DOUBLE-PULSE RESOLUTION IN THE VISUAL FIELD: THE INFLUENCE OF TEMPORAL STIMULUS CHARACTERISTICS

BERNHARD TREUTWEIN and INGO RENTSCHLER
Institut für Medizinische Psychologie, Ludwig-Maximilians-Universität, Goethestrasse 31, 8000 München 2, Germany

(Received 2 May 1991; in revised form 4 February 1992)

Summary—1. It has been suggested that measuring double-pulse resolution in the visual field is more useful than performing flicker perimetry. Yet it is difficult to assess the diagnostic potentials of this technique unless a number of methodological difficulties are overcome.

2. We succeeded to show that double-pulse resolution can be measured efficiently and reliably by varying pulse durations over a wide temporal range, by employing a nine-alternative forced-choice paradigm with nine locations in the visual field, and by employing a maximum likelihood estimate of the threshold parameter. Our results, which have been obtained at the central fovea and at eight locations on the principal meridians with 3.4° eccentricity, reveal three main properties of visual performance.

3. Temporal resolution is worst (i.e. 50-70 ms) at a duration of the leading pulse of 20 ms. It monotonically improves to assume an asymptotic value of about 20 ms beyond pulse durations of say 150 ms. Resolution may also improve if the pulse duration is as brief as 10 ms.

4. The prolongation of the trailing pulse has virtually no effect on double-pulse resolution.

5. Double-pulse resolution in the central fovea is, almost independently of the pulse duration, 10-20 ms better than in the peripheral visual field.

Key words—Double-pulse resolution; maximum likelihood estimate; perimetry; temporal sensitivity; visual field; visual persistence.

INTRODUCTION

It has been claimed that, in case of retrobulbar neuritis, the discrimination of double-pulses of light is a more sensitive test of local damage in the visual field than visual acuity and critical flicker fusion frequency (Galvin et al., 1976a, b, 1977). The same technique holds promise as a diagnostic tool for identifying eyes suspected of having glaucoma (Stelmach et al., 1986). Moreover, Venables (1963) reported different correlations between the level of skin potential and double-pulse resolution in normal controls and schizophrenics although King (1962) found it impossible to distinguish between normal and schizophrenic performance on the basis of double-pulse resolution alone. Given these results, it is surprising that the method of double-pulse resolution has, to our knowledge, not been used in clinical vision research since then. This is possibly a consequence of the apparent contradictions between the results of a number of studies on the foveal discrimination of two successive pulses of light. Another difficulty is that, besides the results of Stelmach et al. (1986), there are no data available which allow a comparison of temporal double-pulse resolution in foveal and eccentric view.

Probably the most important reason for the existence of varying estimates of double-pulse resolution is its dependence on both the absolute and the relative duration of the two pulses. That is, thresholds for the duration of the gap between the two pulses of about 50–70 ms are obtained if pulses of, say, 30 ms or less are delivered to the same location and psycho-physical procedures of the type of limits or constant stimuli are employed (Mahneke, 1958; King, 1962; Venables, 1963). With increasing pulse duration, thresholds decrease to be as low as about 15 ms at pulse durations of 150 ms or more (Mahneke, 1958). In the case of unequal pulse durations, temporal resolution critically depends on whether the duration of the leading or the trailing pulse is varied (Mahneke, 1958). Thus we may note that “simply varying the interpulse interval between two light pulses does not necessarily provide a measure of temporal resolution” (Kietzman and Sutton, 1968, p. 301).
Further problems with measuring double-pulse resolution reside in the dependency of threshold data on psychophysical methodologies. Dunlap (1915) noted that subjects may confuse cues of overall stimulus duration and “doubleness” when the interpulse interval is being varied according to the method of constant stimuli. The use of forced-choice techniques does not prevent such confusion per se (Boynton, 1972) but allows a better understanding of what sort of cues are used provided comparison and target stimuli are appropriately chosen (Kietzman and Sutton, 1968). Generally, the employment of a forced-choice technique leads to a reduction of criterion effects (Blackwell, 1953) and yields considerably lower thresholds (Kietzman, 1967; Lewis, 1967; Kietzman and Sutton, 1968).

As to the influence of pulse energy on temporal resolution, results were conflicting too. Mahneke (1958) explained the improvement of double-pulse resolution with increasing duration of light pulses with the larger amount of energy available, whereas Kietzman (1967) found that energy increments do not change double-pulse resolution. This discrepancy is probably due to the restricted range of energies considered by Kietzman (Lewis, 1967). Another issue is the occurrence of artifacts as a result of Bloch’s law (Kietzman, 1967; see also Brindley, 1970) according to which the product of luminance and duration determines whether a light pulse is seen, and, if it is seen, its brightness.

We think that it is impossible to assess the diagnostic potentials of double-pulse perimetry unless its methodological difficulties are overcome. To achieve this we combined the advantages of several of the earlier studies: we followed Mahneke (1958) by varying the durations of the leading and the trailing pulse in many combinations over a range from 10 to 280 ms. We then used a modified version of the forced-choice paradigm of Stelmach et al. (1986) which allows the quasi-simultaneous measurement of double-pulse resolution at nine locations in the visual field. We also employed an adaptive psychophysical procedure which provides a maximum likelihood estimate of the threshold parameter (Treutwein, 1989, 1991). Our results demonstrate that double-pulse resolution in the visual field can be measured reliably and efficiently.

**METHODS**

**Stimuli**

The stimuli were similar to the type used by Stelmach et al. (1986) who presented their patients with five light patches arranged in the center and at the four corners of a square. The size of the square determined the retinal eccentricity at which the visual field was stimulated. To obtain a more complete set of data at a given eccentricity, we modified this design by using a pattern of nine light patches, one in the center and the other eight on the circumference of a circle. In the present experiments, the diameter of this circle was 6.8 deg of visual angle at a viewing distance of 100 cm, i.e. the eccentricity was kept fixed at 0 and 3.4 deg. The stimulus locations on the circle were its intersections with the vertical, the horizontal, and the two oblique (45 and 135°) meridians. The target location was chosen at random for each stimulus display.

As to the spatial characteristics of the stimuli, each of the nine light patches consisted of a 5 x 5 dot matrix of 0.34 x 0.34 deg angular extension. The substructure of the patches was visible since the dot diameter was about 0.06 deg. Before and after stimulus exposure, a dim fixation pattern of dotted lines was shown on the four principal meridians. These display characteristics are illustrated in Fig. 1(a).

The temporal stimulus characteristics are illustrated in Fig. 1(b-d). The fixation target was switched off 50 ms before stimulus onset and switched on 50 ms after stimulus offset [Fig. 1(b)]. Stimulus onset and offset were identical for both the nontarget [Fig. 1(c)] and the target patches [Fig. 1(d)]. The difference between the target and the nontarget patches was the existence of a transient dark interval in the target patches. At the symmetric stimulus condition, the durations of the leading and the trailing pulses were equal. At the asymmetric conditions, these durations were different.

The synchrony in switching on and off targets and nontargets allowed us to avoid asynchrony cues such as noted by Kietzman and Sutton (1968). Another type of unwanted cues would have been brightness differences between targets and nontargets as predicted by Bloch’s law (e.g. Brindley, 1970) for pulse durations of less than say 100 ms. Such artifacts have been eliminated by rescaling the intensity of the nontargets according to the following rule: let $B = cL$ be the brightness of a light patch with $L$ denoting its luminance, $t$ the corresponding...
Double-pulse resolution in the visual field

pulse duration, and $c$ the constant of proportionality in Bloch's law. In case of the target patch (suffix T), i.e. the double-pulse, this assumes the form

$$B_T = B_1 + B_2 = c(t_1 + t_2)L_T$$

where 1 and 2 index the leading and the trailing pulse. The brightness of the nontarget patch (suffix N) is

$$B_N = c(t_1 + t_g + t_2)L_N$$

since the duration of the nontarget patches is increased with increasing gap duration ($t_g$) to ensure equal overall duration. Equating $B_T$ and $B_N$ yields the scaling rule

$$L_N = L_T \frac{(t_1 + t_2)}{(t_1 + t_g + t_2)}$$

for $t_1 + t_2 < 100$ ms.

With this rule being applied to rescale the luminance of the nontarget patches, there were no apparent brightness differences perceived even for pulse durations of <1 ms. This was established in a preliminary experiment of examining the luminance correction according to Bloch's law.

Luminance rescaling of nontargets was not required once the on-time of the double-pulse was prolonged beyond the critical duration specified by Bloch's law. This was the case with the asymmetric double-pulse configurations, where either the leading pulse or the trailing pulse had a duration of 280 ms, i.e. was longer than summation time.

The stimuli were generated by means of high-speed plotting of points on a Hewlett-Packard CRT (HP 1310, P4 phosphor). This was achieved by employing an electronic interface, provided by G. Finlay (The University of Alberta, Edmonton, Canada), which was under the control of a DEC LSI-11 minicomputer (for technical details see Finlay, 1985). The plotting speed allowed a frame rate of 3.4 kHz.

Procedure

The testing method was designed to measure the duration of the interpulse interval required for the subject to identify the target. This interval, referred to here as "critical gap" duration, is the duration giving a 55.55% probability (i.e. halfway between the guessing probabilities of 11.11 and 100% correct) of correct identification in a nine-alternative forced-choice (9-AFC) paradigm. In our case, the estimation of the nine thresholds at the nine different locations were interleaved in one session.

The general problem with threshold estimation is to find a method which does not waste trials far above and below threshold but concentrates presentations near its (unknown) value. This can be done by applying methodologies of sequential statistics. The method used here is a modification of the maximum likelihood (ML), or Bayesian, estimation. It estimates the threshold and the corresponding confidence interval from all data collected during an experiment up to the current trial.
The next stimulus is then chosen to correspond to the current best threshold estimate. This procedure is repeated until the confidence interval for the threshold estimate is smaller than a preset value for all nine locations of the 9-AFC (for details see Appendix). For the sake of avoiding useless testing, there is no fixed number of trails but the procedure terminates once a certain confidence in the threshold estimate is reached.

In the experimental trials, the subject had to fixate the center of the fixation target. He/she had then to start the stimulus display by pressing any key on the computer keyboard. Depending on where he/she believed that the target stimulus was being displayed, the subject was to press one of the nine number keys. This task was facilitated by assigning the clockwise arranged number keys 8, 9, 6, 3, 2, 1, 4 and 7 to the eight clockwise ordered positions on the circumference of the circle, and the central key 5 to the fixation point. Depending on the correctness of the subject’s response, the computer raised or lowered the gap duration by setting it to the current maximum likelihood estimate of the parameter $\alpha$. The trials continued until the critical gap duration had been estimated to the specified degree of accuracy. A typical session for the determination of a complete set of nine thresholds at given values of pulse durations took about 20–30 min and totalled in about 30–40 presentations per stimulus location.

The stimuli were presented on the dark CRT screen in an otherwise moderately illuminated room. The luminance of the target patches was kept fixed at 21.5 cd/m$^2$. The viewing distance was kept constant at 100 cm in all experiments. Subjects saw the stimulus display monocularly. No additional fixation control was employed because of the random occurrence of the target stimuli at nine symmetrically distributed positions.
Subjects

Two male (HS and KZ) and one female (KL) medical students served as paid subjects in the main experiments. Their age ranged from 19 to 28 yr and they had no known visual defects.

RESULTS

We used the same three types of pulse combinations as have been considered by Mahneke (1958): first, the symmetric condition where both the leading and the trailing pulse were equally prolonged from 10 to 280 ms. Second, an asymmetric condition where the leading pulse was prolonged from 10 to 280 ms while the duration of the trailing pulse was kept constant at 280 ms. Third, an asymmetric condition where the trailing pulse was prolonged from 40 to 280 ms with a fixed duration of the leading pulse (280 ms).

The resulting critical gap durations are plotted in Figs 2–5 as functions of pulse duration. Each of the figures consists of an array of nine plots which reflect the locations of the stimulus patches in the visual field. The data obtained at the fixation point are in the “central” position. “Top” and “bottom” refer to the locations on the vertical meridian, “left” and “right” to the locations on the horizontal meridian, “upper right” and “lower left” to locations on the 45° oblique meridian, and “upper left” and “lower right” to locations on the 135° oblique meridian. The dotted lines in the data plots represent, at a given experimental condition, the grand mean over all nine locations in the visual field and all observers.

1. Leading and trailing pulse equally prolonged

The variability of double-pulse resolution among three healthy, young subjects is illustrated in Fig. 2. The data have been collapsed over the monocular viewing conditions of using the right and the left eye respectively. The
subjects' individual data are quite similar but some discriminating features can be discerned. At pulse durations longer than say 20 ms, the data of K.L. [Fig. 2(b)] show the least deviation from the grand mean. Here, the thresholds of both K.Z. [Fig. 2(a)] and H.S. [Fig. 2(c)] are about 10–20 ms higher. At the shortest pulse durations (i.e. at 10 and 20 ms), the situation is converse in that there is a threshold reduction for K.Z. and H.S., whereas K.L. shows an increase in threshold. Nevertheless, the subjects' individual data are similar enough, and no distortion of the data is caused by averaging them (Fig. 3).

The mean data showed in Fig. 3 have the following general properties: double-pulse resolution is best at the fixation point, where it decreases from 50 ms at 10 ms pulse duration to 10 ms at 280 ms pulse duration. In the peripheral visual field, the temporal resolution is worse.

2. Leading pulse prolonged, while keeping trailing pulse constant

At this condition, the subjects' individual data show the same degree of variability as before. Thus we abstain from showing individual data and present in Fig. 4 the results averaged over subjects and eyes. The main difference between these data (asymmetric condition) and those shown in Fig. 3 (symmetric condition) is a more pronounced increase at peripheral viewing conditions of temporal resolution at the shortest pulse duration of 10 ms.
3. Trailing pulse prolonged, while keeping leading pulse constant

If the leading pulse is kept fixed at 280 ms, there is very little variation of temporal resolution with the prolongation of the trailing pulse (Fig. 5). From this we conclude that there is no effect of backward masking exerted by the trailing pulse on the visibility of the target gap. Otherwise, the data characteristics observed in Figs 3 and 4 can be discerned here too, i.e. the temporal resolution at fixation is superior to the peripheral viewing conditions and there exists virtually no meridional anisotropy of visual performance.

To facilitate the comparison of data, Fig. 6(a) shows the grand means for the symmetric and the two asymmetric conditions. As noted before, there is virtually no difference between the symmetric and the leading pulse conditions besides the sharp increase in double-pulse resolution at the latter condition for 10 ms pulse duration. By contrast, for the trailing pulse condition, resolution is fairly constant over the entire range of pulse durations. Figure 6(b) shows, for the leading pulse condition, the means over observers and eyes for the foveal and the peripheral viewing conditions. The data for the latter conditions were collapsed over positions.

For further analysis we replotted our data in the form of stimulus onset asynchrony SOA (i.e. duration of the leading pulse plus gap duration) against duration of the leading pulse (Fig. 7). It is clear that the SOA hypothesis, i.e. the existence of a fixed value of SOA irrespective of the duration of the leading pulse, would only hold for the very limited range of 20–40 ms. At larger values of the leading pulse duration, a linear
Means of three subjects (KL, KZ, HS)

Fig. 4. Double-pulse resolution as a function of pulse duration. Leading pulse condition; data collapsed over observers, otherwise as in Fig. 2.

relationship between the latter and SOA is a much better approximation.

DISCUSSION

We demonstrated how double-pulse resolution in the visual field can be measured efficiently and reliably and our results reveal three main properties of visual performance:

1. Except for a possible reduction between 10 and 20 ms, temporal resolution improves monotonically with increasing duration of the leading pulse to assume an asymptotic value of about 20 ms beyond say 150 ms.

2. The prolongation of the trailing pulse has virtually no effect on double-pulse resolution.

3. Double-pulse resolution in the central fovea is, almost independently of the pulse duration, 10–20 ms better than in the peripheral visual field.

Clinical assessment of visual function

For the clinical use of double-pulse measures, these results have the following implications. First, there is no point in prolonging pulse durations beyond 150 ms since this leaves visual performance unaltered. Second, it is also not interesting to vary the duration of the trailing pulse because this yields essentially the same constant value of temporal resolution for a range of pulse durations between 40 and 280 ms. Third, if the duration of the leading pulse condition is varied, double-pulse resolution monotonically increases once the pulse duration is increased from 20 to 150 ms. The dependency
is the same no matter whether the trailing pulse is varied in the same way as the leading pulse (symmetric condition) or is kept constant at 280 ms (leading pulse condition). The only difference between the latter two stimulus conditions is the existence of an additional effect of threshold reduction for the leading pulse condition at 10 ms pulse duration. We conclude that all important characteristics of temporal visual processing can be captured by varying the duration of the leading pulse between 10 and 150 ms and keeping that of the trailing pulse fixed at 150 ms (or longer).

**Linear systems approach**

How can these findings be accommodated by existing theoretical frameworks of temporal visual sensitivities? Linear systems theory (LST, see Gaskill, 1975; also Watson, 1986) assumes that the overall response to a linear combination of stimuli is the linear combination of the responses to the individual stimuli (principle of linear superposition; Gaskill, 1975, Fig. 5-2). In the leading pulse condition of the present experiments, the nontargets consisted of an overall, or fundamental, pulse of constant amplitude. The target stimulus can be conceived as consisting of a pulse of negative polarity superimposed onto the same fundamental pulse. This additional pulse provided the discriminative feature of the target. For sake of simplicity, we may assume that the response to the fundamental pulse is a function which raises for some time after stimulus onset, then reaches a peak, and decays with ongoing pulse duration (e.g. Gaskill, 1975, Fig. 5-2). Assuming that linearity holds, the response to the negative pulse would have a similar shape but opposite polarity. The superposition of the two responses yielded, for relatively short gap durations, a notch in

![Graph showing double-pulse resolution as a function of pulse duration.](image)

*Fig. 5.* Double-pulse resolution as a function of pulse duration. Trailing pulse condition; data collapsed over observers, otherwise as in Fig. 2.*
the response function to the fundamental pulse with the lag of the notch depending on the duration of the leading pulse. It is then plausible that, in the sense of Weber's law, the detection of the notch is most difficult if it coincides with the peak and is easier at lower values of the fundamental response.

Thus it seems that, at the leading pulse condition, we probed the visual response to the fundamental pulse with a negative pulse of threshold energy (being proportional to duration). There are interesting analogies between our experiments and a study of transients by Crawford (1947). Crawford's task required the detection of a liminal brightness increment around the onset of a large brightness transient, whereas our experiments may be thought of as tasks of detecting a decrement within a pulse. In both the Crawford transient and the double-pulse method, thresholds are high if the transient (increment or decrement) occurs near the beginning of the pulse, and decreases near the end of the pulse (as shown in Figs 4 and 5).

The idea of probing temporal responses is also the basis of other attempts of measuring the spatio-temporal visual transfer characteristics. Roufs and Blommaert (1975, 1981) determined the influence of subthreshold inducing patterns on the visibility of target patterns. Target and inducing stimuli had identical spatial structure and were displayed at the same location but had variable onset asynchronies. Lupp (1981) and Elsner and Hauske (1984) measured transient visual sensitivities via reaction times to the onset of target sinewave gratings. As did Roufs and Blommaert, they measured reaction times as a function of the onset asynchrony of a superimposed subthreshold grating of identical spatial structure. Both types of studies revealed temporal band-pass characteristics for large stimulus size or low spatial frequency, and low-pass characteristics for small size or high spatial frequency.

It should be clear, however, that we cannot treat double-pulse resolution in terms of LST unless a number of simplifying assumptions are made. Some of them, like that of negligible effects of adaptation, can be justified with respect to experimental conditions, others cannot. Particularly important examples of the latter type are nonlinearities due to information processing in parallel pathways (see Watson, 1986). Here the dichotomies of sustained and transient channels (e.g. Enroll-Cugell and
Robson, 1966; Cleland et al., 1971; Ikeda and Wright, 1974; Kulikowski and Tolhurst, 1973; Breitmeyer and Julesz, 1974) and of on- and off-mechanisms (Kuffler, 1953; Baumgartner, 1961; Schiller, 1982) are of interest. If such caveats of applying LST to the analysis of temporal visual sensitivities did not exist, one might argue that there is no point in clinically measuring double-pulse resolution since it is directly related to the temporal modulation transfer function of the eye (De Lange curve; see Kelly, 1972). This, however, is certainly not the case. Thus it is conceivable that measures of double-pulse resolution reflect transient visual response characteristics, whereas De Lange curves and critical flicker frequency (CFF) exclusively capture steady-state characteristics.

Visual persistence

We shall then turn to considering another approach to spatio-temporal visual sensitivities, namely the study of visual persistence, where the temporal integration (i.e. the complementary function of visual resolution) of figural components in memory is examined. Here, the most remarkable result is the existence of an inverse relationship between the duration of the temporally leading stimulus component and the duration of visual persistence (Di Lollo, 1977, 1980). This finding is not accommodated by models of iconic memory which assume that the contents of this sensory store begin to decay when the inducing stimulus is terminated (interstimulus interval or ISI hypothesis; e.g. Neisser, 1967). Indeed, temporal integration of successive visual displays is a function of stimulus onset asynchrony (SOA; Di Lollo, 1980).

In the context of double-pulse resolution, Boynton (1972) proposed that SOA is a better parameter than ISI for characterizing visual performance. This raises the question of whether double-pulse resolution and visual persistence are closely related phenomena. However, as has been shown in Fig. 7, our data do not support Boynton's claim. Thus we have no reason to assume that visual persistence is a possible explanation for our double-pulse data. A possible reason for this failure is that the stimuli of Di Lollo's experiments were arrays of widely separated dots. In our experiments, the figural components to be integrated (or resolved) were spatially superimposed and contained low spatial frequency information. Under such conditions, the spatio-temporal transfer characteristics of the prestriate visual pathways may cause inhibitory effects between stimulus components (Brettel, 1991). The question of whether these differences in spatial stimulus characteristics are sufficient to quantitatively explain the failure of the SOA hypothesis is currently under investigation.

Visual field differences

The distribution of temporal resolution in the visual field depends critically on all stimulus parameters such as temporal configuration, size, color, luminance and eccentricity. Brooke (1951) measured the CFF with a 2° stimulus at different eccentricities and mean luminances. He found to be for high luminances (3000 cd/m²) an exponential decrease of CFF from 50 Hz in the center to around 20 Hz in the periphery. At the much lower luminance of 0.3 cd/m² the distribution was found to be fairly constant. Below the latter luminance level the peripheral visual field (more than 10° eccentricity) exhibited a clear advantage (13 Hz as opposed to a central 7 Hz). Hartmann et al. (1979) found at specific luminance conditions a non-monotonic distribution of the CFF with a maximum between 10 and 25° eccentricity. The latter observation seems to be supportive of the widely held opinion that the periphery is more sensitive to flicker. Yet we know of only one other study the results of which are consistent with this view: Welde and Cream (1972, as quoted by Sekular et al., 1982) reported visible flicker when presenting a 30° dia stimulus on a standard NTSC-TV display (60 Hz interlaced) with 30 or 60° eccentricity at 10.3 and 30.8 cd/m² mean luminance.

Thus it appears that our findings are in accordance with most data available in the literature when we report a clear central advantage in double-pulse resolution. Furthermore, there is neither a hint at meridional anisotropy (i.e. double-pulse resolution only depends on the amount of eccentricity and not the orientation) nor a hint at differences between the upper and lower hemiretina. It is also of interest that there were no left/right differences in the visual field. The latter observation is consistent with the lack of hemispheric asymmetry in duration of visual persistence (Di Lollo, 1981; Peterzell et al., 1989).

As to the foveal advantage in double-pulse resolution, one might suspect that this advan-
tage is caused by a bias of the subjects to guess an undetected stimulus as foveal. We examined this possibility and found that some of the subjects had indeed personal guessing preferences for specific locations in the visual field. Yet there was no systematic preference for the central location. Moreover, the guessing probability for the preferred stimulus never exceeded 20% and had virtually no influence on the value of the estimated threshold. Thus we are confident that the observed central advantage of double-pulse resolution reflects a genuine property of temporal visual information processing.

Acknowledgements—We thank Hans Brettel, Vince Di Lollo and Hans Strasburger for helpful discussions and comments on the manuscript. An unknown referee drew out attention to the “Crawford transient”. Gary Finlay and Gary Burchett (University of Alberta, Edmonton, Alberta, Canada) kindly provided the point plot buffer. This study was supported by grants InSan I 0784-V-6385/6386, the French National Research Foundation and the Deutsche Forschungsgemeinschaft.

REFERENCES


Welde W. J. and Cream B. W. (1972) Variables influencing the perception of flicker in wide angle CRT displays. AD-766–443 (1472, Mikrofilm).

### APPENDIX

Almost all adaptive psychophysical procedures are based on the application of methods of sequential statistics which were introduced by Wald (1947) for the application in industrial production control. Used in psychophysics, these methods aim at reducing the overhead of measuring points on the psychometric function by concentrating the stimulus presentations near the (generally unknown) threshold. In our study, a computer controlled adaptive procedure, based on Bayes’ estimation of the threshold parameter \( \alpha \), was used to find the critical gap duration independently for each of the nine different stimulus locations. The method we used (YAAP: Yet Another Adaptive Procedure), is described in more detail by Treutwein (1991); the principles of such procedures have been discussed by Pentland (1980), Lieberman and Pentland (1982), Watson and Pelli (1983), and Harvey (1985). The main difference between these three methods and YAAP is the usage of the “direct” likelihood (as opposed to the logarithmic likelihood) and a correct dynamic stopping criterion. The latter allows for stopping a session after the least number of necessary trials given certain confidence requirements for the threshold. YAAP was implemented in Modula-2 by the first author (B.T.) and the source code is available from him.

YAAP varies the gap duration to find the target probability of correct responses. The procedure terminates when the 95% confidence interval for the estimated \( \alpha \) is smaller than ±15 ms. This probability level corresponds to the point of inflection of the logistic function which we used to model the psychometric function:

\[
\psi(x; \alpha) = \frac{1}{1 + \exp[(x - \alpha) / \beta]}. \tag{A1}
\]

Here \( \psi(x) \) is the probability of making a correct response to a stimulus having the gap duration \( x \), \( \alpha \) is the threshold duration [i.e. the duration that yields to a probability for a correct response of \((1 - \delta) / 2 = 0.55\)], \( \delta \) is the probability of a correct response by chance alone (0.11 for our 9-AFC paradigm), and \( \beta \) is the slope of the psychometric function (a value of 20 was used).

During the course of the experiment (trials \( k = 1, 2, \ldots, n \) ) information about the value of the threshold \( \alpha \) is gathered. The possible threshold values \( \alpha \) are the same as the experimentally available stimulus values \( x_i \), \((i = 1, \ldots, I)\). The current best estimate of threshold \( \alpha \) after trial \( k \) is given by the maximum of the normalized joint-probability density function \( P(\alpha) \). \( P(\alpha) \) can be considered as the probability (i.e. normalized likelihood) for the threshold a being equal to gap duration \( x_i \)

\[
P(\alpha) = \frac{1}{c} \sum_{i=1}^{I} \psi(x; \alpha) \cdot P_i^{(1)}(\alpha) \cdot \ldots \cdot P_i^{(n)}(\alpha)
\]

where \( c \) is a normalization factor which, in our case, is \( \sum_i P(\alpha) \), and each \( P_i^{(k)}(\alpha), (k = 1, \ldots, n) \), is the probability for the subject’s actual response on trial \( k \). These probabilities \( P_i^{(k)}(\alpha) \) are given by the psychometric function \( \psi(x; \alpha) \) together with the response of the subject after the \( k \) th presentation of a stimulus with gap duration \( x \) by

\[
p_i^{(k)}(\alpha) = \begin{cases} 
\psi(x; \alpha) & \text{for a correct response at trial } k \\
1 - \psi(x; \alpha) & \text{for an incorrect response at trial } k 
\end{cases}
\]

A crucial point for applying maximum likelihood methods is that a value of the slope for the generally unknown psychometric function has to be assumed. Especially for ML methods which use a dynamic stopping criterion (i.e. methods with a variable number of trials), the assumed value of the slope parameter influences the confidence interval of the ML estimate and therefore also the number of trials needed to terminate a session (Madigan and Williams, 1987). Overestimation of the slope parameter \( \beta \) (with \( \beta \) being inversely related to the slope) leads to a psychometric model function which is too shallow. \( P(\alpha) \) is then too widely spread which in turn leads to a slower convergence of the procedure and to larger confidence intervals for a given response sequence. Underestimation has the opposite effect (a steep psychometric model
function, narrow $P(a)$, narrow confidence intervals and a small number of trials to terminate the session).

In our experiments, we have chosen the large value of 20 for the slope parameter. On the one hand, by fitting logistic functions to the cumulated response data of repeated measurements of the same stimulus configurations, we have previously shown that the true slope is much steeper than our assumed slope (Treutwein, 1991). Thus we are confident, that our ±15 ms confidence intervals are worst case estimates with the actual confidence intervals being much narrower. On the other hand, we do not want to assume a steeper slope in the ML method, since a too steep slope leads to an increased susceptibility for subject lapses and to a systematic bias in the threshold estimates (O'Regan and Humbert, 1989).