Characterizing human perceptual inefficiencies with equivalent internal noise

Zhong-Lin Lu

Department of Psychology, University of Southern California, Los Angeles, California 90089-1061

Barbara Anne Dosher

Department of Cognitive Sciences and Institute of Mathematical Behavioral Sciences, University of California, Irvine, California 92697

Received July 6, 1998; revised manuscript received November 12, 1998; accepted December 1, 1998

A widely used method for characterizing and comparing inefficiencies in perceptual processes is the method of equivalent internal noise—the amount of random internal noise necessary to produce the degree of inefficiency exhibited by the perceptual system in processing [J. Opt. Soc. Am. **46**, 634 (1956)]. One normally estimates the amount of equivalent internal noise by systematically increasing the amount of external noise added to the signal stimulus and observing how threshold—signal stimulus energy required for an observer to maintain a given performance level—depends on the amount of external noise. In a variety of perceptual tasks, a simple noisy linear amplifier model [D. Pelli, Ph.D. dissertation (University of Cambridge, Cambridge, UK 1981)] has been utilized to estimate the equivalent internal noise N_{internal} by fitting of the relation between threshold contrast c_τ and external noise N_{ext} at a single (d') performance level: $c_\tau^2 = (d'/\beta)^2(N_{\text{ext}}^2 + N_{\text{internal}}^2)$. This model makes a strong prediction: Independent of observer and external noise contrast, the ratio between two thresholds at each external noise level is equal to the ratio of the two corresponding d' values. To our knowledge, this potential test for the internal consistency of the model had never been examined previously. In this study we estimated threshold ratios between multiple performance levels at various external noise contrasts in two different experiments: Gabor orientation identification, and Gabor detection. We found that, in both identification and detection, the observed threshold ratios between different performance levels departed substantially from the d' ratio predicted by the simple noisy linear amplifier model. An elaborated perceptual template model [Vision Res. **38**, 1183 (1998)] with nonlinear transducer functions and multiplicative noise in addition to the additive noise in the simple linear amplifier model leads to a substantially better description of the data and suggests a reinterpretation of earlier results that relied on the simple noisy linear amplifier model. The relationship of our model and method to other recent parallel and independent developments [J. Opt. Soc. Am. A **14**, 2406 (1997)] is discussed. © 1999 Optical Society of America [S0740-3232(99)02903-8]

OCIS codes: 330.1800, 330.1880, 330.4060, 330.5510, 330.6100, 330.7310.

1. INTRODUCTION

Limited by various sources of noise such as intrinsic stimulus variability, receptor sampling errors, randomness of neural responses, and loss of information during neural transmission, perceptual processes exhibit various inefficiencies. One can characterize such inefficiencies at an overall system level by conceptualizing perceptual processes as perfect, noise-free computations with separate, equivalent internal noise—random internal noise necessary to produce the degree of inefficiency exhibited by the perceptual system.[1–5] Although this characterization does not distinguish between various sources for the inefficiency, it does allow us to quantify the overall efficiency of the perceptual system and to compare the efficiency of the perceptual system in different perceptual tasks.[3,4,6,7] In fact, specification of internal noise has become a requirement of any realistic model for human perception.[8]

To estimate the amount of equivalent internal noise, psychologists have adopted a method called equivalent input noise that was developed and is frequently used in electrical engineering to measure the properties of noisy amplifiers.[9–11] The method adds systematically increasing amounts of external noise to the signal stimulus and observes how threshold—signal stimulus energy required for an observer to maintain a given performance level—depends on the amount of external noise.[3] The human observer is modeled as a (simplest possible) noisy linear amplifier (perfect linear amplification plus additive noise). The amount of equivalent additive noise can then be scaled in relation to the magnitude of external noise from the threshold versus external noise functions.

The equivalent input noise procedure has been applied to a wide range of perceptual tasks and, in many cases, has produced results with very important implications.[1–7,12–25] At the same time, there have been several attempts to improve the noisy linear amplifier model (LAM) to better accommodate experimental results. For example, Pelli attributed all the inconsistencies between the data and the noisy LAM to stimulus uncertainty in the decision process, asserting that, in detecting a signal, observers have to monitor hundreds of irrelevant channels.[26] Burgess and Colborne found that an additional multiplicative noise source with an amplitude equal to approximately 0.6–0.8 times the added external white noise was necessary to account for their data.[27] Given this additional multiplicative noise source,

0740-3232/99/030764-15$15.00 © 1999 Optical Society of America

they suggested that previous estimates of internal noise had been too high and should be revised.

We first review the simple noisy LAM, because it has been used so widely in perception studies. Although it has been argued that the LAM is an incomplete model of performance,[26,27] it is still extensively used.[23–25] The current approach reveals key deficits of the model. We develop a strong prediction from the model: The ratio between two thresholds at each external noise contrast is equal to the ratio of the corresponding d' values, independent of the observer and the particular external noise contrast. We then critically examine this prediction in two experiments. To reduce stimulus uncertainty effects, a two-alternative forced-choice (2AFC) identification task was used in the first experiment.[28] We then extend the study to two-interval forced-choice (2IFC) detection in the second experiment. We demonstrate that the simple noisy LAM systematically mispredicts the relationship between thresholds at different performance levels. We then describe a perceptual template model (PTM), which was developed by Dosher and Lu[29] and by Lu and Dosher[30,31] in studying attention mechanisms.

The PTM[29–31] is an elaboration of the simple noisy LAM with two additional components: nonlinear transducer function,[32,33] and multiplicative noise.[8,27,32–49] To account for human performance in 2AFC identification, no additional stimulus uncertainty was assumed in the decision process. The PTM provides a much better account of the data than does the simple LAM, suggesting that conclusions and additive noise estimates in earlier studies based on the simple model may require reevaluation.

Parallel to our development of the PTM, Eckstein *et al.*[50] proposed a different model of a human observer, which we shall call the Eckstein–Ahumada–Watson (EAW) model, consisting of additive and multiplicative noise sources and a decision process with stimulus uncertainties. The EAW model is mathematically quite similar (although not identical) to the Lu–Dosher PTM.[29–31] However, the two models have very different conceptual interpretations and could potentially be distinguished experimentally. The relationship between the PTM and the EAW model is discussed in detail in Appendix A and in Section 6.

2. NOISY LINEAR AMPLIFIER MODEL

A. Model

The existence of a sensory threshold for every perceptual process suggests that the perceptual system is limited by an equivalent internal noise source whose amplitude does not depend on the input. Such an additive noise, preceded by a noise-free linear amplification and followed by a decision process, constitutes the simplest model for an observer with sensory thresholds.[2–5]

Figure 1(a) depicts such a model with three stages: (1) a noise-free amplifier that amplifies the signal stimulus by β with a total gain of 1.0 (integrated over space, time, and feature space); (2) an independent, additive random Gaussian noise source with standard deviation N_{add}; and (3) a decision stage.

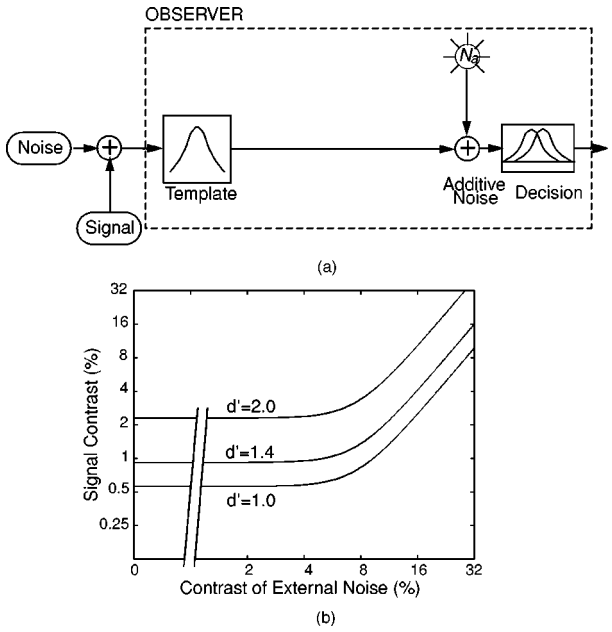

Fig. 1. (a) Simple noisy LAM. The LAM has three major components: (1) a perceptual template, (2) an additive internal noise source, and (3) a decision process. A good example of a perceptual template is a spatial-frequency filter $F(f)$, with a center frequency and a bandwidth such that a range of frequencies adjacent to the center frequency passes through with smaller gains. Limitations of human observers are modeled as equivalent additive internal noise—a noise source whose amplitude does not vary with stimulus contrast. Additive noise is added to the output from template matching, and the noisy signal is the input to a task-appropriate decision process. (b) Contrast threshold of the noisy LAM as a function of the contrast of the external noise for three different performance levels ($d' = 1, 1.41, 2.0$) for a hypothetical case in which $N_{\text{add}} = 0.00393$, $\beta = 2.4$. For each d' level, the contrast threshold is nearly constant when the amplitude of the external noise is small; it increases with the amplitude of the external noise at high noise amplitudes. In that range external noise dominates internal noise; performance is mostly determined by the amount of external noise.

B. Threshold versus External Noise Functions

For a given input consisting of a signal stimulus with rms contrast c plus an experimenter-controlled random noise stimulus whose pixels are drawn from a Gaussian distribution with standard deviation N_{ext}, the total amount of signal at the decision stage is

$$S = \beta c, \tag{1}$$

and the total variance of noise at the decision stage is the summation of the variance of the external and the internal noise:

$$N^2 = N_{\text{ext}}^2 + N_{\text{add}}^2. \tag{2}$$

Thus signal discriminability, d', determined by the signal-to-noise ratio at the decision stage is given by

$$d' = \frac{S}{N} = \frac{\beta c}{\sqrt{N_{\text{ext}}^2 + N_{\text{add}}^2}}. \tag{3}$$

For a fixed threshold d', we can rearrange Eq. (3) to solve for threshold contrast c_t as a function of the amount of external noise:

$$c_\tau^2 = \left(\frac{d'}{\beta}\right)^2 (N_{\text{ext}}^2 + N_{\text{add}}^2). \tag{4}$$

If we define $k = (d'/\beta)$, Eq. (4) can be further simplified:

$$c_\tau^2 = k^2(N_{\text{ext}}^2 + N_{\text{add}}^2). \tag{5}$$

In Fig. 1(b) we plot $\log(c_\tau)$ as a function of $\log(N_{\text{ext}})$ for three fixed threshold levels ($d' = 1, 1.41, 2.0$). Such graphs have three regions: (1) When the external noise N_{ext} is much smaller than the internal noise N_{add}, N_{add} is the limiting factor in performance. Thus threshold contrast c_τ varies minimally with the amount of external noise. (2) When the external noise N_{ext} is much larger than the internal noise N_{add}, external noise becomes the dominant limiting factor. Thus $\log(c_\tau)$ increases directly as a linear function of $\log(N_{\text{ext}})$. (3) When external noise is comparable with internal noise, there is a smooth transition from region 1 to region 2.

C. Threshold Ratio at Two d' Levels

In this section we derive an important prediction from the noisy LAM: The ratio between two thresholds at each external noise contrast is equal to the ratio of the two corresponding d' values, independent of the observer and the external noise level.

Following Eq. (4) above, we find that, at a given external noise contrast N_{ext}, threshold c_1 for d'_1 and threshold c_2 for d'_2 are

$$c_1 = \frac{d'_1}{\beta}\sqrt{N_{\text{ext}}^2 + N_{\text{add}}^2}, \tag{4a}$$

$$c_2 = \frac{d'_2}{\beta}\sqrt{N_{\text{ext}}^2 + N_{\text{add}}^2}. \tag{4b}$$

If we take the ratio between Eqs. (4a) and (4b),

$$\frac{c_1}{c_2} = \frac{d'_1}{d'_2}. \tag{6}$$

That is, the simple noisy LAM predicts that the signal contrast ratio between two threshold performance levels for a given external noise contrast is independent of that particular external noise level. Moreover, it predicts that the ratio is equal to the ratio of the corresponding d' values at the two performance levels.

This prediction [Eq. (6)] forms an important test for the internal consistency of the model that is critically examined in Section 3.

3. EXPERIMENTS

We conducted experiments to measure threshold versus external noise functions at several performance (d') levels in two different perceptual tasks: (1) 2AFC Gabor orientation identification, and (2) 2IFC Gabor detection. The aim was to test the threshold ratio prediction from the simple noisy LAM.

A. Experiment 1: Threshold versus External Noise at Three Performance Levels in a Gabor Orientation Identification Task

In this experiment we used the method of constant stimuli[51] to measure full psychometric functions for each observer at each of nine external noise levels in a 2AFC identification experiment. We report contrast thresholds at three defined performance levels in each external noise condition. This resulted in three threshold versus external noise functions. The family of such functions provides additional ratio tests of the simple noisy LAM.

1. Method

Stimulus and Display. The signals in the task were Gabor patterns tilted θ deg either to the right or left of the vertical with luminance $l(x, y)$ at location (x, y) defined by Eq. (7a):

$$l(x,y) = l_0\left\{1.0 + c_{\text{peak}}\sin[2\pi f(x\cos\theta \pm y\sin\theta)] \times \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)\right\}, \tag{7a}$$

where $f = 0.9$ cycle(s) per degree (c/deg), $\sigma = 1.1$ deg, $\theta = 10$ deg, $l_0 = 27$ cd/m^2. The peak contrast c_{peak} was set by the experimenter according to pilot studies.

Thus local stimulus contrast, defined as $c(x,y) = [l(x, y) - l_0]/l_0$, is

$$c(x,y) = c_{\text{peak}}\sin[2\pi f(x\cos\theta \pm y\sin\theta)] \times \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right). \tag{7b}$$

Each Gabor pattern was rendered on a 128×128 pixel grid, subtending 5.9 deg $\times$ 5.9 deg. External noise frames were made of 2×2 pixel patches each subtending 0.092×0.092 deg^2. The gray level of each pixel patch was sampled from a Gaussian distribution with mean 0 and certain variance depending on the desired amount of external noise. To guarantee that the external noise conformed to the Gaussian distribution, the maximum standard deviation of the noise was 33% or less of the maximum achievable contrast. The size of the noise frames is identical to that of the signal frame.

Apparatus. All the signal and noise frames were generated and displayed in real time by use of programs based on a C++ version of Video Toolbox[52,53] on a 7500/100 Macintosh computer. The stimuli were presented on a Nanao Technology FlexScan 6600 monitor with a P4 phosphor and a refresh rate of 120 frames/s, driven by the internal videocard in the Macintosh. A special circuit[52] was used to combine two 8-bit output channels of the videocard to produce 6144 distinct gray levels (12.6 bits).

A psychophysical procedure was used to generate a linear lookup table that evenly divides the entire dynamic range of the monitor (from 1 to 53 cd/m^2) into 256 levels. The background luminance was set at 27 cd/m^2.

All displays were viewed binocularly with the natural pupil at a viewing distance of approximately 72 cm in a dimly lighted room.

Design. The experiment consisted of six sessions, each with 630 trials. In each session nine different external noise rms contrasts (0, 0.01, 0.02, 0.04, 0.08, 0.12, 0.16, 0.25, and 0.33) were used. A psychometric function, consisting of seven sample points, was generated at each external noise level. The signal contrasts for each noise condition were predetermined through pilot studies.

All the conditions were intermixed. For each observer, data from all six sessions are combined to generate nine psychometric functions, one at each external noise contrast level.

Figure 2 depicts a typical trial. Initialized by the observer with a key press, a trial starts with a 0.25-s fixation cross in the center of the monitor. In the next five refreshes, each lasting 8.3 ms, two independent noise frames, one signal frame (a Gabor patch tilted either to the left or to the right), and another two independent noise frames appear in the center of the display. The observer identified with a key press the orientation of the Gabor. A correct response was followed immediately by a brief beep. The program paused for 1 s before proceeding to the next trial.

Observers. Two University of Southern California students with normal or corrected-to-normal vision, naïve to the purposes of the experiment, and the first author served as observers in the experiment.

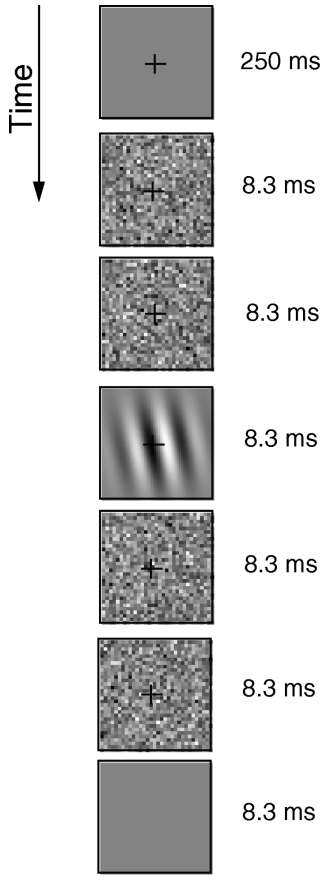

Fig. 2. Trial in the Gabor orientation identification task (experiment 1). A trial starts with a 250-ms fixation cross. In the next five stimulus frames, each lasting 8.3 ms, two external noise frames, one signal Gabor frame, followed by another two noise frames, appear in the center of the display. The Gabor patch is tilted either to the left or to the right; its contrast varied across trials according to a random mixture of nine signal contrasts.

2. Results

For each observer at each external noise contrast, we calculated percent correct (in our notation, Pc) in Gabor orientation identification at seven different signal contrast levels. To compute threshold at various performance levels, we first fitted a Weibull function[54]:

$$\mathrm{Pc} = 1.0 - 0.5 \times 2^{-(c/\alpha)^{\eta}} \tag{8}$$

to each of the nine psychometric functions for each observer, using a maximum-likelihood procedure.[55]

We then computed threshold signal contrast at three performance levels: 65%-, 75%-, and 85%-correct identification. In a 2AFC procedure like this one, these three performance levels correspond to d' values of 0.7706, 1.3490, and 2.0729, respectively.

In the upper panels of Fig. 3, log threshold (rms signal contrast) versus log external noise (rms) contrast functions are plotted at three performance levels for each of the three observers. Error bars[56] indicate the standard deviation of each threshold. For each observer, the threshold versus external noise functions at the three different performance levels are approximately vertical shifts of each other, suggesting that the threshold ratio between two performance levels at a given external noise level is a constant across all the external noise levels. Threshold ratios between 75% and 65% correct and between 85% and 75% correct are plotted at each external noise level for each observer in the lower panels of Fig. 3.

The geometric mean threshold ratios between 75%- and 65%-correct performance levels across all the external noise levels are 1.2579 for subject QL, 1.2413 for subject SM, and 1.2928 for subject ZL. The geometric mean threshold ratio between 85%- and 75%-correct performance levels across all the external noise levels are 1.2101 for QL, 1.1967 for SM, and 1.2379 for ZL. However, the noisy LAM [Eq. (6)] predicts that the two threshold ratios should be 1.7506 and 1.5366 and that they should be the same for all observers. Clearly, the noisy LAM systematically overestimates the threshold ratios.

The simple LAM, with Gaussian d' assumptions, requires that the psychometric functions correspond to a cumulative Gaussian with $\mu = 0$ (signal contrast). This prediction constitutes an additional test of the LAM. However, the observed psychometric functions were very poorly fitted by the cumulative Gaussian form (mean $r^2 = 0.6792$ compared with $r^2 = 0.8818$ for the Weibull estimates). This is an alternative basis for rejecting the LAM.

We chose to evaluate the LAM and the PTM in terms of the threshold ratio tests because (1) the contrast ratios highlight the nature of the model failures; (2) the ratios clearly reveal the prediction of constant threshold ratios over external noise levels, which is a consistency constraint on both models; (3) the Weibull fits and ratio tests provide a direct and simple way to calculate and test the applicability of the LAM without requiring that full model

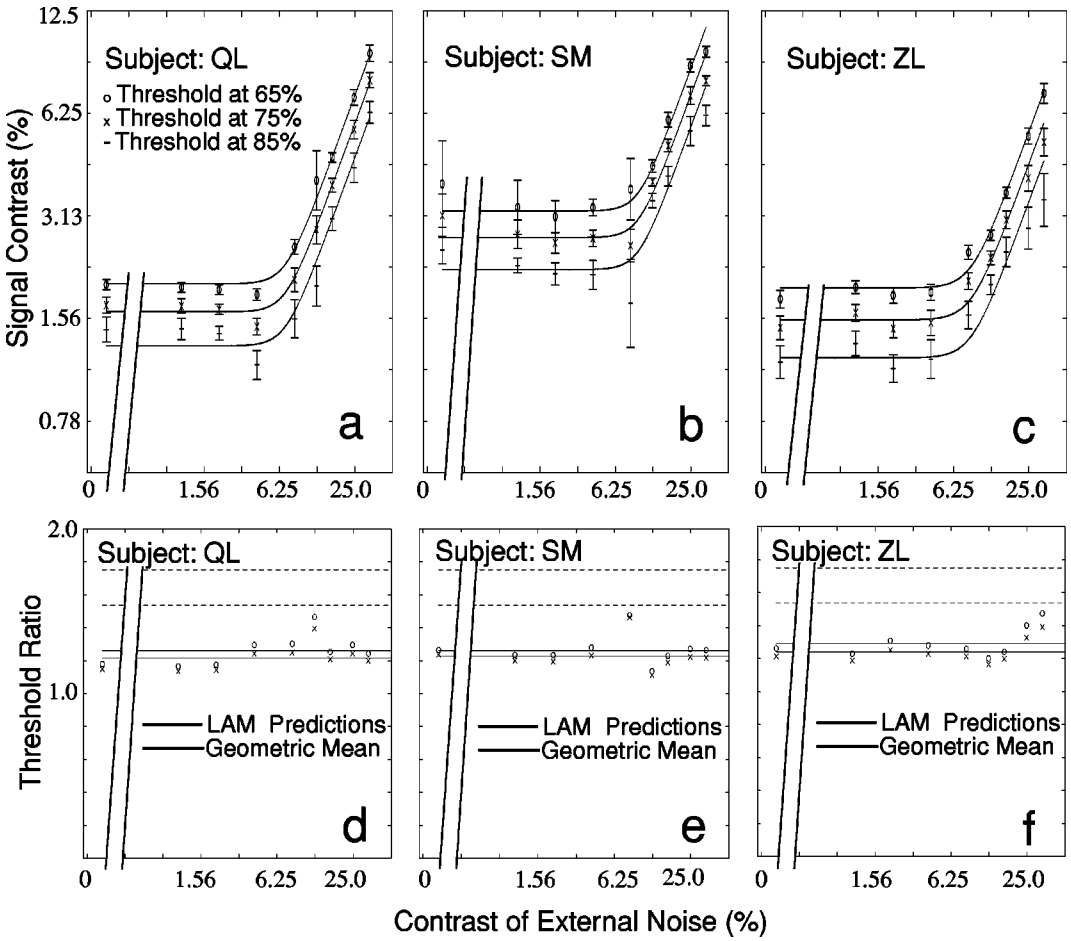

Fig. 3. (a)–(c) Threshold contrast (rms contrast of the Gabor) versus external noise level (rms contrast of the Gaussian random noise) in a central Gabor orientation discrimination task for three observers, each at three different performance levels: –'s indicate thresholds at 65% correct, ×'s indicate thresholds at 75% correct, and ○'s indicate thresholds at 85% correct. Error bars indicate the standard deviation of each threshold estimation. The solid curves are generated with the parameters of the best-fitting PTM model. (d)–(f) Threshold ratios between performance levels versus external noise contrast for each of the three observers. The solid horizontal lines in each graph depict the geometric mean of all the threshold ratios across external noise levels for each observer. The dashed horizontal lines denote predicted ratios from the noisy LAM.

fits be carried out; and (4) the ratios are the basis of constructive solutions for constrained estimates of model parameters.

Alternative direct model tests and a discussion of the constructive solutions appear in Section 5.

B. Experiment 2: Threshold versus External Noise at Three Performance Levels in a Gabor Detection Task

In this experiment we extend our observations from 2AFC orientation identification to 2IFC Gabor detection. Many of the classic experiments measuring equivalent internal noise were detection experiments.[1–5] We again used the method of constant stimuli[51] to measure full psychometric functions for each observer at each of nine external noise contrasts. We computed thresholds at three performance levels in each external noise condition. The resulting three threshold versus external noise functions provide a further test of the simple noisy LAM in the detection domain, where it has been most extensively applied.

1. Method

Unless noted, the method is identical to that of experiment 1.

Stimulus and Display. The signals in the task were vertical Gabor patterns:

$$l(x,y) = l_0\left[1.0 + c\,\sin(2\pi f x)\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)\right], \quad (9)$$

with parameters identical to those of the Gabor patterns used in experiment 1.

Design. Figure 4 depicts a typical trial. Initialized by the observer with a key press, a trial starts with a 0.25-s fixation cross in the center of the monitor. Starting with a brief beep, the first interval, consisting of five refreshes each lasting 8.3 ms, is made of two independent noise frames, one signal or blank frame, and another two independent noise frames in the center of the display. After another 0.50-s delay, starting with another brief beep, the second interval consists of two independent noise frames, one blank or signal frame, and another two independent noise frames. The observer has to report with a key press which of the two intervals contains the Gabor. A correct response is followed immediately by a brief beep. There was a 1-s intertrial interval.

Observers. The observers were the same as those in experiment 1.

2. Results

Following the same data analysis procedure as in experiment 1, we computed threshold signal contrast at three performance levels, 65%, 75% and 85% correct, by first fitting a Weibull function to each of the psychometric functions at each external noise level for each observer. These three performance levels correspond to d' values of 0.7706, 1.3490, and 2.0729, respectively. We also estimated the standard deviations for the thresholds, following the same theoretical resampling procedure as in experiment 1.

In the upper panels of Fig. 5, we plot log threshold (rms signal contrast) versus log external noise (rms) contrast functions at three performance levels for each of the three observers. Error bars[57] indicate the standard deviation at each threshold. For each observer, the threshold versus external noise functions at the three different performance levels were approximately vertical shifts of each other, again suggesting that the threshold ratio between two performance levels at each external noise level was a constant across all the external noise levels. Hence the threshold ratios between 75% and 65% correct and between 85% and 75% correct for each external noise level are plotted for each observer in the lower panels of Fig. 5.

The geometric average threshold ratio between 75%- and 65%-correct performance levels across all the external noise levels was 1.3132 for QL, 1.2793 for SM, and 1.3461 for ZL. The geometric average threshold ratio between 85%- and 75%-correct performance levels across all the external noise levels was 1.2541 for QL, 1.2271 for SM, and 1.2801 for ZL.[58] As before, the noisy LAM [Eq. (6)] predicted that the two threshold ratios should be 1.7506 and 1.5366 for all the observers. Again, the noisy LAM systematically overestimated the threshold ratios.

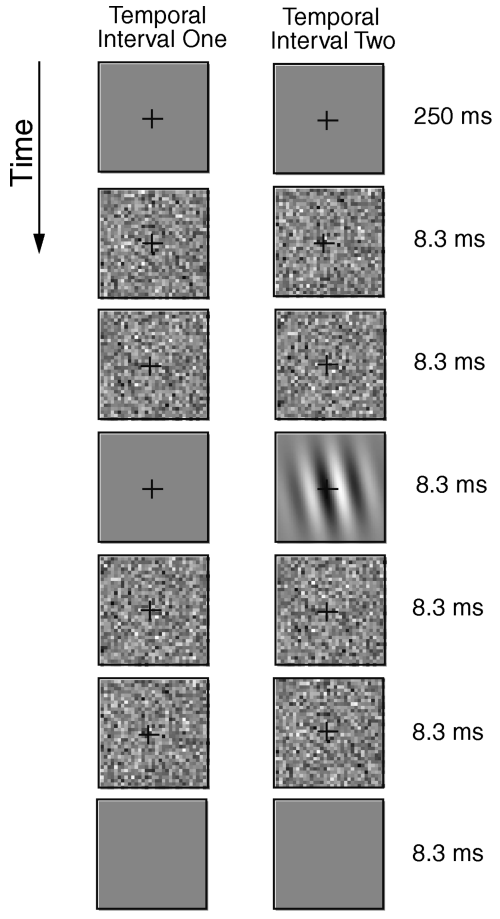

Fig. 4. Trial in the 2IFC Gabor detection task (experiment 2). A trial consists of two intervals. The first interval starts with a 250-ms fixation cross. In the next five stimulus frames, each lasting 8.3 ms, two external noise frames, one signal or blank frame, and then another two noise frames appear in the center of the display. The second interval starts 500 ms after the end of the first interval and is identical to the first, but only one of the intervals contains a signal Gabor. The contrast of the Gabor varied across trials according to a random mixture of nine signal contrasts.

4. NOISY PERCEPTUAL TEMPLATE MODEL

From both the identification and the detection experiments, it is clear that the simple noisy LAM cannot give an adequate account for the data at multiple performance levels. It must be elaborated to allow threshold ratios less than d_1'/d_2' and to allow observer-to-observer variations in ratio. In this section we describe a PTM[29–31] that preserves many of the useful properties of the noisy LAM yet is sufficiently elaborated that it accounts for the observed data pattern. The PTM is related to several of the models and observations in pattern vision and pattern masking.[32,33,35,42,47,49,59,60] Specifically, it incorporates the notion of a nonlinear transducer function. To our knowledge, however, this form has not been previously applied to white-noise masking.

A. Noisy Perceptual Template Model: Overview

We elaborate the noisy LAM by adding two components: a nonlinear transducer function in the signal path, and a multiplicative noise source whose energy is controlled by the stimulus energy. The nonlinear transducer function is necessary to correct the overestimation of threshold ratios between any two performance levels; the multiplicative noise may in some cases be necessary to simultaneously accommodate data at three performance levels.

Figure 6(a) illustrates the PTM. It consists of (1) a perceptual template with certain tuning characteristics, (2) a nonlinear transducer function, (3) a multiplicative internal noise whose amplitude is a monotonic function of the input energy in the signal path, (4) an additive internal noise source, and (5) a decision process.

1. Perceptual Template with Certain Tuning Characteristics

The first component of the PTM is a perceptual processor, termed a perceptual template, which allows inputs with different physical characteristics to go through the processor with different gains. For example, a template matching function could be a spatial-frequency filter $F(f)$, with a center frequency and a bandwidth such that a range of frequencies adjacent to the center frequency passes through the processor with smaller gains. A template matching function might, however, be far more complex (e.g., templates for objects, faces, etc.). It is related to the concept of a matched filter in prior investigations of identification performance.[28]

Without loss of generality,[61] we assume that the total gain (gain integrated over feature space and time) of the template matching function is 1.0 and that the gain for a signal-valued stimulus is β. The gain β is directly related to the covariance of the stimulus and the template. For a signal stimulus of contrast c, the output from the template matching function S has an amplitude

$$S = \beta c. \tag{10a}$$

The external noise—white Gaussian noise added to the stimulus by the experimenter—has a flat Fourier spectrum (equal energy at all the spatial frequencies). For external noise with standard deviation of N_{ext}, the output from the template is directly proportional to N_{ext}. In fact, because the total gain of the template matching function is 1.0, the output from the template has the same standard deviation:

$$\sigma_{\text{ext}} = N_{\text{ext}}. \tag{11a}$$

2. Nonlinear Transducer Function

We need to introduce an expansive nonlinear transducer function in the signal path to correct the problem of overestimating threshold ratios in the LAM. Following the pattern vision literature,[32,33] we choose a power function ($\|\cdot\|^{\gamma_1}$). Thus, after the nonlinear transduction stage, the signal stimulus becomes

$$S' = \beta^{\gamma_1} c^{\gamma_1}, \tag{10b}$$

and the external noise becomes

$$\sigma_{\text{ext}}{}' = N_{\text{ext}}^{\gamma_1}. \tag{11b}$$

3. Multiplicative Internal Noise

There is evidence both from psychophysics[8,27,34–36,42,47–49] and from neurophysiology[37–41,43–46] that perceptual task performance may be limited by a form of noise whose amplitude is directly related to the total amount of stimulus energy. We model this limitation by postulating an independent noise source that is Gaussian distributed with mean 0 and standard deviation σ_{mul}, where

$$\sigma_{\text{mul}}^2 = N_{\text{mul}}^2 (N_{\text{ext}}^{2\gamma_2} + \beta^{2\gamma_2} c^{2\gamma_2}). \tag{12}$$

We assume that the signal stimulus and the external noise both go through a nonlinear transducer function

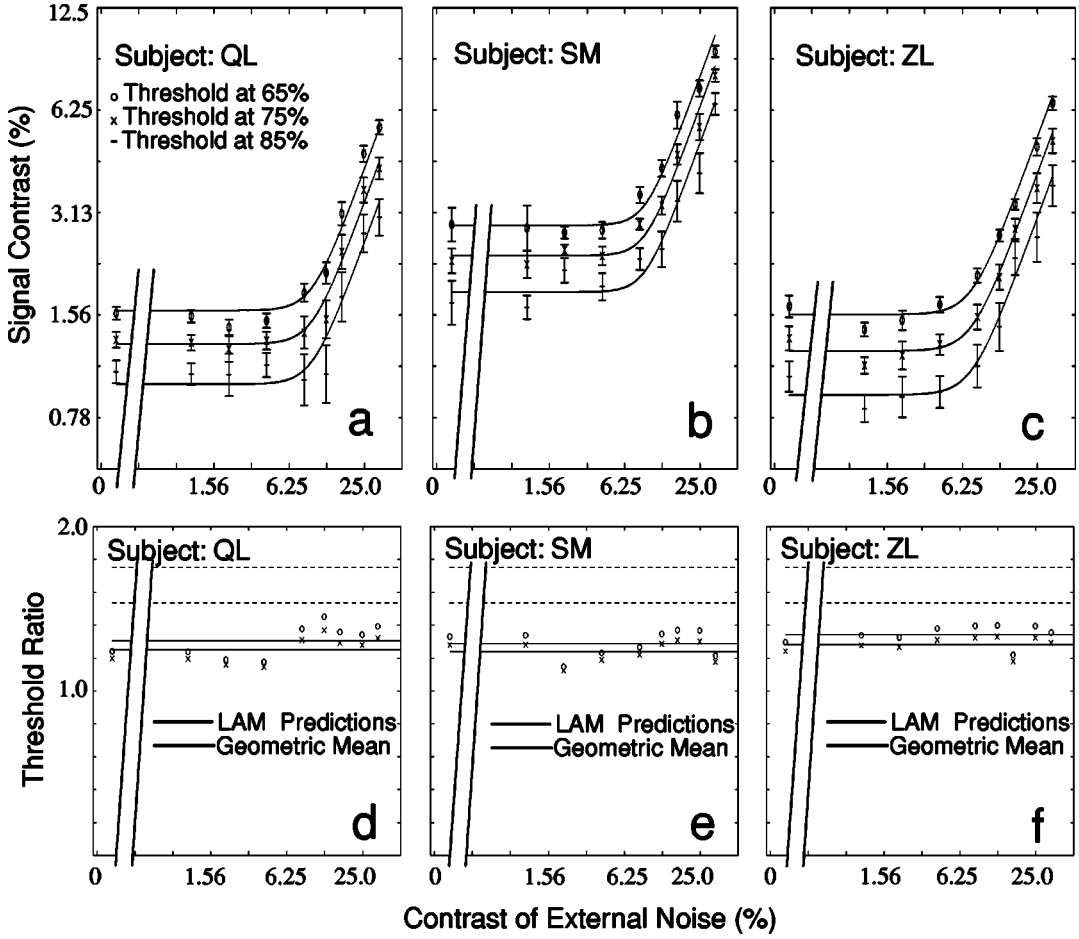

Fig. 5. (a)–(c) Threshold contrast (rms contrast of the Gabor) versus external noise level (rms contrast of the Gaussian random noise) in a central Gabor detection task for three observers each at three different performance levels: −'s indicate thresholds at 65% correct, ×'s indicate thresholds at 75% correct, and ○'s indicate thresholds at 85% correct. Error bars indicate the standard deviation of each threshold estimation. The solid curves are generated with the parameters of the best-fitting PTM model. (d)–(f) Threshold ratio between performance levels versus external noise contrast for each of the three observers. The solid horizontal lines in each graph depict the geometric mean of all the threshold ratios across external noise levels for each observer. The dashed horizontal lines denote predicted ratios from the noisy LAM.

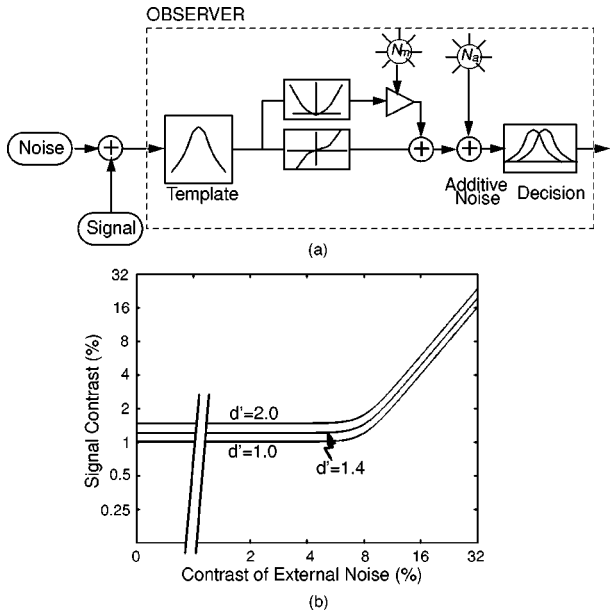

Fig. 6. (a) Noisy PTM. There are five major components: (1) a perceptual template, (2) nonlinear transducer functions, (3) a multiplicative internal noise source, (4) an additive internal noise source, and (5) a decision process. A good example of a perceptual template is a spatial-frequency filter $F(f)$, with a center frequency and a bandwidth such that a range of frequencies adjacent to the center frequency passes through with smaller gains. The nonlinear transducer function takes the form of an expansive power function. Limitations of human observers are modeled as equivalent internal noise. Multiplicative noise is an independent noise source whose amplitude is proportional to the (average) amplitude of the output from the perceptual template. Additive internal noise is another noise source whose amplitude does not vary with signal strength. Both multiplicative and additive noises are added to the output from template matching, and the noisy signal is the input to a task-appropriate decision process. (b) $\mathrm{Log}(c_\tau)$ plotted as a function of the standard deviation of the external noise for three fixed threshold levels ($d' = 1.0, 1.41, 2.0$).

$(\|\cdot\|^{\gamma_2})$ before their total energy is computed. We have chosen this particular form of multiplicative noise without cross terms for computational simplicity. For the data set reported in this paper, including cross terms in multiplicative noise did not lead to substantially different model estimates.

Note that the formulation of multiplicative noise is mathematically equivalent to a theory of contrast gain control.[8,35,62]

4. Additive Internal Noise

As in the simple LAM, to model human limitations at detection threshold we postulate an independent additive noise source whose amplitude does not vary with signal strength. Independent additive noise is modeled as a Gaussian random variable with mean 0 and standard deviation N_{add}.

5. Decision Process

A noisy signal, quantitatively characterized by β, S', $\sigma_{\mathrm{ext}}{}'$, σ_{mul}, σ_{add}, γ_1, and γ_2, is submitted to a decision process. The details of the decision process depend on the particular task, e.g., detection versus identification. These are modeled elsewhere.[63] In this paper we exclude stimulus uncertainty effects in the decision process for identification. This approach follows Burgess's[28] demonstration that identification performance made nearly ideal use of matched filters. We focus on the signal-to-noise ratio.

B. Threshold Predictions for White Gaussian External Noise

In this section we describe how an observer's perceptual threshold depends on the amplitude of the external noise added to the signal in the PTM.

Signal discriminability, d', is determined by the strength of the signal, S', and the standard deviation of the total noise (external and internal), σ_N:

$$d' = S'/\sigma_N. \tag{13}$$

The signal strength is given by Eq. (10b). Since all the noise sources (external, multiplicative, and additive noises) in the PTM are independent,[8,35,62] the total variance of the noise σ_N^2 is the sum of the variances of all the noise sources:

$$\begin{aligned}\sigma_N^2 &= \sigma_{\mathrm{ext}}'^2 + \sigma_{\mathrm{mul}}^2 + \sigma_{\mathrm{add}}^2 \\ &= N_{\mathrm{ext}}^{2\gamma_1} + N_{\mathrm{mul}}^2(\beta^{2\gamma_2}c^{2\gamma_2} + N_{\mathrm{ext}}^{2\gamma_2}) + N_{\mathrm{add}}^2. \end{aligned}\tag{14}$$

Combining these facts [Eqs. (10b), (13), and (14)] gives

$$d' = \frac{\beta^{\gamma_1}c^{\gamma_1}}{[N_{\mathrm{ext}}^{2\gamma_1} + N_{\mathrm{mul}}^2(\beta^{2\gamma_2}c^{2\gamma_2} + N_{\mathrm{ext}}^{2\gamma_2}) + N_{\mathrm{add}}^2]^{1/2}}. \tag{15}$$

The external noise in the stimulus had a Gaussian distribution. After nonlinear transduction, the distribution of the external noise might deviate from the Gaussian distribution. However, spatial and temporal summation in the perceptual system (see Ref. 61) should reduce this deviation. When the external noise is combined with additive and multiplicative noises, both of which are Gaussian distributed, we assume that the sum of the noises is approximately Gaussian. However, we restrict ourselves to performance levels below 90% so as to avoid the tails of the distribution. The Gaussian assumption is not central to the development of the PTM outlined above, but it does simplify the application to signal detection estimation: The Gaussian noise distribution allows us to use the Gaussian form of signal detection calculations.

In the special case in which $\gamma = \gamma_1 = \gamma_2$ (corresponding to the situation in which the rising portion of the threshold versus external noise function has a slope of 1.0), we can rearrange Eq. (15) to express the required threshold signal contrast c_τ as a function of the amount of external noise for a fixed d':

$$c_\tau = \frac{1}{\beta}\left[\frac{(1 + N_{\mathrm{mul}}^2)N_{\mathrm{ext}}^{2\gamma} + N_{\mathrm{add}}^2}{1/d'^2 - N_{\mathrm{mul}}^2}\right]^{1/2\gamma}. \tag{16}$$

Figure 6(b) plots $\log(c_\tau)$ as a function of the log of the standard deviation of the external noise for three fixed threshold levels ($d' = 1$, 1.41, 2.0). The threshold versus external noise functions have properties very similar to those derived from the simple LAM: (1) When the external noise N_{ext} is small, threshold signal contrast c_τ increases only slightly with the amount of external noise;

(2) When the external noise N_{ext} is very large, $\log(c_\tau)$ increases directly as a linear function of log external noise; (3) There is a smooth transition when external noise is comparable with internal noise.

C. Threshold Ratio between Two Performance Levels for a Given External Noise Contrast

The threshold ratio at each external noise level between two performance levels, derived from Eq. (16), is

$$\frac{c_1}{c_2} = \left(\frac{1/d_2'^2 - N_{mul}^2}{1/d_1'^2 - N_{mul}^2}\right)^{1/2\gamma}. \tag{17}$$

Thus, for two performance levels, the ratio between the thresholds at each external noise level is predicted to be a constant, independent of the external noise level. We have introduced two extra free parameters, γ and N_{mul}, in the threshold ratios, which could, in principle, vary across observers.

From Eq. (17) it follows that, to fully characterize the elaborated PTM, we need to estimate at least two threshold ratios: threshold versus external noise contrast functions at three performance levels. This corresponds to the data reported for both experiments 1 and 2. This requirement could also easily be satisfied by some of the existing data in the literature on equivalent internal noise if several points on the psychometric functions can be computed from the data. Thus the PTM and the analytical approach developed in this section provides an easy way to interpret the data in the previous literature without any requirement for new experimentation.

5. APPLICATION OF THE PERCEPTUAL TEMPLATE MODEL

In this section we apply the PTM to the data in experiments 1 and 2.[64] We also compare the quality of the PTM fit to that of the noisy LAM.

A. Estimation Procedure

Our estimation procedure was implemented in MATLAB[65] and was applied separately to each data set for each observer separately for the two experiments. The procedure consists of

1. Solving the following equations [derived from Eq. (17)] for γ and N_{mul}:

$$k_1 = \left(\frac{1/0.7706^2 - N_{mul}^2}{1/1.3490^2 - N_{mul}^2}\right)^{1/2\gamma},$$

$$k_2 = \left(\frac{1/1.3490^2 - N_{mul}^2}{1/2.0729^2 - N_{mul}^2}\right)^{1/2\gamma}, \tag{17a}$$

where k_1 is the geometric mean of threshold ratio between correct-performance levels of 75% (corresponds to $d' = 1.3490$) and 65% (corresponds to $d' = 0.7706$) across external noise levels for a given observer and k_2 is the geometric mean of threshold ratio between correct-performance levels of 85% (corresponds to $d' = 2.0729$) and 75% (corresponds to $d' = 1.3490$) across external noise levels for a given observer[66];

2. Given γ and N_{mul}, using Eq. (16) to compute $\log(c_\tau^{theory})$ from the PTM with guessed parameters (N_{add} and β) for each external noise contrast level;

3. Computing the squared difference between the log threshold prediction from the model and the observed sqdiff $= [\log(c_\tau^{theory}) - \log(c_\tau)]^2$ for each external noise condition[67];

4. Computing L: summation of sqdiff from all the external noise conditions;

5. Using a gradient descent method to adjust N_{add} and β to find the minimum of L;

6. After obtaining the least L, computing the r^2 statistic to evaluate the goodness of the model fit:

$$r^2 = 1.0 - \frac{\sum [\log(c_\tau^{theory}) - \log(c_\tau)]^2}{\sum \{\log(c_\tau) - \text{mean}[\log(c_\tau)]\}^2}, \tag{18}$$

where Σ and mean ( ) run over all the data points for a particular observer in an experiment.

Similar steps were used to fit the noisy LAM to the data. The only difference between the two procedures is that, in the LAM, γ was fixed at 1, and N_{mul} was set to 0. Hence the LAM is a nested (reduced) model with respect to the PTM.

B. Results

Estimated parameters for each observer from both the noisy LAM and the PTM are listed in Tables 1 and 2 for experiments 1 and 2, respectively. The models were fitted to the threshold estimates at the three d' levels.

A special variant of the PTM with γ set to 1.0 was also considered. When $\gamma = 1.0$, from Eq. (17), we have

$$\left(\frac{c_1}{c_2}\right)^2 = \frac{1/d_2'^2 - N_{mul}^2}{1/d_1'^2 - N_{mul}^2}. \tag{17b}$$

When $N_{mul} = 0$, $(c_1/c_2)^2$ has its minimum value: $(d_1'/d_2')^2$. In both experiments, we have shown that the measured (c_1/c_2) value is less than (d_1'/d_2'). Thus, in fitting the PTM with $\gamma = 1.0$, the property of Eq. (17b) requires that N_{mul} be zero to achieve its minimum value. Thus this special variant of the PTM with $\gamma = 1$ is the same as the LAM, with an extra useless free parameter N_{mul} that is constrained to be 0. Computer fits of this model resulted in exactly the same solution as that of the LAM.

The PTM gives a much better account of the data. We used an F-test for nested models to statistically compare the two models. Defining the PTM as the fully saturated model and the LAM as the reduced model, we have

$$F(df_1, df_2) = \frac{(r_{full}^2 - r_{reduced}^2)/df_1}{(1 - r_{full}^2)/df_2}, \tag{19}$$

where $df_1 = k_{full} - k_{reduced}$ and $df_2 = N - k_{full} - 1$. The k values are the number of parameters in each model, and N is the number of predicted data points. For these threshold data estimated from the Weibull, $N = 27$, the number of parameters in the full PTM $k_{full} = 4$, and the number of parameters in the reduced LAM $k_{reduced} = 2$. The r_{full}^2 and $r_{reduced}^2$ values are taken from the PTM fit

Table 1. Best-Fitting LAM and PTM Parameters and Model Comparison (Experiment 1)

Subject	Model	k_1	k_2	N_{mul}	γ	N_{add}	β	r^2	$F(2, 22)$	p
QL	LAM	1.7506	1.5366	0	1	0.07383	6.217	0.8116		
	PTM	1.2579	1.2101	0.1979	2.55	0.001163	4.939	0.9889	175.7	2.96×10^{-14}
SM	LAM	1.7506	1.5366	0	1	0.1077	5.411	0.6540		
	PTM	1.2413	1.1967	0.2002	2.71	0.001756	4.038	0.9595	82.98	5.65×10^{-11}
ZL	LAM	1.7506	1.5366	0	1	0.09431	8.328	0.7923		
	PTM	1.2928	1.2379	0.2149	2.30	0.003734	6.602	0.9672	58.66	1.52×10^{-9}

Table 2. Best-Fitting LAM and PTM Parameters and Model Comparison (Experiment 2)

Subject	Model	k_1	k_2	N_{mul}	γ	N_{add}	β	r^2	$F(2, 22)$	p
QL	LAM	1.7506	1.5366	0	1	0.1035	11.06	0.7844		
	PTM	1.3132	1.2541	0.2004	2.15	0.006622	8.825	0.9751	84.25	4.88×10^{-11}
SM	LAM	1.7506	1.5366	0	1	0.09862	5.798	0.7649		
	PTM	1.2793	1.2271	0.2020	2.38	0.003426	4.545	0.9694	73.51	1.82×10^{-10}
ZL	LAM	1.7506	1.5366	0	1	0.07866	8.816	0.8830		
	PTM	1.3461	1.2801	0.1995	1.97	0.006209	7.339	0.9904	123.1	1.13×10^{-12}

and the LAM fit, respectively. This test evaluates whether the LAM fit is of a significantly worse quality. The values of $F(2, 22)$ were 175.7, 82.98, and 58.66 for subjects QL, SM, and ZL (all $p < 10^{-8}$) in experiment 1. The values of $F(2, 22)$ were 84.25, 73.51, and 123.1 for subjects QL, SM, and ZL (all $p < 2 \times 10^{-10}$) in experiment 2. Hence the noisy LAM can be rejected in favor of the PTM. This conclusion follows from the failure in ratio tests described above.

Across all six data sets reported in this paper, the exponent γ ranges between 1.97 and 2.71. γ values estimated from the data in experiment 1 were, in general, larger (2.30–2.71) than those estimated from the data in experiment 2 (1.97–2.38). This small though systematic variation of γ between the two experiments may reflect some intrinsic bias in the experiments: The threshold values found in identification experiments tend to be larger than those found in detection experiments.

The estimated γ values in both experiments (the mean γ is 2.52 in experiment 1 and 2.17 in experiment 2) would give rise to nonlinear transducer functions that are very similar to those assumed in pattern masking models.[32,33,35,47] This suggests that similar neural mechanisms are involved in white-noise masking and pattern masking.

An alternative approach to testing the LAM and the PTM evaluates their ability to directly predict the full psychometric functions. For both models, the psychometric functions take on the form $G(d'/2)$, where $G(\ )$ is cumulative Gaussian. As noted above, this requires that the psychometric function associated with the LAM be a cumulative Gaussian on c, with $\mu = 0$ (signal contrast). However, the psychometric function associated with the PTM is a cumulative Gaussian on a somewhat complex function of c^{γ}.

Full sets of maximum-likelihood fits were performed on the psychometric functions for both the LAM and the PTM. The results of these model fits corresponded closely to those reported above, on the basis of only three threshold levels. This confirms the claim that three threshold levels provide a sufficient test of the model. Estimated parameter values of the direct fits to the psychometric functions were very close to those listed in Tables 1 and 2. Over experiments and observers, the mean r^2 for the LAM was 0.6756, while the mean r^2 for the PTM was 0.8657. The latter value is remarkable because it approaches the mean r^2 for the Weibull fits of 0.8818. The Weibull fits were free to vary parameters independently for each external noise level, while the PTM must fit all external noise levels with the same parameters. Consistent with Tables 1 and 2, each comparison rejected the LAM in favor of the PTM ($p < 0.0001$).

6. RELATIONSHIP BETWEEN THE PRESENT MODEL AND OTHER RELATED MODELS

Parallel to our development of the PTM model, Eckstein *et al.*[50] independently proposed another elaborated LAM (the EAW model) to account for human performance in M-alternative ($M = 4$ in their experiments) forced-choice signal detection in white Gaussian noise added to various background patterns. The EAW model differs from the PTM in several key respects: (1) The EAW model assumes no nonlinear transducer functions; (2) the amplitude of the multiplicative internal noise in EAW is proportional only to the power of the external noise, while the amplitude of the multiplicative internal noise in the PTM is a function of both the amplitude of the signal and the amplitude of the external noise; (3) the decision process in the EAW model is affected by stimulus uncertainty, while the PTM assumes that, at least in the case of tasks involving a single stimulus location and known signals, the observer uses the best-matching template without additional stimulus uncertainty. In the case of multiple external locations or of an increase in the number of

possible signals, the PTM would require uncertainty extensions consistent with an ideal observer. However, the PTM does not invoke uncertainty as a free parameter of observer inefficiency, as does the EAW model.

Although the PTM and the EAW model are quite different conceptually, the mathematical properties of the two models are very similar.[68] This is not surprising, because it was recognized early in the auditory literature that the energy detection model (a nonlinear transducer) for tone detection can predict performance similar to that of the phase-uncertain observer.[69] The mathematical properties of the nonlinear transducer functions in the PTM are well approximated by the properties of the equivalent uncertainty process in the EAW model over a large range of d' levels. In fact, in the d' range $0.79 < d' < 2.50$, where data were collected, both in the present paper and in Eckstein *et al.*,[50] the two models may account for all the data essentially equally well (see Appendix A for details).

However, the two models do make different predictions when the d' levels involved are very small ($d' \approx 0.5$): While the PTM, in its current form, predicts that the threshold ratio between two d' criteria at each external noise level is independent of the particular external noise level, the EAW model induces changes in ratios at small d' values under stimulus uncertainty. This suggests a potential method for distinguishing the two models. In the d' range in which data were available in three rather different experimental conditions, both the data and the model provided by Eckstein *et al.*[50] confirm the constant-ratio prediction. Although experiments aimed at $d' < 0.5$ are technically demanding, such an experiment might distinguish the two models.

The choice of nonlinear transducer function versus decision uncertainty as a free parameter thus reflects different approaches to the problem. We prefer the nonlinear transducer function approach for the following reasons:

1. Nonlinear transducer functions are a key component of models accounting for pattern masking data.[32,33] Noise masking models should be, in principle, similar to pattern masking models because the same neural substrates are involved in both tasks;
2. The concept of a nonlinear transducer function is consistent with nonlinear properties of visual neurons[48];
3. In stimulus identification by well-practiced observers, previous evidence suggests that stimulus uncertainty does not appear to play a major role.[28] For example, Burgess[28] demonstrated that in ten-alternative forced-choice identification a model utilizing the best-matching

Table 3. r^2 Values for Best-Fitting EAW Models

Data	Uncertainty U								PTM
	$U = 0$	$U = 1$	$U = 2$	$U = 3$	$U = 4$	$U = 5$	$U = 6$	$U = 7$	
ql1	0.8116	0.8764	0.9042	0.9201	0.9302	0.9369	0.9425	0.9456	0.9889
sm1	0.6540	0.7721	0.8234	0.8529	0.8719	0.8845	0.8954	0.9013	0.9595
zl1	0.7923	0.8786	0.9145	0.9345	0.9469	0.9549	0.9615	0.9650	0.9672
ql2	0.7845	0.8655	0.8986	0.9167	0.9276	0.9346	0.9402	0.9432	0.9751
sm2	0.7649	0.8592	0.8989	0.9213	0.9353	0.9444	0.9521	0.9561	0.9694
zl2	0.8831	0.9360	0.9568	0.9677	0.9741	0.9779	0.9809	0.9824	0.9904
	$U = 8$	$U = 9$	$U = 10$	$U = 11$	$U = 12$	$U = 13$	$U = 14$	$U = 15$	
ql1	0.9477	0.9516	0.9527	0.9553	0.9566	0.9569	0.9581	0.9597	0.9889
sm1	0.9054	0.9130	0.9151	0.9203	0.9230	0.9235	0.9260	0.9293	0.9595
zl1	0.9674	0.9717	0.9728	0.9755	0.9769	0.9771	0.9787	0.9797	0.9672
ql2	0.9452	0.9485	0.9494	0.9515	0.9525	0.9526	0.9534	0.9544	0.9751
sm2	0.9590	0.9640	0.9654	0.9687	0.9704	0.9707	0.9721	0.9741	0.9694
zl2	0.9834	0.9848	0.9852	0.9859	0.9862	0.9862	0.9864	0.9866	0.9904
	$U = 20$	$U = 30$	$U = 40$	$U = 50$	$U = 60$	$U = 70$	$U = 80$	$U = 90$	
ql1	0.9622	0.9649	0.9662	0.9669	0.9671	0.9673	0.9674	0.9674	0.9889
sm1	0.9344	0.9406	0.9436	0.9455	0.9463	0.9470	0.9476	0.9479	0.9595
zl1	0.9818	0.9836	0.9839	0.9837	0.9835	0.9831	0.9824	0.9820	0.9672
ql2	0.9556	0.9561	0.9556	0.9547	0.9540	0.9533	0.9522	0.9514	0.9751
sm2	0.9769	0.9797	0.9806	0.9809	0.9809	0.9808	0.9804	0.9801	0.9694
zl2	0.9863	0.9852	0.9838	0.9824	0.9815	0.9805	0.9792	0.9783	0.9904
	$U = 100$	$U = 200$	$U = 300$	$U = 400$	$U = 500$	$U = 600$	$U = 700$	$U = 800$	
ql1	0.9674	0.9667	0.9660	0.9653	0.9647	0.9643	0.9640	0.9637	0.9889
sm1	0.9480	0.9481	0.9476	0.9468	0.9462	0.9457	0.9453	0.9449	0.9595
zl1	0.9816	0.9784	0.9763	0.9742	0.9729	0.9719	0.9711	0.9704	0.9672
ql2	0.9509	0.9464	0.9437	0.9411	0.9394	0.9382	0.9373	0.9363	0.9751
sm2	0.9799	0.9773	0.9756	0.9738	0.9727	0.9717	0.9711	0.9704	0.9694
zl2	0.9777	0.9729	0.9702	0.9677	0.9661	0.9650	0.9641	0.9632	0.9904

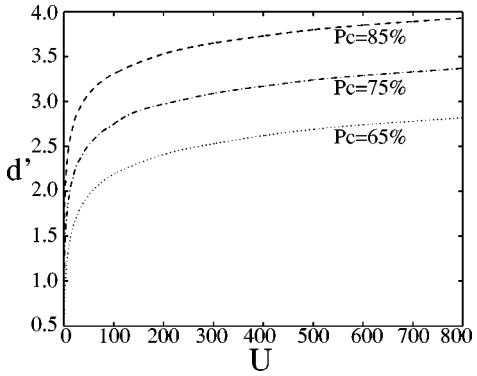

Fig. 7. d' versus number of uncertainty channel U functions at three observed performance levels (65%, 75%, and 85%) in a two-alternative forced-choice decision task.

template for each of the alternative stimuli without any additional uncertainty loss accounted for 94% of the variance in human performance. In a recent perceptual learning experiment studying Gabor orientation identification in peripheral vision, Dosher and Lu[70] demonstrated that the threshold ratio between two d' levels at all the external noise levels for each observer is constant across days, even though the thresholds themselves were improved by a factor of almost 3. This indicates that any hypothetical uncertainty effects, counter to expectations, were unchanged over substantial improvements in performance. In contrast, it is reasonable to assume that nonlinear transducer functions (as in the PTM) may be unaffected by practice;

4. To account for their data, Eckstein *et al.* had to vary the degree of uncertainty for different external noise levels for the same observer in the same experiment in nonsystematic ways. Furthermore, the EAW model fits are essentially equivalent for a wide range of values of uncertainty U; the EAW model is not sharply tuned on U, a consequence of the flatness of the percent correct versus U function (see Fig. 7). The PTM, by contrast, uses a single nonlinearity $\|\cdot\|^{\gamma}$ to account for the data across all levels of external noise;

5. In accounting for the data in the present paper, the best-fitting EAW models (Appendix A) require that U be fairly large, in the range of 20–200 (Table 3). Current models of early visual systems specify fewer visual channels.[71–79]

7. CONCLUSIONS

In this paper we examined the noisy linear amplifier model (LAM), which is widely used to interpret data from equivalent internal noise experiments, by testing its prediction that the ratio between two thresholds at each given external noise contrast should be equal to the ratio of the corresponding d' values for all the observers and noise levels. We demonstrated in two experiments that this prediction was not confirmed. Direct fits of the LAM were also relatively poor, and the observed psychometric functions were incompatible with the required cumulative Gaussian form. The perceptual template model[29–31] (PTM) is an elaboration of the LAM, with two additional components: nonlinear transducer function, and multiplicative noise. The PTM provides a good account of the data. It accommodates the observed threshold ratios while still predicting the equivalence of ratio over external noise levels; it fit the three threshold data and yielded a good direct fit to the full psychometric functions. The PTM is a feasible alternative to the recently developed EAW model. Earlier studies based on the simple noisy LAM may require reinterpretation. Moreover, the PTM and the new estimation techniques developed in this paper provide an easy method of reanalysis that could be applied to all the existing data where three or more performance levels were measured.

APPENDIX A: COMPARING THE PERCEPTUAL TEMPLATE MODEL WITH THE ECKSTEIN–AHUMADA–WATSON MODEL

Following the request of an anonymous referee, here we briefly describe the EAW model and the procedure and the results of fitting the EAW model to our data.

1. Eckstein–Ahumada–Watson Model

The EAW model consists of two equations. The first expresses d' as a function of the rms signal contrast c, the standard deviation of the external noise N_{ext}, and the coefficient for the multiplicative noise component N_{mul}:[68]

$$d' = \frac{c}{(N_{\text{add}}^2 + N_{\text{ext}}^2 + N_{\text{mul}}^2 N_{\text{ext}}^2)^{1/2}}. \tag{A1}$$

Thus, for a given d',

$$c_\tau = d'(N_{\text{add}}^2 + N_{\text{ext}}^2 + N_{\text{mul}}^2 N_{\text{ext}}^2)^{1/2}. \tag{A1a}$$

In log form,

$$\log(c_\tau) = \log(d') + 1/2 \log(N_{\text{add}}^2 + N_{\text{ext}}^2 + N_{\text{mul}}^2 N_{\text{ext}}^2). \tag{A1b}$$

The second equation of the EAW model converts the internal d' measure to percent-correct performance. To detect a signal at one of M locations, intervals, or alternatives, each with U additional hidden channels, we can express observed percent correct (Pc) as a function of M, U, and d':

$$\text{Pc}(M, U, d') = \int_{-\infty}^{\infty} [\, g(x - d')G(x)^{M(1+U)-1} + Ug(x)G(x)^{M(1+U)-2}G(x - d')]\,dx. \tag{A2}$$

2. Procedure

To make a fair comparison of the PTM and the EAW model, we fitted the EAW model to the contrast threshold versus external noise function at three performance levels (65%, 75%, and 85% correct) with one set of parameters for each observer in each experiment. Because the conversion between Pc and d' is a function of U, which is free to vary, we fitted the EAW model with a range of U values.

For a given U, the computational procedure consists of two steps:

1. Compute the corresponding d' value for each of the performance levels. In our particular experiments, $M = 2$. For a fixed Pc and U_0, e.g., $\mathrm{Pc}(2, U_0, d') = 75\%$, we can numerically solve Eq. (A2) for d'. Calculations for a range of U values yield a table that lists d' as a function of U for a given percent-correct performance level. Figure 7 plots d' as a function of U for three fixed levels of Pc values (65%, 75%, 85%). Thus, for each observer in each experiment, we can express the measured threshold contrast as a function of external noise and d' levels;

2. Fit Eq. (A1b) to find the best model parameters N_{add}, N_{mul} for each U, using a least-squares criterion; r^2 is computed with Eq. (18).

3. Results

The purpose of this computation is to compare the performance of the PTM and the EAW model. Table 3 lists the r^2 value of the best-fitting EAW model for each data set for a range of U values. We also list the r^2 value for the corresponding PTM fits.

Several conclusions can be drawn from this computational exercise. First, the PTM and the EAW model account for the data almost equally well. For the explicitly considered U values, the PTM is better in four out of six data sets. Second, the best-fitting EAW model tends to have very large U values, in the range of 20–200. Not only are these U values outside the range considered by Eckstein *et al.* (maximum, $U = 3$), they also seem to be too large to be physiologically meaningful. Third, the EAW model does not sharply constrain the estimate of U in all but the smallest U values. This is a consequence of the flatness of the Pc versus U function for U values above 10.

For a subset of U values, we also performed fits of the EAW model directly to the full psychometric functions. These fits yielded results consistent with those reported for fits to three threshold levels. Additionally, the comparisons of the quality of the EAW model fits and of the PTM fits directly to the psychometric functions preserved the same patterns as those shown in Table 3.

ACKNOWLEDGMENTS

This research is supported by the U.S. Air Force Office of Scientific Research, Life Sciences, Visual Information Processing Program. We thank an anonymous reviewer for his valuable comments and suggestions.

Address correspondence to Z.-L. Lu, Department of Psychology, SGM 501, University of Southern California, Los Angeles, California 90089-1061 or to B. Dosher, Department of Cognitive Sciences, 3151 SSPA, University of California, Irvine, California 92697. The authors can be reached by e-mail at zhonglin@rcf.usc.edu or bdosher @uci.edu, respectively.

REFERENCES AND NOTES

1. H. B. Barlow, "Retinal noise and absolute threshold," J. Opt. Soc. Am. **46**, 634–639 (1956).
2. N. S. Nagaraja, "Effect of luminance noise on contrast thresholds," J. Opt. Soc. Am. **54**, 950–955 (1964).
3. D. G. Pelli, "Effects of visual noise," Ph.D. dissertation (University of Cambridge, Cambridge, UK, 1981).
4. A. E. Burgess, R. F. Wagner, R. J. Jennings, and H. B. Barlow, "Efficiency of human visual signal discrimination," Science **214**, 93–94 (1981).
5. A. J. Ahumada and A. B. Watson, "Equivalent-noise model for contrast detection and discrimination," J. Opt. Soc. Am. A **2**, 1133–1139 (1985).
6. D. G. Pelli, "The quantum efficiency of vision," in *Vision: Coding and Efficiency*, C. Blakemore, ed. (Cambridge U. Press, Cambridge, UK, 1990), pp. 3–24.
7. D. H. Parish and G. Sperling, "Object spatial frequencies, retinal spatial frequencies, noise, and the efficiency of letter discrimination," Vision Res. **31**, 1399–1415 (1991).
8. G. Sperling, "Three stages and two systems of visual processing," Spatial Vision **4**, 183–207 (1989).
9. D. O. North, "The absolute sensitivity of radio receivers," RCA Rev. **6**, 332–344 (1942).
10. H. T. Friis, "Noise figures of radio receivers," Proc. IRE **32**, 419–422 (1944).
11. W. W. Mumford and E. H. Schelbe, *Noise Performance Factors in Communication Systems* (Horizon House-Microwave, Dedham, Mass., 1968).
12. H. Fletcher, "Auditory patterns," Rev. Mod. Phys. **12**, 47–65 (1940).
13. H. B. Barlow, "Incremental thresholds at low intensities considered as signal/noise discrimination," J. Physiol. (London) **136**, 469–488 (1957).
14. J. A. Swets, D. M. Green, and W. P. Tanner, Jr., "On the width of critical bands," J. Acoust. Soc. Am. **34**, 108–113 (1962).
15. U. Greis and R. Rohler, "Untersuchung der subjektiven Detailerkennbarkeit mit Hilfe der Ortsfrequenzfilterung," Opt. Acta **17**, 515–526 (1970).
16. H. Pollehn and H. Roehrig, "Effect of noise on the MTF of the visual channel," J. Opt. Soc. Am. **60**, 842–848 (1970).
17. B. E. Carter and G. B. Henning, "The detection of gratings in narrow-band visual noise," J. Physiol. (London) **219**, 355–365 (1971).
18. C. F. Stromeyer and B. Julesz, "Spatial frequency masking in vision: critical bands and spread of masking," J. Opt. Soc. Am. **64**, 1221–1232 (1972).
19. L. D. Harmon and B. Julesz, "Masking in visual recognition: effects of two-dimensional filtered noise," Science **180**, 1194–1197 (1973).
20. G. B. Henning, B. G. Hertz, and J. L. Hinton, "Effects of different hypothetical detection mechanisms on the shape of spatial-frequency filters inferred from masking experiments. I. Noise masks," J. Opt. Soc. Am. **71**, 574–581 (1981).
21. M. Pavel, G. Sperling, T. Riedl, and A. Vanderbeek, "Limits of visual communication: The effect of signal-to-noise ratio on the intelligibility of American Sign Language," J. Opt. Soc. Am. A **4**, 2355–2365 (1987).
22. T. R. Riedl and G. Sperling, "Spatial-frequency bands in complex visual stimuli: American Sign Language," J. Opt. Soc. Am. A **5**, 606–616 (1988).
23. T. E. Reisbeck and K. R. Gegenfurtner, "Velocity tuned mechanisms in human motion perception," Invest. Ophthalmol. Visual Sci. Suppl. **38**, 376 (1997).
24. A. Gorea and T. V. Papathomas, "Luminance, color and orientation: local vs. global contrasts in texture segregation," Invest. Ophthalmol. Visual Sci. Suppl. **39**, 649 (1998).
25. M. D'Zmura and K. Knoblauch, "Spectral bandwidths for the detection of color," Vision Res. **38**, 3117–3128 (1998).
26. D. G. Pelli, "Uncertainty explains many aspects of visual contrast detection and discrimination," J. Opt. Soc. Am. A **2**, 1508–1532 (1985).
27. A. E. Burgess and B. Colborne, "Visual signal detection. IV. Observer inconsistency," J. Opt. Soc. Am. A **5**, 617–627 (1988).
28. A. E. Burgess, "Visual signal detection. III. On Bayesian use of prior knowledge and cross correlation," J. Opt. Soc. Am. A **2**, 1498–1507 (1985).
29. B. A. Dosher and Z.-L. Lu, "Attention to location mediated by internal noise reduction," Invest. Ophthalmol. Visual

Sci. Association for Research in Vision and Ophthalmology Suppl. **38**, 687 (1997).
30. Z.-L. Lu and B. A. Dosher, "Characterizing attention mechanisms," Bull. Psychonom. Soc. **38**, 58 (1997).
31. Z.-L. Lu and B. A. Dosher, "External noise distinguishes mechanisms of attention," Vision Res. **38**, 1183–1198 (1998).
32. J. Nachmias and R. V. Sansbury, "Grating contrast: discrimination may be better than detection," Vision Res. **14**, 1039–1042 (1974).
33. J. M. Foley and G. E. Legge, "Contrast detection and near-threshold discrimination in human vision," Vision Res. **21**, 1041–1053 (1981).
34. C. F. Stromeyer and S. Klein, "Spatial frequency channels in human vision as asymmetric (edge) mechanisms," Vision Res. **14**, 1409–1420 (1974).
35. G. E. Legge and J. M. Foley, "Contrast masking in human vision," J. Opt. Soc. Am. **70**, 1458–1471 (1980).
36. C. A. Burbeck and D. H. Kelly, "Contrast gain measurements and the transient/sustained dichotomy," J. Opt. Soc. Am. **71**, 1335–1342 (1981).
37. A. F. Dean, "The relationship between response amplitude and contrast for cat striate cortical neurons," J. Physiol. (London) **318**, 413–427 (1981).
38. A. M. Derrington and P. Lennie, "Spatial and temporal contrast sensitivities of neurons in lateral geniculate nucleus of macaque," J. Physiol. (London) **357**, 219–240 (1981).
39. E. Kaplan and R. M. Shapley, "X and Y cells in the lateral geniculate nucleus of macaque monkeys," J. Physiol. (London) **330**, 125–143 (1982).
40. D. G. Albrecht and D. B. Hamilton, "Striate cortex of monkey and cat: contrast response function," J. Neurophysiol. **48**, 217–237 (1982).
41. I. Ohzawa, G. Sclar, and R. D. Freeman, "Contrast gain control in the cat visual cortex," Nature (London) **298**, 266–268 (1982).
42. S. A. Klein and D. M. Levi, "Hyperacuity thresholds of 1 sec: theoretical predictions and empirical validation," J. Opt. Soc. Am. A **2**, 1170–1190 (1985).
43. G. Sclar, J. H. Maunsell, and P. Lennie, "Coding of image contrast in central visual pathways of the macaque monkey," Vision Res. **30**, 1–10 (1990).
44. A. B. Bonds, "Temporal dynamics of contrast gain in single cells of the cat striate cortex," Visual Neurosci. **6**, 239–255 (1991).
45. D. G. Albrecht and W. S. Geisler, "Motion selectivity and the contrast-response function of simple cells in the visual cortex," Visual Neurosci. **7**, 531–546 (1991).
46. D. J. Heeger, "Normalization of cell responses in cat striate cortex," Visual Neurosci. **9**, 181–197 (1992).
47. J. M. Foley, "Human luminance pattern-vision mechanisms: masking experiments require a new model," J. Opt. Soc. Am. A **11**, 1710–1719 (1994).
48. Z.-L. Lu and G. Sperling, "Contrast gain control in first- and second-order motion perception," J. Opt. Soc. Am. A **13**, 2305–2318 (1996).
49. A. B. Watson and J. A. Solomon, "Model of visual contrast gain control and pattern masking," J. Opt. Soc. Am. A **14**, 2379–2391 (1997).
50. M. P. Eckstein, A. J. Ahumada, and A. B. Watson, "Visual signal detection in structured backgrounds. II. Effects of contrast gain control, background variations, and white noise," J. Opt. Soc. Am. A **14**, 2406–2419 (1997).
51. R. S. Woodworth and H. Schlosberg, *Experimental Psychology*, 2nd ed. (Holt, Rinehart & Winston, New York, 1954).
52. D. G. Pelli and L. Zhang, "Accurate control of contrast on microcomputer displays," Vision Res. **31**, 1337–1350 (1991).
53. R. E. Fredericksen's software, "Image manipulation and display library" (1996), was used.
54. In computing d', both the LAM [Eq. (3)] and the PTM [Eq. (15)] assume that the noise at the decision stage has a Gaussian distribution. This implies a simple functional relationship between percent correct and d':Pc $= G(d'/2)$, where $G(\,)$ is cumulative Gaussian. For the LAM, this specifies the cumulative Gaussian with $\mu = 0$ (signal contrast) as the form of psychometric functions, while the PTM specifies the form of psychometric function that is cumulative Gaussian with complex functions of c^γ as its argument owing to the nonlinearity and multiplicative noise. This provides an alternative way to test the models (see Subsection 5.B). In fact, the cumulative Gaussian specified by the LAM fit the psychometric functions very poorly (see Subsection 5.B). Weibull functions were chosen to fit the psychometric functions for the purpose of estimating (interpolating) three threshold contrast levels. The mean r^2 of the Weibull fits to the psychometric data was 0.8818. (A number of other interpolation functions might also have worked. For example, a cumulative Gaussian with guessing correction provided threshold estimates that were essentially identical to the Weibull, $r = 0.9996$, rms difference $= 0.0019$ for a mean threshold of 0.094.) We estimated the bias introduced in the estimation of ratios k_1 and k_2 when a Weibull was fitted to hypothetical psychometric functions specified by the LAM and when the Weibull was fitted to hypothetical psychometric functions specified by the PTM. In both cases, the biases were of the order of 5% or less, too small to significantly influence the ratio estimates.
55. W. L. Hays, *Statistics*, 3rd ed. (CBS College Publishing, New York, 1981).
56. The standard deviation of each threshold was computed with a resampling method.[57] For a given external noise level, the number of correct responses at each of the seven signal contrasts was assumed to have a binomial distribution with a single event probability p equal to the measured percent correct. For each of the seven signal stimulus contrasts in the original psychometric functions at a given external noise level, the number of correct responses at each signal stimulus contrast was independently replaced by a theoretical sample from the corresponding binomial distribution to generate a theoretically resampled psychometric function. Repeating this process 25 times for each external noise condition generated 25 theoretically resampled psychometric functions in every external noise condition. Fitting Weibull functions to each of these psychometric functions and then computing thresholds at 65%, 75%, and 85% performance levels allowed us to estimate the mean and the standard deviations of the parameters in the Weibull function and for the three thresholds.
57. L. T. Maloney, "Confidence intervals for the parameters of psychometric functions," Percept. Psychophys. **47**, 127–134 (1990).
58. Both sets of ratios are closely comparable with those measured for the same observers in the identification task of experiment 1.
59. B. Zenger and D. Sagi, "Isolating excitatory and inhibitory nonlinear spatial interactions involved in contrast detection," Vision Res. **36**, 2497–2513 (1996).
60. R. E. Fredericksen and R. F. Hess, "Estimating multiple temporal mechanisms in human vision," Vision Res. **38**, 1023–1040 (1998).
61. In general, the contrast of a signal can be expressed as a function of space and time: $S(x, y, t) = cS_0(x, y, t)$ rescaled such that $\int\int\int {S_0}^2(x, y, t)\mathrm{d}x\mathrm{d}y\mathrm{d}t = 1.0$. The contrast of the external noise can be expressed as $N(x, y, t) = N_{\text{ext}}G(x, y, t)$, where the value of $G(x, y, t)$ at a particular spatiotemporal point (x, y, t) is drawn from a Gaussian distribution with mean 0 and standard deviation 1.0. For a template matching function $T(x, y, t)$, matching the template to a signal-valued stimulus yields the output $T_S = \int\int\int T(x, y, t)S(x, y, t)\mathrm{d}x\mathrm{d}y\mathrm{d}t = c\int\int\int T(x, y, t)S_0(x, y, t)\mathrm{d}x\mathrm{d}y\mathrm{d}t$; The output of the template matching operation to the external noise is $T_N = \int\int T(x, y, t)N(x, y, t)\mathrm{d}x\mathrm{d}y\mathrm{d}\,t = N_{\text{ext}}\int\int\int T\,(x, y, t)G \times (x, y, t)\mathrm{d}x\mathrm{d}y\mathrm{d}t$. For a fixed template and a fixed signal stimulus, $T_{S_0} = \int\int\int T(x, y, t)S_0(x, y, t)\mathrm{d}x\mathrm{d}y\mathrm{d}t$ is a constant; $T_G = \int\int\int T(x, y, t)G(x, y, t)\mathrm{d}x\mathrm{d}y\mathrm{d}t$ is a Gaussian random variable with mean 0 and a fixed standard deviation σ_{T_G}. Because, mathematically, T_S and T_N can be

known only up to a constant, without losing any generality, we set σ_{T_G} to 1.0. Thus, after template matching, $T_N = N_{\text{ext}}G(0,1)$, $T_S = \beta c$, where $\beta = T_{S_0}/\sigma_{T_G}$.

62. C. R. Carlson and R. W. Klopfenstein, "Spatial-frequency model for hyperacuity," J. Opt. Soc. Am. A **2**, 1747–1751 (1985).
63. N. A. MacMillan and C. D. Creelman, *Detection Theory: A User's Guide* (Cambridge U. Press, New York, 1991).
64. In experiment 2, where only one temporal interval contains the signal, the standard deviation of the noise at the decision stage is not identical for the two intervals. This is not an issue in the identification task of experiment 1. To simplify our discussion, we made the approximation of using the standard deviation of the signal-present interval in both situations. In practice, the inaccuracy caused by this approximation is small because noise is dominated by N_{add} in the region of low external noise and is dominated by N_{ext} in the region of high external noise.
65. MATLAB integrated technical computing environment (The MathWorks, Inc., 24 Prime Park Way, Natick, Mass. 01760-1500).
66. The PTM is a highly nonlinear model with complex dependence on its parameters. The set of analytical constraints on the model parameters developed here allows us to solve for two important parameters independent of the rest in the model. This approach is most advantageous when the solution space of the model parameters is complicated, which is certainly true in the case of the PTM.
67. The log approximately equates the standard error over large ranges in contrast thresholds, corresponding to weighted least squares, an equivalent to the maximum-likelihood solution for continuous data.
68. The two models do not yield identical functions, and, furthermore, they partition noise, especially stimulus noise, in a different manner. The EAW model distinguishes between contrast gain control noise associated with structured noise backgrounds in their displays, σ_{cgc}, and that associated with additional white-noise variation, σ_{var}. The EAW model assumes the isolation of signal from background noise in the stimulus since contrast-modulated noise depends only on noise, not signal contrast. The PTM considers the power in the stimulus, both signal and external noise, as controlling multiplicative noise.
69. D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (Krieger, Huntington, New York, 1974).
70. B. Dosher and Z.-L. Lu, "Mechanisms of perceptual learning," submitted to Vision Res.
71. F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," J. Physiol. (London) **197**, 551–566 (1968).
72. A. Pantle and R. Sekuler, "Size detecting mechanisms in human vision," Science **162**, 1146–1148 (1969).
73. C. Blakemore and F. W. Campbell, "On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images," J. Physiol. (London) **203**, 237–260 (1969).
74. J. P. Thomas, "Model of the function of receptive fields in human vision," Psychol. Rev. **77**, 121–134 (1970).
75. D. H. Hubel and T. N. Wiesel, "Uniformity of monkey striate cortex: a parallel relationship between field size, scatter and magnification factor," J. Comp. Neurol. **158**, 295–306 (1974).
76. S. P. Mckee and G. Westheimer, "Improvement in vernier acuity with practice," Percept. Psychophys. **24**, 258–262 (1978).
77. H. R. Wilson and J. R. Bergen, "A four mechanism model for threshold spatial vision," Vision Res. **19**, 19–32 (1979).
78. N. Graham, "Spatial frequency channels in human vision: detecting edges without edge detectors," in *Visual Coding and Adaptability*, C. S. Harris, ed. (Erlbaum, Hillsdale, N.J., 1980), pp. 215–262.
79. M. A. Webster and R. L. de Valois, "Relationship between spatial-frequency and orientation tuning of striate-cortex cells," J. Opt. Soc. Am. A **2**, 1124–1132 (1985).