New look at Bloch's law for contrast

Andrei Gorea* and Christopher W. Tyler

Smith-Kettlewell Institute of Visual Sciences, 2232 Webster Street, San Francisco, California 94115

Received December 8, 1983; accepted August 10, 1985

It has been commonly reported that the temporal integration of grating contrast proceeds more slowly as spatial frequency is increased. Such results have been based on the critical duration for sensitivity to contrast pulses varying in duration, but the analyses have not assumed full integration at short durations and have neglected the effects of probability summation over time. To take such effects into account, we discuss a class of analytical models based on nonlinear temporal integration. On the assumption that the temporal impulse response of the visual system determines contrast integration over time, we develop both a high-threshold model and a signal-detection approach involving multiple and independent nonlinear signal detectors with a time-limited integration span. The redefined critical durations predicted by the models and verified by the data are about 35 msec and vary by no more than 10 msec across spatial frequency. This variation is entirely attributable to a change in the strength of inhibition with spatial frequency, and the analysis implies that the excitatory component is constant at all spatial frequencies, contrary to previous accounts.

1. INTRODUCTION

Bloch's law states that for sufficiently short stimulus durations, detection threshold decreases inversely with the duration of the stimulus.[1] For longer durations, empirical thresholds in all known stimulus domains progressively depart from this law and have typically been assumed to approach a constant value independent of the stimulus duration. The threshold–duration function has thus often been characterized by two asymptotic slopes of −1 and 0 on double logarithmic coordinates whose intercept defines the *critical duration* of the system.

Several investigators have determined threshold–duration functions for detection of contrast increments in sinusoidal-grating stimuli.[2–5] These studies all agree in fitting a zero-slope segment to the datum points obtained with very long durations, although segments of arbitrary slopes have often been fitted to the transition region of the threshold–duration function.[5] However, the assumption that the slope reaches zero ignores the effects of probability summation of decision events over time.[6] Under certain conditions to be analyzed in this paper, such probability summation implies that the threshold should continue to decrease as duration increases, requiring a reevaluation of the definition of critical duration.

The studies cited did not use sufficiently short durations to obtain the full reciprocity between time and contrast required by Bloch's law and instead fitted their data in this region with a straight segment of arbitrary slope in double logarithmic coordinates. Critical duration was then estimated from the intercept of the first two fitted segments. Under this loose definition, critical durations for contrast integration increased by as much as half of a log unit with spatial frequency. However, the lack of theoretical assumptions constraining the fitting procedures leads to difficulty in comparing critical durations obtained in different studies and across spatial frequency. This calls into question the implied temporal characteristics of sustained (*sluggish*) and transient (*fast*) mechanisms that might be processing such stimuli.[3,5,7]

We present here a quantitative model to account for Bloch's law for contrast and propose a new interpretation of the critical duration. This model derives the form of the threshold–duration function by nonlinear integration of theoretical neural responses to stimulus events, with an additional stage of probability summation among temporally separated integrators. As opposed to previous concepts of probability summation over time[6] and over space,[8,9] which used Quick's[10] high-threshold formulation, our development of the probability-summation process is based on a signal-detection approach.

The temporal modulation transfer function changes its bandpass characteristics with spatial frequency.[11] The neural impulse response as derived from it by inverse Fourier transformation[12] will therefore be altered correspondingly. The variations in the temporal integration function with spatial frequency may be due solely to the impulse-response characteristics, without implying the presence of separate transient and sustained mechanisms as previously suggested.[3] We therefore test whether models that do not include such a dichotomy will correctly predict threshold–duration functions at low and high spatial frequencies.

2. THEORETICAL ANALYSIS

A. Empirical Constraints

The previous data quoted require a model whose threshold–duration predictions exhibit the following characteristics: (1) They must level off for durations longer than about 80 msec at low spatial frequencies but (2) still be time dependent for high spatial frequencies; (3) they must be nonmonotonic, showing a threshold increase somewhere between 500 and 800 msec for low spatial frequencies and (4) exhibit a decelerating decrease in threshold beyond this limit for high spatial frequencies. All these constraints are empirically based, but only the first two have been unanimously taken into consideration in previous studies.

The late dip in sensitivity of item (3) is not always observed but is very evident in data from some laboratories.

0740-3232/86/010052-10$02.00 © 1986 Optical Society of America

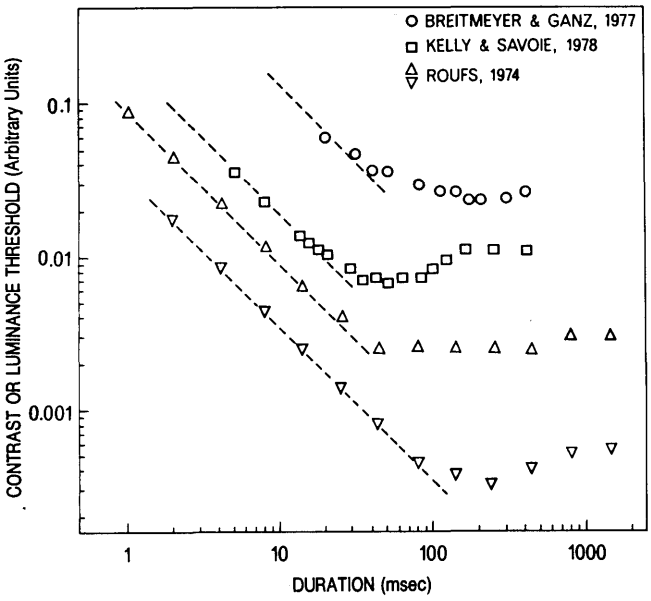


Fig. 1. Data from previous experiments showing sensitivity dips. Typical experimental variables controlling the strength of the dip are the size, the spatial structure, and the average luminance of the inspection field (see text).

To illustrate this point, Fig. 1 shows four sets of data reproduced from previous work: Breitmeyer and Ganz,[3] 4 × 6° field; Kelly and Savoie,[13] 8° field; Roufs,[14] 1200 and 1 Td, 1° field. Note that the sensitivity dip appears to depend on both the spatial characteristics and the average luminance of the inspection field. Whereas the former may affect the slope of the psychometric function,[15] the latter was shown to determine the shape of the impulse response.[16] We therefore require a model that takes these factors into account.

B. The Linear Stage

To the extent that the visual system behaves as a unique temporal filter at each given spatial frequency,[17] its temporal impulse response $h(\tau)$ is the Fourier transform of the modulation transfer function (MTF) measured at the given spatial frequency. Depending on the characteristics of the MTF,[11] the impulse response will be either monophasic or biphasic. The energy under its negative lobe, usually associated with the amount of lateral inhibition in the visual system,[12] depends directly on the spatial structure of the stimulus. It approaches zero when the spatial frequency is higher than about 3–4 cycles/degree (c/deg). Since our intention is to illustrate two extreme cases of Bloch's law for contrast, namely, for low and high spatial frequencies, biphasic (full inhibition) and monophasic (no inhibition) impulse responses were used to fit our data (Fig. 2). They were derived from Watson's[18] theoretical impulse response which appeared to be a sufficiently accurate estimate of the neural impulse response.

The first step of the model is to assume that the visual filter responding to a pulse stimulus is linear. As a consequence, its pulse response $H(t)$ will be given by the convolution of its impulse response $h(\tau)$ with the temporally modulated stimulus $f(t)$:

$$H(t) = \int_0^\infty f(t-\tau) \times h(\tau)\mathrm{d}\tau. \qquad (1)$$

The lower 0 limit of the integral is required for the system to be causal. Figure 3 displays numerical estimations of $H(t)$ obtained with monophasic and biphasic impulse responses for short and long rectangular pulses (43- and 193-msec duration, respectively).

Whereas for monophasic impulse responses the absolute energy under $H(t)$ increases indefinitely with the duration of the pulse, for biphasic impulse responses it eventually stops increasing. This occurs at the duration where the two lobes of the pulse response begin to separate, as determined by the overall duration of the impulse response.

C. High-Threshold Models

We first consider four classes of model based on the concept of probability summation over time derived from the high-threshold assumption of the detection process.[10] In this subsection we will show that each class may be rejected by its failure to predict one or more of the empirical constraints in Subsection 2.A.

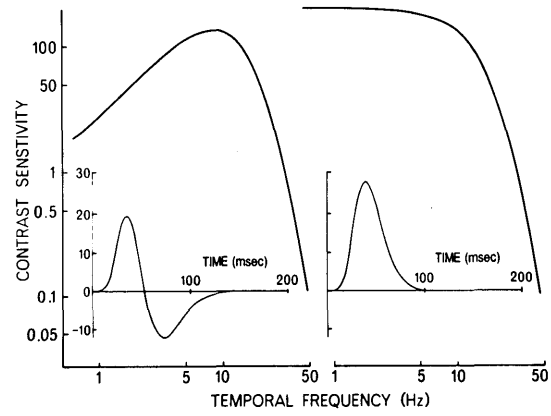


Fig. 2. Typical temporal MTF's obtained with low (left) and high (right) spatial frequencies. The corresponding impulse responses (insets) are from Watson.[6]

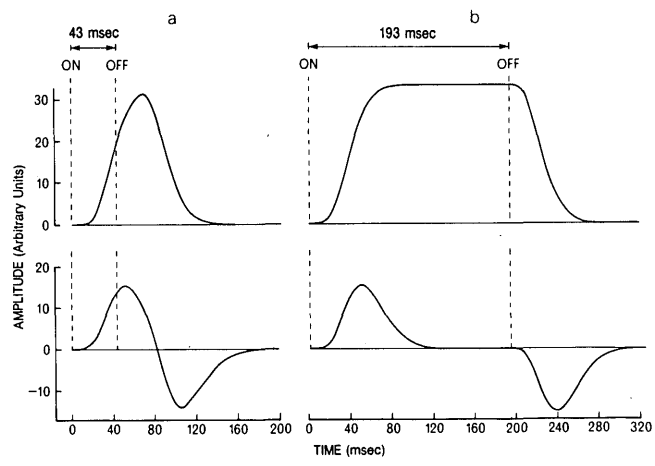


Fig. 3. Pulse responses obtained by convolution of the monophasic (upper panels) and the biphasic (lower panels) impulse responses of Fig. 2 with a, short and b, long stimulus pulses [Eq. (1)]. Note that for short pulses both responses are similar in form to the respective impulse response but are more extended in time. For long pulses, the area under the monophasic pulse response progressively increases, while it reaches a constant value for the biphasic impulse response. In this latter case, only the separation between the two lobes is time dependent.

1. *Indefinite-Integration Model*

The most recent model that is applicable to the temporal integration function is that of Watson,[6] in which the detection process can be described as a power integration of the pulse response, with an exponent being determined by the probabilistic nature of the integration process. At threshold, the response of the system is then given by

$$1 = \int_0^{t_D} |H(t)|^\beta \mathrm{d}t. \tag{2}$$

In this equation the absolute expression can be interpreted as the internal response of time-independent mechanisms whose outputs are summated probabilistically according to Quick's[10] formulation, where the β exponent is the measured slope of the psychometric function. According to Watson, the integral in this expression is infinite in both directions of time. This is trivially implausible for the future direction, since it implies that the observer waits forever before making a response. We therefore regard the upper limit of this indefinite integral as corresponding to t_D, the time of the decision, where the observer is instructed to wait until all response events appear to be over before making a decision. Therefore t_D is equal to the duration of the stimulus plus a constant sufficiently long to include most of the energy of the decaying pulse response. Integration will thus continue throughout the duration of the internal response, in contrast with other models[19–21] that assume a short, finite integration limit. Moreover, it is reasonable to assume that the integration starts at the beginning of a given trial ($t = 0$) since it will otherwise include internal responses to previous trials.

Given the presence of the absolute operator in Eq. (2), the response energy to be integrated will always be positive. Thus increasing the duration of the stimulus will never decrease the value of the indefinite integral. In the case of impulse responses with an imbalance between the positive and the negative lobes (high spatial frequencies—monophasic impulse responses), the indefinite integral will continue to increase with increasing stimulus duration. As discussed below, the rate of this increase at long durations is incompatible with data obtained with high spatial frequencies. For impulse responses that are perfectly balanced (low spatial frequencies—biphasic impulse responses), the indefinite integral will reach a constant value at the point where the stimulus is longer than the impulse-response duration. Therefore the indefinite-integral model cannot account for the empirical requirement of a reduction in sensitivity at longer durations (Fig. 1).

2. *Limited-Integration Model*

The simplest way to obtain a dip in sensitivity (rise in threshold) at longer durations is, as suggested above, to limit the integral in Eq. (2) to the appropriate time T:

$$1 = \int_0^{T} |H(t)|^\beta \mathrm{d}t. \tag{3}$$

Consider first the case of the biphasic impulse response of Fig. 3. If T is more than twice as long as the impulse response, then the integral will reach its full value when the stimulus has the same duration as the impulse response. As the stimulus duration approaches T, the second lobe of the pulse response (Fig. 3b, lower panel) will progressively tend to fall outside the integration limit. Eventually, the only portion of the pulse response, $H(t)$, within the integration limit will be its positive lobe, which contains half of the total energy under $H(t)$. The value of the nonlinear integral at long durations will therefore drop by a factor of $2^{1/\beta}$ compared with that at the optimum duration. Note that the strength of the dip depends on the power of the nonlinearity β so that it may be undetectable if β is large.

Whereas the limited integral can account for the empirical requirement of a dip for biphasic impulse responses, it does not fare so well in the case of monophasic impulse responses. Although the pulse response for this case increases indefinitely (Fig. 3b, upper panel), for the limited-integration window the value of the integral will reach a plateau when stimulus duration exceeds the integration limit. Values of the integration limit of 200 and 500 msec were used by previous authors,[19–21] while the data of Fig. 1 require values betwen 200 and 800 msec. However, previous data[2,5] obtained with high spatial frequencies (i.e., for a monophasic impulse response) can show improvement in sensitivity up to 2 sec. Moreover, the theoretical interpretation of the power nonlinearity as reflecting probability summation requires that this improvement continue indefinitely. We can therefore reject the limited-integration model on both empirical and theoretical grounds.

3. *Double-Integration Model*

An alternative approach that retains the sensitivity dip for biphasic impulse responses while allowing integration to continue for monophasic impulse responses is to include two levels of integration in the model:

$$1 = \int_0^{t_D} \int_t^{t+T} |H(t - t')|^\beta \mathrm{d}t'\mathrm{d}t, \qquad t + T \le t_D. \tag{4}$$

In this equation the first integration level is limited and corresponds to a discrete sensitivity window of a single detection mechanism. This formulation assumes the existence of an indefinite number of such time-limited, independent detectors each of which starts to integrate at an arbitrary time with respect to the stimulation period. The visual system is then assumed to summate over the responses of all these mechanisms up to the time of the decision (t_D) in the outer integral. The idea that each detection mechanism may show a variable responsiveness over time is physiologically plausible, given the sensitivity fluctuations observed in single-unit recordings (e.g., Tolhurst *et al.*[22]). The time-limited integral of equal length for all mechanisms is a first approximation to the idea of these continuous fluctuations of varying length, but it will produce a similar overall output when integration occurs over a large number of mechanisms.

The limited inner integral will produce a sensitivity dip for biphasic impulse responses when the stimulus duration exceeds the limit of the integral, because each mechanism will then be able to integrate only the first or the second lobe of the pulse response, which are now separated by more than the integration length. Intermediate sensitivity windows will receive less than the full energy from either lobe and will contribute to the overall integral to a smaller extent.

The problem with the double-integration model lies in its behavior for high spatial frequencies at durations longer than the integration window. For these long durations, the model predicts (as is also the case with Watson's[6] formula-

tion) a rate of decrease in threshold equal to $-1/\beta$ (see below). Although both the precise rate of threshold decrease and the form of the psychometric function may be difficult to measure experimentally, the data to be presented below as well as previous data[5] indicate that the decrease may be substantially shallower than predicted. Extensive evidence obtained in the spatial domain[23] actually suggests that the threshold-decrease rate with stimulus spatial extent is well described by a $-1/2\beta$ slope.

4. *Diode Models*

The family of models just described used rectification of the pulse response before its nonlinear processing. However, it is also physiologically plausible that positive and negative internal responses are detected through independent (on and off) channels[14,24] whose output is subsequently added probabilistically. All the preceding models can then be rewritten in this diode form.

Nevertheless, to the extent that both on and off channels are equally sensitive, none of such diode-type models would account for the empirical requirement of a sensitivity dip observed with low spatial frequencies. Indeed, whether the sign-specific integrators are time limited or not, they will exclusively include only the appropriate lobe (positive or negative) of the pulse response. Therefore the increasing separation of the two lobes with increasing duration will not produce the output change required, as discussed above, to simulate the empirical sensitivity dip illustrated in Fig. 1.

On the other hand, if the two sign-specific integrators have sufficiently unequal sensitivities, the diode solution can produce an appropriate sensitivity dip.[13] However, the required asymmetry is a factor of 2.7, which is far beyond the asymmetry of a factor of only 1.1 measured directly in increment–decrement-threshold studies.[25] Given this small asymmetry, none of the diode models can produce a sensitivity dip of the appropriate magnitude.

We conclude that none of the models discussed has the appropriate form to account for all the empirical requirements specified initially.

D. The Signal Detection Approach

In all the above models the β power can be looked on either as reflecting a probability-summation parameter within the high-threshold formulation[10] or as a hard-wired nonlinearity followed by nonprobabilistic linear summation.[20] In either case, as discussed, they do not fit all the empirical requirements specified initially. Furthermore, the high-threshold formulation has been acknowledged as physiologically implausible by Quick himself and rejected on different grounds since.[15] For the single-integrator models, which imply the existence of a single processing channel, the interpretation of β as a hard-wired nonlinearity is also inconsistent with our knowledge of the visual system as made up of multiple parallel channels fluctuating in sensitivity over time.

This leaves the double-integrator model of Eq. (4) that allows for multiple channels whose responses are of the form of the inner integral but then assumes that their outputs are integrated linearly over time. Such integration is consistent with signal-detection theory[26] only under the assumption that the noise in all channels is correlated over time. Given the activity of peripheral and cortical visual neurons, it is unlikely that the neural noise maintains a significant degree of correlation over periods of the order of 1 sec. We therefore assume that the noise is independent in each of the channels whose responses R_i are given by the inner integral of Eq. (4). The overall response R of the system is then obtained as described by Eq. (5), which assumes ideal signal detectors with uncorrelated noise[26]:

$$R = \left(\sum R_i^2\right)^{1/2}. \tag{5}$$

At threshold, where $R = 1$, this type of summation can be incorporated into Eq. 4, which becomes

$$1 = \int_0^{t_D}\left[\int_t^{t+T} |H(t-t')|^\beta dt'\right]^2 dt, \qquad t + T \le t_D. \tag{6}$$

In the signal-detection formulation the psychometric function has the form of a cumulative Gaussian curve.[26] In the two-alternative forced-choice paradigm such a function can be well approximated by the function[10]

$$P = 1 - 0.5 \exp[-(SC)^\beta], \tag{7}$$

where C is stimulus contrast and S is a sensitivity parameter.

Wilson[27] has recently shown that for low contrasts C, the internal response R_i of a given signal detector i is well approximated by $(S_iC)^\beta$, with S_i its sensitivity parameter. S_i can be chosen such that the internal response is scaled in units of the standard deviation of the internal noise, in which case

$$R_i = d'_i = (S_iC)^\beta \tag{8}$$

and

$$P_i = 1 - 0.5 \exp(-R_i). \tag{9}$$

In this expression β appears as a hard-wired nonlinearity of the transducer function. Note that, although Eq. (7) is identical to Quick's expression for the psychometric function, the definition of R_i in Eq. (9) differs from that of Quick, and they must now be summed according to Eq. (5) to obtain

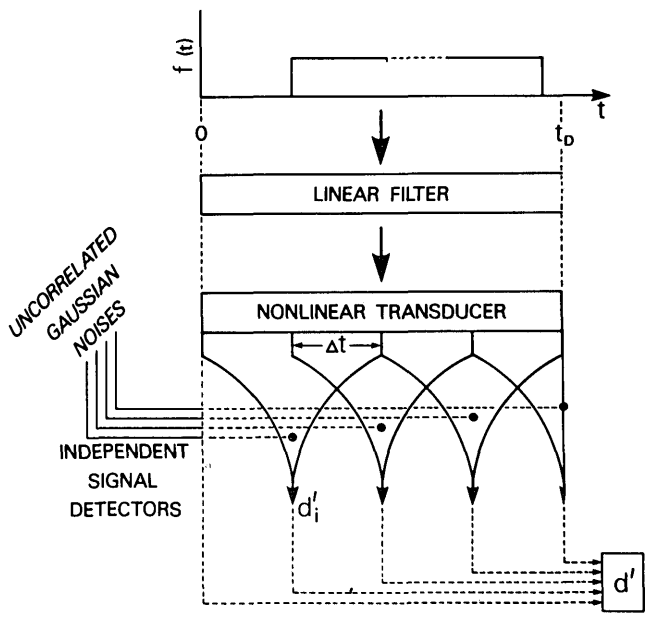
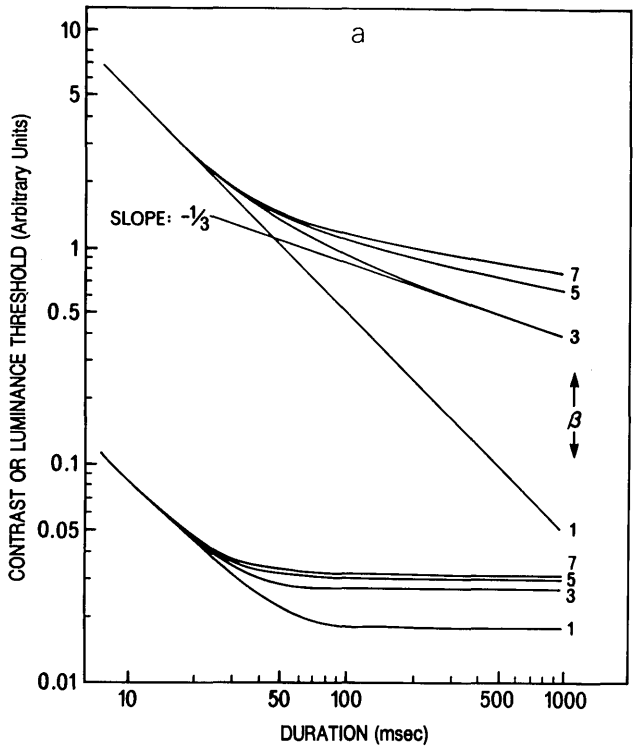
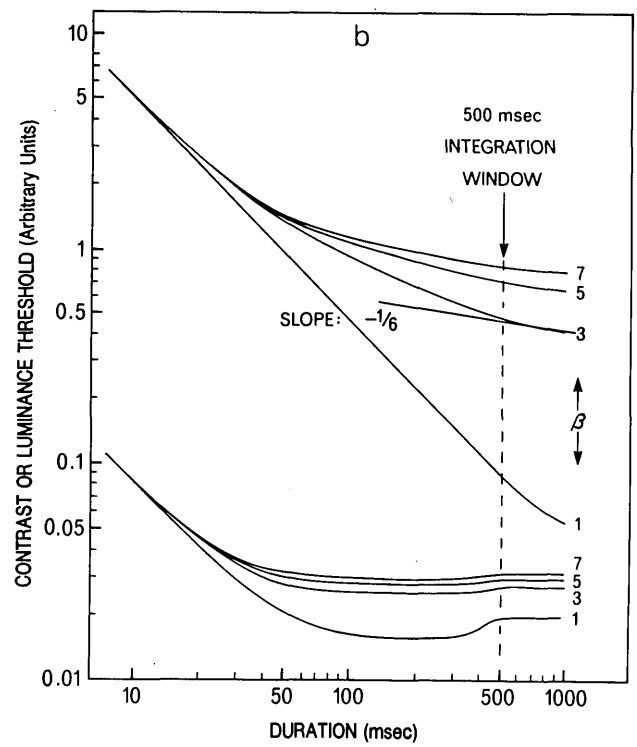


Fig. 4. Block diagram of the model discussed in the text [Eq. (6)]. Note that the operation of the model is illustrated as a function of time t, as specified by the abscissa in the uppermost part of the figure.





the overall response R. Therefore Eq. (7) can no longer be derived from Eq. (9) by applying the standard high-threshold assumption, according to which $P = 1 - \pi\ (1 - P_i)$. Consequently, although the slope parameter of the psychometric function given by Eq. (7) remains β, overall sensitivity S will not vary as the $1/\beta$ root number of detectors N, as predicted by the high-threshold formulation. If the detectors all have equal sensitivity, it can easily be shown from Eqs. (5) and (8) that

$$S = N^{1/2\beta} S_i. \tag{10}$$

Figure 4 displays a block diagram of the model described by Eq. (6), where the processing of the signal at each integration level is specified analytically.

3. PREDICTIONS OF THE MODELS

We present here numerical simulations of the contrast thresholds as predicted by Watson's model [Eq. (2)] to be contrasted with predictions of the model that we propose [Eq. (6)]. All other models discussed in Subsection 2.D lead to predictions quite similar to those of Eq. (2) and will not be discussed further. The simulations were run for the monophasic and the biphasic impulse responses described in Subsection 2.B using a range of values of the β parameter (Figs. 5a and 5b). The double-integral model also allowed for variations in the length of the integration window (Fig. 5c).

Although full integration (−1 slope) always occurs for short stimulus durations, at longer durations the behavior of the function depends markedly on the type of the impulse response. For the monophasic type (upper curves in Figs. 5a and 5b), the threshold in the simple integrator model

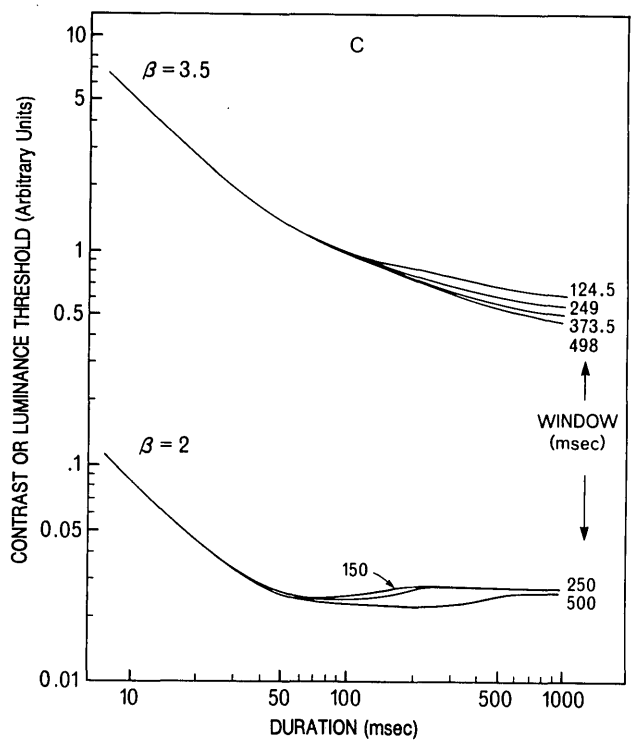


Fig. 5. Threshold–duration functions as predicted by a, Eq. (2) and b, Eq. (6). Predictions of the two models (a and b) were made for different β values. Panel c displays predictions obtained with Eq. (6) for different window lengths (as specified in the figure) at a constant β. Upper and lower curves in each panel were obtained with monophasic and biphasic impulse responses, respectively. The straight lines show the asymptotic slopes of the threshold–duration functions as derived from the models.

attains a decrease rate of $-1/\beta$ (Fig. 5a). Introduction of the integration window in the double-integrator model produces a second asymptotic slope of $-1/2\beta$ beyond the window duration (Fig. 5b), so that thresholds in this condition will now decrease more slowly. Varying the duration of the window simply shifts the point of intersection of the $-1/\beta$ and $-1/2\beta$ asymptotes.

For biphasic impulse responses (lower curves in Figs. 5a and 5b) the simple integrator model now shows an asymptotic approach to a constant value, which varies with β (Fig. 5a). Since we cannot measure absolute sensitivity, the variation in this constant level is best expressed as a variation in the intersection point of the two asymptotes. When the double integrator is introduced, the effect of the window is to produce an increase in threshold as the end of the window is approached (Fig. 5b). The strength of this sensitivity dip is maximum for $\beta = 1$, and it decreases as $2^{1/2\beta}$.

To estimate the extent to which the proposed model fits the experimental data, we measured contrast thresholds for low- and high-spatial-frequency gratings as a function of exposure duration. The β parameter was independently estimated from psychometric functions measured at different presentation times.

4. METHODS

The stimuli were 0.8- and 8-c/deg sinusoidal gratings displayed on a Hewlett-Packard cathode-ray tube (HP/1332A) with green phosphor and were generated by an Apple-II+ computer. They were viewed through a circular aperture of 2.5° diameter at 115 cm from the observer. The display had a large equiluminant surround of matched chromaticity, with 40-cd/m^2 average luminance. A tiny dark spot was used to help fixation. To match the conditions of a related experiment where phase sensitivity was measured,[23] the relative phase of the stimuli was randomly changed between trials. The two authors served as observers in all experiments.

Contrast thresholds were measured by means of a two-alternative forced-choice staircase method. Contrast was always increased by 0.06 log unit for a wrong response and decreased by the same amount with a probability $p = 0.3$ for a correct response. This produces an average detection level of 80% correct on the psychometric function.[27] After 10 preliminary trials, the threshold was computed as the mean contrast of the following 40 trials. The stimuli were presented in a rectangular temporal window of variable length. Up to 13 presentation times ranging from 7.5 to 1000 msec were used. Spatial frequency and presentation times were randomly varied from run to run. At least three thresholds were measured for each experimental condition.

Another set of experiments was designed to measure the psychometric function for the two spatial frequencies. Probability p was used as a parameter of the staircase procedure and was randomly varied from one run to the next so as to span the whole range of detection levels describing the psychometric function. Typically, variations of p from 0.1 to 0.95 resulted in detection levels ranging from 0.55 to 0.98.[28] The psychometric functions were obtained with durations ranging from 50 to 1000 msec.

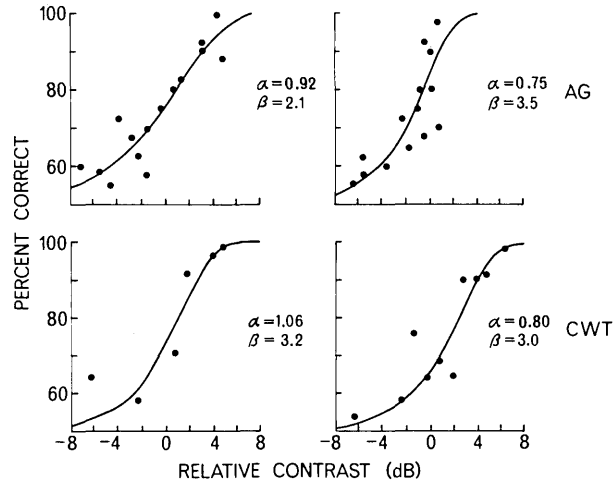


Fig. 6. Percentages of correct responses as a function of contrast obtained with 0.8- (left panels) and 8-c/deg (right panels) stimuli presented for 400 msec. Observers AG and CWT, upper and lower panels, respectively. The continuous lines show maximum-likelihood estimates of the psychometric function.

5. RESULTS

A. Psychometric Functions

Figure 6 shows psychometric functions for the 0.8- and 8-c/deg gratings and for the two observers with 400-msec stimuli. The smooth curves drawn through the datum points are from Eq. (7), and their slope was estimated by maximizing the likelihood ratio of the theoretical binomial distribution of the observed number of correct responses.[6] Although observer AG shows different slopes for low and high spatial frequencies (2.1 and 3.5, with standard deviations of ±0.35 and 0.46, respectively), observer CWT has similar slopes in both conditions (3.2 at low and 3 at high spatial frequencies with standard deviations of ±0.45 and 0.34, respectively). Variations in the slope of the psychometric function with the spatial frequency of the stimulus have been previously described.[15] From Fig. 4 it can thus be predicted that observer CWT should show a shallower dip than observer AG in the 0.8-c/deg condition.

For high spatial frequencies one might expect a change in slope of the threshold–duration function (beyond the integration window) to be accompanied by a corresponding change in the slope parameter of the psychometric function β. However, our formulation of the double-integrator model [Eqs. (6) and (7)] predicts that β should remain invariant with duration for all stimulus durations. To test this prediction we measured β for the 8-c/deg grating at durations ranging from 50 to 1000 msec and obtained values of 3.3–3.5 for observer AG and 3.0–3.3 for observer CWT. These values are constant within experimental error (±0.4).

B. Comparison of the Two Models for Temporal Integration of Contrast

Although absolute threshold values would be predicted by the models from temporal MTF's at the two spatial frequencies studied, such MTF's were not measured in the present study. We therefore found the vertical shift required to fit the data by a standard least-squares procedure.

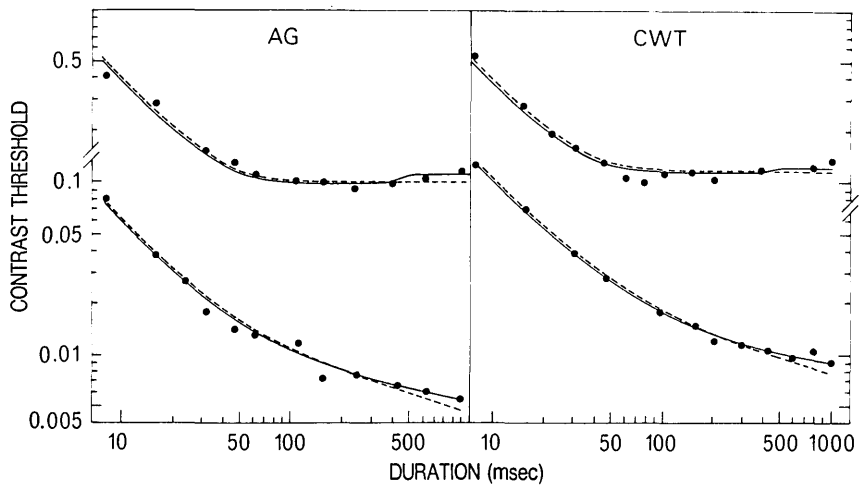


Fig. 7. Forced-choice threshold–duration data obtained for the two observers with 0.8- (upper curves) and 8-c/deg (lower curves) stimuli. Dashed and continuous curves are predictions obtained from Eqs. (2) and (6), respectively (see text for details).

Figure 7 shows the empirical data and predictions derived from Eq. (2) (with parameters from Watson[6]—dashed curve) with only sensitivity as a free parameter. For Eq. (6) (the signal-detection variant—continuous curve) the same procedure was used, together with a 500-msec integration limit determined by inspection, to provide the best fit for both observers. Given that the impulse responses were derived from frequency-response data from another laboratory in different observers, both models provide an excellent fit to the temporal integration functions for both low- and high-spatial-frequency gratings. In logarithmic units the average standard deviations of the data from the single-integrator model are 1.53 and 1.32 dB for the two observers, whereas the deviations from the double-integrator model are 1.32 and 1.18 dB (where 1 dB = 1/20 $\log_{10}$ unit). Overall, the single-integrator model accounts for about 94% of the variance for both observers, whereas the double-integrator model accounts for 96% of the variance. Note that in the biphasic condition, a change in the impulse-response time constant by only 16% or in the β used by 38% would produce a statistically significant increase in the standard deviation of the fit.[29]

We next address the question of whether the double-integrator model statistically improves the fit to the data. This can be tested for the residual variance of the deviations from the best-fitting prediction for the single-integrator model. The double-integrator model has one additional free parameter of window length of the second integration, since the magnitude of the changes is predicted by the measured β. The χ^2 for goodness of fit of the predictions with the residual variances shows a significant improvement with the use of the double integrator for one condition for observer CWT and with both conditions combined for each observer.[30] Furthermore, the second integrator provides a parameter appropriate to improve the fit of the model to the data, i.e., threshold at low spatial frequencies and decreased slope at high spatial frequencies. Note that because the strength of the dip is strongly dependent on the β parameter, the shallow sensitivity dip for observer CWT (biphasic condition) is consistent with his relatively high β obtained in this condition.

6. DISCUSSION

A. Impulse-Response Assumptions

We have described a general class of models intended to account for contrast integration over time by the visual system. On the assumption that the visual system is linear at a first level of integration, its temporal integration behavior was inferred from typical impulse responses. The use of the biphasic (fully developed inhibition) and monophasic (absence of inhibition) impulse responses to account for contrast integration for low (0.8-c/deg) and high (8-c/deg) spatial frequencies, respectively, was justified on two grounds. The bandpass characteristics of the temporal MTF are practically invariant below 1 c/deg,[31] suggesting that within this spatial-frequency region inhibition is fully developed. Above 3 c/deg the temporal MTF has an invariant low-pass form,[30] implying the absence of inhibition in this region. Had the 0.8-c/deg grating not produced a fully balanced inhibitory component, the measured contrast-integration functions would not have reached a plateau of 0 slope but continued to decrease. Conversely, for a 8-c/deg grating, the presence of a small inhibitory component would have produced slopes shallower than $-1/\beta$ for durations below 300 msec. Neither type of error is evident in the data (Fig. 7).

B. Estimation of the Critical Duration

Most previous measurements of threshold–duration functions have been used to infer a critical duration for the mechanisms under study. Because such functions show a smooth transition between the two limiting asymptotes, it has been convenient to define the critical duration as the intercept of these asymptotes. By using *empirically defined* slopes for the limiting asymptotes, several authors have suggested that the critical duration is strongly dependent on the spatial frequency of the stimulus. However, our analysis implies that the definition should be constrained so as to allow only for an asymptote of -1 slope for sufficiently short durations and for asymptotes of either $-1/\beta$ (monophasic impulse responses) or 0 (fully biphasic impulse responses) slope for longer durations not exceeding the inte-

Table 1. Estimated Critical Durations (τ_c)

Spatial Frequency	β AG	β CWT	Model τ_c (msec) Single Integrator AG	Single Integrator CWT	Double Integrator AG	Double Integrator CWT	Mean (τ_c)
Low (0.8 c/deg)	2.1	3.2	36	32	34	31	33
High (8 c/deg)	3.2	3.0	38	43	38	42	40

gration window. The -1 and $-1/\beta$ or 0 asymptotes specify the upper and lower efficiency limits of the integration process. Their intersection can therefore be looked on as the duration at which the two limits are equally distant from the smooth curve, and the efficiency of the full integration process is at half of its maximum. We therefore define critical duration as the intersection between these theoretical asymptotes.

Under this new definition the critical duration may be determined for the two spatial frequencies tested. The result should be insensitive to the use of a single- versus a double-integrator model, since both models are equivalent up to the duration of the second integration window (500 msec), as can be seen in Fig. 7. (Use of the complete data set for this fit, up to 1000 msec, perturbs the parameters for short durations only slightly.) The estimates of critical duration for both model fits are given in Table 1, together with the empirical β values for the model fits. As predicted, the estimates from the two models are very similar (within 2 msec). They also differ by less than 5 msec between observers.

Combining the estimates for both observers and both models gives average estimates for critical durations of 33 msec at low and 40 msec at high frequencies. Thus, when probability summation is taken into account, the critical duration varies by less than 10 msec as spatial frequency is varied, despite the large difference in impulse response assumed in the model and its effects on summation behavior seen in the data. This analysis requires a radical reevaluation of the temporal processing of spatial stimuli to have a relatively invariant integration time across spatial frequency but to differ in the effects of probability summation according to changes in temporal impulse response. Note that this conclusion is independent of the presence of a second level of integration and of the small differences between the two models at the β values produced by our observers.

The conclusion that critical duration is almost invariant with spatial frequency at threshold is consistent with evoked potential latency data. Kulikowski[32] showed that the peak latency remains constant with spatial frequency from 5 to 20 c/deg when stimulus contrast is scaled to an equal ratio with respect to contrast threshold. This is in accord with our psychophysical result that at threshold the integration time is almost invariant with spatial frequency.

C. Interpretation of Temporal Integration Functions

To provide a fuller insight into the interpretation of temporal integration functions, we ask what implications their form carries for the shape of the temporal impulse responses under the assumptions of our model. Once again, these considerations are based on the short-duration results and are independent of the presence of the second level of integration, although its effects are shown for completeness.

Figure 8 shows the simulations already displayed in Fig. 5b together with the intercepts of their asymptotic values and the impulse responses used to generate them. It can be seen that, while the critical duration is insensitive to the β parameter for monophasic impulse responses (Fig. 8a), it varies with β for biphasic impulse responses (Fig. 8b). For a

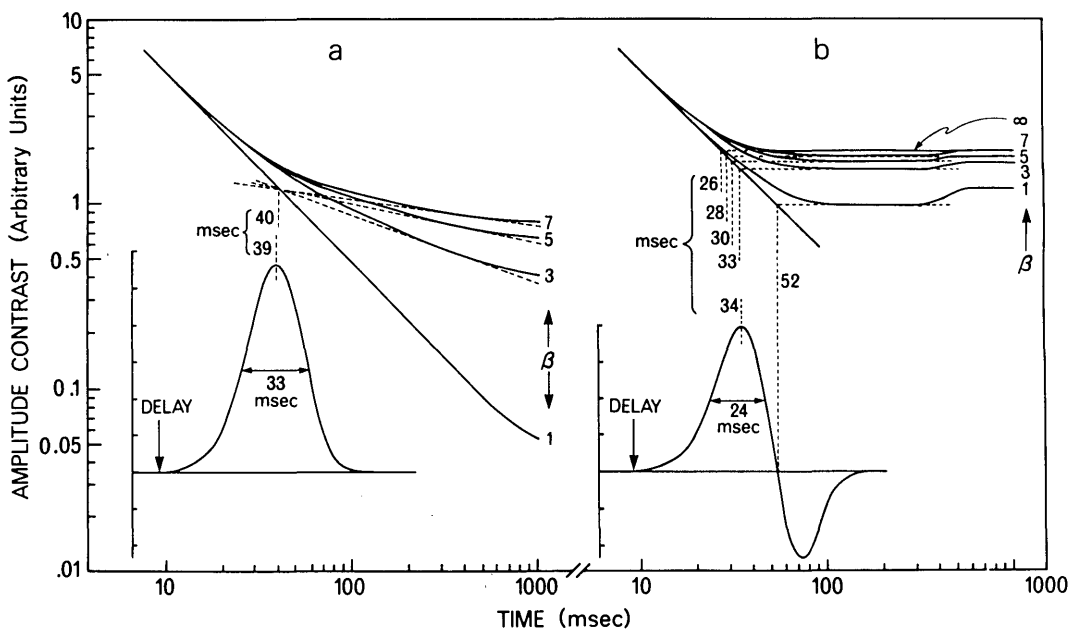


Fig. 8. Illustration of the relation between critical-duration estimates and features of the impulse responses on which they depend (see text). Note that impulse responses are now displayed on a logarithmic abscissa, which causes an apparent distortion in their form.

perfectly symmetrical impulse response having no delay with respect to the onset of the stimulus, the intersection point would correspond both to the width at half-height of this impulse response and to the time at which it reaches its peak. In fact, the monophasic and biphasic impulse responses displayed in Fig. 2 show a period of no apparent response of approximately 9 msec, which must be taken into account when the time to peak is considered. This impulse-response delay is implicit in the form of the impulse response and is not related to an absolute transmission delay of the impulse response relative to the stimulus, which is unmeasurable by threshold techniques. The impulse-response delay forms one of the many components of the reaction time.

As a consequence of the impulse-response delay, the peak of the impulse response should be longer than its width by 9 msec. Figure 8a shows that this is approximately the case for the monophasic impulse response: critical duration = 40 msec, width = 33 msec, peak at 39 msec. The asymmetry of the monophasic impulse response is such that, in practice, the critical duration τ_c can be expressed as

$$\tau_c(\text{monophasic}) \simeq \text{width} + \text{delay} \simeq \text{peak time}. \quad (11a)$$

For a biphasic impulse response the situation is more complicated. The weight given to the positive and the negative lobes in the integration process will depend heavily on the nonlinearity of this integration. For very high values of β the low-amplitude, negative lobe of the pulse response will become negligible with respect to the high-amplitude, positive lobe. The critical duration should thus again be equal to the width at half-height of the first lobe of the impulse response. Figure 8b shows a rather good agreement with this hypothesis: critical duration = 26 msec; width = 24 msec.

However, at low and intermediate β (more typical of the actual slopes of measured psychometric functions[6,15]) the asymmetry of the biphasic impulse response will be such that the critical duration will depend on the weighted contribution of its positive and negative lobes through the nonlinear integration stage. For the extreme case of $\beta = 1$, the zero crossing of the impulse response (at 52 msec) will coincide with the peak of its integral, and hence the critical duration will be equal to 52 msec (Fig. 8b) by the same logic as for the monophasic case but applied to each lobe of the pulse response. The critical duration for our biphasic impulse response can thus vary between a minimum of 24 msec and a maximum of 52 msec when β varies from infinity to 1:

$$\text{width } (\beta = \infty) < \tau_c(\text{biphasic}) < \text{zero crossing } (\beta = 1). \quad (11b)$$

Although for intermediate values of β the intercept will be difficult to interpret precisely, it is possible to use the empirical data to estimate both the peak time and the zero crossing if an independent measure of β is available. The data may then be fitted to one of the family of curves in Fig. 8b. The width of the positive lobe may be read from the curve corresponding to the highest β, while the zero crossing is given by the intercept of the curve for $\beta = 1$.

D. The Double-Integrator Model

The double-integrator model that we have developed accounts satisfactorily for many features of the data. Addition of the second stage of integration significantly improves the fit to the data for all four conditions.[31] Even without the present result, development of the double integrator is justified *a priori* by the much larger discrepancies from a single-integrator model found by previous investigations under low-spatial-frequency conditions (Fig. 1). The need to account for these discrepancies led several investigators[13,14,24] to propose more complex models, but their approach was inadequate, as is discussed in Subsection 2.C (Diode Models).

E. Conclusion

The contrast thresholds measured for short-duration gratings conformed to the complete summation hypothesis of Bloch's law for contrast up to about 20-msec duration. This confirms that it is necessary for the short-duration asymptote to have a slope of −1 in double logarithmic coordinates, as opposed to the arbitrary slopes used by most previous investigators. At medium durations, the behavior of the data is well predicted by nonlinear integration models in which the power of the nonlinearity is derived from the steepness of the psychometric function. These observations lead to a theoretical definition of the critical duration that bears a direct relationship to several features of the visual impulse response.

Our analysis shows that the critical durations at the extremes of the spatial-frequency spectrum differ only to the extent that the inhibition present at low spatial frequencies is absent at high spatial frequencies. No variation in the excitatory component of the impulse response is required to fit the data, which therefore do not support the contention that transient, low-spatial-frequency-selective mechanisms are significantly faster than sustained, high-frequency-selective ones. We conclude that the only change with spatial frequency that influences the form of the threshold–duration function for contrast is the ratio of excitation to inhibition in the initial impulse respone of the system. For our experimental conditions the critical duration changes by less than 10 msec (from 33 to 40 msec on average) from low to high spatial frequencies.

ACKNOWLEDGMENTS

This research was supported by grants from Fondation pour la Recherche Medicale, Philippe Foundation, UNESCO, and Centre National de la Recherche Scientifique to A. Gorea and National Institutes of Health grants EY3884, RR5566, and EY1186 to Christopher W. Tyler.

Parts of this paper were communicated at the 6th European Conference on Visual Perception, Lucca, Italy, 1983.

* Present address, Laboratoire de Psychologie Experimentale, Université René Descartes, 28 rue Serpente, 75009 Paris, France.

REFERENCES AND NOTES

1. A. M. Bloch, "Expériences sur la vision," C. R. Seances Soc. Biol. Paris **37**, 493–495 (1985).
2. U. Tulunay-Keesey and R. M. Jones, "The effect of micromovements of the eye and the exposure duration on the contrast sensitivity," Vision Res. **16**, 481–488 (1976).
3. B. G. Breitmeyer and L. Ganz, "Temporal studies with flashed gratings: inferences about human transient and sustained systems," Vision Res. **17**, 861–865 (1977).

4. A. Gorea, "Effects stationnaires et dynamiques dans le traitement des frequences spatiales," Ph.D dissertation (Université René Descartes, Paris, 1978).
5. G. E. Legge, "Sustained and transient mechanisms in human vision: temporal and spatial properties," Vision Res. **18**, 69–81 (1978).
6. A. B. Watson, "Probability summation over time," Vision Res. **19**, 515–522 (1979).
7. J. J. Kulikowski and D. J. Tolhurst, "Psychophysical evidence for sustained and transient detectors in human vision," J. Physiol. **232**, 149–162 (1973).
8. G. E. Legge, "Space domain properties of a spatial frequency channel in human vision," Vision Res. **18**, 959–970 (1978).
9. H. R. Wilson and J. R. Bergen, "A four mechanism model for threshold spatial vision," Vision Res. **19**, 19–32 (1979).
10. R. F. Quick, "A vector magnitude model of contrast detection," Kybernetik **16**, 65–67 (1974).
11. J. G. Robson, "Spatial and temporal contrast sensitivity functions of the visual system," J. Opt. Soc. Am. **56**, 1141–1142 (1966).
12. D. H. Kelly, "Theory of flicker and transient responses. I. Uniform fields," J. Opt. Soc. Am. **61**, 537–546 (1971).
13. D. H. Kelly and R. E. Savoie, "Theory of flicker and transient responses. III. An essential nonlinearity," J. Opt. Soc. Am. **68**, 1481–1490 (1978).
14. J. A. J. Roufs, "Dynamic properties of vision—IV. Threshold of decremental flashes, incremental flashes and doublets in relation to flicker fusion," Vision Res. **14**, 831–851 (1974).
15. J. Nachmias, "On the psychometric function for contrast detection," Vision Res. **21**, 215–223 (1981).
16. J. A. J. Roufs and F. J. J. Blommaert, "Temporal impulse and step responses of the human eye obtained psychophysically by means of a drift-correcting perturbation technique," Vision Res. **21**, 1203–1221 (1981).
17. H. R. Wilson, "Spatiotemporal characterization of a transient mechanism in the human visual system," Vision Res. **20**, 443–452 (1980).
18. A. B. Watson, "Derivation of the impulse response: comments on the method of Roufs and Blommaert," Vision Res. **22**, 1335–1337 (1982).
19. C. Rashbass, "The visibility of transient changes of luminance," J. Physiol. (London) **210**, 165–186 (1970).
20. C. Rashbass, "Unification of two contrasting models of the visual increment threshold," Vision Res. **16**, 1281–1283 (1976).
21. J. J. Koenderink and A. J. van Doorn, "Detectability of power fluctuations of temporal visual noise," Vision Res. **18**, 191–195 (1978).
22. D. J. Tolhurst, J. A. Movshon, and A. F. Dean, "The statistical reliability of signals in single neurons in cat and monkey visual cortex," Vision Res. **23**, 775–786 (1983).
23. C. W. Tyler and M. J. Mayer, "Probability summation over space: predictions from a nonlinear signal-detection approach," J. Opt. Soc. Am. A **1**, 1301 (A) (1984).
24. J. Krauskopf, "Discrimination and detection of changes in luminance," Vision Res. **20**, 671–677 (1980).
25. G. E. Legge and D. Kerstein, "Light and dark bars: contrast discrimination," Vision Res. **23**, 473–483 (1983).
26. D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966).
27. H. R. Wilson, "A transducer function for threshold and suprathreshold human vision," Biol. Cybern. **38**, 171–178 (1980).
28. Decreasing contrast with a probability $p = 0.33$ for each correct response is equivalent to decreasing the contrast after three correct responses in a row. Any p chosen can similarly be expressed as the number n of in-row correct responses required to elicit a step down. The probability of a correct response at a given contrast level i, $P(i)$, is then given by

$$P_n(i) = [1 - P(i)] * [1 + P(i) + P(i)^2 + P(i)^3 + \ldots + P(i)^n];$$

see R. A. Schumer and B. Julesz, "Binocular disparity modulation sensitivity to disparities offset from the plane of fixation," Vision Res. **24**, 533–542. (1984).
29. A significant increase in the error variance of a fit may be assessed by the usual F test for the ratio of two variances [J. J. Roscoe, *Fundamental Research Statistics for the Behavorial Sciences* (Holt, New York, 1975)]:

$$F = \frac{\sigma_1^2}{\sigma_2^2}, \qquad \sigma_2 > \sigma_1.$$

For a fit to 12 points a significant increase is defined by $F = 2.82$ at $p = 0.05$, $N - 12$, degrees of freedom = 11,11. Thus parameter variations, such as to increase the variance of the fit by a factor of 2.82, cause a significantly poorer fit. This was found by separate variation of the parameters to occur with a 16% change of the impulse-response time constant or a 38% change in the value of β.
30. The appropriate approach to this nonlinear regression is to estimate the χ^2 statistic for the goodness of fit of the residual variance from the one-parameter model to fit with the extra parameter introduced in the two-parameter model [P. R. Bevington, *Data Reduction and Error Analysis for the Behavioral Sciences* (McGraw-Hill, New York, 1969)]. The relevant equation is given by the F ratio

$$F_\chi = \frac{(\chi_m^2 - \chi_{m+k}^2)(n - k - 1)}{\chi_{m+k}^2},$$

where n is the number of points to be fitted, m is the number of free parameters in the simpler model, and k is the number of extra parameters in the full model.
This F ratio may also be expressed in terms of the residual error variance σ_ϵ^2, from the best fit to each model:

$$F = \frac{(\sigma_{\epsilon:m}^2 - \sigma_{\epsilon:m+k}^2)(n - k - 1)}{\sigma_{m+k}^2},$$

since the error terms for individual measures from each mean in the denominator of each χ^2 are all identical and cancel. Note that this statistic may be used for any nonlinear model with any number of fixed or free parameters.
The degrees of freedom for to be used for the F test are $(1, n - k - 1)$. In the present case the full model has one extra parameter for individual conditions and two for the combined conditions since the duration is assumed to be the same for both. In addition, an extra degree of freedom was allowed since we selected the model from a number of candidates.
The F ratios are given in the table below (after Bevington) and are significant at $P < 0.05$ (two-tailed) only for the sustained condition for observer CWT and when combined over both conditions for each observer. We may conclude that the additional parameter can account for a significant improvement in the model fit for both observers.

Observer	Condition	$F\chi$	Degrees of Freedom	$F_{p<0.05}$
AG				
	T	1.38	8	5.32
	S	3.26	9	5.72
	Combined	5.47[a]	19	4.39
CWT				
	T	2.98	10	4.96
	S	24.54[a]	9	5.72
	Combined	13.53[a]	21	4.33

[a] F significant at $p < 0.05$.

31. D. H. Kelly, "Visual contrast sensitivity," Opt. Acta **24**, 107–129 (1977).
32. J. J. Kulikowski, "Visual evoked potentials as a measure of visibility," in *Visual Evoked Potentials in Man: New Developments*, J. E. Desmedt, ed. (Clarendon, Oxford, 1977), pp. 168–183.