that formed by the trailing stimulus, and the integration of the two. Whether one or more of these configurations is seen, depends on the display conditions. For example, as described above, at very long ISIs two successive configurations of 12 dots each are seen. It may be said that, under these conditions, the representation of the integrated parts has zero strength, while the separate represent-

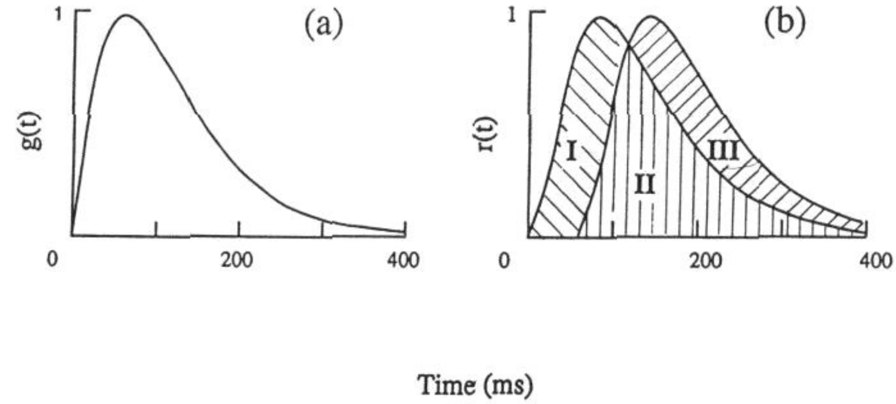

Figure 2. The panel on the left (a) shows the temporal impulse response function used in the model. The panel on the right (b) shows the response of the system to two stimuli displayed successively for 40 ms each, separated by an ISI of 20 ms. The region designated as II represents the area of temporal overlap between the two response functions. The regions designated as I and III represent the non-overlapping areas of the responses to the leading and trailing stimuli, respectively.

ations of the two component stimuli have maximum strengths. By contrast, at an ISI of zero, only one configuration of 24 dots is seen. In this case, the representation of the integrated parts is at maximum strength, while the other two are at zero.

Estimating the strengths of the representations over time requires several steps, as follows. Let us consider a hypothetical single display. The stimulus is held to produce a peripheral neural response defined by $r(t)$ as described in Eq. (1). Based on this peripheral activity, a more central representation is extracted whose strength, at first approximation, can be described in terms of the integral of $r(t)$ over time.

Let us now consider the integration of two successive representations. It is plausible to assume that the extent to which a common representation can be extracted is bounded by the weaker of the two successive neural responses at any point in time. Although there are several viable alternatives, we use the area of overlap of the system's responses to the two pulses to estimate the strength of the common representation. More specifically, we assume that the strengths of the combined representation and of the individual components stand in a countervailing relation: increments in the strength of the combined representation are matched by corresponding decrements in the strength of the individual components. In practice, we subtract the strength of the combined representation from that of the two components. These assumptions are illustrated in Fig. 2b, where the strength of the leading display is given by Area I, the strength of the trailing display is given by Area III, and the strength of the combined representation is given by Area II.

In terms of signal detectability theory, the integrated representation may be regarded as the signal (with strength proportional to Area II in Fig. 2b), and the two individual representations may be regarded as noise (with strength proportional to area I + III in Figure 2b). Hence, the signal-to-noise ratio of the temporally integrated representation is given by the ratio II/(I + III), or, more precisely, by

$$\mathrm{SNR} = \frac{\int_\eta^\infty r_1(t)\mathrm{d}t + \int_0^\eta r_2(t)\mathrm{d}t}{\int_0^\infty r_1(t)\mathrm{d}t + \int_0^\infty r_2(t)\mathrm{d}t - 2\{\int_\eta^\infty r_1(t)\mathrm{d}t + \int_0^\eta r_2(t)\mathrm{d}t\}} \tag{3}$$

where $r_1(t)$ and $r_2(t)$ are the system's responses to the first and second parts of the stimulus pattern, respectively, and η is defined such that $r_1(\eta) = r_2(\eta)$.

The third part of the temporal integration model is concerned with mapping the unbounded signal-to-noise ratio (3) into the interval $[c, 1]$ for predicting the probability of correct identification, where c is the performance at chance level. Previous studies (e.g., Di Lollo, 1980) have shown that observers almost never report a dot if the second flash as missing. Hence, we assume $c = 1/12$. For the transformation mapping SNR into probability we used.

$$p_{\text{correct}} = c + (1-c)\frac{2}{\sqrt{2\pi}}\int_0^{\alpha} \exp(-z^2/2)\,dz \tag{4}$$

where $\alpha = \text{SNR}^d$ and d is a free parameter. This transformation is equivalent to a single-channel version of the continuous-threshold model suggested by Legge and Foley (1980) with the parameter d controlling the acceleration of the response function (see Olzak and Thomas, 1986).

The model's predictions for Experiment 1 are shown in Fig. 1. The parameter values for observers VDL, LCT, and MTG, respectively, were as follows: $a = 0.0104\,\text{ms}^{-1}$, $0.0185\,\text{ms}^{-1}$, $0.0148\,\text{ms}^{-1}$; $b = 0.0200\,\text{ms}^{-1}$, $0.0215\,\text{ms}^{-1}$, $0.0193\,\text{ms}^{-1}$; $d = 3.071$, 1.815, 1.835. All parameter estimates were obtained using a steepest-descent algorithm known as STEPIT. For each of the three observers, respectively, the goodness of fit, $\chi^2(5)$ was 0.046, 0.057, 0.074 (in each case, $P > 0.5$). The proportion of the variance accounted for by the model was 0.99, 0.98, and 0.98 for each observer, respectively.

Although the model contains three free parameters, the family of curves generated in the parameter space is quite restricted. Thus, the model could have been easily refuted as has been the case with several alternative models that we considered. The estimated parametric values of a and b are reasonably close to those obtained by Caelli *et al.* (1985) for predicting data on flicker frequency, subthreshold summation, and masking ($a = 0.030$ ms; $b = 0.040$ ms). Furthermore, the a/b ratio in the present work was very close to that reported by Caelli *et al.* (1985).

It should be stressed that, while the present model uses information from sustained mechanisms only (as was done in the model of Caelli *et al.*, 1985), the present model could be expanded to encompass the additional activity of transient mechanisms. We have not implemented such an expanded model because the present data did not demand it.

EXPERIMENT 2

Goodness of temporal integration in relation to the duration of ISI was examined in Experiment 1, and was explained in terms of a mathematical model of visible persistence. Experiment 2 examined the effect of exposure duration in terms of the same model.

It has been reliably demonstrated that goodness of temporal integration diminishes as stimulus duration is increased (Allport, 1968; Efron, 1970; Bowen *et al.*, 1974; Di Lollo, 1980). From these results it was inferred that the duration of visible persistence varied inversely with the duration of the inducing stimulus. This effect has come to be known as the *inverse duration effect.*

In the present experiment, the duration of the temporally leading stimulus was systematically varied to assess the model's ability to account for the inverse duration effect. In addition, the duration of the temporally trailing stimulus was varied in the same manner as the first. This was done to examine the model's assumption that