Uncertainty explains many aspects of visual contrast detection and discrimination

Denis G. Pelli

Institute for Sensory Research, Syracuse University, Syracuse, New York, 13210

Received December 1, 1983; accepted May 7, 1985

More than 20 years ago, Tanner [Ann. N.Y. Acad. Sci. **89**, 752 (1961)] noted that observers asked to detect a signal act as though they are uncertain about the physical characteristics of the signal to be detected. The popular assumptions of probability summation and decision variable, taken together, imply this uncertainty. This paper defines an uncertainty model of visual detection that assumes that the observer is uncertain among many signals and chooses the likeliest. With only four parameters, the uncertainty model explains why d' is approximately a power function of contrast ("nonlinear transduction") and accurately predicts effects of summation, facilitation, noise, subjective criterion, and task for near-threshold contrasts. Thus the uncertainty model offers a synthesis of much of our current understanding of visual contrast detection and discrimination.

1. INTRODUCTION

This paper is about an explanation of visual contrast detection and near-threshold contrast discrimination, especially how they depend on the signal's contrast and extent and on the observer's subjective criterion. It seems likely that a closely analogous explanation could be made for auditory detection and discrimination, but that will not be attempted here.

Attempts to model human detection performance have often first considered the ideal observer of an exactly known signal. For example, the fluctuation theory of Rose[1] and de Vries[2] assumed that the observer would count photons in the area and duration of the disk to be detected and ignore everything else. This is possible only if the observer knows, in advance, the time and place and the size and duration of the signal. Signal-detection theory introduces the concept of an *ideal observer* that uses the available information to detect the signal with maximum reliability. Rose and de Vries assumed that the observer knew the signal exactly and was ideal except for having a low transduction efficiency.

Detectability of a signal by an observer is measured by d', which is computed from the observer's performance. If the human observer acted ideally (as suggested by Rose and de Vries), then detectability, d', would be proportional to signal contrast.[3] However, it is now well established that d' arises as a power greater than 1 (typically 2) of signal contrast whether the signal is a spot[4-9] or a sinusoidal grating.[10-12] Since d' has often been thought of as the mean of an internal effect of the stimulus, the finding that d' is nonlinearly related to signal contrast has often been called nonlinear transduction. However, this name may be misleading. Tanner[5] noted that an ideal observer of one of many possible signals has the nonlinear psychometric functions (i.e., $d' \propto c^b, b > 1$) observed experimentally for human detection of exactly known signals. Tanner suggested that the ideal-detection-of-a-signal-known-exactly model of human performance might be wrong in assuming exact knowledge of the signal:

> An examination of our procedure indicates there might be uncertainty both as to the time and location of the signal presentation. . . . The uncertainty may be introduced at a central rather than retinal level, suggesting that one should not try to account for nonlinearities at the level of the end-organ, to the exclusion of central hypotheses.

This paper is an examination of Tanner's suggestion.

The hypothesis to be tested is that the observer expects one of many possible signals and chooses the likeliest. Uncertainty increases threshold for detection of a given signal because the subject's decision must be based on many noisy measurements, one for each possible signal.

The cause of the uncertainty is not relevant here. The uncertainty may be *intrinsic* (in the observer), *extrinsic* (in the stimulus), or both. The experimenter introduces extrinsic uncertainty by randomly presenting one of many possible signals. The subject is intrinsically uncertain if he acts as though one of many possible signals is to be presented even when the experimenter has indicated which is to be presented. Furthermore, one can imagine two kinds of intrinsic uncertainty. An observer might be unable to remember the set of possible signals, or he might know them but be unable to apply that knowledge to the detection task. All that matters here is the total uncertainty: the set of signals that the observer seems to expect. The dimensions of the uncertainty are irrelevant. The observer may be uncertain about when the signal will be presented, or where, or may be uncertain along any other physical dimension, such as spatial frequency. It makes no difference to performance in the experiments considered here and thus will not be discussed here. What is important is how much uncertainty there is.

The uncertainty model (also called multiple-band filter model with a largest-of detector, disjunctive receiver, and channel uncertainty model) is the maximum-likelihood receiver for signals in white noise. It has been repeatedly suggested for visual[5,7,11,13-18] and auditory[17,19,20] detection. Nevertheless, I believe that most of its properties as a detector in the relevant psychophysical tasks are documented here for the first time.

Section 2 will show that, taken together, the probability-summation and decision-variable assumptions imply uncertainty and a maximum-of decision rule, strongly suggestive of a maximum-likelihood receiver. Section 3 reviews classical signal-detection theory, beginning with the optimal strategy for signal detection and showing how successively more de-

0740-3232/85/091508-25$02.00

tailed specifications of the signals eventually yield a model whose performance can be analyzed: the uncertainty model. Section 4 defines the uncertainty model and derives the equations that define its performance. Each subsequent section examines one aspect of human detection performance, showing that in each case the uncertainty model is at least as successful as all previous explanations. Effects of contrast are examined in Section 5, extent in Section 6, subjective criterion in Section 7, task in Sections 8 and 9, and noise in Section 10.

Notation

A, contrast-normalization factor, $A = c/c'$.
b, log–log steepness parameter in expression $d' \propto c^b$.
c, Michelson contrast, $c = (L_{max} - L_{min})/(L_{max} - L_{min})$.
c', normalized signal contrast, $c' = c/A$; see Subsection 4.C.
d', detectability, for two-alternative forced choice, $d' = \sqrt{2}\Phi^{-1}(P)$.
G, probability of an irrelevant hit: irrelevant-hit rate.
K, number of receptive fields (expected signals) stimulated by the actual signal.
M, amount of uncertainty: the number of expected signals.
$P(c)$, probability of a hit: hit rate.
$P(0)$, probability of a hit at zero contrast: false-alarm rate.
$P^*(c)$, probability of a relevant hit: relevant-hit rate.
α, Weibull function parameter, the threshold contrast at which the hit rate is $1 - (1 - \gamma)/e$.
α', normalized threshold contrasts, $\alpha' = \alpha/A$.
β, Weibull function parameter that controls the log–log steepness.
γ, Weibull function parameter, the false-alarm rate.
λ, subjective criterion.
$\Phi(x)$, the cumulative normal probability distribution.
$\Phi^{-1}(P)$, the inverse cumulative normal distribution, i.e., $x = \Phi^{-1}[\Phi(x)]$.

2. PROBABILITY SUMMATION AND THE DECISION VARIABLE

The uncertainty model is little more than the combination of two popular assumptions about visual detection: the probability-summation assumption and the decision-variable assumption. Together these two assumptions imply uncertainty.

A. Probability-Summation Assumption

A yes–no trial consists of a single presentation that is either a signal or a blank. The subject responds either "yes," the signal was presented, or "no," it was not. A hit is a yes given in response to a signal presentation. A false alarm is a yes given in response to a blank presentation.

A two-alternative forced-choice (2afc) trial consists of two presentations, one a signal and the other blank, in random order. The subject must choose which presentation was the signal. A hit is a correct choice.

For either kind of trial, a zero-contrast hit is a hit at zero signal contrast. A zero-contrast signal is physically identical to a blank. Thus the probability of a zero-contrast hit is the same as the probability of a false alarm.

The idea of probability summation is that there are many independent reasons for detecting a signal and any of these reasons is sufficient to detect the signal successfully, i.e., cause a hit. Thus two quite different stimuli presented simultaneously are slightly more detectable than either is alone, because the observer has two independent chances at detection. This idea has been successful in explaining summation effects, that is, experiments that measured detectability or threshold as a function of number of signals or signal extent.

Probability summation assumes that the relevant reasons never, or at least hardly ever, cause hits at zero signal contrast. However, since observers do occasionally make zero-contrast hits, it is necessary also to suppose irrelevant reasons for hits. Thus, altogether, probability summation requires that there be independent reasons for hits, some relevant and some irrelevant.

Figure 1 represents the probability-summation assumption. The assumption has four parts:

First, it asserts that hits occur for either (or both) of two types of reason: relevant and irrelevant. By definition, the probability of a *relevant* hit (a "true" hit) depends on signal contrast, and the probability of an *irrelevant* hit (a "guess") is independent of signal contrast.

Second, these two types of hit are stochastically independent. Thus the hit *rate* (i.e., probability) $P(c)$ as a function of contrast c is the "probability sum" of the irrelevant-hit or guess rate G, which is contrast independent, and the relevant-hit rate $P^*(c)$, which is contrast dependent:

$$P(c) = 1 - (1 - G)[1 - P^*(c)]. \tag{2.1}$$

(Note that this assumption does allow the hit rates to depend on nonstochastic factors, such as the task, or a subjective criterion.)

Third, if there is more than one relevant or irrelevant reason for a hit, then they are all stochastically independent, so that

$$P(c) = 1 - \prod_{i=K+1}^{M} (1 - G_i) \prod_{i=1}^{K} (1 - P_i^*), \tag{2.2}$$

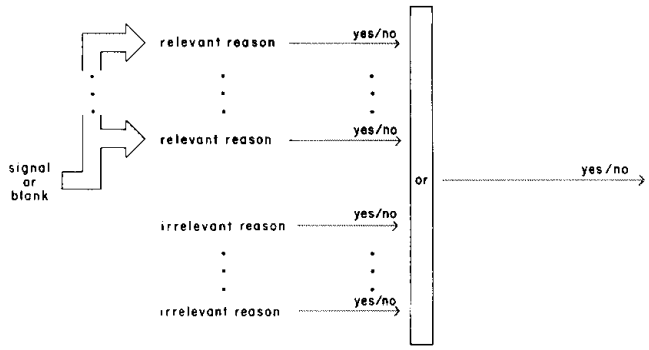


Fig. 1. The probability-summation assumption. The stimulus, either a signal or a blank, is a function of space and time. The arrows on the left are broad to indicate the transmission of an entire function. The thin arrows on the right designate transmission of simple scalar quantities, in this case yes or no. The stimulus may cause a yes for any of many relevant reasons. A yes may also occur for any of many irrelevant reasons. The long box labeled "or" will cause the observer to say "yes" if any of the reasons produces a yes.

which reduces to Eq. (2.1), where G is the total irrelevant-hit rate

$$G = 1 - \prod_{i=K+1}^{M} (1 - G_i) \tag{2.3}$$

and $P^*(c)$ is the total relevant-hit rate

$$P^*(c) = 1 - \prod_{i=1}^{K} [1 - P_i^*(c)]. \tag{2.4}$$

Finally, and most important, assume that we may *correct for guessing*,[3] that is, recover the relevant-hit-rate function $P^*(c)$ from the hit-rate function $P(c)$ by the following formula:

$$P^*(c) \approx 1 - [1 - P(c)]/[1 - P(0)]. \tag{2.5}$$

Glancing at Eq. (2.1), we see that this is equivalent to asserting that the zero-contrast-hit rate $P(0)$ is approximately equal to the irrelevant-hit rate G:

$$P(0) \approx G. \tag{2.6}$$

Given Eq. (2.1), this is equivalent to requiring that when the irrelevant-hit rate G is not close to 1, then the zero-contrast relevant-hit rate $P^*(0)$ is approximately zero:

$$\text{if } G \not\approx 1 \text{ then } P^*(0) \approx 0. \tag{2.7}$$

This completes the probability-summation assumption.

Expressions (2.5)–(2.7) are written as approximations because it would be physically impossible to satisfy them exactly. However, most measurements are not sufficiently precise to distinguish a very small zero-contrast relevant-hit rate $P^*(0)$, say, less than 3%, from zero.

The physical impossibility of satisfying expressions (2.5)–(2.7) exactly results from the photon nature of light. All visual stimuli are mediated by photon absorptions in the retina. Since photon absorptions are a random process, any pattern of absorptions produced at nonzero contrast can also occur at zero contrast, although generally with much lower probability. Thus any observer state (e.g., a relevant hit) that can occur at nonzero contrast can also occur at zero contrast. The occurrence of even a single relevant hit when the contrast is zero and the irrelevant-hit rate is less than 1 would violate the exact form of expressions (2.5)–(2.7).

B. Decision-Variable Assumption

Now consider effects of the subjective criterion. It is well known that observers can be induced to increase or reduce their caution in saying yes and that the degree of caution affects the probability of saying yes on both signal and blank trials. If the data are corrected for guessing, it is found that the relevant- and irrelevant-hit rates covary. A convenient explanation of this is the decision-variable assumption, which comes from signal-detection theory. The key idea is that observers combine everything that they sense about the stimulus into a scalar quantity called a decision variable and make their decisions on the basis of this number.[3,4,10,21–23]

Figure 2 represents the decision-variable assumption. The assumption has three parts. First, the observer's responses in all contrast-detection and -discrimination tasks (e.g., rating scale, yes–no, 2afc) depend on each stimulus presentation only by means of a scalar quantity, a hypothetical internal effect, called the decision variable. Second, the likelihood ratio that a signal, as opposed to a blank, was presented increases monotonically with the decision variable. Third, the decisions are based on the decision variable in the ideal way.

The ideal way to use a decision variable is specified in the theory of signal detectability, in which the decision variable is always the likelihood ratio of signal versus blank presentation. In yes–no, as illustrated in Fig. 2, the observer says "yes" only if the decision variable exceeds some subjective criterion; in rating scale the observer gives ratings monotonically related to the decision variable; and in multiple-interval forced-choice experiments the observer chooses the interval that produced the largest value of the decision variable.

C. Combined Assumption

The two assumptions are appealing. Probability summation explains summation effects, and the decision variable explains effects of the subjective criterion. Together they explain even more and imply uncertainty.

Figure 3 shows the combined assumption. Again the stimulus is a signal or a blank, a function of space and time. It affects the relevant variables, $L_1, \ldots, L_K$. Irrelevant variables, $L_{K+1}, \ldots, L_M$, also produce values on each presentation. The maximum value is passed by the long box to become the decision variable L. The decision-variable assumption is satisfied because the stimulus is reduced to a single number, which is used as a decision variable. The independence requirement of probability summation is satisfied because any variable will cause a yes response if it exceeds the subjective criterion λ.

If the formal equivalence of Fig. 3 to the combination of Figs. 1 and 2 is not apparent, consider replacing each of the reasons in Fig. 1 with a variable that is to be compared with λ. The probability-density function of the variable must be such that its probability of exceeding λ equals the probability

Fig. 2. The decision-variable assumption. The stimulus is either a signal or a blank, a function of space and time. All the information about the stimulus is somehow reduced to a scalar quantity L, which is the decision variable. The observer says "yes" if the decision variable L exceeds the subjective criterion λ and says "no" otherwise.

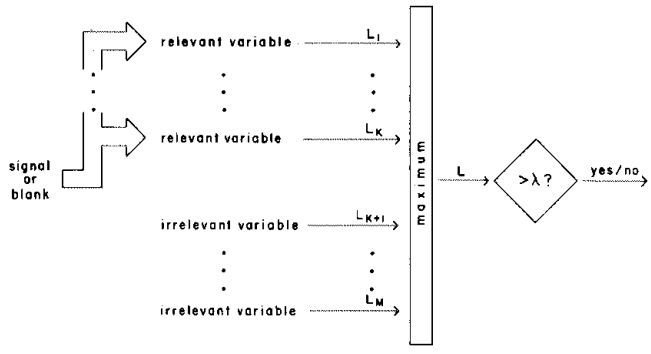


Fig. 3. The probability-summation and decision-variable assumptions. The stimulus is a signal or a blank. It affects the relevant variables, $L_1, \ldots, L_K$. Irrelevant variables, $L_{K+1}, \ldots, L_M$, also produce values on each presentation. The maximum value is passed by the long box to become the decision variable L.

of the reason it replaces. Now note that one or more of these variables will exceed λ if and only if their maximum exceeds λ.

Specifically, the combined assumption states that the observer uses a decision variable equal to the maximum of M stochastically independent variables, some of which are relevant (i.e., contrast dependent) and some of which are irrelevant (i.e., contrast independent). Furthermore, although there is no way to indicate it on Fig. 3, the correction-for-guessing part of the probability-summation assumption requires [expression (2.6)] that the zero-contrast-hit rate be almost entirely due to irrelevant hits. In other words, at zero contrast the maximum variable is almost always irrelevant.

So where is the uncertainty? Well, why does the observer monitor the irrelevant variables? Why does he or she not ignore them? The observer acts as though he were uncertain about which variables were relevant. He acts as though, not knowing which variables are sensitive to the signal, he monitors all the variables that he thinks might be relevant.

D. Discussion

The derivation of the combined assumption may seem strange to some readers since the two assumptions are usually presented as incompatible. That is because typical definitions of probability summation also make the high-threshold assumption.[24] The high-threshold assumption attributes the irrelevant hits to intentional guessing by the observer, such that the observer can change either the relevant- or the irrelevant-hit rate without affecting the other. In other words, the high-threshold assumption asserts that the irrelevant-hit rate G is independent of the subjective criterion, contrary to the decision-variable assumption, which states that all yes–no responses are based on comparisons of the decision variable and the subjective criterion. Criterion independence of the irrelevant-hit rate has been repeatedly disproved by showing that when subjects are induced to reduce their subjective criterion (be less cautious), the irrelevant- and relevant-hit rates both go up.[3,7,12,25–27] These data reject the high-threshold assumption but not probability summation. The probability-summation assumption is not concerned with any dependence between the irrelevant-hit rate and the subjective criterion because probability summation deals with data presumably collected at a constant irrelevant-hit rate and subjective criterion. Thus the high-threshold assumption is disproved, but probability summation stands.

Nachmias[12] considered the combined assumption too, noting that it could satisfy the evidence against the high-threshold assumption and the evidence for probability summation in yes–no experiments.

This section has shown that the probability-summation and decision-variable assumptions together imply uncertainty. In the combined assumption illustrated by Fig. 3 it is tempting to suppose that each of the independent variables L_i is a measure of the likelihood that a particular signal s_i was present. In other words, Fig. 3 is strongly suggestive of a maximum-likelihood receiver, the subject of the next section.

3. SIGNAL-DETECTION THEORY

This section reviews the classical detection theory from which the uncertainty model arose. People more interested in the relation of the theory to the data may wish to skip this section on their first reading.

A. Optimal Decisions

Even the earliest papers on signal-detection theory acknowledged that it would often be unreasonable to assume exact knowledge of the signal to be detected. One alternative is detection of one of many possible signals.

In many paradigms we ask the observer to classify the stimulus. The observer entertains several exclusive hypotheses (e.g., signal or blank) and chooses one. The optimal strategy—to minimize the expected error—depends on the weights assigned to the various types of error that the observer may make. If the errors (i.e., choosing a wrong hypothesis) are all equally weighted, then the optimal strategy is to choose the hypothesis with maximum posterior probability. Furthermore, if the hypotheses have equal prior probability, then this is equivalent to choosing the hypothesis with maximum likelihood.[28]

B. Equally Probable Signals

So far the discussion has emphasized simple yes–no detection of a single possible signal. Now consider presentation of one of two possible signals, $s_1(x, y, t)$ or $s_2(x, y, t)$, in one of two intervals, 1 or 2. (The restriction to two signals and two intervals is merely for clarity of presentation. All statements in this section may be generalized in the obvious way to M signals $s_1, s_2, \ldots, s_M$ and N intervals $1, 2, \ldots, N$.) First, we could ask the observer to identify the signal as s_1 or s_2. Second, we could ask the observer to detect which interval the signal appeared in, 1 or 2. Or, third, we could ask both, counting the answer as wrong unless both the identity and the interval are correct. The optimal strategy is different in all three cases.

Physically, there are four distinct possibilities for each trial: signal s_1 in interval 1, s_1 in 2, s_2 in 1, or s_2 in 2. Assume that the four possibilities have equal prior probability so that we may work with likelihoods instead of posterior probabilities. Let $l_{i,j}$ represent the likelihood of the simple hypothesis, signal i in interval j, so we have four likelihoods to compute: $l_{1,1}$, $l_{2,1}$, $l_{1,2}$, and $l_{2,2}$. Let $l_{*,j}$ represent the likelihood of the compound hypothesis: any signal in interval j. Let $l_{i,*}$ represent the likelihood of the compound hypothesis, signal s_i in any interval. The likelihoods of the compound hypotheses are obtained by summing the likelihoods of simple hypotheses:

$$l_{*,j} = \sum_i l_{i,j}, \tag{3.1}$$

$$l_{i,*} = \sum_j l_{i,j}. \tag{3.2}$$

The optimal strategies are easily stated. To identify the signal, choose the signal i with largest likelihood $l_{i,*}$, i.e., largest sum $l_{i,1} + l_{i,2}$. To detect the signal (i.e., choose the interval), choose the interval j with largest likelihood $l_{*,j}$, i.e., largest sum $l_{1,j} + l_{2,j}$. To detect *and* identify (getting both right), choose the signal and the interval (i, j) with largest likelihood $l_{i,j}$. Note that the separate choices of signal and interval are not the same as the combined choice. This is because the separate choices do not have to weigh the cost of

an error in choosing the other parameter, whereas the combined choice must weigh the cost of an error in both parameters.

The third rule will preoccupy us for the rest of this paper. It chooses the simple hypothesis that has maximum likelihood. Therefore I will call it a *simple maximum-likelihood receiver.* The uncertainty model to be presented in the next section is a simple maximum-likelihood receiver. The simple maximum-likelihood receiver always chooses the likeliest simple hypothesis, so it is optimal in classifying simple hypotheses but suboptimal in classifying composite hypotheses.

Now consider 2afc detection, since that and yes–no are the two tasks that this paper will deal with. From above, the optimal decision rule is to compare $l_{*,1}$ and $l_{*,2}$ and to choose interval 1 if $l_{*,1}$ is larger or 2 if $l_{*,2}$ is larger. This may be represented symbolically:

$$l_{*,1} \underset{\text{"2"}}{\overset{\text{"1"}}{\gtrless}} l_{*,2}. \tag{3.3}$$

The greater-or-less-than sign means: Say 1 if greater than and say 2 if less than. Substituting from (Eq. 3.2), this becomes

$$\sum_i l_{i,1} \underset{\text{"2"}}{\overset{\text{"1"}}{\gtrless}} \sum_i l_{i,2}. \tag{3.4}$$

Any rule that yields the same decisions as relation (3.4) is equivalent to relation (3.4) and therefore also optimal. Any rule not equivalent to relation (3.4) will be suboptimal for this task. Thus the simple maximum-likelihood receiver,

$$\max_i l_{i,1} \underset{\text{"2"}}{\overset{\text{"1"}}{\gtrless}} \max_i l_{i,2}, \tag{3.5}$$

is suboptimal for detection of one of many possible signals, although it may not be experimentally distinguishable from the optimal rule. When there is only one possible signal, $M = 1$, the two rules are equivalent. As in the combined assumption of Section 2, the simple maximum-likelihood receiver has many variables combined by a maximum-of rule. Although this is the proposed model for human performance, its performance depends on details of the signals that have yet to be specified. Thus to evaluate its performance it is necessary to specify enough properties of the signals to determine the model's performance.

C. Equally Probable Signals in White Noise

If the stimulus is the sum of a possible signal and white noise, then the likelihood l is given, except for a proportionality constant, by

$$l_{i,j} \propto \exp(c_i' L_{i,j}), \tag{3.6}$$

where $L_{i,j}$ is the cross correlation of signal s_i with the stimulus from interval j, normalized by the standard deviation of that cross correlation, and c_i' is the normalized contrast of signal s_i.

Normalized contrast c' of a signal is defined as the ratio of mean to standard deviation of the cross correlation of that signal with the stimulus presentation of that signal. The normalized contrast c' is proportional to the physical contrast c. c_i' is the mean value of $L_{i,j}$ when signal i is presented on interval j.

In practice, it is usual to compute the cross correlations $L_{i,j}$ and unusual to evaluate the likelihoods $l_{i,j}$. Thus it is more common to see decision rules expressed in terms of $L_{i,j}$. The optimal detector [relation (3.4)] is

$$\sum_i \exp(c_i' L_{i,1}) \underset{\text{"2"}}{\overset{\text{"1"}}{\gtrless}} \sum_i \exp(c_i' L_{i,2}). \tag{3.7}$$

The simple maximum-likelihood receiver [relation (3.5)] is

$$\max_i \exp(c_i' L_{i,1}) \underset{\text{"2"}}{\overset{\text{"1"}}{\gtrless}} \max_i \exp(c_i' L_{i,2}), \tag{3.8}$$

or, equivalently, since $\exp x$ is a strictly increasing function of x,

$$\max_i c_i' L_{i,1} \underset{\text{"2"}}{\overset{\text{"1"}}{\gtrless}} \max_i c_i' L_{i,2}. \tag{3.9}$$

D. Equally Probable, Equal-Energy Signals in White Noise

Define the contrast function $s(x, y, t)$ of any visual stimulus as the luminance function $L(x, y, t)$ divided by its mean L_{av} minus 1 (Ref. 29):

$$s(x, y, t) = L(x, y, t)/L_{av} - 1.$$

The contrast energy E of a signal is the integral of the squared contrast function over space and time[30]:

$$E = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s^2(x, y, t) \mathrm{d}x \mathrm{d}y \mathrm{d}t. \tag{3.10}$$

If all the signals have the same contrast energy, E_i, then they will all have the same normalized contrast c', and relation (3.9) reduces to

$$\max_i L_{i,1} \underset{\text{"2"}}{\overset{\text{"1"}}{\gtrless}} \max_i L_{i,2}. \tag{3.11}$$

Graham *et al.*[31] and Yager *et al.*[32] called relation (3.11) a maximum-output rule and compared its performance with an adding-of-outputs rule,

$$\sum_i L_{i,1} \underset{\text{"2"}}{\overset{\text{"1"}}{\gtrless}} \sum_i L_{i,2}. \tag{3.12}$$

Detection and identification of one of two possible signals has been successfully modeled both by the maximum-output rule[33] and by the adding-of-outputs rule.[34] The adding-of-outputs rule is suboptimal. Substitution from relation (3.6) (which assumes white noise) shows that the adding-of-outputs rule [relation (3.12)] is equivalent to multiplying the likelihoods $l_{i,j}$, whereas the optimal rule [expression (3.4)] adds them.

E. Equally Probable, Equal-Energy, Orthogonal Signals in White Noise

As noted above, an optimal way to detect 1 of M equally probable signals is to evaluate the likelihood of each signal and sum the likelihoods to determine the overall likelihood of a signal. Peterson *et al.*[35] analyzed the performance of this ideal (i.e., optimal) detector [relation (3.7)] for M equally probable equal-energy (all c_i' equal to c') orthogonal signals

in white noise. Since they were considering the yes–no task, there is only one interval, and the relation for the optimal detector is

$$\sum_i \exp(c' L_{i,1}) \underset{\text{"no"}}{\overset{\text{"yes"}}{\gtrless}} \lambda. \tag{3.13}$$

Unfortunately the resulting performance equations are difficult to manipulate, and Peterson *et al.* were able to provide only an approximate solution for the ideal's performance as a function of the number M and normalized contrast c' of the signals:

$$(d')^2 \sim \ln\left[1 - \frac{1}{M} + \frac{1}{M}\exp(c'^2)\right]. \tag{3.14}$$

Nolte and Jaarsma[36] solved the original equations numerically for several values of M and c' and published receiver operating characteristic (ROC) curves. They found that the approximation of Peterson *et al.* was accurate when c' was less than 1 but extremely inaccurate when c' was as large as 4. The interesting cases for the purposes of this paper turn out to have large values of M (i.e., high uncertainty). When M is greater than 100 the approximation is inaccurate for values of d' greater than 0.5. Human psychophysical performance can be measured with reasonable accuracy and effort only over the range of d' extending from approximately 0.2 to 2, and the approximation of Peterson *et al.* is inaccurate over most of this range. Nolte and Jaarsma's solutions are accurate, but their paper could include solutions only for a modest number of pairs of values of M and c'.

Nolte and Jaarsma also calculated the yes–no performance of the simple maximum-likelihood receiver [relation (3.11)], which they called a "disjunctive receiver" (after Wainstein and Zubakov[37]). They called it a disjunctive receiver because in yes–no detection it says yes if any of the L_i exceeds the subjective criterion λ:

$$\max_i L_{i,1} \underset{\text{"no"}}{\overset{\text{"yes"}}{\gtrless}} \lambda. \tag{3.15}$$

We noted that the simple maximum-likelihood receiver is suboptimal, i.e., it performs less reliably than the ideal detector of 1 of M signals, but Nolte and Jaarsma found that the performance graphs (ROC curves) were nearly superimposed, showing that the simple maximum-likelihood receiver is virtually ideal for detection of equally probable equal-energy signals in white noise. (The range that they examined is $1 \leqq M \leqq 5000$ and $0.7 \leqq c' \leqq 4$.)

F. Generality of the Results

Analysis of maximum-likelihood receivers is usually restricted to stochastically independent measures of likelihood L_i, because the analysis is otherwise much more difficult. However, there are some powerful results in the theory of asymptotic distributions of maxima that show that the statistics of the maximum are not affected by correlations of the individual variables as long the correlation is less and less as the variables are more and more separated (see Galambos,[38] especially Theorems 2.1.3 and 3.8.2). Thus the assumption of independence should be understood as an analytic convenience and not as a major restriction on the applicability of the results. Even so, for any physiologically plausible modeling it will be necessary to understand the consequences of dropping the independence assumption.

Instead of assuming white additive Gaussian noise, we could instead assume photon noise, which is Poisson. In that case too, the simple maximum-likelihood receiver may be implemented with linear receptive fields and a maximum rule. The difference is that, whereas in Gaussian noise the weighting functions of the receptive fields were proportional to the expected signals $s_i(x, y, t)$, in Poisson noise the weighting functions are $\ln[1 + s_i(x, y, t)]$.[39]

G. Discussion

Section 2 has shown that two attractive assumptions about human detection imply a maximum-of decision rule applied to many independent variables. Here we saw that the optimal decision strategy is generally a maximum-of rule and that when the signals are equally probable it is equivalent to a maximum-likelihood rule. Further specifying that the equally probable signals are also equal energy and orthogonal and in white noise yields a concrete model whose performance can be analyzed and compared with human performance.

4. UNCERTAINTY MODEL

A. Definition of the Uncertainty Model

In the preceding section the scope was broadened to consider arbitrary detection–identification tasks. This paper is not concerned with identification experiments, so the scope will be narrowed again to the tasks of interest. However, it is worth remembering that the simple maximum-likelihood receiver (i.e., the uncertainty model) can do identification tasks (by choosing the expected signal s_i corresponding to the largest hypotheses L_i) and is optimal if all the choices are simple.

The models developed in Sections 2 and 3 are close approximations to what we need, but, for clarity, I will begin with a clean slate and define the uncertainty model by seven assumptions:

Assumption 1. The decision-variable assumption (see Section 2).

Consider what the relevant and irrelevant variables in Fig. 3 might represent. When an observer makes a decision about a stimulus we usually consider that he or she monitors the activity of various sensors, or channels, and uses those data as the basis for the decision. Each variable in the model takes on a single value when a stimulus is presented, so I will call the variables *samples* from the activity of whatever sensors or channels the subject is monitoring. As before, the samples may be contrast dependent, and therefore relevant, or contrast independent and therefore irrelevant.

Assumption 2. The decision variable is the maximum of M samples.

Assumption 3. At zero contrast, all M samples are identically distributed and stochastically independent.

These assumptions imply the first three parts of the probability-summation assumption but not the final part: correction for guessing. Of the M samples, some number K will be relevant (i.e., contrast dependent). Appendix A shows that when fewer than about 1% of the samples are relevant, $K/M \leqq 1\%$, then the uncertainty model permits correction for guessing of both yes–no and 2afc hit rates.

For most of the predictions in this paper it would suffice at this point to assume that for repeated presentations of any given stimulus the mean values of the samples are linear functions of the stimulus and that the samples are normally distributed with a variance independent of contrast. However, I prefer to make stronger assumptions, 4–6 below, which will allow the model to make predictions of noise-masking experiments. (The reader may skip to Assumption 7 and return here later if necessary.)

Define the *effective* stimulus as the sum of the actual stimulus and an equivalent noise. Equivalent noise is the spatiotemporal noise that would have to be added to the stimulus to model the observer as a black box that is otherwise noise free.

Assumption 4. Each sample is a weighted sum over space and time of the effective stimulus. In other words, each sample has a linear receptive field.

Assumption 5. All M samples' receptive fields are orthogonal, and the equivalent noise is white.

For two functions to be orthogonal means that their integrated product is zero. For example, $\sin(x)$ and $\cos(x)$ are orthogonal because their product $\sin(x)\cos(x) = \sin(2x)/2$ has an average value of zero. The orthogonality requirement would be satisfied by, for example, expected signals that do not overlap in space, or do not overlap in time, or that do not share any spatial frequencies.

White noise is Gaussian and uncorrelated over space and time. A uniform retinal illuminance produces photon absorptions that are uncorrelated over space and time at the retina.

Appendix B proves that Assumptions 4 and 5 imply that the samples are normally distributed and stochastically independent.

Assumption 6. The equivalent noise is independent of the stimulus.

This makes the sample variance independent of contrast. The next assumption is needed for the modeling of summation experiments in Section 6.

Assumption 7. The K relevant samples are identically distributed.

The uncertainty model is illustrated in Fig. 4. On the left is the effective stimulus: the sum of the stimulus and the equivalent noise. The effective stimulus is a function of space and time. Each pathway represents a receptive field. Mathematically, the action of a receptive field is called cross correlation. The effective stimulus is multiplied by the weighting function of the receptive field, and the resulting product function is integrated over space and time, yielding a single variable L_i. The samples whose receptive fields are sensitive to the signal will be relevant, and the samples whose receptive fields are insensitive to the signal will be irrelevant. As before, the largest value is passed by the long box to become the decision variable L.

Each sample is a single value for each stimulus presentation. Think of the receptive fields $s_1(x, y, t), \ldots, s_M(x, y, t)$ as expected signals. Each sample tells us about the likelihood that its expected signal was presented. To achieve optimal performance, the observer should use one receptive field per possible signal, each with a weighting function proportional to the expected signal. The model shows M receptive fields. M is the degree of uncertainty of the observer. With no uncertainty, M would be 1.

As before, on a yes–no trial the observer would say "yes" if the decision variable exceeded the subjective criterion and would say "no" otherwise. On a 2afc trial the observer would choose the interval that resulted in the larger value of the decision variable.

B. Discussion

The uncertainty model does not distinguish between *intrinsic* uncertainty (i.e., the observer's inability to use prior information about the signal's identity) and *extrinsic* uncertainty (i.e., the presentation of a signal that is a random sample from many possible signals). Experiments with extrinsic uncertainty are called uncertainty experiments.

Some of the attempts to explain human performance by models like the uncertainty model did not allow for intrinsic uncertainty and assumed that M would equal the extrinsic uncertainty.[13,17–20] Some did allow for an unknown intrinsic uncertainty.[5,7,8,11,14–16] Allowing for an unknown intrinsic uncertainty, I will take M as a free parameter in modeling human performance.

M expresses the subject's degree of uncertainty. With no uncertainty, M would be 1. We will see that estimates of M range from 30 to 10,000. Large changes in M are required to make small changes in behavior of the model, so think of log M as the parameter to be estimated.

Questions arise almost immediately. Which M samples? Along what dimension(s) is the observer uncertain? These questions will not be answered in this paper, because they have little effect on the model's predictions. As a guess, I would say that the observer is uncertain about time and place.

C. Normalized Contrast c'

For the experiments to be considered here the equivalent noise level was unknown. Furthermore, the samples' expected signals were unknown. However, these unknowns add only a single degree of freedom to the uncertainty model. Define normalized contrast c' as the ratio of the mean (in response to a signal) to the standard deviation (in response to a blank) of any of the K relevant samples.

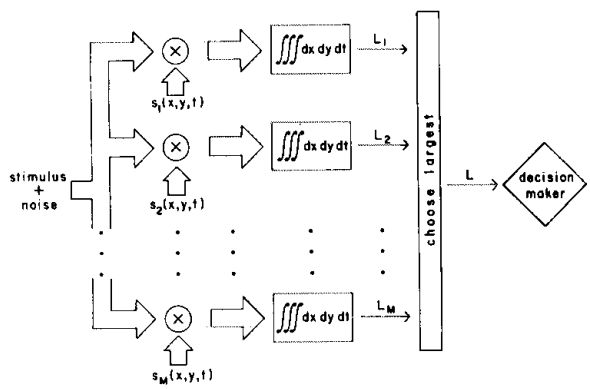


Fig. 4. The uncertainty model. Each receptive field multiplies the input (stimulus plus noise) with an expected signal $s_i(x, y, t)$ and integrates the result (i.e., it evaluates the cross correlation), yielding a single number L_i on each presentation. The largest L_i is passed to the decision maker on the right. The box says "choose largest" to indicate that the identity as well as the value of the most active receptive field is preserved. This allows the model to be applied to identification experiments. Adapted from van Trees[28] Fig. 4.22.

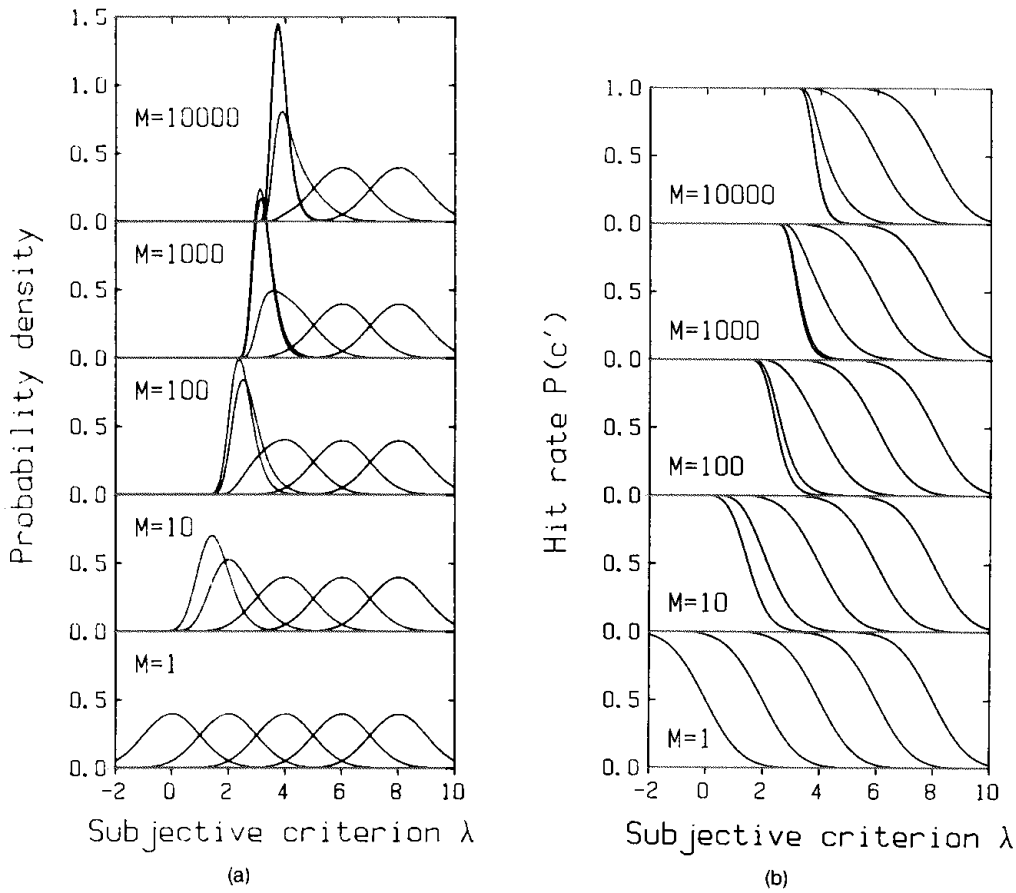


Fig. 5. The probability distribution of the decision variable L in the uncertainty model. Each curve in (a) shows the probability density for the decision variable L equaling the subjective criterion λ, $L = \lambda$, as a function of λ. Each curve in (b) shows the yes–no hit rate $P(c')$, which is the probability of the decision variable's exceeding the subjective criterion λ, $L > \lambda$, as a function of λ. Each figure has separate panels for five degrees of uncertainty, M = 1, 10, 100, 1,000, 10,000. Each panel has separate curves for five normalized contrasts, c' = 0, 2, 4, 6, 8. K is 1.

Physical contrast c is proportional to normalized contrast c':

$$c = Ac', \tag{4.1}$$

where the proportionality constant A is a parameter in the uncertainty model. A depends on the signal, the expected signals of the relevant samples, and the equivalent noise level.

D. Probability Distribution of the Decision Variable

The distribution of the decision variable depends on only three things: the normalized contrast c', the number M of samples, and the number K of the samples that are relevant. (The samples $L_1, \ldots, L_M$, the decision variable L, and the subjective criterion λ are all measured in units of the standard deviation of the samples.) Figures 5(a) and 5(b) show the distribution of the decision variable in different ways. Each curve in Fig. 5(a) shows the probability density for the decision variable L equaling the subjective criterion λ, $L = \lambda$, as a function of the subjective criterion λ. Each curve in Fig. 5(b) shows the yes–no hit rate $P_{yn}(c')$, which is the probability of the decision variable's exceeding the subjective criterion λ, $L > \lambda$, as a function of the subjective criterion λ. Each figure has separate panels for five degrees of uncertainty M = 1, 10, 100, 1000, 10,000. Each panel has separate curves for five normalized contrasts, c' = 0, 2, 4, 6, and 8. Figure 5 assumes that only one sample is affected by the signal (K = 1), but the conclusions of our inspection of it will be general to all K.

The bottom row in each figure has no uncertainty, M = 1. The distributions are Gaussian with unit variance and mean equal to the normalized contrast c'. This is the ideal detector for a signal known exactly, such as those considered by Rose[1] and De Vries.[2] (They assumed the Poisson noise, but, by the central-limit theorem, the distribution of the samples would be approximately Gaussian anyway.)

The leftmost curve on each row shows the distribution at zero contrast, i.e., a blank presentation. The hit rate at zero contrast is called the false-alarm rate. On a blank presentation the decision variable is the maximum of M zero-mean Gaussians. The zero-mean Gaussian distribution is shown by the leftmost curve (c' = 0) in the bottom row (M = 1) of each figure. Increasing the uncertainty M makes this noise-only distribution narrower and shifts it to the right, approximately in proportion to the log of M.

Recall that the correction-for-guessing assumption is satisfied if at zero contrast the maximum irrelevant variable is almost always greater than the maximum relevant variable. That is, the irrelevant distribution must be at higher values and have little overlap with the relevant distribution. In Fig. 5(a), as the uncertainty increases the relevant distribution shifts further and further to the right of the irrelevant distribution, and the overlap becomes vanishingly small. For uncertainties M of at least 100, to keep the false-alarm rate below 0.9 the subjective criterion λ must be at least 2. For K

= 1, the probability of a relevant hit is shown by the leftmost curve in the bottom panel. At $\lambda = 2$ the probability of a zero-contrast relevant hit is only 0.023.

So far we have considered only detection, which is really the discrimination of nonzero from zero contrast. Uncertainty makes the zero-contrast decision-variable distribution indistinguishable from its distribution at low contrast, making low contrast undetectable. Contrast discrimination also includes discrimination of two nonzero contrasts. A glance at Figs. 5(a) and 5(b) shows that at suprathreshold contrasts the decision variable is unaffected by the uncertainty. Thus the uncertainty has no effect on discrimination of suprathreshold contrasts. Contrast discrimination will be considered further in Section 9.

E. Yes-No Performance

It is given that the samples are independent and normally distributed with equal variance. For convenience the samples are measured in units of the common standard deviation, so a sample responding to noise alone has probability $\Phi(\lambda)$ of not exceeding a value λ, where $\Phi(\lambda)$ is the cumulative normal integral:

$$\Phi(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} \exp(-t^2/2td). \tag{4.2}$$

A sample responding to its expected signal as well as the noise has probability $\Phi(\lambda - c')$ of not exceeding λ.

In a yes–no detection task the probability of a hit as a function of contrast is $P_{yn}(c')$:

$$P_{yn}(c') = 1 - (1 - G_{yn})[1 - P^*_{yn}(c')], \tag{4.3}$$

where K may have any integer value from 1 to M. The relevant-hit rate $P^*_{yn}(c')$ is the probability that any of the K contrast-dependent samples will exceed the subjective criterion:

$$P^*_{yn}(c') = 1 - \Phi^K(\lambda - c'). \tag{4.4}$$

The irrelevant-hit rate G_{yn} is the probability that noise in the $M - K$ non-signal-bearing samples will cause one of them to exceed the subjective criterion,

$$G_{yn} = 1 - \Phi^{M-K}(\lambda). \tag{4.5}$$

Substituting Eqs. (4.4) and (4.5) into Eq. (4.3) yields the hit rate, the probability of a "yes" on a signal presentation,

$$P_{yn}(c') = 1 - \Phi^{M-K}(\lambda)\Phi^K(\lambda - c'). \tag{4.6}$$

At zero contrast, $c' = 0$, we have the false-alarm rate, the probability of a "yes" on a blank presentation,

$$P_{yn}(0) = 1 - \Phi^M(\lambda). \tag{4.7}$$

These equations parameterize the ROC curve and have been plotted by Nolte and Jaarsma[36] for $K = 1$ and several values of c' and M. Nachmias and Kocher[7] noted that their ROC curves for detection of a spot were similar to curves produced by Eqs. (4.6) and (4.7) with $K = 1$ and $M \approx 32$.

F. Two-Alternative Forced-Choice Performance

The proportion correct in a 2afc is the probability that the largest value in the signal interval will exceed the largest value in the other interval and is given by the area under the ROC curve[3]:

$$P_{fc}(c') = \int_{\lambda=+\infty}^{\lambda=-\infty} P_{yn}(c')\mathrm{d}P_{yn}(0). \tag{4.8}$$

The integration goes from plus to minus infinity because the false-alarm rate $P_{yn}(0)$ increases as the subjective criterion λ decreases. Substituting our expressions for $P_{yn}(0)$ and $P_{yn}(c')$ from Eqs. (4.6) and (4.7) yields

$$P_{fc}(c') = 1 - M \int_{-\infty}^{+\infty} \Phi^K(\lambda - c')\Phi^{2M-K-1}(\lambda) \times \frac{\mathrm{d}\Phi}{\mathrm{d}\lambda}(\lambda)\mathrm{d}\lambda. \tag{4.9}$$

The integral is evaluated by numerical integration. $\Phi(\ )$ is computed by Eqs. (26.2.17) and (26.2.13) from Ref. 40.

G. Summary

Assumptions 1–7 make it unnecessary to know the actual and expected signals. The model's performance depends only on the signal contrast c and four model parameters: the subjective criterion λ (only in yes–no), the number M of samples, the number K of the samples that are relevant, and the proportionality constant A between physical and normalized contrast.

5. CONTRAST DETECTION

This section examines 2afc and yes–no contrast detection by the uncertainty model.

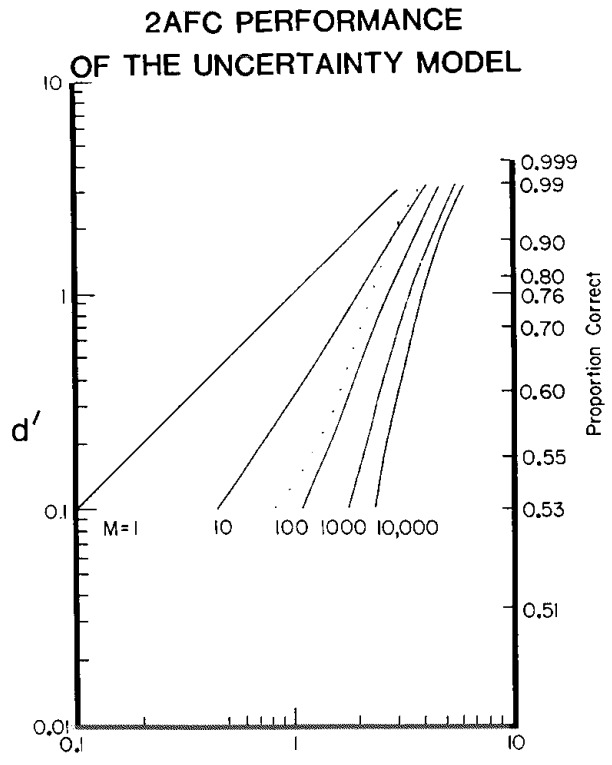
From here on, the paper takes an empirical approach, relying primarily on Monte Carlo simulations of the model performance, analyzed in the same way as human data. Again and again, this will reveal the similarity of the performance of the model and human observers. In addition, a few model properties will appear that make strong predictions for future experiments.

A. Psychometric Functions

The solid curves in Figs. 6(a) and 6(b) show psychometric functions for 2afc contrast detection by the uncertainty model. The curves are given by Eq. (4.9) with $K = 1$. Figure 6(a) has logarithmic scales, and Fig. 6(b) has linear scales. The horizontal axis is the normalized contrast c', and the vertical axis is d'. [d' is $\sqrt{2}$ times the normal deviate $\Phi^{-1}(P)$ corresponding to the proportion correct.[3]] Proportion correct is also indicated, on a scale at the right. The functions are shown for values of M ranging from 1 to 10,000. (The dotted curves are the approximation (3.14) of Peterson *et al.* for $M = 100$ and are obviously inaccurate for these values. The curves are nearly straight in the log–log coordinates of Fig. 6(a). Increasing the uncertainty M increases threshold and steepens the log–log slope of the psychometric function, although large changes in uncertainty are required to produce a small change in slope.

Several features of Fig. 6(b) are noteworthy. First, at high uncertainties, $M \geq 100$, the curves are parallel; each further increase of M by a factor of 10 simply slides the previous curve over to the right. The finding that increasing the log uncertainty is equivalent to subtracting from the contrast is derived analytically in Appendix D.

Second, note that each curve is asymptotically straight with unit slope. This reflects the fact, illustrated in Figs. 5(a) and 5(b), that the decision-variable distributions for supra-



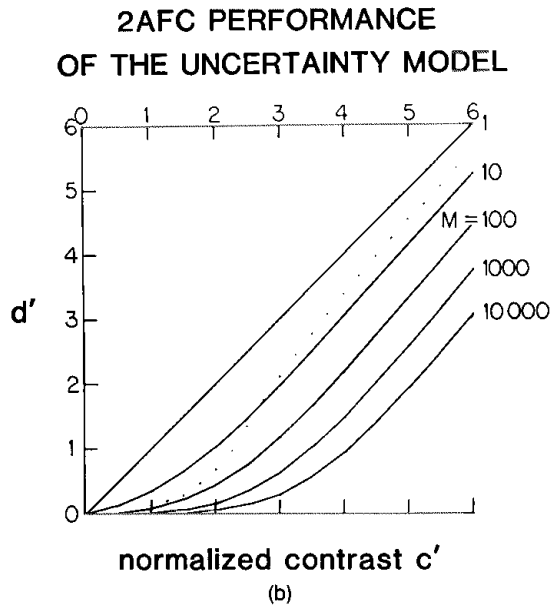


Fig. 6. 2afc detection by the uncertainty model. The solid curves show the detectability d' as a function of normalized contrast c' and uncertainty M. K is 1. The dotted lines show the approximation of Peterson *et al.*[35] for $M = 100$. (a) Shows that the curves are nearly straight in log–log coordinates. (b) Shows that in linear coordinates the curves tend to translate along the contrast axis and that above threshold ($d' > 1$) they all have unit slope.

threshold contrasts are unaffected by the uncertainty M. Indeed, Fig. 6(b) would be well described by supposing ideal detection of an effective contrast equal to the actual contrast minus a threshold linearly related to the log of M (see Appendix D).

Yes–no psychometric functions of the uncertainty model appear in Fig. 5(b) and were discussed in Section 4.

B. Weibull Fits

For the past 10 years psychometric functions for visual contrast detection have usually been described by the best-fitting parameters of a Weibull function

$$P(c) = 1 - (1 - \gamma)\exp[-(c/\alpha)^{\beta}], \tag{5.1}$$

where γ is the false-alarm rate $P(0)$, α is the threshold contrast at which $P(c) = 1 - (1 - \gamma)/e$, and β controls the log–log steepness. Therefore I have simulated an observer performing 2afc and yes–no detection by the rules of the uncertainty model, and fitted Weibull functions to the performance data, to obtain maximum-likelihood estimates of the parameters of the Weibull function. The resulting estimates of α, β, and γ may be directly compared with corresponding estimates for human observers. Additionally, since human 2afc data are occasionally fitted by a d' power law, $d' = (c/c_{0.76})^b$, I have determined the parameters of that fit as well.

All the results of the model relate to a normalized contrast scale, indicated by a prime, as in c' and α'. Both yes–no and 2afc share the same scale factor A between physical threshold contrast α and normalized threshold contrast α':

$$\alpha_{yn}/\alpha'_{yn} = \alpha_{fc}/\alpha'_{fc} = A. \tag{5.2}$$

The results appear in Table 1 (2afc) and Table 2 (yes–no). There are three dependent variables α', β, and γ, and two model parameters, M and λ (only in yes–no). The model parameter K was set to 1. The fourth model parameter, A, does not appear, as it merely provides scaling of contrast to match human data.

C. Normalized Threshold Contrast α'

In both Tables 1 and 2, a linear regression of α' on β shows that α' and β are nearly equal:

$$\alpha' \approx \beta. \tag{5.3}$$

This is for a K of 1. Section 6 will develop a more general formula for all K small relative to M. I do not have an intuitive explanation for the equality, but it can be shown that proportionality of α' and β is implied by the fact that high uncertainty merely subtracts from contrast (see Appendix D).

D. Log–Log Steepness β

Applying a linear regression to the data in Table 1 shows that β is a linear function of the log of the uncertainty M:

$$\beta = 1.358 + 0.792 \log M \pm 0.043. \tag{5.4}$$

The nearly straight curves in Fig. 6(a) suggest that we might characterize each curve by two numbers indicating its log–log slope b and its threshold $c'_{0.76}$, at which $P_{fc}(c') = 76\%$, and $d' = 1$. These two values, b and $c'_{0.76}$, determine a line in our log–log coordinates:

$$\log d' \approx b \log(c'/c'_{0.76}),$$

or, equivalently,

$$d' \approx (c'/c'_{0.76})^b. \tag{5.5}$$

Table 1. 2afc Performance of the Uncertainty Model[a]

M	Weibull α'	Weibull β	$d' = (c'/c'_{0.76})^b$ $c'_{0.76}$	$d' = (c'/c'_{0.76})^b$ b
1	1.25	1.41	1.00	1.01
3	1.72	1.73	1.44	1.26
10	2.22	2.13	1.94	1.56
30	2.65	2.50	2.36	1.85
100	3.07	2.93	2.79	2.20
300	3.41	3.28	3.14	2.51
1,000	3.75	3.72	3.49	2.89
3,000	4.04	4.16	3.79	3.29
10,000	4.31	4.55	4.06	3.70
30,000	4.54	4.97	4.29	4.11
100,000	4.77	5.26	4.52	4.45

[a] Each trial was a Monte Carlo simulation of an observer with a psychometric function given by Eq. (4.9), with parameters M (as listed) and $K = 1$. The threshold α' and steepness β are the parameters of the Weibull function for a maximum-likelihood fit[42]; the threshold $c'_{0.76}$ and steepness b are the parameters of the d' power law [i.e., a straight line in Fig. 6(a)] for a maximum-likelihood fit. Each fit was made to the result of 100,000 Monte Carlo trials distributed over 10 contrasts. The table entries are the average of 10 replications. The Weibull parameter γ was fixed at 0.5 since these are 2afc trials.

Table 2. Yes–No Performance of the Uncertainty Model[a]

λ	α'			β			γ		
0	0.9			1.5			0.50	1.00	1.00
1	1.5			1.8			0.16	1.00	1.00
2	2.3			2.4			0.025	1.00	1.00
3	3.4	3.5		3.3			0.0006	0.75	1.00
4	4.4	4.4		4.4	4.5		0.0000	0.032	1.00
5	5.4	5.4	5.4	5.6	5.7	5.1	0.0000	0.000	0.24
6	6.4	6.4	6.4	6.6	6.7	6.7	0.0000	0.0000	0.001
7	7.5	7.4	7.4	7.6	7.6	7.7	0.0000	0.0000	0.0000
8	8.4	8.5	8.4	8.9	8.8	8.7	0.0000	0.0000	0.0000
9	9.5	9.5	9.5	9.6	9.9	9.8	0.0000	0.0000	0.0000
$M = 1$	10^3	10^6	1	10^3	10^6	1	10^3	10^6	

[a] Each trial was a Monte Carlo simulation with a hit rate given by Eq. (4.6), with parameters λ and M, as listed, and $K = 1$. The subjective criterion λ is listed in the left-hand column. The uncertainty M is listed in the bottom row. The threshold α', steepness β, and false-alarm rate γ are the parameters of the Weibull function for a maximum-likelihood fit.[42] Each fit was made to the result of 2,500 Monte Carlo trials distributed over five contrasts, including one near zero. Each table entry is the average of 10 replications. Only the significant digits are shown; α' and β have no significant digits when γ is near 1.

Table 1 lists the threshold $c'_{0.76}$ and slope b at each uncertainty M. On the basis of the similarity of form of the Weibull-function Eqs. (5.1)–(5.5), we would expect the two steepness measures β and b to be proportional. A fit to the data in Table 1 yields

$$b = 0.80\beta \pm 0.15. \tag{5.6}$$

Equations (5.4) and (5.6) can be used to estimate the uncertainty M directly from the measured steepness β or b of the 2afc psychometric function.

Table 2 shows the yes–no results. β is strongly dependent on the subjective criterion λ and independent of M. Applying a linear regression to the data in the table yields

$$\beta \approx \lambda + 0.6. \tag{5.7}$$

This subsection has found that in both yes–no and 2afc tasks the model performance is characterized primarily by β. However, β responds to different things in the two tasks. In yes–no, β is a linear function of the subjective criterion λ, yet, because of the decision-variable assumption, λ plays no role in 2afc. In 2afc, β is a linear function of the log uncertainty, yet, because of correction for guessing, the uncertainty affects only the false-alarm rate in yes–no.

E. False-Alarm Rate γ

In the 2afc fits of the Weibull function the false-alarm rate γ was fixed at 0.5.

In the yes–no fits of the Weibull function, correction for guessing assigns the false-alarm rate to γ and leaves the relevant-hit rate to be fitted with the parameters α and β. As a result, α and β are independent of the uncertainty M. The false-alarm rate γ depends only on the subjective criterion λ and the uncertainty M, by Eq. (4.7). Compare this with the irrelevant-hit rate G [Eq. (4.5)]. The false-alarm rate γ will be approximately equal to the irrelevant-hit rate G when K is small relative to M. This will satisfy the conditions for correction for guessing.

F. Two-Alternative Forced-Choice Experimental Results

Figure 7 shows two typical psychometric functions $P(c)$ for human 2afc contrast detection. The O's are from Nachmias and Sansbury,[10] and the ×'s are from Legge.[41] The horizontal axis is a logarithmic scale showing contrast of the grating to be detected, and the vertical axis on the left is a logarthmic

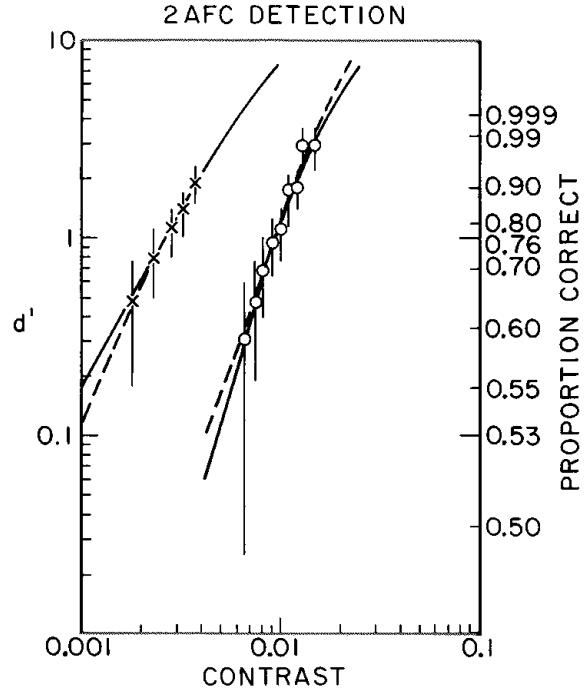


Fig. 7. 2afc detectability d' as a function of contrast c of a grating. Proportion correct are indicated on the right. The O's (from Nachmias and Sansbury[10]) are for detection of a 9-c/deg grating, and the ×'s (from Legge[41]) are for detection of a 3-c/deg grating. There are 160 trials per point. The vertical bars represent 95% confidence intervals. The solid curves are fits by the uncertainty model with $K = 1$ and $M = 800$ (O's) and $M = 18$ (×'s). The dashed lines are maximum-likelihood fits by Weibull functions with $\beta = 3.66$ (O's) and $\beta = 2.36$ (×'s).

scale of d'. Proportion correct is indicated on the right. The vertical bar through each point indicates the 95% confidence interval. The dashed curves show maximum-likelihood fits by the Weibull function.[42] The values of β are 3.66 for the O's and 2.36 for the ×'s. The solid curves show fits by the uncertainty model. K was assumed to be 1, and the best-fitting values of uncertainty M were 18 for the ×'s and 800 for the O's. The Weibull and uncertainty model fits are experimentally indistinguishable. Tanner[5] concluded that an uncertainty of 100 expected signals would account for the steep psychometric function that he measured for detection of a light increment. There is a large range in estimates of the uncertainty M, e.g., from 18 to 800, because $\log M$ is linearly related to β in 2afc, and estimates of β are variable.

The fits in Fig. 7 are nearly straight, with slopes greater than 1. This implies that d' is proportional to a power greater than 1 of contrast, $d' \propto c^b$, $b > 1$. In fact, Nachmias and Sansbury[10] originally fitted their data (O's) by a straight line with slope 2.9, representing the relation $d' \propto c^{2.9}$. In general, Weibull functions and 2afc predictions of the uncertainty model are essentially straight in these coordinates with a slope of approximately 0.8β, implying that d' is proportional to the contrast raised to the 0.8β power.

G. Discussion

The way in which the parameters of the Weibull function depend on the parameters of the uncertainty model is instructive. It seems likely that a reasonable correspondence could be established between the samples in the model and some unit of physiological organization, such as retinal or cortical neurons (see, e.g., Ref. 43). Then the parameters—M, the number of samples, and K, the number sensitive to the signal—might have a physiological interpretation.

The principal finding of this section was the importance of the log–log steepness β. In both yes–no and 2afc tasks the model performance is largely determined by β. 2afc β is a linear function of $\log M$, and yes–no β is a linear function of the subjective criterion λ. Thus the two tasks, analyzed in the same way, yield different kinds of information about the observer. Yes–no β tells us about the observer's subjective criterion. 2afc β tells us about the observer's uncertainty.

6. SUMMATION EFFECTS

A. Background

Over the past 30 years there has been some importance attached to studies of summation. A compound grating containing two disparate frequencies is slightly more detectable than either is alone.[24,44–46] Over a wide range, the contrast theshold for a sinusoidal grating gradually falls if the grating is extended in time[41,42] or space[47,48] (in a region of uniform sensitivity). Over this range, threshold α is inversely proportional to the extent raised to a small exponent, $-1/\beta$, usually called the index of summation:

$$\alpha \propto \text{extent}^{-1/\beta}. \tag{6.1}$$

These summation effects have been successfully attributed to probability summation assuming a Weibull distribution. The uncertainty model incorporates probability summation for yes–no (provided that $K \ll M$; see Appendix A) and 2afc [provided that $K \ll M$ and $\log(K + 1) \ll \log(2M)$; see Appendixes A and C] but assumes a Gaussian distribution.[12,17,31] Is the Gaussian distribution similar enough to a Weibull for the uncertainty model to survive?

The probability-summation assumption was introduced in Section 2. However, by itself it does not predict the results of a summation experiment. The utility of the probability-summation assumption is greatly enhanced by the *Weibull assumption*, which has two parts. First, the relevant-hit rate $P_i^*(c)$ of each contrast-dependent reason is a Weibull function [Eq. (5.1)] with zero false-alarm rate γ:

$$P_i^*(c) = 1 - \exp[-(c/\alpha_i)^\beta], \tag{6.2}$$

where c is contrast and α_i and β are constants. α_i is a threshold parameter, being the value of c for which $P_i^*(c) = 0.63$, and β is a steepness parameter. Second, β is the same for all of the contrast-dependent reasons. (Nachmias[12] calls this the homogeneity assumption.) This completes the Weibull assumption.

Probability summation with the Weibull assumption has a long history before its introduction to vision by Brindley[49] and Quick,[50] preceding even Weibull[51] himself.[52,53] Galambos[38] points out that one of the reasons for the wide applicability of the Weibull function is that it has enough parameters to describe a wide range of distribution shapes adequately.

To understand the utility of the Weibull assumption, note that $P_i^*(c)$ in Eq. (6.2) is zero at zero contrast, more than satisfying the probability-summation assumption, expression (2.7). This reduces Eq. (2.1) to a Weibull function [Eq. (5.1)], where α is the observer's threshold:

$$\alpha = \left(\sum_{i=1}^{K} \alpha_i^{-\beta}\right)^{-1/\beta}. \tag{6.3}$$

The Weibull function is unique in that a probability sum,

$$P(c) = 1 - \prod_i 1 - P_i(c), \tag{6.4}$$

over many functions $P_i(c)$ that are identical except possibility for position on the log contrast axis (i.e., α_i) yields a function $P(c)$ that is identical except for position on the log contrast axis.[49,54] Thus the thresholds α_i of the individual contrast-dependent reasons affect only the observer's threshold α and not the form of the psychometric function $P(c)$. If all the thresholds α_i are equal (as is often assumed), then the observer's threshold α will be

$$\alpha = K^{-1/\beta}\alpha_i. \tag{6.5}$$

That is, the observer's threshold α should fall with increasing K, in proportion to $K^{-1/\beta}$.

In order to predict summation effects, the number K of contrast-dependent samples is assumed to be proportional to the signal extent, $K \propto$ extent, so Eq. (6.5) becomes relation (6.1). As mentioned above, relation (6.1) is an empirical fact (where β is the result of a fit of a Weibull function to the human psychometric function), so if the uncertainty model is to survive, it must at least approximately satisfy the equation, despite its lacking the Weibull assumption.

B. Threshold α

Section 5 noted that 2afc and yes–no psychometric functions of the uncertainty model are well fitted by Weibull functions

with parameters α, β, and γ. The dependencies of β and γ on the parameters of the uncertainty model were discussed there. This subsection will discuss α. Yes–no and 2afc will be discussed together because the same conclusions apply to α in both tasks.

The excellent fits of the Weibull function to the 2afc and yes–no psychometric functions means that the hit rate $P(c)$ of the uncertainty model is approximately equal to a Weibull function:

$$P(c') \approx 1 - (1 - \gamma)\exp[-(c'/\alpha')^{\beta}]. \quad (6.6)$$

For yes–no the hit rate $P(c')$ is given by Eq. (4.6). For 2afc the hit rate $P(c')$ is given by Eq. (4.9). Now consider just the case where K is 1,

$$P_1(c') \approx 1 - (1 - \gamma_1)\exp[-(c'/\alpha_1')^{\beta_1}], \quad (6.7)$$

where the subscript 1 indicates that $K = 1$. The Weibull function may be corrected for guessing. Appendix A shows that both yes–no and 2afc hit rates of the uncertainty model may be corrected for guessing if $K/M \leqq 1\%$. Applying correction for guessing [relation (2.5)] to both sides of relation (6.7) reveals the relevant-hit rate $P_1^*(c')$:

$$P_1^*(c') \approx 1 - \exp[-(c'/\alpha_1')^{\beta_1}]. \quad (6.8)$$

Subtracting both sides from 1, raising to the power K, and subtracting again from 1 yields

$$1 - [1 - P_1^*(c')]^K \approx 1 - \exp[-(K^{1/\beta_1}c'/\alpha_1')^{\beta_1}]. \quad (6.9)$$

We recognize the left-hand side as the observer's relevant-hit rate $P^*(c')$ when there are K independent and equally probable relevant reasons for a hit. [As is shown in Appendix C, probability summation among many independent reasons does hold for 2afc, provided that $K/M \leqq 1\%$ and that $\log(K + 1) \ll \log(2M)$.] The right-hand expression is a Weibull function,

$$P^*(c') \approx 1 - \exp[-(c'/\alpha')^{\beta}], \quad (6.10)$$

where

$$\alpha' = \alpha_1' K^{-1/\beta_1} \quad (6.11)$$

and

$$\beta = \beta_1. \quad (6.12)$$

This shows that β is approximately independent of K, provided that K and M satisfy the appropriate conditions of Appendixes A and C. As conclusions, Eqs. (6.11) and (6.12) are only approximations, since they are based on approximation (6.10). Note that Eq. (6.11) is the standard summation equation (6.5) and that Eq. (6.12) is one aspect of the homogeneity assumption.

Usually these results are derived exactly by assuming probability summation and the Weibull assumption. Here they were derived approximately given probability summation and that the Weibull function provides a good fit to the psychometric function when $K = 1$. However, any particular value of K could have been used. These results are not unique to the uncertainty model. In general, any model that incorporates probability summation and provides a good fit to human psychometric functions will approximately satisfy summation Eqs. (6.11) and (6.12), since human psychometric functions are well fitted by the Weibull function.

Returning to the particular case of the uncertainty model, we know from relation (5.3) that α_1' and β are approximately equal. Thus our formula [Eq. (6.11)] for α' becomes

$$\alpha' \approx \beta K^{-1/\beta}. \quad (6.13)$$

This applies to both yes–no and 2afc. This formula also shows that the key parameter to describe the state of the observer is the log–log steepness β. Note that this formula is stronger than the simple proportionality usually derived [relation (6.1) or Eq. (6.5)], as it predicts the effect of a change in β. Furthermore, since we already know how β depends on the subjective criterion and task, we can now predict how α depends on subjective criterion and task. These predictions are tested in Sections 7 and 8.

Relation (6.13) allows us to solve for the normalization factor $A = \alpha/\alpha'$ in terms of α, β, and K:

$$A \approx \alpha/(\beta K^{-1/\beta}). \quad (6.14)$$

α and β are directly measurable, and Section 7 shows a way to estimate K.

C. Experimental Findings

While in principle it would be more elegant to predict the summation effects from measurements of the psychometric function (i.e., β), in practice it is always done the other way around because the summation effects can be measured with considerably more precision. Therefore we will ask whether thresholds at several signal extents can predict the β of the psychometric function.

First we must, somewhat arbitrarily, specify how K depends on signal extent. As in standard summation, I assumed it to be proportional to signal extent and arbitrarily set it to be 1 for the smallest signal extent:

$$K = \text{duration}/(100 \text{ msec}). \quad (6.15)$$

Both 2afc and yes–no summation will be examined.

The ×'s in Fig. 8 are the threshold contrasts reported by Legge[48] for 2afc detection of a 3-cycle/deg (c/deg) grating as a function of duration, from 100 to 2000 msec. The dashed line shows Legge's fit by the standard summation equation [relation (6.1)], yielding an estimate of β of 2.60. Of course, Eq. (6.11) would make the same prediction, but it is only an approximation for the uncertainty model, and its derivation put conditions on K and M. Therefore, I fitted the uncertainty model directly to the data, using its two free parameters, M and A, for best fit by eye. In these log–log coordinates, the uncertainty M controls the slope of the fit, and the scaling factor A controls its vertical position. The result is the solid curve showing the performance of the uncertainty model with $M = 60$ and $A = 0.00133$. (Note that the performance of the model is defined only at integer values of K; the continuous curve connects those points.) Following Legge, I did not include the point at 1000 msec in the fit. Both fits are reasonably close to the data and are similar to each other.

Are these fits consistent with the psychometric functions? Figure 7 shows the psychometric function[48] for detection of the 180-msec-duration signal. It is well fitted by a Weibull function [Eq. (5.1)] with exponent $\beta = 2.44$, shown by the

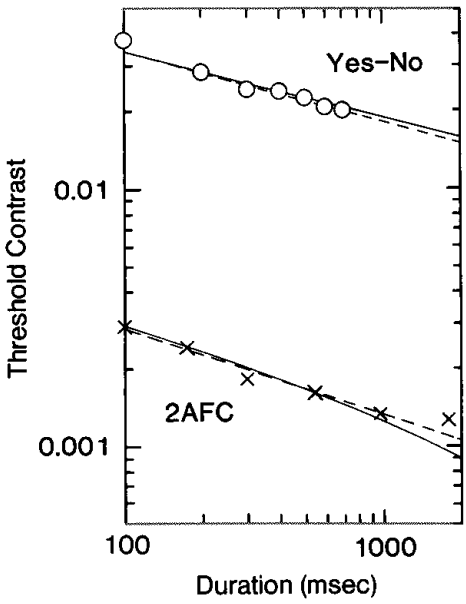


Fig. 8. Threshold contrast for a sinusoidal grating as a function of duration. The ×'s are 2afc thresholds (from Legge[41]), and the O's are yes–no thresholds (from Watson[42]). The dashed lines show the fit by the standard summation formula [relation (6.1)] as made by the original authors. The solid curves show the fit by the uncertainty model. For the ×'s, $M = 60$ and $A = 0.00133$. For the O's, $\lambda = 3.5$, $A = 0.00304$, and $M = 3000$. Both yes–no fits ignore the point at 100 msec, and both 2afc fits ignore the point at 1096 msec.

dashed curve. Recall that the fit of the standard summation equation to Fig. 8 yielded a β estimate of 2.60. A fit (not shown) to the psychometric function of the uncertainty model [Eq. (4.9)] for $K = 2$ (the nearest point to 180 msec) also yields $\beta = 2.60$. These values are not significantly different from 2.44. The important point is that the uncertainty model describes the 2afc human data just as well as the Weibull assumption does.

Now consider yes–no detection. The O's in Fig. 8 are yes–no contrast thresholds from Watson[42] for detecting a 4-c/deg grating, as a function of duration, from 100 to 800 msec. His fit by the standard summation formula [relation (6.1)] is shown by the dashed line running through most of the points. He (and I) excluded the 100-msec point from the fit. The fit by the uncertainty model had two degrees of freedom, λ (controlling slope) and A (controlling vertical position). The best fit ($\lambda = 3.5$, $A = 0.00122$) is shown by the solid line. The parameter M does not affect α, and therefore does not affect this fit, and was arbitrarily set to 3000. The fits are superimposed and experimentally indistinguishable.

The fit by the standard summation formula yielded a β estimate of 3.69. For values of K from 1 to 8, the uncertainty model, with parameters determined by the fit, yields values of β with a mean of 4.0. Watson reported 52 estimates of β with a mean of 4.65 and a standard deviation of 1.4. Once again, the standard summation formula is in reasonable agreement with the data, and the uncertainty model does at least as well.

Thus the 2afc and yes–no summation found by Legge and Watson are well accounted for by probability summation with either the Weibull assumption or the uncertainty model.

The reader may wonder whether the observer might not vary his uncertainty M or subjective criterion λ as the signal extent is varied. This complication is ruled out for the yes–no experiments, since Watson took the precaution of randomly interleaving all the different signal durations.

D. Conclusions

Experimentally, it is well known that the small index of summation $-1/\beta$ found in summation experiments accurately predicts the high steepness β of the psychometric function. This has usually been explained by assuming probability summation and the Weibull assumption. However, the Weibull assumption cannot be exactly true; the physics of light imply that the relevant-hit rate cannot be exactly zero at zero contrast. Furthermore, Nachmias[12] noted that assuming probability summation and the Weibull assumption for yes–no implies that they cannot hold exactly for 2afc.

Given the experimental finding that human psychometric functions are well fitted by the Weibull function, we can expect any acceptable model of detection to have psychometric functions well fitted by the Weibull function. It was shown here that this, combined with probability summation, is sufficient to predict the standard summation equation, at least approximately.

Although the uncertainty model and the Weibull assumption assume different psychometric functions, it seems that the range over which they can be measured is too narrow to distinguish them. Thus the probability summation of the uncertainty model does not appear to be experimentally distinguishable from probability summation with the Weibull assumption.

The uncertainty model differs from popular probability-summation models principally in assuming a Gaussian distribution rather than a Weibull distribution for each sample. The Weibull distribution is physically implausible in that it excludes the possibility of a detect state at zero contrast, requiring the supposition of a separate guessing process to account for the false-alarm rate. This high threshold assumption is untenable. However, the Weibull distribution is easy to work with; summation calculations are particularly easy. The Gaussian distribution is physically plausible but unwieldy; even the simplest calculations require computer-assisted numerical integrations. What I have shown here is that we can have the best of both worlds. By assuming the Gaussian distribution we can avoid the high-threshold assumption and incorporate a decision variable, resulting in a good treatment of noise, task, and subjective criterion effects. However, we can also approximate the model's psychometric function by the best-fitting Weibull distribution and then easily and accurately estimate summation effects. It is likely that comparable results could be obtained with other physically plausible distributions.[31,32]

7. EFFECTS OF THE SUBJECTIVE CRITERION

The decision-variable assumption implies that yes–no performance will be strongly dependent on the observer's subjective criterion λ and that, if the hit rate is analyzed into relevant- and irrelevant-hit rates, then both rates will covary with the subjective criterion. Contrary to this, the high-threshold assumption claims that the irrelevant- and relevant-hit rates may be increased or reduced independently of each other. Nachmias[12] dramatized the failure of the high-threshold assumption by fitting the Weibull function to

Table 3. Data from Nachmias[a]

Condition	Yes–No			2afc	
	α	β	γ	α	β
A	0.0380	6.20	0	0.0196	2.50
	0.0254	3.69	0.006		
	0.0209	2.78	0.121		
B	0.0292	4.84	0	0.0167	2.89
	0.0200	3.24	0.1		
	0.0184	2.97	0.57		
C	0.1470	4.17	0.003	0.0726	2.31
	0.0909	3.01	0.021		
	0.0609	2.15	0.215		
D	0.0470	7.59	0.00193	0.0338	3.54
	0.0395	5.43	0.0737		
	0.0377	4.44	0.461		
E	0.0239	4.07	0.0566		
	0.0179	3.87	0.32		
	0.0169	3.05	0.63		
F	0.0782	6.18	0.0532		
	0.0688	5.61	0.1712		
	0.0669	4.43	0.4775		
G	0.0976	5.53	0.012		
	0.0823	4.49	0.0954		
	0.0754	3.74	0.38		

[a] Ref. 12. By conducting a rating-scale experiment at several contrasts, Nachmias obtained yes–no psychometric functions for three subjective criterion levels simultaneously. Furthermore, for some of the conditions, a 2afc psychometric function was measured under the same conditions. The table lists maximum-likelihood estimates of the parameters of a Weibull function fit to each psychometric function. For 2afc, γ was fixed at 0.5.

yes–no data collected over a wide range of false-alarm rates. The Weibull function describes the relevant-hit rate by the parameters α and β and irrelevant-hit rate by the false-alarm rate γ. Nachmias found that both α and β fell as the false-alarm rate γ increased, contrary to the high-threshold assumption but qualitatively consistent with the decision-variable assumption. However, Nachmias did not present any quantitative explanation for his finding. This section will compare the uncertainty model's predictions with Nachmias's results.

Nachmias's data are listed in Table 3. There were seven conditions, A–G. For each condition, Nachmias measured rating responses to several contrasts. He then summed the rating frequencies to obtain three yes–no hit-rate functions for each condition. Columns 2–4 of Table 3 list the parameters of a Weibull function fitted by maximum likelihood to each yes–no hit-rate function.

By using the results of Sections 5 and 6 it is easy to estimate the parameters of the uncertainty model from the yes–no data. First consider α. We have relation (6.13), that is, $\alpha \approx A\beta K^{-1/\beta}$. There are three (α, β) pairs for each condition, so the appropriate fitting procedure is to find values of the parameters A and K to minimize the least-squares error in estimating $\log \alpha$ by $\log(A\beta_i K^{-1/\beta_i})$:

$$\min_{A,K} \sum_{i=1,3} [\log(\alpha_i) - \log(A\beta_i K^{-1/\beta_i})]^2. \tag{7.1}$$

This is simplified by the change of variables

$$a_i = \log(\alpha_i/\beta_i),$$

$$b_i = -1/\beta_i$$

to

$$\min_{A,K} \sum_{i=1,3} [a_i - \log(A) - b_i \log(K)]^2, \tag{7.2}$$

where the a_i and b_i are known and $\log(A)$ and $\log(K)$ are unknown constants. This is solved by doing a linear regression of the b_i on the a_i, yielding $\log(A)$ and $\log(K)$. Surprisingly, the resulting estimates of K were mostly less than 1, even though the model requires K to be at least 1. This arose because, in five of the seven conditions, $\log(\alpha/\beta)$ has a negative correlation with $-1/\beta$. The model [see relation (6.13)] can account for positive correlation (when $K > 1$), or zero correlation (when $K = 1$), but not negative correlation because K cannot sensibly be less than 1. The other two conditions (E and C) yielded estimates for K of 3 and 1, respectively. Since K cannot be less than 1, and since even case C suggests a small value, I set K to 1 for all the conditions, so relation (6.13) reduces to $\alpha \approx A\beta$. For least-squares error in estimating $\log \alpha$, the quantity $\log A$ is then the mean of $\log \alpha/\beta$. Parameter A is listed in the second column of Table 4.

Nachmias's data from Table 3 are plotted in Figs. 9(a)–9(c) as letters A–G. Each condition is represented by two or three points connected by dashed lines. The model's predictions are plotted as solid curves for M = 1, 10, 100, 1,000, 10,000, 100,000, 1,000,000. Since the parameter K was set to 1, the model has two free parameters, M and A, in fitting each condition. The figures plot the normalized threshold contrast α' instead of the physical threshold contrast α, where $\alpha' = \alpha/A$ and the values of A are listed in Table 4.

Figure 9(a) shows Nachmias's data (A–G) plotted as normalized threshold contrast α' versus β. The solid curves show the predictions of the uncertainty model for $K = 1$ and values of M from 1 to 1,000,000. Note that all the curves for different values of M are virtually the same, reflecting the independence of α' and β from M. Nachmias's data exhibit the predicted proportionality of α and β. In the uncertainty model, the yes–no β is linearly related to the subjective criterion λ [relation (5.7)]. So, according to the uncertainty model, this graph shows that threshold is linearly related to the subjective criterion λ.

The solid curve in Fig. 9(a) has unit slope because K was set to 1. The slope would be greater for larger K. The fit cannot be improved by increasing K because the data have slopes mostly less than 1. If the reader examines Fig. 9(a) closely it will be apparent that five of the seven conditions would be better fitted with slightly shallower lines, that is, with a K less

Table 4. Yes–No Fit Predicts 2afc Results[a]

Condition	Yes–No Fit		2afc Prediction/Data	
	A	M	$\hat{\alpha}/\alpha$	$\hat{\beta}/\beta$
A	0.0069	10	0.77	0.9
B	0.0061	100	1.08	1.0
C	0.031	10	0.94	1.0
D	0.0073	100,000	1.15	1.5
E	0.0053	1,000		
F	0.013	1,000,000		
G	0.019	10,000		

[a] The second and third columns show estimates of the parameters A and M of the uncertainty model (with $K = 1$) to fit the yes–no data in Table 3. The last two columns show the ratios of the resulting 2afc predictions, $\hat{\alpha}$ and $\hat{\beta}$, to the actual 2afc data, α and β.

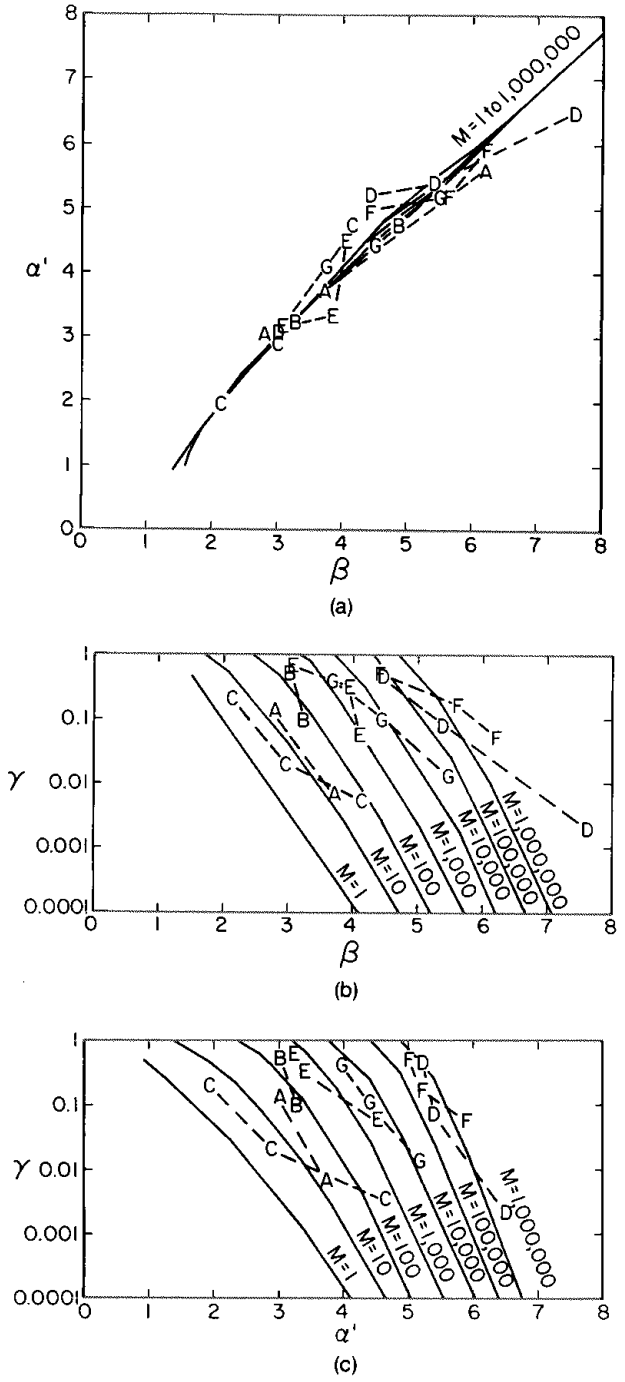


Fig. 9. Analysis of yes–no performance. Each condition from Table 3 is represented in each figure by two or three points connected by dashed lines. The solid curves show the performance of the uncertainty model for values of M from 1 to 1,000,000. K is 1. (a) Normalized threshold contrast α' versus β. Note that the curve is essentially independent of M. (b) False-alarm rate γ versus β. (c) False-alarm rate γ versus normalized threshold contrast α'.

than 1, if that were possible. Even so, the standard deviation of the fit shown, i.e., the standard deviation of $\log \alpha'/\beta$, is only 0.039, or 0.8 dB, which seems to be approximately the standard deviation of the data themselves.

Figure 9(b) shows the false-alarm rate γ versus β. Recall that in the uncertainty model, β is linearly related to the subjective criterion λ by relation (5.7). Here we see that the false-alarm rate γ drops precipitously as the subjective criterion λ rises. Increasing β by 2, e.g., from 5 to 7, reduces the false-alarm rate by nearly 4 orders of magnitude. The solid curves represent the model's performance for values of M from 1 to 1,000,000.

I have estimated by eye the power of 10 for M that best fits each condition. These values are listed in column three of Table 4. There is fair agreement between model and data, requiring values of M ranging from 10 (data sets A and C) to 1,000,000 (data set F). Note that the parameter A does not affect this plot. The data are shallower than the model at the higher values of β (especially data set D). This may be an artifact resulting from testing the observers (by Nachmias) and the model (by me) at different normalized contrasts. In particular, estimates of γ and β tend to be lower if trials at zero or near-zero contrast are not included. Furthermore, at least part of the deviation, although striking, is not significant. For example, the rightmost D, at a false-alarm rate of 0.002, appears to be much too high, yet it probably is nonzero because of a single false alarm by the subject.

Figure 9(c) shows normalized threshold contrast α' versus the false-alarm rate γ. As expected, the model's curves are very similar to those in Fig. 9(b), but the agreement with the data is even better. The model and most of the data show the same slopes throughout the figure, although the fit for data set C is no better than in Fig. 9(b).

The deviation between model and data in Figs. 9(b) and 9(c) does not appear to be significant. Future studies should test model and observer at the same normalized contrasts.

In sum, with only two free parameters, A and M, the uncertainty model gives a good account for the effects of the subjective criterion on all three parameters of the yes–no hit rate: α, β, and γ.

8. TASK EFFECTS

Yes–No versus Two-Alternative Forced Choice

Nachmias[12] also documented the dependence of β and α on task. The high-threshold assumption incorrectly predicts that β and α (and the relevant-hit rate) should be independent of task. If fact, β and α are higher (and the relevant-hit rate is lower) when estimated from yes–no (yn) data than when estimated from 2afc (fc) data.[12,55]

The uncertainty model can predict the ratio of α on the two tasks from the ratio of β on the two tasks. Nachmias suggested a typical value for β_{yn} of 4.2 and for β_{fc} of 3. The results of the preceding section suggest that it is appropriate to assume that K is 1. We may estimate α from β by our relation $\alpha \approx A\beta K^{-1/\beta} = A\beta$, since K is 1. This predicts that the ratio of α's on the two tasks should equal the ratio of β's:

$$\alpha_{yn}/\alpha_{fc} \approx \beta_{yn}/\beta_{fc} = 4.2/3 = 1.4, \qquad (8.1)$$

which is in reasonable agreement with the average ratio $\alpha_{yn}/\alpha_{fc} = 1.3$ that Nachmias reported.

Table 3 lists both yes–no (derived from ratings) and 2afc data obtained under similar conditions. It is reasonable to expect the parameters of the uncertainty model to be the same in both tasks. The second and third columns of Table 4 list the values of A and M for the best fit of the uncertainty model to the yes–no data in Table 3. As noted above, K was set to 1. These parameters completely determine the model, permitting a prediction of the 2afc data α and β. For 2afc, γ is

always 0.5. The predicted values $\hat{\alpha}$ and $\hat{\beta}$ are obtained by using Eq. (5.4) and relation (6.13) as follows:

$$\hat{\beta} = 1.358 + 0.792 \log M, \tag{8.2}$$

$$\hat{\alpha} = A\hat{\beta}K^{-1/\hat{\beta}} = A\hat{\beta}, \tag{8.3}$$

since K is 1. The fourth and fifth columns of Table 4 show the ratios of the predictions to the actual values of α and β. Most of the predictions are excellent. The ratios are mostly close to 1 and are as often higher than 1 as lower than 1.

In these sorts of data, estimates of α typically have a 95% confidence interval of ±20%, and estimates of β typically have a 95% confidence interval of ±50%. Thus, of the eight predictions, there were only two significant errors: the predicted α in condition A is slightly too low, $\hat{\alpha}/\alpha = 0.77$, and the predicted β in condition D is slightly too high, $\hat{\beta}/\beta = 1.5$.

The fact that the model has the nearly same task dependence as the observer is probably a result of incorporating the decision-variable assumption. Green and Swets[3] proved that the decision-variable assumption implies that the proportion correct in a 2afc will equal the area under the ROC curve, i.e., a graph of hit rate versus the false-alarm rate, for the same stimulus. This prediction has been tested by Tanner and Norman,[55] who confirmed it, and, recently, in a more careful study, by Nachmias,[12] who found that the area under the ROC curve was slightly less than the proportion correct. Possible explanations for the discrepancy include trial-to-trial variations of the observer's subjective criterion (contrary to the decision-variable assumption). The criterion variations would reduce the area under the ROC curve but not affect the 2afc performance, which is criterion free. Since this small discrepancy is the only evidence against the decision-variable assumption, the assumption will be retained for its great explanatory value.

In sum, presumably as a consequence of incorporating the decision-variable assumption, the uncertainty model predicts the relation of yes–no and 2afc performance with reasonable although not perfect accuracy.

9. CONTRAST DISCRIMINATION: FACILITATION AND d' ADDITIVITY

A. Experimental Results

We can discriminate contrasts that differ by less than the smallest contrast that we can detect. This phenomenon is sometimes called facilitation or negative masking. The uncertainty model exhibits this facilitation effect. Looking at Fig. 5(a) one can see that when the uncertainty is high, the distribution of the decision variable for contrast $c' = 2$ is nearly superimposed upon that for zero contrast, so these two contrasts, 0 and 2, are indiscriminable. In other words, a contrast of $c' = 2$ is undetectable. The same difference in contrast at a high pedestal level (e.g., $c'_1 = 4$ and $c'_2 = 6$) produces distributions with only modest overlap, so these two contrasts, 4 and 6, are readily discriminated.

Nachmias and Kocher[7] and Nachmias and Sansbury[10] showed that facilitation may be related to the steepness of the psychometric function for detection, i.e., the fact that d' is a positively accelerating function of contrast. I will consider only 2afc experiments, so d' is $\sqrt{2}\Phi^{-1}(P_{fc})$, where P_{fc} is the proportion correct. Nachmias and Sansbury[10] showed that they could predict the d' for discriminating two contrasts by the difference in d values for detecting them. I call this finding d additivity, although it has also been called z additivity[11,15] and constancy of discriminal dispersion.[23]

Think of d' as the distance between two contrasts along a discriminability continuum, as illustrated below. Let $d'_{i,j}$ represent the d' for discriminating contrast c_i and c_j, where c_i is less than c_j. Nachmias *et al.*[7,10] found that discriminabilities add:

$$d'_{1,3} = d'_{1,2} + d'_{2,3},$$

$$c_1 \leftarrow d'_{1,2} \rightarrow c_2 \leftarrow d'_{2,3} \rightarrow c_3$$

$$c_1 \longleftarrow d'_{1,3} \longrightarrow c_3. \tag{9.1}$$

Nachmias and Sansbury examined values of c_1 ranging from 0 to 0.0014, but they used only one value of c_2. Foley and Legge[11] confirmed d' additivity using several values of c_2.

Nachmias and Sansbury[10] explained both d' additivity and the dependence of detectability on contrast by the decision-variable assumption with the additional condition that the decision variable be Gaussian with constant variance. In that case d' will be proportional to the mean of the decision variable on the signal presentation minus the mean on the blank presentation. Any relation between d' and contrast will imply an identical relation between the mean of the decision variable and contrast. Since the relation is typically nonlinear, they called this model nonlinear transduction, although Tanner[5] has pointed out that this name may be misleading (see Section 1).

Nachmias and Sansbury[10] showed that, if the decision variable is Gaussian with constant variance, then d' additivity must hold. However, they did not establish in general what conditions are required for d' additivity. Recall that any strictly increasing transformation of the decision variable will leave performance unchanged, including the property of d' additivity. Thus Nachmias and Sansbury's condition is sufficient, but certainly not necessary, for d' additivity. Given the decision-variable assumption, I find the notion of d' additivity more useful than assumptions about the distribution of the decision variable, because d' additivity is a directly measurable empirical property, whereas the distribution of the human observer's decision variable is unknowable.

Although the uncertainty model incorporates the decision-variable assumption, inspection of Fig. 5(a) shows that, for near-threshold contrasts at high uncertainty, its decision variable is far from Gaussian, and its variance increases with contrast, at least at low contrasts. There is some evidence that if human observers use a decision variable, then its variance must increase with its mean.[3,7] Thus human observers and the uncertainty model do not satisfy Nachmias and Sansbury's condition and therefore might not obey d' additivity.

B. Predictions of the Uncertainty Model

By the decision-variable assumption (Assumption 1) the observer chooses the interval producing the larger value of the decision variable. The contrast-discrimination performance of the uncertainty model may be calculated from our formula for $P_{yn}(c')$ [Eq. (4.6)]. Let the two signal strengths to be discriminated be c'_1 and c'_2. Assume that only one sample is stimulated by the signal, $K = 1$. Then

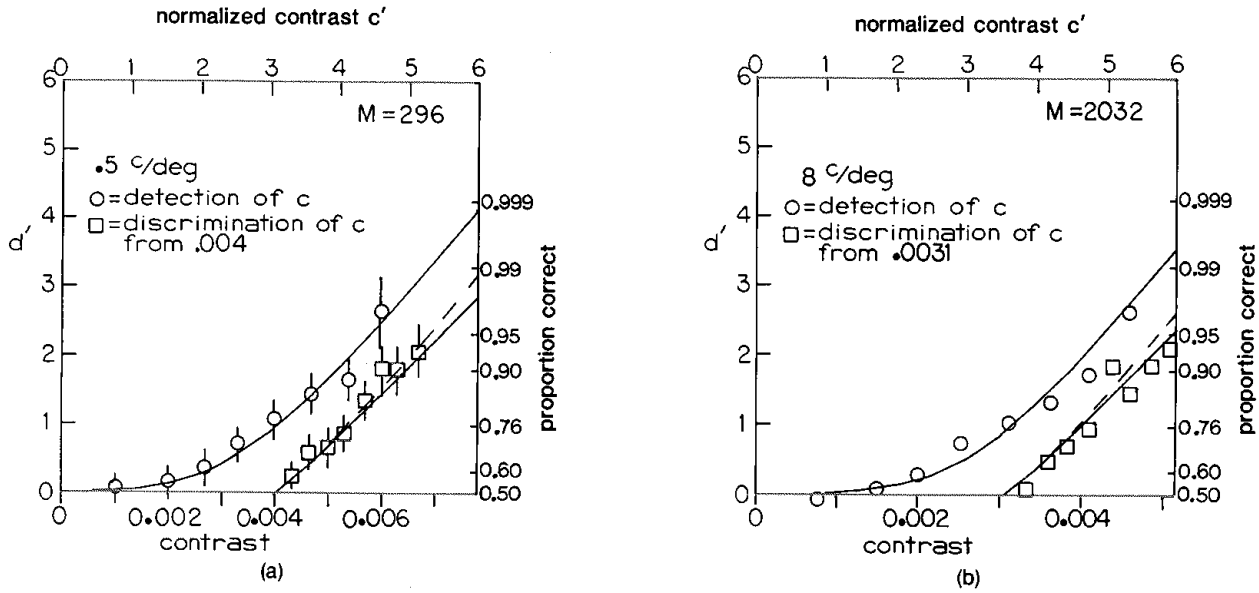


Fig. 10. 2afc detection and discrimination of contrast of sinusoidal gratings. The solid curves are a fit of the uncertainty model to the detection data (circles) and a prediction of the discrimination data (squares). The dashed curves show the prediction of d' additivity. The data are from Foley and Legge.[11] The vertical bars in (a) indicate the 95% binomial confidence interval. (a) and (b) Show the 0.5- and 8-c/deg results, respectively. A similarly successful prediction was also made for Foley and Legge's 2-c/deg results.

$$P_{yn}(c_1') = 1 - \Phi(\lambda - c_1')\Phi^{M-1}(\lambda),$$
$$P_{yn}(c_2') = 1 - \Phi(\lambda - c_2')\Phi^{M-1}(\lambda). \quad (9.2)$$

These equations describe a ROC curve, and, by assumption 1, $P_{fc}(c')$ is given by the area enclosed:

$$P_{fc}(c') = \int_{\lambda=+\infty}^{\lambda=-\infty} P_{yn}(c_1')\mathrm{d}P_{yn}(c_2') \quad (9.3)$$

$$= 1 - \int_{-\infty}^{+\infty} \Phi(\lambda - c_1')\Phi^{2M-3}(\lambda)\left[\Phi(\lambda)\frac{\mathrm{d}\Phi}{\mathrm{d}\lambda}(\lambda - c_2') + (M-1)\Phi(\lambda - c_2')\frac{\mathrm{d}\Phi}{\mathrm{d}\lambda}(\lambda)\right]\mathrm{d}\lambda. \quad (9.4)$$

The data of Foley and Legge confirmed d' additivity for human observers. Does this reject the uncertainty model? Figures 10(a) and 10(b) show d' as a function of the contrast c_1' in the signal interval for several values of the contrast c_2' in the noise interval. The circles in Figs. 10(a) and 10(b) represent proportions correct for observer JMF in detection of spatial frequencies of 0.5 and 8 c/deg, respectively. (A similar figure was prepared for 2 c/deg but is not shown.) Each point represents about 200 to 300 2afc trials. In Fig. 10(a) error bars are shown for each point indicating the 95% confidence interval based solely on the binomial distribution. Foley and Legge also determined the best values (producing least square error) of M and scale factor A (between c and c') to fit the uncertainty model to their detection data. That fit is shown by the continuous line running through the circles. As they noted, it is a good fit. The squares represents proportions correct in contrast discrimination.

The prediction of d' additivity is indicated by the dashed curves, which are just vertical translations of the detection curve, i.e., $d_{2,1}' = d_1' - d_2'$. The predictions of the uncertainty model, using Legge and Foley's estimates of M and c/c', are shown as solid curves. There were no degrees of freedom in either prediction. These data cannot reject either hypothesis. The fact that the predictions are not identical shows that the uncertainty model does not obey d' additivity exactly. However, the uncertainty model predicts these data as successfully as d' additivity does, and the extent to which the uncertainty model fails to satisfy d' additivity seems too small to distinguish in a feasible experiment.

C. Conclusions

Facilitation is quantitatively described by d' additivity. The predictions of the uncertainty model and of d' additivity are so similar that they cannot be distinguished by available data.

10. SUPRATHRESHOLD CONTRAST DISCRIMINATION AND NOISE MASKING

A. Suprathreshold Contrast Discrimination

Section 9 considered discrimination for near-threshold contrasts. Experimentally, graphs of threshold contrast increment Δc versus pedestal contrast c have a dipper shape.[10,56] The triangles in Fig. 11(a) show the typical result. (The O's and X's will be explained below.) At subthreshold pedestal contrasts they fall, bottoming near the threshold contrast of the pedestal (indicated by the vertical arrow at 0.007), and then rise at suprathreshold contrasts. The same thing is shown more clearly in Fig. 11(b), where the increment and pedestal contrasts have been normalized by the threshold contrast α for detection. Again, consider only the triangles. A fit by the uncertainty model ($M = 1000, K = 1$) is shown by the solid curve. It follows the data as it drops from the detection threshold of 1 down to about 0.3, but then it flattens out and fails to rise as the data do at suprathreshold contrasts.

The uncertainty model fails to predict the rise because the uncertainty has almost no effect on discrimination of suprathreshold contrasts. At suprathreshold contrasts $c > \alpha$ the single relevant sample almost always[57] produces the maximum

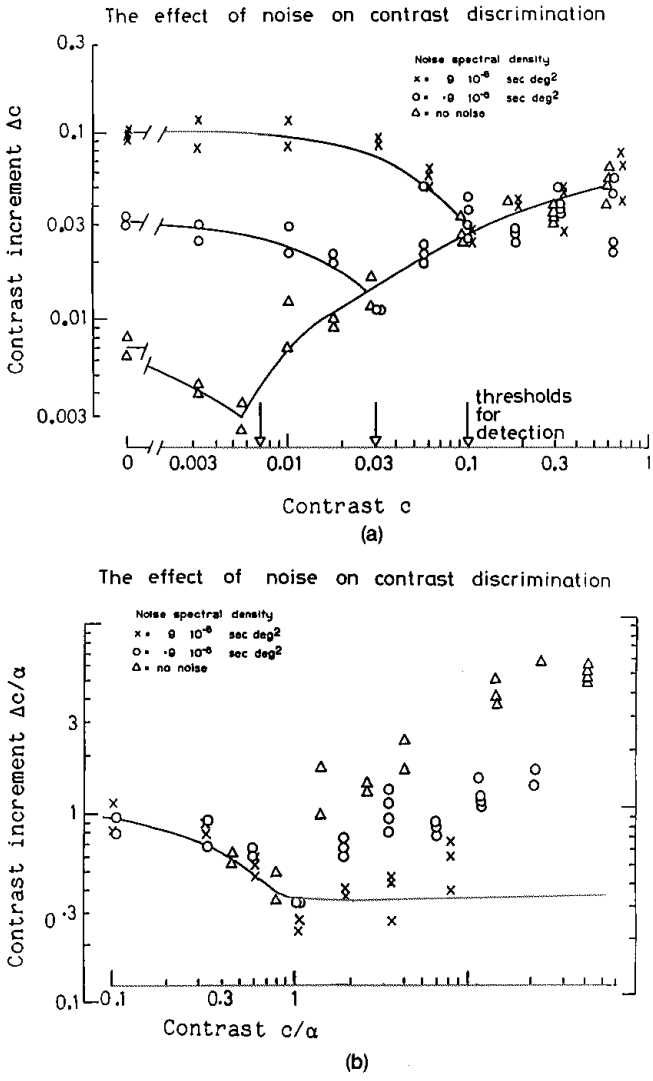


Fig. 11. 2afc contrast discrimination in noise. (a) The vertical scale is the just-detectable (82%-correct) increment in contrast Δc from a pedestal contrast of c, represented by the horizontal scale. The pattern was a 4-c/deg grating. The thresholds for detection are indicated by vertical arrows along the contrast axis. In two of the conditions (O and ×) white noise, uncorrelated over space and time, was added to the display. (b) Same as (a), but both scales have been normalized by the threshold contrast α.

response on both intervals, so the $M - 1$ irrelevant samples have almost no effect. This is apparent in Fig. 5(a); e.g., the probability distribution function of the decision variable for a normalized contrast of 6 is the same for uncertainties of 1 to 10,000. At suprathreshold contrasts the model performs nearly the same as if there were no uncertainty, $M = 1$, that is, nearly the same as the ideal detector of a signal known exactly. It is well known that for this ideal the threshold contrast increment is independent of the pedestal.

Previous authors[10,56,58] have attributed the growth of the increment threshold with the pedestal to an increase in internal noise level with pedestal contrast or to a nonlinear compression of the signal before combination with a constant noise level. Both of these explanations imply an increase in the equivalent noise level with signal contrast, contrary to Assumption 6. I believe that is the correct explanation: that the equivalent noise rises with physical contrast. A good way to study the equivalent noise of many systems is to measure the effects of adding more noise.

B. Method

The threshold contrast increment Δc over a pedestal contrast c for 82%-correct discrimination of c and $c + \Delta c$ was measured in a two-interval forced-choice paradigm, guided by the QUEST[59] adaptive staircase procedure. The signal was a 4-c/deg grating vignetted in contrast by a Gaussian envelope with a $1/e$-space constant of 4° and a $1/e$-time constant of 70 msec. The display always maintained a mean luminance of 300 cd/m^2.

In two of the conditions, white noise was added to the contrast function of the signal, producing random variations of luminance uncorrelated over time and space. Noise level is measured as power spectral density, that is, the contrast power per unit spatiotemporal bandwidth. Contrast power is the average of the square of the contrast function, i.e., contrast power equals the square of rms contrast.

C. Effect of Noise Masking

The model performance does not depend directly on either the signal or noise but only on their ratio, as embodied in c'. Thus the only effect of adding noise is to increase the scale factor A between physical and normalized contrasts. Perhaps the easiest way to measure the change in scale factor is to measure the ratio of thresholds for detection. Thus our graph of increment-threshold contrast versus pedestal contrast should scale in proportion to detection threshold α. This prediction will test Assumption 6, that the effective noise level is independent of the signal contrast.

Figure 11(a) shows the increment contrast Δc versus the pedestal contrast c in the presence of three noise levels: zero (triangles), 0.9 μsec deg^2, and 9 μsec deg^2. First note that the prediction is at least qualitatively correct. The noise increases the threshold by more than a factor of 10, and the dipper shape is preserved but scaled to the new threshold, so that the minimum detectable contrast increment is now at pedestal contrasts near the new threshold for detection. For a more precise test it is helpful to normalize all contrasts by the detection threshold α.

Figure 11(b) shows the result. As noted above, the solid curve shows the fit of the uncertainty model ($M = 1000$, $K = 1$). At subthreshold contrasts ($c/\alpha < 1$) the data points are all in excellent agreement with one another and with the fit by the uncertainty model. At suprathreshold contrasts, $c/\alpha > 1$, the data are scattered and do not agree well with the uncertainty model.

D. Discussion

At subthreshold contrasts the prediction was successful, supporting Assumption 6 that the equivalent noise is independent of signal contrasts. However, the prediction failed miserably at suprathreshold contrasts.

Return to Fig. 11(a) and examine the results at suprathreshold pedestal contrasts, i.e., the handles of the dippers. They all coincide. Although the three different noise levels changed threshold dramatically, and strongly affected performance at subthreshold pedestal contrasts, they had no effect at suprathreshold pedestal contrasts. The easiest way

to explain this is to accept the failure of Assumption 6 and suppose that at suprathreshold contrasts the dominant component of the effective noise is dependent only on the physical contrast c of the pedestal.

To account for the approximately 0.5 log–log slope of Δc versus c (see Ref. 60), the equivalent noise spectral density would have to increase as approximately the unit power of c. (I.e., equivalent noise rms contrast is proportional to the square root of signal contrast.) This is intriguing, since the noise would have to be of neural origin (photon noise is not contrast dependent at low contrasts). Neural spike rate, usually being an approximately Poisson process, has a variance proportional to its mean.

11. SUMMARY AND CONCLUSIONS

Here are the conclusions by section:

Section 2. The combination of the probability-summation and decision-variable assumptions implies uncertainty and a maximum-of decision rule.

Section 3. The motivation from signal-detection theory was presented for a simple maximum-likelihood decision rule. First, if the signals are equally probable, then this strategy is optimal or nearly optimal, depending on the task. Second, if the noise is white or Poisson, then the likelihoods may be computed by linear receptive fields, so this model is physiologically feasible.

Section 4. A simple maximum-likelihood receiver (the uncertainly model) was defined by seven assumptions. The model's performance depends only on the signal contrast and four model parameters: uncertainty, signal extent, subjective criterion, and a contrast-normalizing factor proportional to noise amplitude.

Section 5. The uncertainty model exhibits the nonlinear psychometric functions typical of human observers. For yes–no psychometric functions the log–log steepness is linearly related to the subjective criterion. For 2afc psychometric functions the log–log steepness is linearly related to the log of the uncertainty.

Section 6. The uncertainty model exhibits the same yes–no and 2afc summation effects as human observers. A new, stronger summation formula is derived that predicts the effects of both signal extent and the log–log steepness of the psychometric function.

Section 7. The uncertainty model exhibits approximately the same effects of the subjective criterion on the yes–no psychometric function as human observers. This is the first successful quantitative explanation of these effects.

Section 8. Characterizing human observers' yes–no psychometric functions by the best-fitting parameters of the uncertainty model successfully predicted the human observers' 2afc psychometric functions.

Section 9. For contrast discrimination, the uncertainty model exhibits the same facilitation effect as human observers.

Section 10. Because the uncertainty model is nonparametric, its performance is unaffected by proportionate scaling of the signal and effective noise amplitude. A contrast-discrimination-in-noise experiment is analyzed to show that human performance, for near-threshold signal contrasts, is consistent with this prediction. Accounting for suprathreshold performance requires a modification of the model to assume that the equivalent input noise power spectral density increases in proportion to signal contrast, as has been suggested by previous authors for other models.

This paper has argued that human observers asked to detect or discriminate near-threshold contrasts act as though they are uncertain among many signals and choose the likeliest. That is, they make the same calculation that a simple maximum-likelihood receiver makes. This idea accurately predicts effects of contrast, extent, subjective criterion, task, and noise in visual contrast detection and discrimination.

APPENDIX A: HIGH UNCERTAINTY JUSTIFIES CORRECTION FOR GUESSING

It will be shown that when fewer than about 1% of the samples are relevant, $K/M \leqq 1\%$, then Assumptions 1–3 of the uncertainty model satisfy the requirements for correction for guessing of both yes–no and 2afc hit rates.

Correction for guessing asserts that the relevant-hit rate $P^*(c)$ may be recovered from the hit rate $P(c)$ by the formula $P^*(c) \approx 1 - [1 - P(c)]/[1 - P(0)]$. This is proved by first establishing that $P = 1 - (1 - G)(1 - P^*)$, which is given only for yes–no, and then showing that the false-alarm rate $P(0)$ is approximately equal to the irrelevant-hit rate G.

Proof for Yes–No

Let the false-alarm rate for each sample be F. The irrelevant-hit rate G is the probability sum of $M - K$ samples,

$$G = 1 - (1 - F)^{M-K}. \tag{A1}$$

The false-alarm rate $P(0)$ is the probability sum of all M samples,

$$P(0) = 1 - (1 - F)^{M}. \tag{A2}$$

By substituting for F and rearranging we obtain

$$1 - P(0) = (1 - G)^{M/(M-K)}. \tag{A3}$$

It is easily shown that if $1 - p = (1 - g)^{1+\epsilon}$, where $0 \leqq p \leqq 1$, $0 \leqq g \leqq 1$, and $0 \leqq \epsilon$, then $1 \leqq p/g \leqq 1 + \epsilon$. Using this result, and taking $1 + \epsilon = M/(M - K)$, we can conclude that

$$1 \leqq P(0)/G \leqq M/(M - K) \leqq 1.0101\ldots, \tag{A4}$$

since it was given that $K/M \leqq 1\%$. Since with even a thousand trials the measured false-alarm rate will have a standard deviation of about 3%, relation (A4) is more than sufficient to guarantee that the false-alarm rate and the zero-contrast irrelevant-hit rate will not be significantly different. Thus correction for guessing will be valid whenever at most 1% of the samples are relevant, $K/M \leqq 1\%$. This completes the proof.

Proof for Two-Alternative Forced Choice

First I will show that there are two types of reasons for a correct response: relevant (i.e., contrast dependent) and irrelevant (i.e., contrast independent). The reasons are not exclusive; a given hit may occur for several reasons. As illustrated in Fig. 3, let there be M variables, of which K are contrast dependent and $M - K$ are contrast independent. There

will be two presentations: a signal presentation and a blank presentation. The entire trial will generate $2M$ samples, M from each presentation. All the samples generated on the blank presentation are independent of the contrast of the signal. Therefore, of the $2M$ samples, K are contrast dependent and $2M - K$ are contrast independent. The contrast-dependent reason for getting the trial right, a relevant hit, is that one of the K contrast-dependent samples exceeds all the $2M - K$ contrast-independent samples. The contrast-independent reason for getting the trial right, an irrelevant hit, is that the largest of the contrast-independent samples happens to occur on the signal presentation rather than on the blank presentation. Second, since the two types of reason are stochastically independent, the probability of a correct response $P(c)$ is the probability sum of the relevant-hit rate $P^*(c)$ and the irrelevant-hit rate G,

$$P(c) = 1 - (1 - G)[1 - P^*(c)]. \tag{A5}$$

The probability G of an irrelevant hit is simply the proportion of irrelevant samples that arise from the signal interval,

$$G = (M - K)/(2M - K). \tag{A6}$$

Third, since the zero-contrast hit rate $P(0)$ is always 0.5 in 2afc, it follows that

$$1 < P(0)/G \leqq 1.00505\ldots, \tag{A7}$$

since it was given that $K/M \leqq 1\%$. This completes the proof.

APPENDIX B: ORTHOGONAL RECEPTIVE FIELDS HAVE UNCORRELATED RESPONSES TO WHITE NOISE

It will be shown that orthogonal receptive fields have uncorrelated responses to white noise. White noise is uncorrelated over space and time.

Before beginning the proof, note that although zero correlation is a weaker condition than stochastic independence, the additional assumption (Assumption 5) that the noise is Gaussian makes them equivalent. It is well known that linear functions of a normally distributed function are jointly normally distributed. Furthermore, if two variables are jointly normally distributed and uncorrelated, then they are stochastically independent.[61] Thus all orthogonal linear functions of white noise are stochastically independent.

It will now be shown that samples produced by orthogonal receptive fields stimulated by noise that is uncorrelated among photoreceptors (e.g., photon noise) are uncorrelated. The restriction to a spatial array of photoreceptors is merely to make the language familiar; the result is general to all dimensions.

Proof. Let us number the photoreceptors in the retina from 1 to k. Designate the quantum catch in receptor i by n_i. Consider two linear receptive fields with outputs A and B, which have been called samples. Each receptive field assigns a weight (which may be positive, negative, or zero) to each photoreceptor. Sample A assigns weight a_i to the ith photoreceptor, and sample B assigns weight b_i to it. Thus the samples are

$$A = a_1n_1 + a_2n_2 + \ldots + a_kn_k,$$
$$B = b_1n_i + b_2n_2 + \ldots + b_kn_k. \tag{B1}$$

It is given, first, that the receptive fields are orthogonal, i.e.,

$$a_1b_1 + a_2b_2 + \ldots + a_kb_k = 0. \tag{B2}$$

Second, the noise is uncorrelated from receptor to receptor and is stationary, having mean μ and variance σ^2 in every receptor. Now it must be shown that A and B are uncorrelated, i.e., that the expected value of their product equals the product of the expected values. The expected value of the product is $\mathrm{Av}\{AB\}$:

$$\mathrm{Av}\{AB\} = \mathrm{Av}\{(a_1n_1 + a_2n_2 + \ldots + a_kn_k)(b_1n_i + b_2n_2 + \ldots + b_kn_k)\}$$
$$= \mathrm{Av}\left\{\sum_{\substack{i=1,k\\j=1,k}} a_in_ib_jn_j\right\}. \tag{B3}$$

The product of the expected values is $\mathrm{Av}\{A\}\mathrm{Av}\{B\}$:

$$\mathrm{Av}\{A\}\mathrm{Av}\{B\} = \mathrm{Av}\{a_1n_1 + a_2n_2 + \ldots + a_kn_k\}\mathrm{Av}\{b_1n_i + b_2n_2 + \ldots b_kn_k\}$$
$$= \mathrm{Av}\left\{\sum_{i=1,k} a_in_i\right\}\mathrm{Av}\left\{\sum_{j=1,k} b_jn_j\right\}. \tag{B4}$$

Showing that the difference is zero will prove that A and B are uncorrelated. We have

$$\mathrm{Av}\{AB\} - \mathrm{Av}\{A\}\mathrm{Av}\{B\} = \mathrm{Av}\left\{\sum_{\substack{i=1,k\\j=1,k}} a_in_ib_jn_j\right\} - \mathrm{Av}\left\{\sum_{i=1,k} a_in_i\right\}\mathrm{Av}\left\{\sum_{j=1,k} b_jn_j\right\}. \tag{B5}$$

The average of a sum is the sum of the averages, and since the a's and b's are constant they can pass outside the average:

$$= \sum_{\substack{i=1,k\\j=1,k}} a_ib_j\,\mathrm{Av}\{n_in_j\} - \sum_{i=1,k} a_i\,\mathrm{Av}\{n_i\}\sum_{j=1,k} b_j\,\mathrm{Av}\{n_j\}. \tag{B6}$$

It is given that $\mathrm{Av}\{n_i\} = \mu$:

$$= \sum_{\substack{i=1,k\\j=1,k}} a_ib_j\,\mathrm{Av}\{n_in_j\} - \sum_{\substack{i=1,k\\j=1,k}} a_ib_j\mu^2$$
$$= \sum_{\substack{i=1,k\\j=1,k}} a_ib_j(\mathrm{Av}\{n_in_j\} - \mu^2). \tag{B7}$$

It is useful to treat the terms for which $i = j$ separately from those for which $i \neq j$:

$$= \sum_i a_ib_i(\mathrm{Av}\{n_i^2\} - \mu^2) + \sum_{i\neq j} a_ib_j(\mathrm{Av}\{n_in_j\} - \mu^2). \tag{B8}$$

It was given that the noise was uncorrelated from receptor to receptor, so when $i \neq j$ we have $\mathrm{Av}\{n_in_j\} = \mathrm{Av}\{n_i\}\mathrm{Av}\{n_j\} = \mu^2$, showing the right-hand sum to be zero. It was given that the noise produced the same variance σ^2 in every receptor, so $\mathrm{Av}\{n_i^2\}$ equals σ^2:

$$= (\sigma^2 - \mu^2)\,\Sigma\, a_ib_i + 0. \tag{B9}$$

It was given that the receptive fields were orthogonal, so $\Sigma a_i b_i$ is zero:

$$= 0. \tag{B10}$$

This completes the proof that the outputs of the two receptive fields are uncorrelated.

If the proof seems counterintuitive, consider the simple case of $a_1 = b_1 = 1$ and $a_2 = 1, b_2 = -1$, and all other a_i and b_i equal to zero. The contributions to A and B from a_1 and b_1 are positively correlated; the contributions from a_2 and b_2 are negatively correlated. The result is zero correlation between A and B.

APPENDIX C: TWO-ALTERNATIVE FORCED-CHOICE PROBABILITY SUMMATION BY THE UNCERTAINTY MODEL

Does the uncertainty model satisfy the requirements for 2afc probability summation? Appendix A has already shown that the uncertainty model (Assumptions 1–3) with high uncertainty ($K/M \leqq 1\%$) satisfies three of the four parts of the probability-summation assumption for 2afc: (1) there are relevant and irrelevant reasons for a hit, (2) they are stochastically independent, and (3) we may correct for guessing. What remains to be shown is that there are many independent relevant reasons for a hit.

It will be shown that the relevant-hit rate is approximately equal to the probability that any of K independent relevant events will occur. This is proven for the full uncertainty model (Assumptions 1–7) with high uncertainty, $K/M \leqq 1\%$ and for the further condition that $K + 1$ is only a small root of $2M$:

$$\log(K + 1) \ll \log(2M). \tag{C1}$$

Proof. As in Appendix A, the probability of a relevant hit $P^*(c')$ is the probability that X is greater than Y:

$$1 - P^*(c') = 1 - P(X > Y), \tag{C2}$$

where X is the maximum of the K relevant samples X_i and Y is the maximum of the $2M - K$ irrelevant samples Y_i. The variables X and Y are independent. If the variance of X is much greater than the variance of Y, then we may approximate the probability that X is greater than Y by the probability that X is greater than the mean of Y:

$$\text{If } \sigma_X{}^2 \gg \sigma_Y{}^2, \text{ then } P(X > Y) \approx P(X > \mu_Y). \tag{C3}$$

In Fig. 5(a) the left-hand curve in each panel shows the probability-density function of the maximum of M samples from a normal distribution. As M rises from 1 to 10,000, the curve becomes narrower and narrower, indicating that the variance is falling. The variance of the maximum of n independent identically distributed normal samples with unit variance is well approximated by $\pi^2/[12 \ln(n + 1)]$. (The error is less than 20% at $n = 1$ and asymptotically approaches zero with increasing n.) Thus the ratio of the variances of X and Y is

$$\sigma_X{}^2 \, \sigma_Y{}^2 \approx \ln(2M - K + 1)/\ln(K + 1). \tag{C4}$$

Given high uncertainty, $K/M \leqq 1\%$, and relation (C1), it follows that $\log(2M - K + 1) \gg \log(K + 1)$, so $\sigma_X{}^2 \gg \sigma_Y{}^2$. By assertion (C3) we may now write

$$1 - P(X > Y) \approx 1 - P(X > \mu_Y) \tag{C5}$$

and

$$= \prod_{i=1}^{K} 1 - P(X_i > \mu_Y), \tag{C6}$$

since X will fail to exceed μ_Y only if all the K-independent samples X_i fail to exceed μ_Y. Noting Eq. (C2), we have the desired result:

$$1 - P^*(c') \approx \prod_{i=1}^{K} 1 - P(X_i > \mu_Y). \tag{C7}$$

We have shown that the probability that the observer will make a relevant hit is approximately the probability that any of the relevant samples will exceed the mean of the maximum of the irrelevant samples. This completes the proof.

APPENDIX D: LOG UNCERTAINTY SUBTRACTS FROM CONTRAST

For the uncertainty model, it will be shown that increasing the uncertainty M beyond $100 + K$ is approximately equivalent to subtracting $0.330 \ln[(M - K)/100]$ from the normalized contrast c'. This allows us to predict the effect on 2afc detection performance of increasing the uncertainty by a known fraction, e.g., in an uncertainty experiment.

The desired result follows easily from a theorem concerning the asymptotic distribution of extreme order statistics.

Theorem 1 (From Ref. 38, p. 65). Let $X_1, \ldots, X_n$ be independent Gaussian samples with zero mean and unit variance. Let Z_n be the maximum of $X_1, \ldots, X_n$. The cumulative distribution functions are

$$P(X_i < x) = \Phi(x) = 1/\sqrt{2\pi} \int_{-\infty}^{x} \exp(-x^2/2xd), \tag{D1}$$

$$P(Z_n < x) = P^n(X < x) = \Phi^n(x). \tag{D2}$$

Define

$$a_n = (2 \ln n)^{1/2} - 0.5(\ln \ln n + \ln 4\pi)/(2 \ln n)^{1/2}, \tag{D3}$$

$$b_n = 1/(2 \ln n)^{1/2}. \tag{D4}$$

Then, in the limit as $n \to \infty$, the variable $(Z_n - a_n)/b_n$ has a double exponential distribution,

$$\lim_{n \to \infty} P(Z_n b_n + a_n < x) = H(x), \tag{D5}$$

or, equivalently,

$$\lim_{n \to \infty} P\{(Z_n - a_n)/b_n < x\}, \tag{D6}$$

where

$$H(x) = \exp(-e^{-x}). \tag{D7}$$

See Ref. 38 for proof.

It is easily shown that

$$H(x) = H^n\{x + \ln(n)\} \text{ for all } n. \tag{D8}$$

Unfortunately the convergence in Eq. (D5) to $H(x)$ is slow. However, for n as large as 100 the approximation is quite good, except in the tails[38,52]:

$$H\{(x - a_n)/b_n\} \approx P(Z_n < x) \text{ for } n \geqq 100. \quad \text{(D9)}$$

Substituting from Eq. (D2),

$$= P^n(X < x) \quad \text{(D10)}$$

$$= [P^{100}(X < x)]^{n/100}. \quad \text{(D11)}$$

Substitute, using relations (D9) and (D10),

$$\approx [H\{(x - a_{100})/b_{100}\}]^{n/100}. \quad \text{(D12)}$$

Now, by Eq. (D8), we have

$$= H\{(x - a_{100})/b_{100} - \ln(n/100)\},$$

that is,

$$H\{(x - a_n)/b_n\} \approx H\{(x - a_{100})/b_{100} - \ln(n/100)\}. \quad \text{(D13)}$$

If this approximation were exact, then b_n/b_{100} would equal one and $a_n - [a_{100} + b_{100} \ln(n/100)]$ would equal zero. This is nearly true; over a three-decade range, $100 \leqq n \leqq 100{,}000$, we have

$$1 \geqq b_n/b_{100} \geqq 0.64 \quad \text{(D14)}$$

and

$$|a_n - a_{100} - b_{100} \ln(n/100)| \leqq 0.08. \quad \text{(D15)}$$

Applying relations (D9) and (D2) to relation (D13), we have

$$\Phi^n(x) \approx \Phi^{100}\{x - b_{100} \ln(n/100)\}. \quad \text{(D16)}$$

If $M - K$ is larger than 100 we may use relation (D16) to rewrite Eqs. (4.6) and (4.7) as

$$P_{yn}(c') \approx 1 - \Phi^{100}(\hat{\lambda})\Phi^K(\hat{\lambda} - \{c' - b_{100} \ln[(M - K)/100]\}), \quad \text{(D17)}$$

$$P_{yn}(0) \approx 1 - \Phi^{100}(\hat{\lambda}), \quad \text{(D18)}$$

where $b_{100} = 0.330$ [by Eq. (D4)], and the new subjective criterion is $\hat{\lambda} = \lambda - b_{100} \ln[(M - K)/100]$. Since the subjective criterion is not measurable, the change of variables from λ to $\hat{\lambda}$ is innocuous; it does not matter which one the subject uses. Note that M appears only once, and it is in a term that subtracts from the normalized contrast c'. This shows that increasing $M - K$ beyond 100 is equivalent to subtracting $b_{100} \ln[(M - K)/100]$ from the normalized contrast c'. Yes–no performance curves will translate unchanged along the contrast axis only if the false-alarm rate $P_{yn}(0)$ is kept constant by raising the subjective criterion λ when the uncertainty is increased. However, the subjective criterion plays no role in 2afc, so any graph of 2afc performance versus c' will translate to the right as $M - K$ is increased beyond 100. This is apparent in Fig. 6(b). Furthermore, it implies that the contrast threshold should be linearly related to the log of M, as is confirmed by direct calculation [see relations (5.3) and (5.4)].

It follows that if $P_1(c')$ is the 2afc hit rate for some uncertainty M_1 (assumed greater than 100) and some experimental procedure increases the uncertainty to M_2, then the new hit rate will be

$$P_2(c') \approx P_1[c' - 0.330 \ln(M_2/M_1)], \quad \text{(D19)}$$

where for clarity I have assumed that $M_1 \gg K$ so that K may be neglected. Note that this prediction depends only on the ratio of the uncertainties, M_2/M_1. Alternatively, we may rewrite this in terms of physical contrast,

$$P_2(c) \approx P_1[c - 0.330\, A \ln(M_2/M_1)], \quad \text{(D20)}$$

where $P_1(c)$ and $P_2(c)$ are now the 2afc hit rate as a function of *physical* contrast c and A is the contrast normalization factor, which may be estimated by means of relation (6.14).

There is a strong parallel between the double exponential [Eq. (D7)] and the Weibull [Eq. (5.1)] distributions. Both arise as asymptotic distributions of the maximum of many independent identically distributed random variables. As the number of variables is increased, the Weibull distribution shifts unchanged along the log contrast axis, while the double exponential shifts unchanged along the linear contrast axis.[38,52]

ACKNOWLEDGMENTS

I thank the many people who have contributed to the refinement of these ideas by discussion or criticism of earlier drafts, in particular, Jack Nachmias (especially on signal detection and high threshold theories and d' additivity), Norma Graham (exposition of the model), Al Ahumada, John Robson, Gordon Legge, Roy Patterson (uncertainty), Gary Rubin, John Foley, Dan Kersten (equivalent noise), Art Burgess (signal-to-noise ratio), Pat Kramer, Arthur Bradley, and an anonymous referee. I thank Jack Nachmias for providing the data that appear in Table 3 and Figs. 9(a)–9(c). This research was supported during 1979 by British Ministry of Defense contract "Spatial noise spectra and target detection/recognition" to F. W. Campbell, during 1979–1981 by National Institutes of Health grant EY02934 to G. E. Legge, and during 1982–1985 by National Institutes of Health grant EY04432 to me.

Some of these results appeared in my Ph.D. dissertation,[62] and some were presented at the 1980 and 1984 Annual Meetings of the Optical Society of America in Chicago[15] and San Diego[63] and at the 1984 meeting of the European Conference on Visual Perception in Cambridge, England.[64]

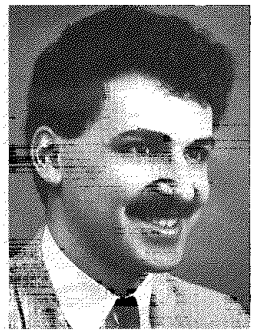
REFERENCES

1. A. Rose, "The sensitivity performance of the human eye on an absolute scale," J. Opt. Soc. Am. **38,** 196–208 (1948).
2. H. de Vries, "The quantum character of light and its bearing upon threshold of vision, the differential sensitivity and visual acuity of the eye," Physica **10,** 553–564 (1943).
3. D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966).
4. W. P. Tanner, Jr., and J. A. Swets, "The human use of information. I. Signal detection for the case of the signal known exactly," IRE Trans. Inf. Theory **PGIT-4,** 213–221 (1954).
5. W. P. Tanner, Jr., "Physiological implications of psychophysical data," Ann. N.Y. Acad. Sci. **89,** 752–765 (1961).
6. B. Leshowitz, H. B. Taub, and D. H. Raab, "Visual detection of signals in the presence of continuous and pulsed backgrounds," Percept. Psychophys. **4,** 207–213 (1968).
7. J. Nachmias and E. C. Kocher, "Visual detection and discrimination of luminous increments," J. Opt. Soc. Am. **60,** 382–389 (1970).
8. T. E. Cohn, L. N. Thibos, and R. N. Kleinstein, "Detectability of a luminance increment," J. Opt. Soc. Am. **64,** 1321–1327 (1974).
9. C. F. Stromeyer III and S. Klein, "Spatial frequency channels in human vision as asymmetric (edge) mechanisms," Vision Res. **14,** 1409–1420 (1974).

10. J. Nachmias and R. V. Sansbury, "Grating contrast: discrimination may be better than detection," Vision Res. **14,** 1039–1042 (1974).
11. J. M. Foley and G. E. Legge, "Contrast detection and near-threshold discrimination in human vision," Vision Res. **21,** 1041–1053 (1981).
12. J. Nachmias, "On the psychometric function for contrast detection," Vision Res. **21,** 215–223 (1981).
13. T. E. Cohn and D. J. Lasley, "Detectability of a luminance increment: effect of spatial uncertainty," J. Opt. Soc. Am. **64,** 1715–1719 (1974).
14. T. E. Cohn, "Detectability of a luminance increment: effect of superimposed random luminance fluctuation," J. Opt. Soc. Am. **66,** 1426–1428 (1976).
15. D. G. Pelli, "Channel uncertainty as a model of visual detection," J. Opt. Soc. Am. **70,** 1628–1629 (1980).
16. D. J. Lasley and T. E. Cohn, "Why luminance discrimination may be better than detection," Vision Res. **21,** 273–278 (1981).
17. D. M. Green and T. G. Birdsall, "Detection and recognition," Psychol. Rev. **85,** 192–206 (1978).
18. E. T. Davis, P. Kramer, and N. Graham, "Uncertainty about spatial frequency, spatial position, or contrast of visual patterns," Percept. Psychophys. **33,** 20–28 (1983).
19. C. D. Creelman, "Detection of signals of uncertain frequency," J. Acoust. Soc. Am. **32,** 805–810 (1960).
20. D. M. Green, "Detection of auditory sinusoids of uncertain frequency," J. Acoust. Soc. Am. **33,** 897–903 (1961).
21. J. Nachmias, "Signal detection theory and its applications to problems in vision," in *Handbook of Sensory Physiology VII/4, Psychophysics*, D. Jameson and L. Hurvich, eds. (Springer-Verlag, Berlin, 1972), pp. 56–77.
22. L. L. Thurstone, "A law of comparative judgement," Psychol. Rev. **34,** 273–286 (1927).
23. L. L. Thurstone, "Psychophysical analysis," Am. J. Psychol. **38,** 368–389 (1927).
24. M. B. Sachs, J. Nachmias, and J. G. Robson, "Spatial frequency channels in human vision," J. Opt. Soc. Am. **61,** 1176–1186 (1971).
25. H. B. Barlow, "Retinal noise and absolute threshold," J. Opt. Soc. Am. **46,** 634–639 (1956).
26. J. A. Swets, W. P. Tanner, Jr., and T. G. Birdsall, "Decision processes in perception," Psychol. Rev. **68,** 301–340 (1961).
27. J. Nachmias and R. M. Steinman, "Study of absolute visual detection by the rating-scale method," J. Opt. Soc. Am. **53,** 1206–1213 (1963).
28. H. L. van Trees, *Detection, Estimation, and Modulation Theory* (Wiley, New York, 1968).
29. E. H. Linfoot, *Fourier Methods in Optical Image Evaluation* (Focal, New York, 1964).
30. A. B. Watson, H. B. Barlow, and J. G. Robson, "What does the eye see best?" Nature **302,** 419–422 (1983).
31. N. Graham, P. Kramer, and D. Yager, "Explaining uncertainty effects and probability summation," Invest. Ophthal. Vis. Sci. Suppl. **24,** 186 (A) (1983).
32. D. Yager, P. Kramer, M. Shaw, and N. Graham, "Detection and identification of spatial frequency: models and data," Vision Res. **24,** 1021–1036 (1984).
33. A. B. Watson, "Detection and recognition of simple spatial forms," NASA Tech. Memorandum 84353 (1983).
34. J. P. Thomas, J. Gille, and R. A. Barker, "Simultaneous visual detection and identification: theory and data," J. Opt. Soc. Am. **72,** 1642–1651 (1982).
35. W. W. Peterson, T. G. Birdsall, and W. C. Fox, "Theory of signal detectability," IRE Trans. Inf. Theory **PGIT-4,** 171–212 (1954).
36. L. W. Nolte and D. Jaarsma, "More on the detection of one of M orthogonal signals," J. Acoust. Soc. Am. **41,** 497–505 (1967).
37. L. A. Wainstein and V. D. Zubakov, *Extraction of Signals from Noise* (Prentice-Hall, Englewood Cliffs, N.J., 1962).
38. J. Galambos, *The Asymptotic Theory of Extreme Order Statistics* (Wiley, New York, 1978).
39. W. S. Geisler, "Physical limits of acuity and hyperacuity," J. Opt. Soc. Am. A **1,** 775–782 (1984).
40. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1964).
41. G. E. Legge, "Sustained and transient mechanisms in human vision: temporal and spatial properties," Vision Res. **18,** 69–81 (1978).
42. A. B. Watson, "Probability summation over time," Vision Res. **19,** 515–522 (1979).
43. 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).
44. N. Graham, "Visual detection of aperiodic stimuli by probability summation among narrow band channels," Vision Res. **17,** 637–652 (1977).
45. H. R. Wilson and J. Bergen, "A four-mechanism model for threshold spatial vision," Vision Res. **19,** 19–32 (1978).
46. N. Graham, J. G. Robson, and J. Nachmias, "Grating summation in fovea and periphery," Vision Res. **18,** 815–826 (1978).
47. G. E. Legge, "Space domain properties of a spatial-frequency channel in human vision," Vision Res. **18,** 959–969 (1978).
48. J. G. Robson and N. Graham, "Probability summation and regional variation in contrast sensitivity across the visual field," Vision Res. **21,** 409–418 (1981).
49. G. S. Brindley, "The order of coincidence required for visual thresholds," Proc. Phys. Soc. London Ser. B **67,** 673–676 (1954).
50. R. F. Quick, "A vector magnitude model of contrast detection," Kybernetik **16,** 65–67 (1974).
51. W. Weibull, "A statistical distribution function of wide applicability," J. Appl. Mech. **18,** 292–297 (1951).
52. H. A. David, *Order Statistics*, 2nd ed. (Wiley, New York, 1981).
53. J. Lieblein, "Two early papers on the relation between extreme values and tensile strength," Biometrika **41,** 559–560 (1954).
54. D. M. Green and R. D. Luce, "Parallel psychometric functions from a set of independent detectors," Psychol. Rev. **82,** 483–486 (1975).
55. W. P. Tanner, Jr., and R. Z. Norman, "The human use of information: II. Signal detection for the case of an unknown signal parameter," IRE Trans. Inf. Theory **PGIT-4,** 222–227 (1954).
56. G. E. Legge and J. M. Foley, "Contrast masking in human vision," J. Opt. Soc. Am. **70,** 1458–1471 (1980).
57. The probability of the maximum sample's being a relevant sample can be estimated from the measured proportion correct P. In a 2afc trial a "relevant hit" means that a relevant sample on the signal interval was larger than all the other samples on both the signal and the blank intervals. If the uncertainty is high, $K/M < 1\%$, we may correct for guessing to estimate the relevant-hit rate $P^* \approx 2(P - 0.5)$. In general, for any uncertainty, the relevant-hit rate is $P^* = 1 - (1 - P)(2M - K)/M$, which is greater than or equal to $2(P - 0.5)$. Threshold α is the contrast at which the hit rate is $P = 82\%$, which corresponds to a relevant-hit rate P^* greater than or equal to 64%. Thus most of the time ($\geq$64%) at threshold contrast $c = \alpha$ the decision variable will be a relevant sample. At suprathreshold contrasts $c > \alpha$ the proportion will be even higher, and the irrelevant samples will have little effect on performance.
58. J. P. Thomas, "Underlying psychometric function for detecting gratings and identifying spatial frequency," J. Opt. Soc. Am. **73,** 751–758 (1983).
59. A. B. Watson and D. G. Pelli, "QUEST: A Bayesian adaptive psychometric method," Percept. Psychophys. **33,** 113–120 (1983).
60. G. E. Legge, "A power law for contrast discrimination," Vision Res. **21,** 457–467 (1981).
61. A. Papoulis, *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York, 1965).
62. D. G. Pelli, "Effects of visual noise," Ph.D. dissertation (Cambridge University, Cambridge, 1981).
63. D. G. Pelli, "Uncertainty in visual detection and identification," J. Opt. Soc. Am. A **1,** 1240 (A) (1984).
64. D. G. Pelli, "Intrinsic uncertainty in visual detection," Perception **13,** 14 (A) (1984).

(see overleaf)

Denis G. Pelli



Denis G. Pelli received the B.A. degree in applied mathematics *magna cum laude* from Harvard University in 1975 and the Ph.D. degree in physiology from Cambridge University in 1981. He spent two years as a research fellow at the psychology department of the University of Minnesota. He is now an assistant professor at the Institute for Sensory Research, Syracuse University. His current research areas include the effects of visual noise and the visual requirements of everyday tasks.

Erratum

There is a printer's error in Eq. 4.2 on page 1516. The correct formula for the cumulative normal integral is:

$$\Phi(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} \exp(-t^2/2)\, dt. \qquad (4.2)$$

Pelli, D.G. (1985) Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America* A2:1508-1532.