Psychological Review
2008, Vol. 115, No. 1, 44–82

0033-295X/08/$12.00 DOI: 10.1037/0033-295X.115.1.44

Characterizing Observers Using External Noise and Observer Models: Assessing Internal Representations With External Noise

Zhong-Lin Lu
University of Southern California

Barbara Anne Dosher
University of California, Irvine

External noise methods and observer models have been widely used to characterize the intrinsic perceptual limitations of human observers and changes of the perceptual limitations associated with cognitive, developmental, and disease processes by highlighting the variance of internal representations. The authors conducted a comprehensive review of the 5 most prominent observer models through the development of a common formalism. They derived new predictions of the models for a common set of behavioral tests that were compared with the data in the literature and a new experiment. The comparison between the model predictions and the empirical data resulted in very strong constraints on the observer models. The perceptual template model provided the best account of all the empirical data in the visual domain. The choice of the observer model has significant implications for the interpretation of data from other external noise paradigms, as well as studies using external noise to assay changes of perceptual limitations associated with observer states. The empirical and theoretical development suggests possible parallel developments in other sensory modalities and studies of high-level cognitive processes.

Keywords: signal detection theory, internal representation, observer model, perceptual template, internal noise

Human decisions are based on internal representations of information. Understanding how stimuli are represented internally is one of the classic problems in psychology. This article examines how modifying an external stimulus with external noise can provide insight into how the stimulus is processed by the human observer. We conducted a systematic and comprehensive review of the external noise paradigms and observer models widely used in characterizing the internal response properties of human observers. The observer approach builds on the broadly applicable framework of signal detection theory (SDT) by elaborating the relationships between external stimuli and the internal response distributions that form the basis for decision. The empirical tests introduce external noise—either masking noise or variation in the relevant stimulus dimension—to provide a reference for characterizing and quantifying the limiting factors in perceptual sensitivity. The current review, analysis, and empirical test focus on visual perception. Some of the model properties, especially the empirical findings, may be modality specific. Still, this framework and the findings could serve as an example for parallel development of the empirical methods and theoretical models in other sensory modalities. The model development and testing also have major implications for applications of the external noise paradigms in understanding the mechanisms underlying changes of perceptual sensitivity in different cognitive, disease, and/or developmental states.

Zhong-Lin Lu, Laboratory of Brain Processes (LOBES), Dana and David Dornsife Cognitive Neuroscience Imaging Center, Departments of Psychology and Biomedical Engineering, and Neuroscience Graduate Program, University of Southern California; Barbara Anne Dosher, Memory, Attention, & Perception (MAP) Laboratory, Department of Cognitive Sciences and Institute of Mathematical Behavioral Sciences, University of California, Irvine.

This research was supported by the Air Force Office of Scientific Research, the National Institute of Mental Health, the National Science Foundation, and the National Eye Institute.

Correspondence concerning this article should be addressed to Zhong-Lin Lu, Department of Psychology, SGM 501, University of Southern California, Los Angeles, CA 90089-1061. Email: zhonglin@usc.edu

Internal Response Distributions

SDT provides a general framework for analyzing human decision making in perceptual and cognitive tasks (Green & Swets, 1966; Macmillan & Creelman, 1991). In a simple yes–no task, an observer is presented with a single input stimulus, which either contains or does not contain a signal, and must decide whether the signal is present ("yes") or absent ("no"). According to the SDT, signal-present and signal-absent trials generate internal perceptual representations characterized by two different internal response distributions (see Figure 1A). The observer makes the decision on the basis of a subjective criterion: If the internal response is greater than the criterion, the observer reports that the signal is present; otherwise, the observer reports that the signal is absent. The internal response distributions and the criterion jointly determine the probabilities of all four possible outcomes: hit, false alarm, miss, and correct rejection. Critically, d', the sensitivity of the observer, which is independent of the subjective criterion, can be obtained by measuring the receiver operating characteristics (ROCs; see Figure 1B)—the hit-versus-false-alarm-rate function (see Appendix A). In another paradigm, the two-alternative forced-choice (2AFC) task, an observer is presented with two input stimuli, one from each of two stimulus categories, and must sort the stimuli into the two categories. The SDT postulates that stimuli from the two categories generate two different internal response distributions (see Figure 1C). In each trial, the observer compares the magnitudes of the two internal responses and decides that the

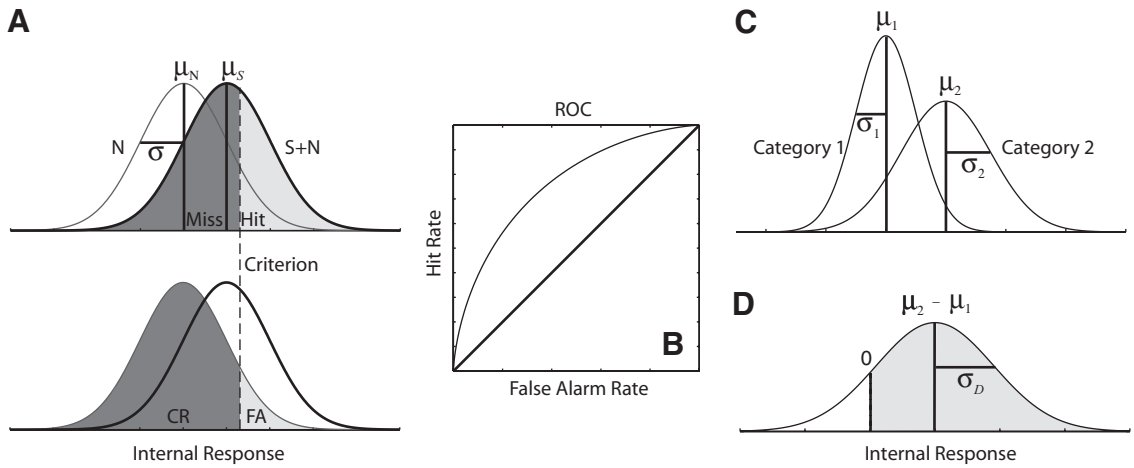


Figure 1. Two example applications of the signal detection theory (SDT). A: SDT in a yes–no task. The two bell curves in each panel represent internal response distributions for signal-absent (N) and signal-present (S + N) trials. To decide whether the signal is present ("yes") or not ("no"), the observer first chooses a subjective criterion response. If the (single) input stimulus generates an internal response greater than the criterion response, the observer decides that the signal is present; otherwise, she or he decides that the signal is absent. The internal response distributions and the criterion jointly determine the probabilities of hit, false alarm (FA), miss, and correct rejection (CR), denoted by the four shaded areas. B: Receiver operating characteristics (ROCs)—hit rate as a function of the false-alarm rate as the observer varies his or her criterion response. C: SDT in a two-alternative forced-choice (2AFC) task. A 2AFC task presents two stimuli to an observer in each trial, one from each of two categories, and forces the observer to decide the correspondence between the stimuli and the categories. The two bell curves present the internal response distributions for the two stimulus categories with different means and standard deviations. For an unbiased observer, the probability of making the correct choice is equal to the area of the shaded region in the difference distribution (Panel D)—the distribution of a random variable that is the difference between the two internal responses.

stimulus that generates the greater internal response belongs to the category with a higher expected internal response. For an unbiased observer, the probability of making a correct choice can be calculated from the distribution of the differences between the internal responses to the two stimulus categories (see Figure 1D); one can derive the sensitivity of the observer directly from measurements of performance accuracy (see Appendix A).

The analysis of the yes–no and 2AFC tasks represents two rudimentary applications of the SDT framework. The most important concept in both applications is the noisy internal response distribution. Over the past 5 decades, the formal development of the SDT framework has included multidimensional internal response distributions (Ashby, 1992); general classification paradigms in which observers use *M* responses to sort *N* stimuli into categories, of which the yes–no and 2AFC tasks are two special cases (Green & Swets, 1966; Macmillan & Creelman, 1991); comparison paradigms (e.g., same or different) and compound tasks (G. Sperling & Dosher, 1986); and decision under uncertainty (Graham, 1989). Empirically, the SDT framework has been applied extensively not only in studying all the sensory modalities but also in studying high-level cognitive tasks such as object recognition, memory, and language processing, as well as diagnosis and assessment (Swets, 1996). Theoretically, many models of human behavior are motivated and based on the SDT framework (see Logan, 2004, for a review of an alternative approach). Over decades of development, noisy internal response distributions have remained the central concept in the framework.

The enormous success of the SDT framework has highlighted the critical role of internal response distributions in understanding human behavior. Yet the SDT, as a theoretical framework, does not specify the internal representations. Additional assumptions about the functional relationship between the internal representations and the physical characteristics of external stimuli are required to make specific predictions. Gaussian internal response distributions, although not necessary for the framework, prove to suffice in most SDT applications (Wilkens, 2002). The mean and the variance fully specify a Gaussian distribution. The success of the SDT framework illustrates the power of studying not just the mean but also the variance of representations underlying human performance. In most SDT applications, however, the variances of the internal response distributions are theoretical constructs that are not referenced to the physical characteristics of external stimuli (Wilkens, 2002).

Characterizing Internal Responses

The critical role of internal response distributions for understanding human performance has prompted major research efforts in both psychology and neurophysiology to independently specify and model the internal responses (Barlow, 1956; Burgess, Wagner, Jennings, & Barlow, 1981; Legge, Kersten, & Burgess, 1987; Nagaraja, 1964; Pelli, 1981). One approach is to attempt to identify, measure, and interpret the brain responses to relevant stimuli. An alternative approach is to construct observer models that spec-

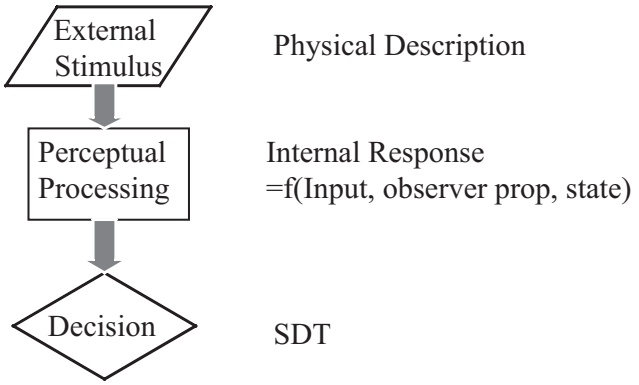


Figure 2. A complete observer model must include a perceptual module that specifies the functional relationship between the internal responses and the external stimuli (inputs) and a decision module that maps internal responses to perceptual decisions (e.g., the signal detection theory [SDT]). prop = property.

ify the functional relationship between external stimuli and internal responses, as well as the decision process (e.g., SDT) in human behavior (see Figure 2). On the basis of the internal processes and intrinsic limitations of the observer, observer models provide the theoretical basis for generalizing the results of a particular experiment to predict the performance of the same observer in other tasks.

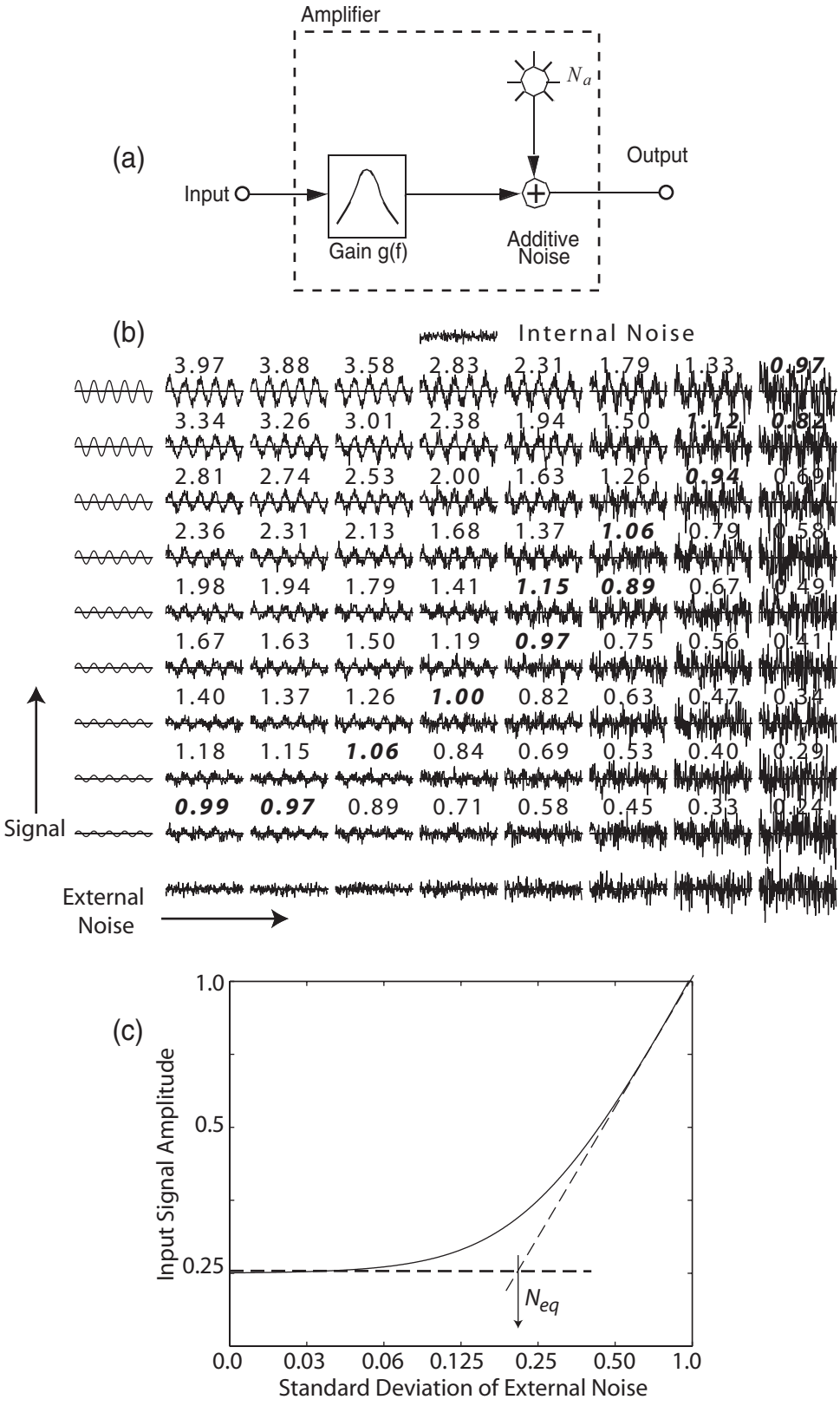
Neurophysiology and functional imaging studies can in principle provide measures of the internal responses of the perceptual and cognitive systems at various stages of processing. The brain response to a certain stimulus is measured, but it must be further determined which aspect and brain location of the measured responses are relevant to the behavioral choice. Behavioral approaches, including many psychological paradigms, have also been developed to reveal the internal response distributions at the overall system level. All these paradigms involve adding external noise to the signal stimuli to externalize the internal responses. These include various procedures related to critical band masking (Fletcher, 1940), the equivalent input noise method (Barlow, 1956; Burgess et al., 1981; Legge et al., 1987; Nagaraja, 1964; Pelli, 1981), the double-pass consistency test (Burgess & Colborne, 1988; Green, 1964), and the classification image method (Ahumada & Lovell, 1971). Direct measurement of physiological responses and quantitative modeling of the observer properties expressed in behavioral responses are converging approaches to understanding internal representations. As investigation of the physiological responses advances, the identified responses should have properties that are consistent with those identified through the observer model characterizations.

The basic paradigms and models of the observer approach to understanding system limitations have been developed and applied mostly in the study of sensory and perceptual systems. Although our focus in this article is on visual perception, we hope a systematic review of the various external noise methods and the observer models will illustrate some important theoretical considerations that may stimulate new theoretical developments and empirical tests in other domains of research.

External Noise Methods

The basic principle of the external noise paradigms can be best exemplified by the equivalent input noise method, originally developed by engineers to measure the response properties of electronic amplifiers (Friis, 1944; Mumford & Schelbe, 1968; North, 1942). Two response properties are important for electronic amplifiers: large amplification and low intrinsic noise (see Figure 3a). Amplifiers with high intrinsic noise are undesirable because their outputs are noisy. The equivalent input noise method used by engineers to estimate the intrinsic noise of amplifiers is demonstrated in Figure 3b: Mixtures of signal and external noise (both generated and known by the engineer) of various combinations of amplitudes are passed through the amplifier. Outputs of the amplifier are analyzed to extract the signal-to-noise ratio (the average amplitude of the output over its standard deviation) in each signal and external noise condition. There are two sources of variability in the output of the amplifier, the known external noise and the unknown intrinsic noise. When the external noise is much less than the intrinsic noise, the variability in the output of the amplifier and therefore the signal-to-noise ratio for a given signal condition are mostly determined by intrinsic noise. A relatively constant amount of signal is required to maintain a constant signal-to-noise ratio across the external noise conditions. When the external noise is much greater than the intrinsic noise, the variability in the output of the amplifier and therefore the signal-to-noise ratio for a given signal condition are mostly determined by external noise. Increasing amounts of signal are required to maintain a constant signal-to-noise ratio as external noise increases. At the transition point of these two regimes, the elbow of the constant signal-to-noise ratio contour in Figure 3c, the intrinsic and external noises are equally damaging. Therefore, the intrinsic noise is equivalent to the input external noise at the elbow of the contour.

Figure 3 (opposite). a: Model diagram of a linear electronic amplifier. b: The equivalent input noise procedure. An internal noise sample is shown on the top. Signal sine waves with increasing amplitude are shown in the left column; waveforms of external noise with increasing standard deviation are shown in the bottom row. The waveforms in the rest of the panel are constructed by summing the signal in the corresponding row, an independent sample of external noise with the same standard deviation in the corresponding column, and an independent sample of internal noise. The numerical value above each waveform indicates the mean signal-to-noise ratio. Signal-to-noise ratio values close to 1.0 are highlighted with a bold italic font. When connected, they trace a contour of signal and external noise conditions with a constant signal-to-noise ratio of 1.0. c: A smoother version of the constant signal-to-noise ratio contour with finer samples of the signal amplitudes and external noise standard deviations. The abscissa of the elbow of the function provides an estimate of the intrinsic noise—the equivalent input noise N_{eq}.


Amplifier
(a)
N_a
Input
Gain g(f)
Additive Noise
Output
(b)
Internal Noise
3.97 3.88 3.58 2.83 2.31 1.79 1.33 0.97
3.34 3.26 3.01 2.38 1.94 1.50 1.12 0.82
2.81 2.74 2.53 2.00 1.63 1.26 0.94 0.69
2.36 2.31 2.13 1.68 1.37 1.06 0.79 0.58
1.98 1.94 1.79 1.41 1.15 0.89 0.67 0.49
1.67 1.63 1.50 1.19 0.97 0.75 0.56 0.41
1.40 1.37 1.26 1.00 0.82 0.63 0.47 0.34
1.18 1.15 1.06 0.84 0.69 0.53 0.40 0.29
0.99 0.97 0.89 0.71 0.58 0.45 0.33 0.24
Signal
External Noise
(c)
Input Signal Amplitude
1.0
0.5
0.25
N_{eq}
0.0 0.03 0.06 0.125 0.25 0.50 1.0
Standard Deviation of External Noise

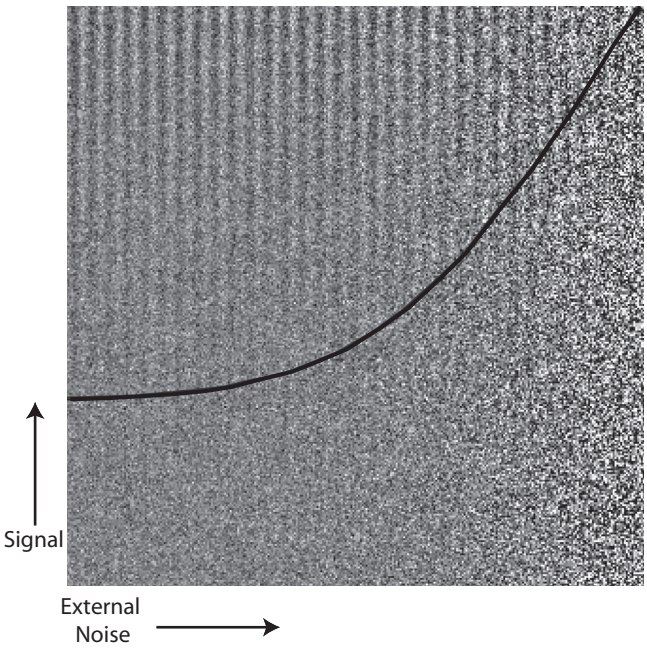


Figure 4. An illustration of the equivalent input noise method in vision. The image is made from the superposition of three images—a vertical sine wave (signal) with increasing contrast in the vertical direction, an external noise image with increasing variance in the horizontal direction, and a simulated internal noise image with a constant variance. To help the reader, we have traced out an equal visibility contour of the signal grating. The contour is flat in low-external-noise conditions and rises with external noise in high-external-noise conditions. The amplitude of the external noise at the elbow of the contour provides an estimate of the variance of the internal noise.

Nagaraja (1964) first suggested that the equivalent input noise method could be adopted to measure the internal noise of the perceptual system. The approach was popularized by Pelli and others in the 1980s (Burgess et al., 1981; Legge et al., 1987; Pelli, 1981). The basic idea is that the perceptual system of the observer functions as the noisy amplifier. We must infer the signal-to-noise ratio in response to the input stimulus from the behavioral response. The behavioral detectability of the sine-wave grating is determined by the signal-to-noise ratio at the decision stage. The method is illustrated in Figure 4 in which a vertical signal sine-wave grating with increasing contrast in the vertical direction is combined with simulated internal noise with a constant standard deviation and increasing amounts of external noise in the horizontal direction.[1] When the external noise is relatively low, the visibility of the grating is not affected by the amount of external noise. Relatively constant signal amplitude is sufficient for the grating to be seen. When the external noise is relatively high, the visibility of the grating is greatly affected by the amount of external noise. The signal amplitudes required for the grating to be visible increase with the amount of external noise. Similar to the situation with electronic amplifiers, the transition point of the two regimes reveals the magnitude of the internal noise.

The external noise paradigms have been used to analyze human sensitivity and reveal observer characteristics in a wide range of auditory (Ahumada & Lovell, 1971; Bos & Deboer, 1966; Eijkman, Thijssen, & Vendrik, 1966; Hartmann & Pumplin, 1988; Humes & Jesteadt, 1989; Moore, 1975; Osman, 1971; Richards, Heller, & Green, 1991) and visual tasks (Ahumada, 1987; Ahumada & Watson, 1985; Barlow, 1956; Burgess et al., 1981; D'Zmura & Knoblauch, 1998; Gegenfurtner & Kiper, 1992; Geisler, 1989; Hay & Chesters, 1972; Legge et al., 1987; Lu & Dosher, 1999, 2001; Nagaraja, 1964; Pelli, 1981, 1990; Rose, 1948; Tanner & Birdsall, 1958; Tjan, Braje, Legge, & Kersten, 1995; Van Meeteren & Barlow, 1981). One important discovery is that many observer characteristics are invariant across different perceptual tasks within a modality (Pelli & Farell, 1999). The paradigms have also been further developed to investigate mechanisms underlying the effects of various cognitive, developmental, and disease states on the perceptual system (Lu & Dosher, 1998). By combining the external noise paradigms with manipulations of cognitive, developmental, or disease states of the observer, we can estimate changes of the internal observer characteristics associated with performance in those states. Domains of applications of the general approach include attention (Dosher & Lu, 2000a, 2000b; Lu & Dosher, 1998, 2000; Lu, Liu, & Dosher, 2000; Talgar, Pelli, & Carrasco, 2004), perceptual learning (Chung, Levi, & Tjan, 2005; Dosher & Lu, 1998, 1999; Gold, Bennett, & Sekuler, 1999; R. W. Li, Levi, & Klein, 2003; Lu, Chu, Dosher, & Lee, 2005; Lu & Dosher, 2004), adaptation (Dao, Lu, & Dosher, 2006), amblyopia (Huang, Tao, Zhou, & Lu, 2007; Levi & Klein, 2003; Xu, Lu, Qiu, & Zhou, 2006), perceptual interaction (Yu, Levi, & Klein, 2001), dyslexia (A. Sperling, Lu, Manis, & Seidenberg, 2005), and visual memory (Gold, Murray, Sekuler, Bennett, & Sekuler, 2005). In many cases, the approach found modifications of only one or two of the observer characteristics in different cognitive, developmental, or disease states and a striking invariance of other observer characteristics across very different states and wide performance ranges. Thus, the observer model, together with estimated parameters, provides a compact characterization of the observer that can precisely predict performance in previously untested task and stimulus conditions. By identifying the modified components of the observer model associated with a change in the cognitive, developmental, or disease state, the method provides insights into the mechanisms underlying their effects on the perceptual system.

Observer Models

Several noisy observer models have been proposed to interpret the empirical results from the external noise paradigms and model the internal responses of the human observers (for parallel developments in pattern masking, see Foley & Chen, 1999). Because human perception exhibits many inefficiencies due to various sources of noise in the perceptual process, the observer models specify the noises in the perceptual system in addition to other noise-free computations in the perceptual process (G. Sperling, 1989). All the models are based on a number of component processes derived from both sensory psychology and physiology,

[1] Most visual phenomena related to detection and discrimination are relatively independent of the absolute luminance level for an extremely wide range of luminance (Hood & Finkelstein, 1986). Therefore, it is convenient to define and discuss visual stimuli in terms of their contrast: $c(x, y) = (L(x, y) - L_0)/L_0$, where $L(x, y)$ is the luminance at point (x, y) and L_0 is the mean luminance of the display.

including a perceptual template, a nonlinear transducer, additive noise, multiplicative noise, contrast-gain control, and a decision process (Burgess et al., 1981; Eckstein, Ahumada, & Watson, 1997; Lu & Dosher, 1999; Pelli, 1981, 1985). Different models include different subsets of these component processes. The goal of these models is to predict the signal-to-noise ratio of behavioral performance from the driving stimulus and estimated values of several observer-specific properties.

The most prominent observer models include the *linear amplifier model* (LAM; Pelli, 1981), the *induced noise model* (INM; Burgess & Colborne, 1988), the *linear amplifier model with decision uncertainty* (LAUM; Pelli, 1985), the *induced noise and uncertainty model* (INUM; Eckstein et al., 1997), and the *perceptual template model* (PTM; Lu & Dosher, 1999). Often, a particular observer model was developed to handle the results from a particular experimental procedure and has not been tested with data from other procedures (see below). It has become pressingly important for the field as a whole to engage in a comprehensive review and comparison of the various observer models, especially in their ability to simultaneously account for results from all the major external noise methods. This is also important for a conceptual understanding of which attributes of behavioral data constrain model processes. That is the goal of this article.

We derive the predictions of the five major observer models in relation to the three most general external noise paradigms: the equivalent input noise method, the triple-threshold-versus-external-noise-contrast (triple-TvC) method, and the double-pass procedure. Although the properties of each particular model may exist in the literature for one particular empirical test, here we translate the various previous theoretical analyses into a common formalism and derive predictions of each model for a common set of behavioral tests. The predictions make it possible to evaluate and compare these models against existing data in the literature as well as their fit to the data in a new experiment later in the article. As we show, the joint or convergent results from the three major external noise methods provide very strong constraints on the observer models while identifying key relationships in detection or discrimination data. The choice of the observer model has significant implications for the interpretations of data from other external noise paradigms, such as those involving critical band masking, as well as applications of the methods to the study of mechanisms of cognitive, developmental, and disease processes.

The remaining external noise methods, that is, the various methods related to critical band masking and classification images, have been mostly used to characterize the detailed properties of the perceptual templates such as sensitivity to spatial or temporal frequencies or to the most diagnostic spatial regions. In critical masking, the characteristics (spatial or temporal frequency) of the external noise are varied to reveal the tuning properties of the perceptual template—the stimulus evidence selected to support a specific task or decision. Masking damages performance if the mask has energy that is within the tuning of the template. The classification image method typically compares the external noise images between correct or incorrect trials to find the spatial regions or temporal segments of external noise most correlated with performance and therefore reveals spatial or temporal properties of the perceptual template (see Abbey & Eckstein, 2006, for an attempt to use the classification image method to investigate nonlinearities in the perceptual system). We focus on those aspects of the observer models that do not depend on these details of the perceptual templates. The development of the observer model may in turn be extended to refine the methods for the measurement of templates (e.g., Lu & Dosher, 2001).

Overview

We first describe the three external noise paradigms and the mathematical properties of the five observer models, focusing on the 2AFC and two-alternative forced-identification tasks.[2] We then present existing empirical evidence and a new experiment to compare the observer models. We conclude that the five-component PTM, with a perceptual template, a nonlinear transducer function, internal additive noise, internal multiplicative noise, and a decision structure, provides the best account of all the existing data in visual tasks. The INUM,[3] which substitutes the nonlinear transducer function with decision uncertainty, has qualitatively similar properties but provides worse accounts of the data. All the other observer models, essentially various reduced forms of these two models, can be clearly rejected. The combination of the triple-TvC and double-pass methods provides critical constraints on the observer models. Finally, we discuss the implications of our results for the study of other sensory modalities and the study of observer state changes and high-level cognitive processes.

External Noise Paradigms

Equivalent Input Noise; Single TvC

The existence of an absolute threshold for every perceptual task suggests that the perceptual system is limited by some form of internal noise arising from intrinsic stimulus variability (Rose, 1948), receptor sampling errors (Geisler, 1989), randomness of neural responses (Tolhurst, Movshon, & Dean, 1983), loss of information during neural transmission (Barlow, 1957), and variation in absolute and comparative judgments as well as decision criteria (Wickelgren, 1968). Sensory psychologists long ago adapted the equivalent input noise method and the LAM to characterize the internal noise in the perceptual system (Ahumada & Lovell, 1971; Barlow, 1956; Bos & Deboer, 1966; Burgess et al., 1981; Eijkman et al., 1966; Moore, 1975; Nagaraja, 1964; Osman, 1971; Pelli, 1981; Tanner & Birdsall, 1958). In a typical applica-

[2] The choice to focus on 2AFC and two-alternative forced-identification tasks is made to simplify the presentation. The single-interval two-alternative forced-identification paradigm is equivalent to the 2AFC paradigm if (a) the stimuli are so far apart that each excites a different detector (or group of detectors) and (b) the outputs of the detectors are well enough labeled (Graham, 1989). However, if the identification is based on the output magnitudes instead of on detector identity, the single-interval identification paradigm is similar to a yes–no paradigm. The 2AFC identification paradigm used in this article and many of our previous publications (Dosher & Lu, 2000a, 2000b; Lu & Dosher, 1998, 1999, 2000) uses stimuli that are far apart, and each excites a different labeled detector. They are equivalent to 2AFC paradigms, not yes–no paradigms. The development can be extended to cover other paradigms that require more extensive treatment of the decision process.

[3] The label *induced noise* is related to that of multiplicative noise; the magnitudes of both induced and multiplicative noises depend upon the stimulus. The differences in the precise formulation of the two forms are addressed in the presentation of the PTM and the new experiment.

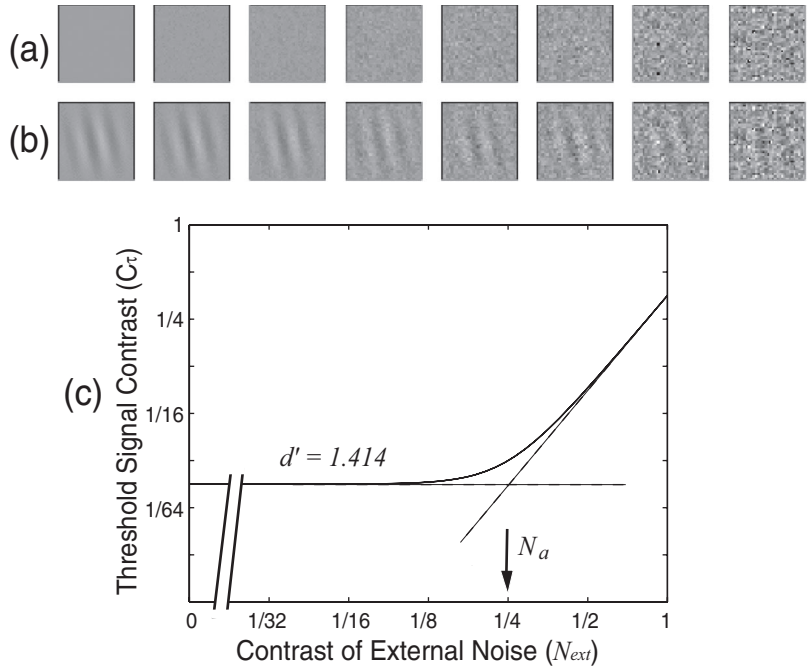


Figure 5. a: From left to right, eight external noise images with increasing contrast. b: A signal Gabor patch embedded in increasing levels of external noise. c: The input signal amplitude required to maintain threshold performance at a particular criterion level ($d' = 1.414$), as a function of the contrast (standard deviation) of external noise.

tion, contrast thresholds in a two-alternative forced-identification task or two-interval forced-choice detection task—signal stimulus energy required for an observer to maintain a predetermined performance level—are measured in a range of external noise conditions with either an adaptive procedure (e.g., the staircase procedure) or the method of constant stimuli. The external noise images in each external noise condition are made of pixels whose contrasts are drawn from independent and identically distributed Gaussian distributions with a particular standard deviation, which varies across external noise conditions. The external noise images generated with independent and identically distributed pixel contrasts are white, that is, with a flat power spectrum over a range of spatial frequencies. White-noise images are preferred in estimating the magnitude of internal noises because they have equal variance in all the spatial (or temporal) frequencies and can therefore be used to provide a good reference for the internal noise in all the frequencies. Often, six to nine external noise conditions, indexed by the variance of the contrast of the external noise images, are used to sample the TvC functions (see Figure 5). As illustrated in Figures 3 and 4, the elbow of the TvC function provides a direct estimate of the amount of equivalent internal noise if the system behaves like a linear electronic amplifier.[4] The TvC functions are normally graphed in log–log plots to facilitate viewing of data over a large dynamic range (a factor of 10 to 100 in both signal and external noise contrast) and to display thresholds with approximately equal error bars in log units.

Triple TvC

It has been well established that perceptual sensitivity (d') increases as a nonlinear function of signal contrast (Cohn, Thibos, & Kleinstein, 1974; Foley & Legge, 1981; Leshowitz, Taub, & Raab, 1968; Nachmias, 1981; Nachmias & Kocher, 1970; Nachmias & Sansbury, 1974; Stromeyer & Klein, 1974; Tanner, 1961). Some researchers (e.g., Foley & Legge, 1981; Lu & Dosher, 1999; Nachmias & Sansbury, 1974) have attributed the nonlinear relation between d' and signal contrast to some form of nonlinear transformation, or transducer function, acting on the stimulus strength; others have attributed the nonlinearity to statistical uncertainty in the decision process (Eckstein et al., 1997; Pelli, 1985). Regardless of the theoretical interpretation, measurements of the nonlinear properties of the perceptual system are essential for constructing and constraining observer models.

We (Lu & Dosher, 1999) introduced and incorporated the triple-TvC method into the equivalent input noise paradigm to measure the nonlinear properties of the perceptual system. In this method, TvC functions at three separate criterion performance levels (e.g., 65%, 75%, and 85% correct in a two-alternative forced-identification task) are measured for each observer (see Figure 6A).[5] From the three TvC functions, two threshold ratios at each external noise level can be obtained (see Figure 6B). Indicative of observer nonlinearities in perceptual tasks (Pelli, 1985), these ratios provide very strong empirical constraints on the nonlinear components of the observer models. We (Lu & Dosher, 1999)

[4] If the system is not a simple linear system, the interpretation is related but may depend on other factors.

[5] An alternative is to measure full psychometric function across all the external noise conditions. Using three widely separated performance criteria, the triple-TvC method provides an excellent proxy to the full psychometric method (Lu & Dosher, 1999).

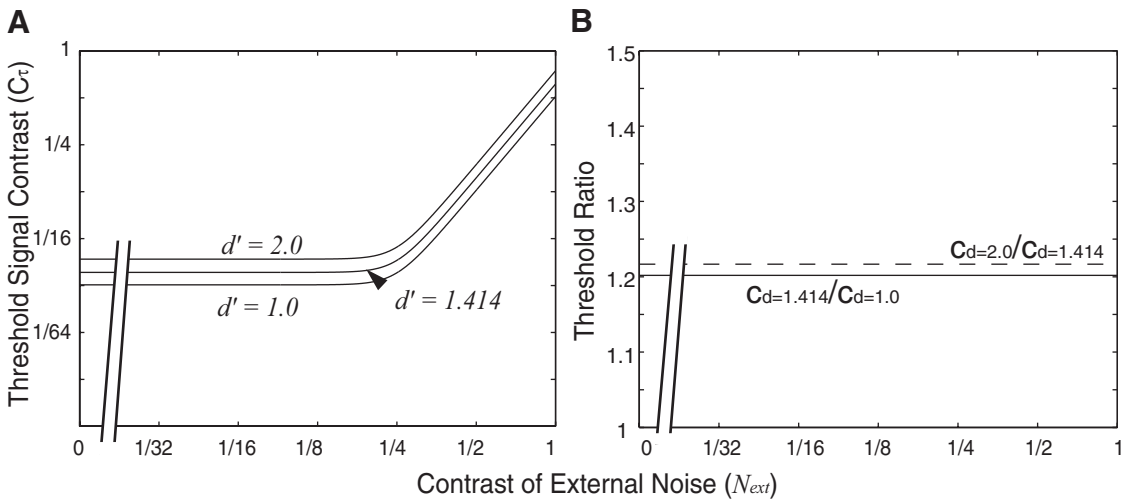


Figure 6. A: Threshold-versus-external-noise-contrast functions at three performance levels, corresponding to $d' = 1.0$, 1.414, and 2.0. B: Threshold ratios between two performance levels in each external noise condition: ratios between thresholds at $d' = 2.0$ and $d' = 1.414$ and ratios between thresholds at $d' = 1.414$ and $d' = 1.0$.

showed that, mathematically, measuring TvC functions at three performance criterion levels is necessary to fully constrain the PTM. We found that three reasonably widely spaced TvC functions provide sufficient constraints to solve for the necessary free parameters and a good approximation to a set of full psychometric functions over the same range of external noise conditions.

In Dosher and Lu (1998, 1999), TvC functions at two different performance levels were measured throughout multiple days of perceptual learning. The two TvCs were used to distinguish state-dependent changes due to an improved template or reduction of additive internal noise from changes in multiplicative noise or nonlinearity. We found that threshold ratios remained constant throughout training, despite large performance improvements from practice. Strong regularities in these ratio properties have been observed in a large set of experiments. These regularities rule out certain models and help to disambiguate different kinds of changes in behavior that result from changes of state in the observer.

Double-Pass Agreement

The Weber's law-type behavior of *difference thresholds* in perceptual tasks—that the just noticeable difference between two stimuli is proportional to the amplitude of the comparison stimulus—suggests that there is another noise source in the perceptual system, a multiplicative noise, whose amplitude is a function of the contrast energy of the input stimulus. In fact, there is ample evidence both from psychophysics (Burbeck & Kelly, 1981; Burgess & Colborne, 1988; Foley, 1994; Klein & Levi, 1985; Legge & Foley, 1980; Lu & Sperling, 1996; G. Sperling, 1989; Stromeyer & Klein, 1974; Watson & Solomon, 1997) and from neurophysiology (e.g., Albrecht & Geisler, 1991; Albrecht & Hamilton, 1982; Bonds, 1991; Derrington & Lennie, 1981; Heeger, 1993; Kaplan & Shapley, 1982; Ohzawa, Sclar, & Freeman, 1982; Sclar, Maunsell, & Lennie, 1990) that the perceptual system is limited by a form of noise whose amplitude is directly related to the total amount of contrast energy in the stimulus.

The double-pass procedure was developed to directly estimate the total amount of internal noise, both additive and multiplicative, relative to external noise, in the perceptual system for each stimulus (signal and external noise) condition (Ahumada, 1967; Burgess & Colborne, 1988; Gilkey, Frank, & Robinson, 1978, 1981; Green, 1964; Spiegel & Green, 1981). In comparison, the equivalent input noise method with a single-TvC function was developed to estimate the magnitude of a single fixed (additive) noise source across all the external noise conditions. In the double-pass procedure, the same sequence of stimulus trials (signal + external noise) is repeated twice for each observer. Repeating each particular sample of external noise provides an assessment of the relative influence of the external and internal noises. Both response accuracy and response consistency (whether the response is or is not the same on the two identical tests) are measured across different passes of the same stimulus condition. As summarized by Green (1964, p. 397), "On an operational level, internal noise is equivalent to the observation that the same physical stimulus may elicit different responses. In a sense, then, internal noise is the limiting factor in a trial-by-trial prediction of the subject's response."

The principle is illustrated in Figure 7a, where we show how signal-to-noise ratio and probability of agreement depend on the amplitudes of the input signal and internal noise for external noise at a single level. Signals with increasing amplitudes are mixed with external noise of a fixed standard deviation. Each pair of the waveforms is made of the same external noise sample, the same signal, and independent samples of internal noise, mimicking the double-pass procedure. When the internal noise is zero (the leftmost column), the correlation between each pair of waveforms is 1.0, independent of the signal-to-noise ratio (or percent correct in behavioral tests). When the internal noise increases, it decreases the correlation between each pair of waveforms at a given signal amplitude.

The percent-agreement results are traditionally (Burgess & Colborne, 1988) displayed for conditions of varying signal contrast and measure probability correct (P_C) as a function of the probability that

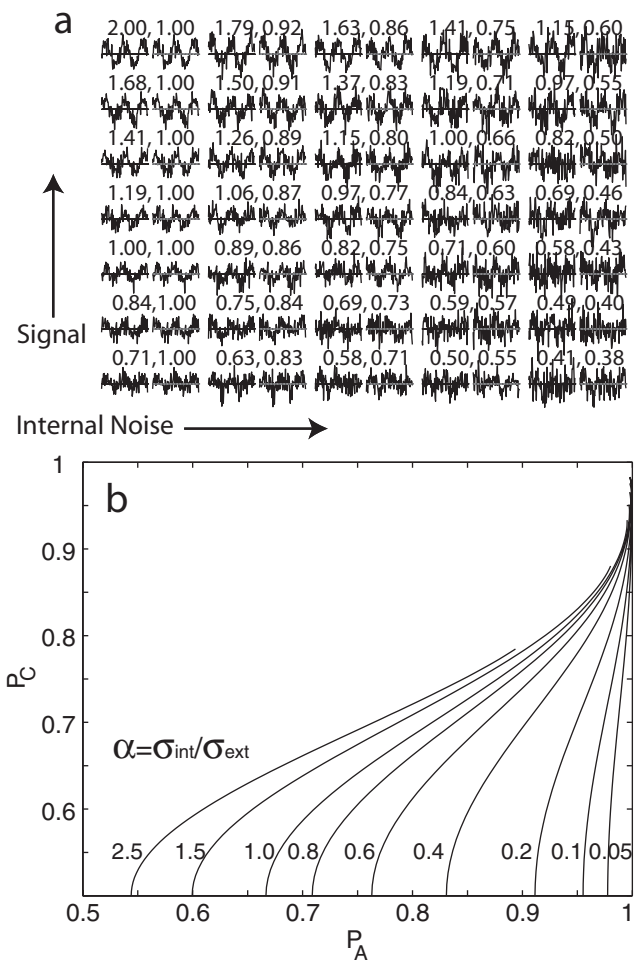


Figure 7. a: An illustration of the principle behind the double-pass method. Signals with increasing amplitudes are mixed with external noise of a fixed standard deviation. Each pair of waveforms is made of the same external noise sample, the same signal, and independent samples of internal noise. The pair of numbers above each pair of waveforms denotes the signal-to-noise ratio in the waveforms and their correlations. The standard deviation of the internal noise is varied. When the internal noise is zero (the leftmost column), the correlation between each pair of waveforms is 1.0, independent of the signal-to-noise ratio, which determines percent correct. When the internal noise increases, it decreases the correlation between each pair of waveforms at a given signal amplitude. b: Probability correct (P_C) versus probability consistent (P_A) for a range of internal-to-external-noise ratio αs.

the two responses to the two presentations of the same stimulus are in agreement (P_A), regardless of being right or wrong (see Figure 7b). The double-pass method is designed to estimate the ratio, α, of the standard deviations of the total internal noise and of the external noise. α completely determines the shape of the P_C versus P_A function in each external noise condition (see Appendix B). The functional form of the relationship between P_C and P_A is derived by extending the basic SDT equations to the double-pass procedure.

A family of P_C versus P_A functions for a range of internal-to-external-noise ratio αs is illustrated in Figure 7b. When $\alpha \to 0$, $P_A \to 1.0$ (the curve approaches a vertical line), even though P_C still goes from 50% (chance) to 100% (perfect performance). This is because, without any internal noise ($\alpha \to 0$), the performance of the observer is completely determined by external noise. A particular external noise sample may cause the observer to make a wrong response, yet its impact on performance is the same in the two passes, and therefore, the responses are consistent. As α increases, the observer is (relatively) more affected by internal noise and becomes less consistent in her or his response. Therefore, the P_C versus P_A curves become more slanted.

In actual experiments, the ratio of the standard deviation of internal-versus-external-noise α is estimated from the P_C versus P_A function in each external noise condition (Burgess & Colborne, 1988). To provide a reliable estimate of α for a given external noise contrast, the P_C versus P_A functions need to be measured over a range of signal contrast levels such that the data can be fit with the theoretical P_C versus P_A curves. The estimated standard deviation of (the total) internal noise for a given external noise condition is then $\sigma_{int} = \alpha\sigma_{ext}$. So, through this method, the total internal noise is benchmarked to a physical quantity that is controlled by the experimenter. The total internal noise may include both multiplicative and additive internal noises.

Mathematical Properties of the Observer Models

In this section, we introduce and present properties of the five most prominent observer models. The components of the models are summarized in Table 1. We start with the model that has the fewest components and then gradually build up to models with more components. The mathematical symbols used in the various models are defined in Table 2. We seek to identify the simplest model(s) consistent with the key properties of the behavioral data.

Linear Amplifier Model

The LAM of a human observer (see Figure 8a) is essentially a direct analogue of the linear electronic amplifier model (see Figure 3a). The terminology is however rather different. The amplification in Figure 3a is replaced by a perceptual template (see Figure 8a). The perceptual template is essentially a task-specific filter tuned to the relevant signal stimulus, to which it responds with a contrast gain of β to a signal stimulus of contrast 1. The value c is the contrast of the signal stimulus. The additive noise, N_{add}, represents the impact of the aggregate of all the intrinsic additive noise sources in the perceptual system. A decision stage is also added to model the human decision process (Green & Swets, 1966; Macmillan & Creelman, 1991).

In the LAM, signal discriminability, d', is determined by the signal-to-noise ratio (see Appendix C for details):

$$d' = \frac{T_S}{N_{\text{total1}}} = \frac{T_S}{N_{\text{total2}}} = \frac{\beta c}{\sqrt{\sigma_{\text{ext}}^2 + \sigma_{\text{add}}^2}}. \quad (1)$$

For a 2AFC task, we can simply substitute d' in Equation A10 (see Appendix A) with Equation 1:

$$P_C = \int_{-\infty}^{+\infty} g(x - d', 0, 1)G(x, 0, 1)dx$$

$$= \int_{-\infty}^{+\infty} g\left(x - \frac{\beta c}{\sqrt{\sigma_{\text{ext}}^2 + \sigma_{\text{add}}^2}}, 0, 1\right)G(x, 0, 1)dx. \quad (2)$$

Table 1
Components of the Five Most Prominent Observer Models

Model	Perceptual template	Additive, internal noise	Decision process	Decision uncertainty	Induced noise	Multiplicative noise	Nonlinear transducer
LAM	√	√	√				
LAUM	√	√	√	√			
INM	√	√	√		√		
INUM	√	√	√	√	√		
PTM	√	√	√			√	√

Note: LAM = linear amplifier model; LAUM = linear amplifier model with uncertainty; INM = induced noise model; INUM = induced noise with uncertainty model; PTM = perceptual template model.

The probability that the observer responds to two passes of the same stimulus consistently follows directly from Equation B3a (see Appendix B) after substituting S with the template response to the signal stimulus βc:

$$P_A = \int_{-\infty}^{+\infty} g(x - \beta c, 0, \sqrt{2}\sigma_{ext})\{G^2(x, 0, \sqrt{2}\sigma_{add}) + [1 - G(x, 0, \sqrt{2}\sigma_{add})]^2\}dx. \quad (3)$$

By inverting Equation 1, we can express the threshold signal contrast energy c_τ^2 required for the observer to maintain a given performance criterion level, that is, a fixed P_C or d', as a function of external noise contrast:

$$c_\tau^2 = k(\sigma_{ext}^2 + \sigma_{add}^2), \quad (4)$$

where $k = (d'/\beta)^2$. This is the efficiency relation between threshold and external noise in LAM. The parameter k has historically been called *calculation efficiency,* thought to reflect the efficiency of a human observer in utilizing the information in the stimulus relative to an ideal observer who can make optimal use of all the information in the stimulus in making a perceptual judgment (Pelli, 1981). Because k is proportional to d'^2, it is obvious that it (and thus observer efficiency) depends on the particular criterion performance level at which the threshold is defined. That is, the efficiency parameter k depends on the criterion performance level (i.e., 75%) selected by the experimenter and is therefore not a fundamental property of the human observer. The more fundamental parameter in this model is the gain of the perceptual template to the signal stimulus β, which is independent of the performance criterion.

The TvC relation expressed in Equation 4 is illustrated in Figure 8b for a hypothetical LAM at three performance levels, 65%, 75%, and 85% correct. The TvC functions, typically shown in log–log plots, have three distinct regions: (a) When $\sigma_{ext} \ll \sigma_{add}$, internal noise is the limiting factor of performance, threshold contrast $\log(c_\tau)$ is almost invariant to $\log(\sigma_{ext})$, and the TvC function is almost flat. (b) When $\sigma_{ext} \gg \sigma_{add}$, external noise is the limiting factor of performance, threshold contrast $\log(c_\tau)$ increases linearly with $\log(\sigma_{ext})$, and the TvC function has a nearly constant slope. (c) When $\sigma_{ext} \approx \sigma_{add}$, internal and external noises are both important in determining performance level, there is a smooth transition from Region 1 to Region 2 on the TvC function.

Empirically, to estimate σ_{add} and β for an LAM in a particular experiment, a TvC function is obtained by measuring signal contrast thresholds at a single performance criterion level (e.g., 75% correct) using an adaptive procedure (e.g., staircase) or the method of constant stimuli over a range of external noise levels. The optimal σ_{add} and β are then estimated from the TvC function using Equation 4.

Although measurement of a single-TvC function is sufficient to constrain the parameters of an LAM and is conventionally performed in many empirical studies, the LAM does make a very simple yet highly constraining prediction on the ratio between thresholds at different performance criteria: In any given external noise condition, the ratio between contrast thresholds at two criterion performance levels is

$$\frac{c_{\tau 1}}{c_{\tau 2}} = \frac{d_1'}{d_2'}. \quad (5)$$

In other words, the LAM predicts that the contrast threshold ratio between two criterion performance levels for a given external noise contrast is equal to the ratio of the corresponding d's. This strong relationship holds at all external noise levels. As we discuss

Table 2
Definitions of Common Symbols

Symbol	Definition
μ	Mean of a distribution.
σ	Standard deviation of a distribution.
C	Response criterion.
d'	Sensitivity, often expressed as the ratio of signal and noise.
U	Number of hidden detectors.
P_C	Probability of correct response.
P_A	Probability two responses to the two passes of the same input stimulus agree.
S	Internal response to the signal stimulus.
β	Gain of the perceptual template to the signal stimulus.
γ	The exponent of the nonlinear transducer function.
N_{ext}	Internal response to an external noise stimulus.
N_{mul}	Proportional constant of induced or multiplicative noise.
$g(x, \mu, \sigma)$	Probability density function of a Gaussian random variable x, with mean μ, and standard deviation σ.
$G(x, \mu, \sigma)$	Cumulative probability density function of a Gaussian random variable x, with mean μ, and standard deviation σ.
$G^{-1}(p, \mu, \sigma)$	Inverse cumulative Gaussian probability density function of probability p, with mean μ, and standard deviation σ.

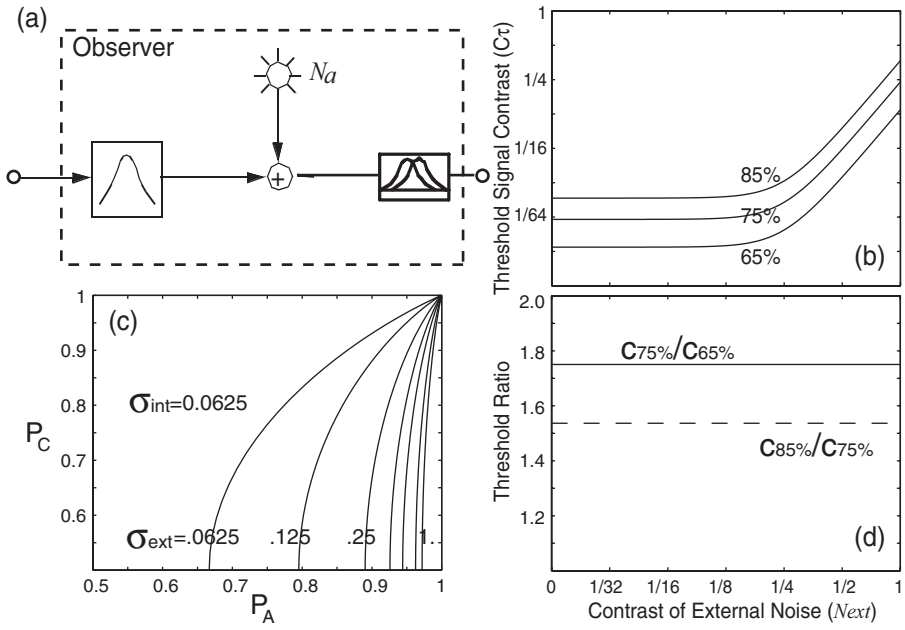


Figure 8. a: The linear amplifier model of a human observer. b: Threshold contrasts required to maintain performance at 65%, 75%, and 85% correct (corresponding to $d' = 0.54$, 0.95, and 1.47) as functions of the contrast (standard deviation) of external noise. c: Probability correct (P_C) versus probability consistent (P_A) functions for a range of external noise levels. d: Threshold ratios between different performance criterion levels: 75% and 65% correct (solid line) and 85% and 75% correct (dashed line).

in the next section, the prediction that the threshold contrast ratio equals the d' ratio fails all tests to date. This failure is related to the fact that, for the LAM, the parameters estimated from TvC curves measured at different performance levels are inconsistent.

Linear Amplifier Model With Uncertainty

The ratio relationship between contrast thresholds and the corresponding d's in the LAM (see Equation 5) implies a linear relation between d' and signal contrast. However, it has been well established in the visual domain that the observed d'_{obs} increases as a power function of signal contrast (Cohn et al., 1974; Foley & Legge, 1981; Leshowitz et al., 1968; Nachmias, 1981; Nachmias & Kocher, 1970; Nachmias & Sansbury, 1974; Stromeyer & Klein, 1974; Tanner, 1961). Pelli (1985) proposed that the nonlinear relationship is due to statistical uncertainty in the decision process, that is, the observer is uncertain about some aspects of the signal and therefore makes decisions based not only on task-relevant decision variables but also on task-irrelevant decision variables. For example, perhaps the exact spatial frequencies or orientations of the signal are sampled, but so too are inputs from other spatial frequencies or orientations.

Within the SDT framework, decision under uncertainty is modeled with task-irrelevant *hidden detectors,* which add sources of false alarms to the decision process. In a 2AFC task, the observer is presented with two input stimuli, one from each of two stimulus categories. However, the observer uses the outputs of $(U + 1)$ detectors in determining the response. Only one of those detectors is task relevant, but the observer cannot identify it and has to make a decision based on the responses of all the detectors. The observer therefore has to monitor a total of $2(U + 1)$ internal responses, of which one is from the detector relevant to the stimulus from Category 1, one is from the detector relevant to the stimulus from Category 2, and $2U$ are from task-irrelevant detectors (see Appendix A). The inability to identify the task-relevant detector in making perceptual decisions is termed *decision uncertainty.*

We refer to the LAM augmented with decision uncertainty as the LAUM (see Figure 9). This is typically implemented with a maximum decision rule (see Appendix A) with U hidden detectors (Eckstein et al., 1997). The maximum rule is not the optimal Bayesian rule in decision uncertainty but approximates it in many cases (Nolte & Jaarsma, 1967; G. Sperling & Dosher, 1986).

The basic d' function in the LAUM is the same as that of the LAM (see Equation 1). We can simply substitute d' in the SDT with the uncertainty equation (see Equation A12 in Appendix A) with Equation 1 to compute P_C:

$$P_C = \int_{-\infty}^{+\infty} [g(x - d', 0, 1)G^{2U+1}(x, 0, 1) + Ug(x, 0, 1)G^{2U}(x, 0, 1)G(x - d', 0, 1)]dx$$

$$= \int_{-\infty}^{+\infty} \left[g\left(x - \frac{\beta c}{\sqrt{\sigma_{ext}^2 + \sigma_{add}^2}}, 0, 1\right) G^{2U+1}(x, 0, 1) + Ug(x, 0, 1)G^{2U}(x, 0, 1)G\left(x - \frac{\beta c}{\sqrt{\sigma_{ext}^2 + \sigma_{add}^2}}, 0, 1\right)\right] dx. \quad (6)$$

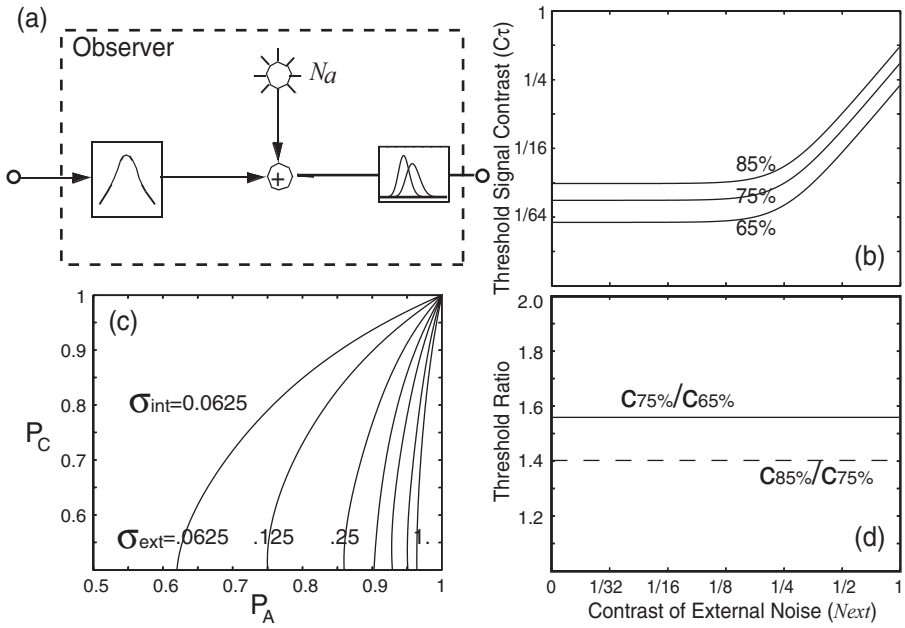


Figure 9. a: The linear amplifier with uncertainty model of a human observer ($U = 2$). b: Threshold contrasts required to maintain performance at 65%, 75%, and 85% correct (corresponding to $d' = 0.90$, 1.40, and 1.97 and $d'_{obs} = 0.54$, 0.95, and 1.47) as functions of the contrast (standard deviation) of external noise. c: Probability correct (P_C) versus probability consistent (P_A) functions for a range of external noise levels. d: Threshold ratios between different performance criterion levels: 75% and 65% correct (solid line) and 85% and 75% correct (dashed line).

In the LAM, the internal response distributions at the decision stage are Gaussian. In the LAUM with a maximum decision rule, the relevant internal distributions are derived from the maximum of the $U + 1$ internal responses generated by stimuli in the two categories (see Appendix A). In Equation 6, d' is the signal-to-noise ratio in the two task-relevant detectors. It must be clearly distinguished from the observed d'_{obs}, which is normally converted from observed probability correct (P_C) using a table that does not consider decision uncertainty. The observed signal-to-noise ratio, d'_{obs}, is less than the underlying signal-to-noise ratio in the task-relevant detectors, d', because decision uncertainty, that is, including the activities of the task-irrelevant detectors in the decision process, increases the level of false alarms. We illustrate the relationships between P_C and d' and between d'_{obs} and d' in Figure 10 for a range of Us. In the LAM, $U = 0$, and $d'_{obs} = d'$. However, in the LAUM, $U > 0$, and d'_{obs} is a nonlinear function of d'.[6]

A number of properties of the LAUM are illustrated in Figure 9. In Figure 9b, we plot TvC functions at 65%, 75% and 85% correct for a 2AFC task with $U = 2$, corresponding to $d'_{obs} = 0.54$, 0.95, and 1.47, but $d' = 0.90$, 1.40, and 1.97 in the task-relevant detectors. This is because, as stated above, d'_{obs} is calculated from percent correct assuming $U = 0$.

Similar to the LAM, the threshold ratios between different performance levels in the LAUM are equal to the corresponding d' ratios, independent of the external noise levels. Because d'_{obs} and d' are related nonlinearly, the threshold ratios are therefore nonlinear functions of the corresponding d'_{obs}s. This is why decision uncertainty could be a potential explanation for nonlinear psychometric functions. We plot two threshold ratios at each external noise level for a 2AFC task with $U = 2$ in Figure 9b. The threshold ratios for a range of Us, $c_{75\%}(N_{ext}, U)/c_{65\%}(N_{ext}, U)$, and $c_{75\%}(N_{ext}, U)/c_{65\%}(N_{ext}, U)$, are plotted in Figure 11.

The derivation and the resulting analytic relationship between probability correct and probability agreement for the LAUM are too long to present here. Instead, we used a Monte Carlo simulation procedure to compute the functions. As in the LAM, the functional relationship between P_C and P_A is determined by the ratio of (total) internal and external noise in the LAUM. The MATLAB program used in the Monte Carlo simulation is presented in Appendix D. We illustrate P_C as functions of P_A for a range of external noise conditions for a given LAUM with $U = 2$ in Figure 9c. The relationship between P_C and P_A is illustrated for a range of Us in Figure 12. One important characteristic of these functions is that the spread of the P_C versus P_A functions for different αs increases with U. As we discuss later, relatively large Us are required to account for the empirical threshold ratios between multiple performance levels, yet the empirical P_C versus P_A functions tend to collapse as external noise increases. The two opposite demands on the value of U undermine the LAUM.

Induced Noise Model

In both the LAM and the LAUM, the internal noise consists of a single fixed additive component that does not vary with stimulus

[6] d' is determined by the signal-to-noise ratio in a detector, which is not directly observed in an experiment. The ratio test in triple TvC is always performed on d'_{obs}, which is converted from measured percent correct.

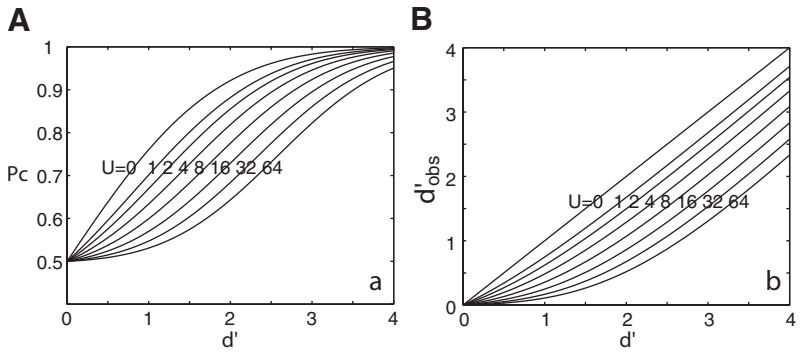


Figure 10. A: A plot of the relationship between probability correct (P_C) and the d' in the task-relevant detector. B: A plot of the relationship between the observed d'_{obs} and the d' in the task-relevant detector for a two-alternative forced-choice decision for a range of number of hidden detectors (U).

conditions. Using the double-pass procedure, Burgess and Colborne (1988) concluded that the internal noise has two components: one constant additive noise component and the other induced noise component with a standard deviation proportional to that of the external noise: $\sigma_{induced} = N_{mul}\sigma_{ext}$. The latter component is necessary to account for the fact that the estimated amplitude of the total internal noise is proportional to that of the external noise once the external noise level exceeds a certain level. A superset of the LAM can be made to incorporate the induced noise by adding another noise term to Equation 1:

$$d' = \frac{\beta c}{\sqrt{\sigma_{ext}^2 + \sigma_{induced}^2 + \sigma_{add}^2}} = \frac{\beta c}{\sqrt{(1+N_{mul}^2)\,\sigma_{ext}^2 + \sigma_{add}^2}}. \tag{7}$$

The corresponding TvC function becomes

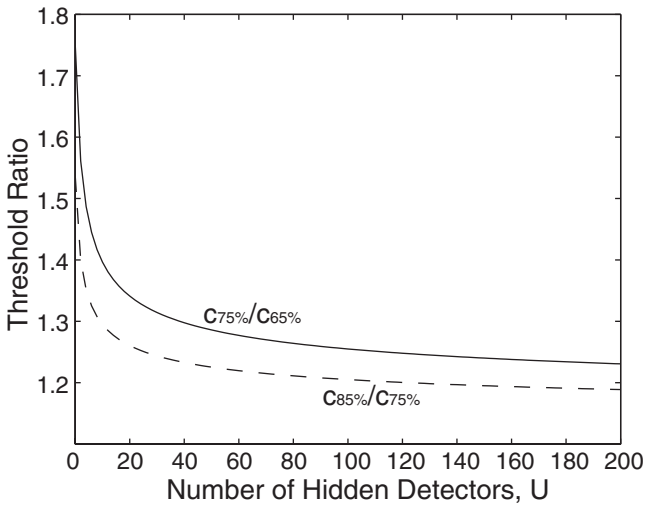


Figure 11. Threshold ratios between two performance levels in each external noise condition as functions of number of hidden detectors in the linear amplifier model with uncertainty.

$$c_\tau^2 = \left(\frac{d'}{\beta}\right)^2 [\sigma_{ext}^2 + \sigma_{induced}^2 + \sigma_{add}^2] = \left(\frac{d'}{\beta}\right)^2 [(1 + N_{mul}^2)\sigma_{ext}^2 + \sigma_{add}^2]. \tag{8}$$

The INM is illustrated in Figure 13a. There is no decision uncertainty in the INM: $U = 0$, $d'_{obs} = d'$, and the noise and signal distributions are both assumed to be Gaussian. The TvC functions for an INM at three performance levels, 65%, 75%, and 85% correct, are shown in Figure 13b.

For the INM, the probability correct in a 2AFC task is computed by substituting d' (see Equation 7) into Equation A10 (see Appendix A):

$$P_c = \int_{-\infty}^{+\infty} g(x - d', 0, 1)G(x, 0, 1)dx = \int_{-\infty}^{+\infty} g\left(x - \frac{\beta c}{\sqrt{(1 + N_{mul}^2)\sigma_{ext}^2 + \sigma_{add}^2}}, 0, 1\right)G(x, 0, 1)dx \tag{9}$$

Equation B3a (see Appendix B) is also elaborated to include the induced noise in calculating probability agreement for the double-pass procedure:

$$P_A = \int_{-\infty}^{+\infty} g(x - \beta c, 0, \sqrt{2}\sigma_{ext})\{G^2(x, 0, \sqrt{2(\sigma_{add}^2 + N_{mul}^2\sigma_{ext}^2)} + [1 - G(x, 0, \sqrt{2(\sigma_{add}^2 + N_{mul}^2\sigma_{ext}^2)})]^2\}dx. \tag{10}$$

In the INM, as σ_{ext} increases, the total amount of internal noise ($\sigma_{total} = \sqrt{\sigma_{add}^2 + N_{mul}^2\sigma_{ext}^2}$) is increasingly dominated by the induced noise, that is, $\sigma_{total} \rightarrow N_{mul}\sigma_{ext}$; therefore, the ratio between the amplitudes of the internal noise and external noise approaches a constant, and the family of P_C versus P_A functions in all the external noise conditions, determined by the ratio of (total) internal and external noise, approaches a single curve. This point is illustrated in Figure 13c.

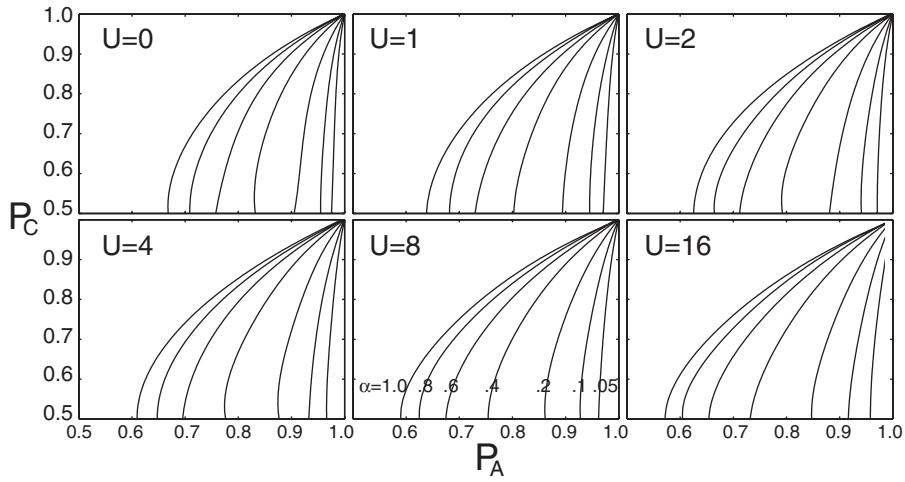


Figure 12. Relationship between probability correct (P_C) and probability agreement (P_A) for a range of number of hidden detectors (U) in the linear amplifier model with uncertainty.

Following Equation 8, we show that the contrast threshold ratio between two threshold performance levels in a given external noise condition is equal to the ratio of the corresponding d's in the INM, independent of the contrast of the particular external noise level (see Figure 13d):

$$\frac{c_{\tau 1}}{c_{\tau 2}} = \frac{d'_1}{d'_2}. \quad (11)$$

Thus, this model has the same failure as the LAM and is rejected by the same failure of the contrast threshold ratio predictions. The next two models address this shortcoming by reintroducing decision uncertainty (INUM) or, alternatively, a nonlinear transducer function (PTM).

Induced Noise With Uncertainty Model

Eckstein et al. (1997) constructed another observer model by adding decision uncertainty to the INM. They suggested that induced noise is necessary to account for the component of internal noise that increases with background contrast and that decision uncertainty is necessary to account for the nonlinear d' psychometric functions.[7] We refer to the model proposed by Eckstein et al. as the INUM.

The INUM (see Figure 14a) is an extension of the INM, with the addition of decision uncertainty. As in the INM, Equations 7 and 8 (duplicated below) describe the signal-to-noise ratio and threshold function in the INUM:

$$d' = \frac{\beta c}{\sqrt{\sigma^2_{ext} + \sigma^2_{induced} + \sigma^2_{add}}} = \frac{\beta c}{\sqrt{(1 + N^2_{mul})\sigma^2_{ext} + \sigma^2_{add}}}, \quad (7)$$

and

$$c_\tau^2 = \left(\frac{d'}{\beta}\right)^2 [\sigma^2_{ext} + \sigma^2_{induced} + \sigma^2_{add}]$$

$$= \left(\frac{d'}{\beta}\right)^2 [(1 + N^2_{mul})\sigma^2_{ext} + \sigma^2_{add}]. \quad (8)$$

However, the d's in Equations 7 and 8 represent the signal-to-noise ratio in the task-relevant detectors. As is the case with the LAUM, they must be clearly distinguished from the observed d'_{obs}, which is often converted from the observed percent correct (P_C) using a table that assumes no decision uncertainty ($U = 0$). In Figure 14b, we plot TvC functions at 65%, 75% and 85% correct for a 2AFC task with $U = 2$, corresponding to $d'_{obs} = 0.54$, 0.95, and 1.47, but $d' = 0.90$, 1.40, and 1.97 in task-relevant detectors.

A Monte Carlo simulation procedure was used to compute the probability correct (P_C) versus probability agreement (P_A) functions using the MATLAB program in Appendix D with $\sigma_S = \sigma_N = \sqrt{N^2_{mul}\sigma^2_{ext} + \sigma^2_{add}}$. We illustrate P_C as functions of P_A for a range of external noise conditions for the INUM described above (see Figure 14c). Similar to the INM, as σ_{ext} increases, the total amount of internal noise ($\sigma_{total} = \sqrt{\sigma^2_{add} + N^2_{mul}\sigma^2_{ext}}$) is increasingly dominated by the induced noise, that is, $\sigma_{total} \rightarrow N_{mul}\sigma_{ext}$; the P_C versus P_A functions approach a single curve.

Similar to the LAUM, the threshold ratios between different performance levels in the INUM are equal to the corresponding d' ratios, independent of the external noise levels. However, because d'_{obs} and d' are related nonlinearly, the threshold ratios are therefore nonlinear functions of the corresponding d'_{obs}s. This is why decision uncertainty could be a potential explanation of nonlinear psychometric functions. We plot two threshold ratios at each external noise level for a 2AFC task with $U = 2$ in Figure 14b.

Perceptual Template Model

The PTM was proposed to explicitly model nonlinear psychometric functions and Weber's law in perceptual tasks

[7] Some applications of this model to data allow each condition to have an independently estimated degree of uncertainty—allowing a special adjustment for each condition (Eckstein et al., 1997). Our illustrations impose a single value of uncertainty, U, across multiple conditions of a basic task.

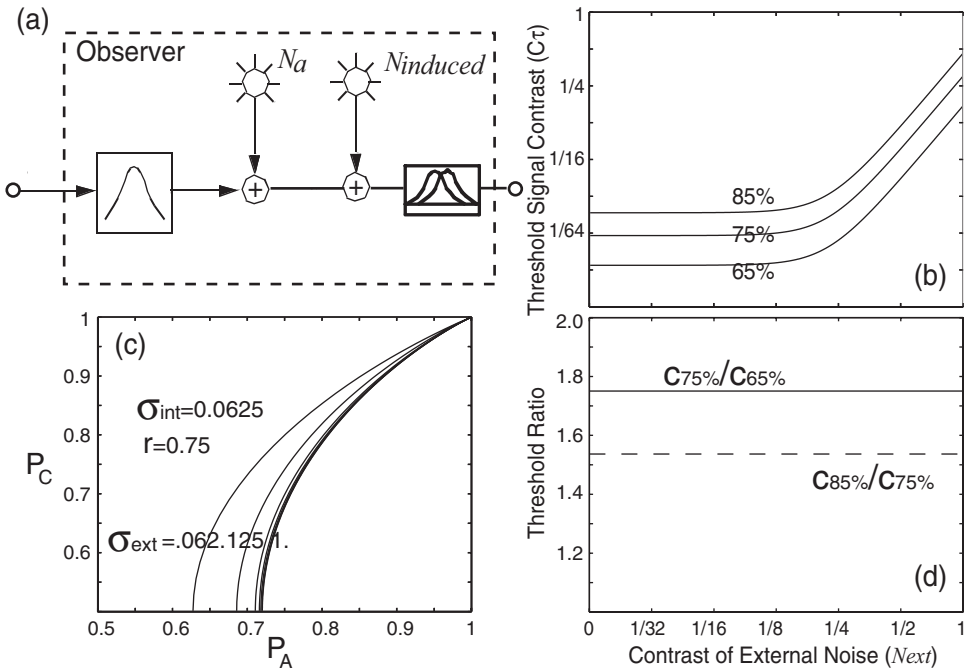


Figure 13. a: The induced noise model of a human observer. b: Threshold contrasts required to maintain performance at 65%, 75%, and 85% correct (corresponding to $d' = 0.5449$, 0.9539, and 1.4657) as functions of the contrast (standard deviation) of external noise. c: Probability correct (P_C) versus probability consistent (P_A) functions for a range of external noise levels. d: Threshold ratios between different performance criterion levels: 75% and 65% correct (solid line) and 85% and 75% correct (dashed line).

(Woodworth, 1938). Following the tradition in pattern vision (Foley, 1994; Foley & Legge, 1981; Fredericksen & Hess, 1997; Gorea & Sagi, 2001; Klein & Levi, 1985; Kontsevich, Chen, & Tyler, 2002; Legge & Foley, 1980; Nachmias & Sansbury, 1974; Watson & Solomon, 1997), the PTM includes a nonlinear transducer function instead of decision uncertainty to model nonlinear d' psychometric functions. Two noise sources, an additive noise and a multiplicative noise, produce internal noise in the PTM. The multiplicative noise accounts for Weber's law in perceptual tasks. Unlike the induced noise in INM and INUM, which is only related to external noise contrast, the magnitude of multiplicative noise in the PTM is determined by the total amount of contrast energy in the input stimulus, including contributions from both the signal stimulus and external noise stimulus. It is necessary to include signal contributions to the multiplicative noise to (a) induce the classic Weber-like phenomena and (b) make the model equivalent to contrast-gain control models. Dao et al. (2006) showed that the PTM with multiplicative noise is mathematically equivalent to a contrast-gain control model with two fixed noise sources, one before and the other after the gain control. Additionally, excluding a signal contribution to the multiplicative noise implies an ability to perfectly segregate signal from noise in processing, which presupposes a solution to the signal–noise problem.

In the PTM, input stimuli are processed in two pathways (see Figure 15a and Appendix E). In the signal pathway, input stimuli pass through a perceptual template with certain selectivity for stimulus characteristics (e.g., color, spatial frequency, orientation, temporal/spatial windowing). As in the LAM, the gain of the template to white Gaussian external noise is 1.0 because the total gain of the template is normalized to 1.0. The contrast gain of the template to the matched signal stimulus is β and to the nonmatched signal stimulus is 0. A template matching function might, however, be far more complex, for example, templates for objects, faces, and so on. It is related to the concept of matched filter in investigations of object recognition (Burgess, 1985). The outputs of the perceptual template are then processed by an expansive nonlinear transducer function
(Output = sign(Input)|Input|$^{\gamma_1}$), typically used in the pattern vision literature (Foley & Legge, 1981; Nachmias & Sansbury, 1974). In the gain-control pathway, input stimuli also pass through a perceptual template, and the output of this pathway determines the amplitude of the multiplicative noise. The template in this pathway may or may not differ from the template in the signal pathway; the response to the signal stimulus through this template is $\beta_2 c$ if the signal matches the template and 0 if it does not. Similarly, the nonlinearity in this path is parameterized by γ_2.

At the decision stage, using the standard formula (see Equation A10 in Appendix A) with the expanded definitions of internal noise and nonlinearity, probability correct for a 2AFC task is:

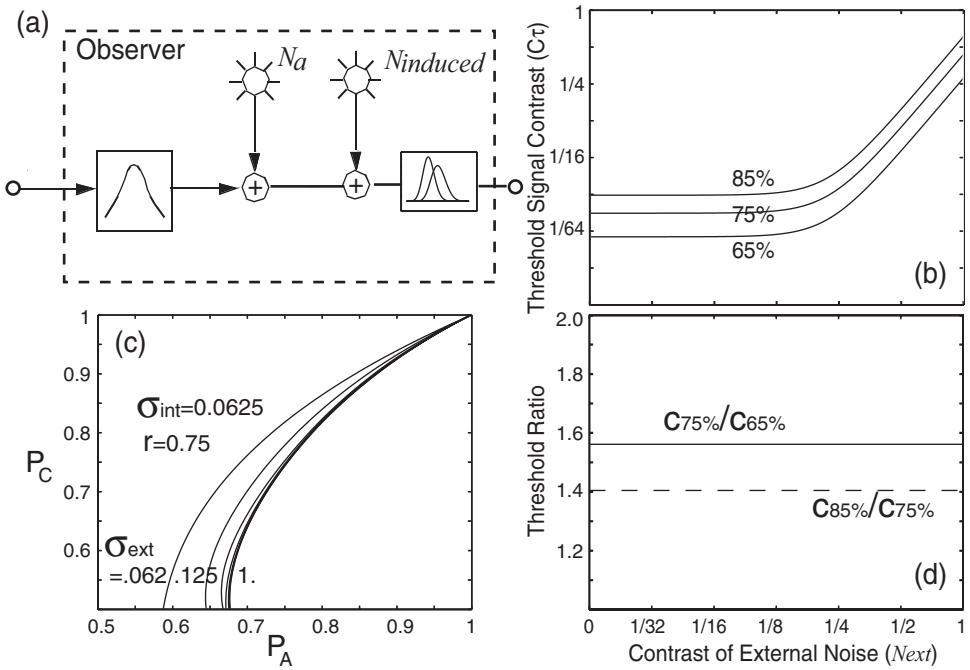


Figure 14. a: The induced noise with uncertainty model of a human observer. b: Threshold contrasts required to maintain performance at 65%, 75%, and 85% correct (corresponding to $d' = 0.90$, 1.40, and 1.97) as functions of the contrast (standard deviation) of external noise. c: Probability correct (P_C) versus probability consistent (P_A) functions for a range of external noise levels. d: Threshold ratios between different performance criterion levels: 75% and 65% correct (solid line) and 85% and 75% correct (dashed line).

$$P_C = \int_{-\infty}^{+\infty} g(x - \beta^{\gamma_1} c^{\gamma_1}, 0, \times \sqrt{\sigma_{ext}^{2\gamma_1} + N_{mul}^2[\sigma_{ext}^{2\gamma_2} + (\beta_2 c)^{2\gamma_2}] + \sigma_{add}^2}) \times G(x, 0, \sqrt{\sigma_{ext}^{2\gamma_1} + N_{mul}^2 \sigma_{ext}^{2\gamma_2} + \sigma_{add}^2}) dx. \quad (12)$$

The average signal-to-noise ratio (d') in the PTM can be calculated[8]:

$$d' = \frac{\sqrt{2} S_1}{\sqrt{\sigma_{total1}^2 + \sigma_{total2}^2}} = \frac{(\beta c)^{\gamma_1}}{\sqrt{\sigma_{ext}^{2\gamma_1} + N_{mul}^2 \left[\sigma_{ext}^{2\gamma_2} + \frac{(\beta_2 c)^{2\gamma_2}}{2} \right] + \sigma_{add}^2}}. \quad (13)$$

The special case $\gamma = \gamma_1 = \gamma_2$ corresponds to the situation where the rising portion of the TvC function in log–log plot has a slope of 1.0; this property is consistent with many observed TvC functions in the literature. If $\gamma = \gamma_1 = \gamma_2$, we can obtain an analytical relationship between threshold signal contrast c_τ and external noise contrast σ_{ext} at a given performance criterion (i.e., d')[9]:

$$c_\tau = \left\{ \frac{d'^2[(1 + N_{mul}^2)\sigma_{ext}^{2\gamma} + \sigma_{add}^2]}{\beta^{2\gamma} - N_{mul}^2 \beta_2^{2\gamma} d'^2/2} \right\}^{\frac{1}{2\gamma}}. \quad (14)$$

The functional relation between c_τ and σ_{ext} in Equation 14 is illustrated in Figure 15b for a PTM at three performance levels: 65%, 75% and 85% correct (corresponding to $d' = 0.54$, 0.95, and 1.47).

When the same stimulus (signal + external noise) is passed to the PTM twice, the probability that the two responses are consistent can be derived from Equation B3a (see Appendix B):

[8] When the variances of the signal and noise distributions are not equal, the ROC curve in a yes–no paradigm is not symmetric about the diagonal line. There are, however, several ways to define a d'. Our choice of the d' definition is based on one important mathematical property of the ROC curve in 2AFC, which is symmetric about the diagonal even when the signal and noise distributions have unequal variance (Green & Swets, 1966; Macmillan & Creelman, 1991). Consistent with the difference rule in 2AFC, we define d' as the ratio of the mean and the standard deviation of the difference distribution (the difference between the internal responses of the two detectors); the variance of the difference distribution is equal to the sum of the variances of the component distributions. In previous PTM applications (Dosher & Lu, 1999; Lu & Dosher, 1999), we assumed that the template in the gain-control pathway is broadly tuned; independent of whether the signal stimulus matches the template in the signal pathway, the response of the template in the gain-control pathway is $\beta_2 c$. Therefore, the variance of the multiplicative noise is the same in the two detectors, one matched to the stimulus and the other not matched to the stimulus, eliminating the factor of 1/2 in Equation 13. Each is a reasonable approximation given the ability of the data to constrain the model.

[9] In the case where $\gamma_1 \neq \gamma_2$, the theoretical TvC functions of the PTM can be numerically derived using iterative methods (see Lu & Dosher, 1998).

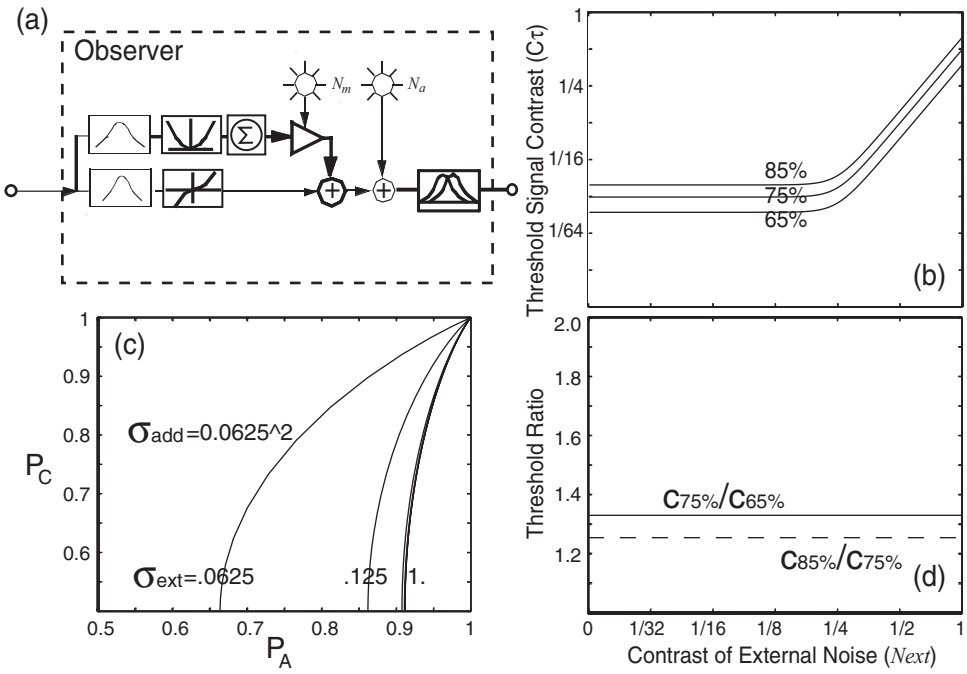


Figure 15. a: The perceptual template model of a human observer. b: Threshold contrasts required to maintain performance at 65%, 75%, and 85% correct (corresponding to $d' = 0.54$, 0.95, and 1.47) as functions of the contrast (standard deviation) of external noise. c: Probability correct (P_C) versus probability consistent (P_A) functions for a range of external noise levels. d: Threshold ratios between different performance criterion levels: 75% and 65% correct (solid line) and 85% and 75% correct (dashed line).

$$P_A = \int_{-\infty}^{+\infty} g(x - (\beta c)^{\gamma_1}, 0, \sqrt{2}\sigma_{ext}^{\gamma_1}) \times \{G^2(x, 0, \sqrt{N_{mul}^2[2\sigma_{ext}^{2\gamma_2} + (\beta_2 c)^{2\gamma_2}] + 2\sigma_{add}^2}) + [1 - G(x, 0, \sqrt{N_{mul}^2[2\sigma_{ext}^{2\gamma_2} + (\beta_2 c)^{2\gamma_2}] + 2\sigma_{add}^2})]^2\} dx. \quad (15)$$

We illustrate P_C as functions of P_A in a range of external noise conditions for the PTM described above in Figure 15c. Similar to the INM and INUM, as σ_{ext} increases, the total amount of internal noise is increasingly dominated by multiplicative noise, and the P_C versus P_A functions collapse into a single curve.

In all applications of the PTM to empirical data so far, we have found that the PTM with $\gamma = \gamma_1 = \gamma_2$ provides an excellent description of the empirical data. In the rest of this article, we restrict our discussion to this simplified set of PTMs. The same logic could be followed to understand the properties of PTMs with $\gamma_1 \neq \gamma_2$.

It follows directly from Equation 14 that, for any given external noise contrast $\forall N_{ext}$, the threshold signal contrast ratio between two performance criterion levels (corresponding to d_2' and d_1'), is

$$\frac{c_{\tau_2}}{c_{\tau_1}} = \left[\frac{d_2'^2}{d_1'^2} \frac{\beta^{2\gamma} - N_{mul}^2 \beta_2^{2\gamma} d_1'^2/2}{\beta^{2\gamma} - N_{mul}^2 \beta_2^{2\gamma} d_2'^2/2}\right]^{\frac{1}{2\gamma}}. \quad (16)$$

Thus, the PTM predicts that the threshold signal contrast ratio between two performance criterion levels for any given external noise contrast is a nonlinear function of the corresponding d's, independent of the particular external noise level (see Figure 15d).

Summary

In this section, we have derived theoretical predictions of the five most prominent observer models for a set of behavioral tests by translating them into a common formalism. The predictions of the models for three behavioral tests, TvC functions, threshold ratios, and double-pass agreement, are summarized in Table 3. The models have a number of qualities in common but several critical qualities that differ. All five models predict the same general shape for TvC functions: a relatively flat portion when internal noise dominates external noise and a rising portion when external noise dominates internal noise. In addition, all five models predict that the threshold ratio between any two performance criterion levels in a given external noise condition is invariant across different external noise levels. The models differ qualitatively in two major ways: (a) Models without decision uncertainty or a nonlinear transducer, that is, the LAM and INM, predict that the threshold ratio between two performance criterion levels in a given external noise condition is equal to the ratio of the corresponding d's, while models with decision uncertainty or a nonlinear transducer, that is, the LAUM, INUM, and PTM, predict that the threshold ratio between two performance criterion levels in a given external noise condition is a nonlinear function of the corresponding d's. (b) Models with induced noise or multiplicative noise, that is, the INM, INUM, and PTM, predict that as external noise contrast

Table 3
Summary of Model Properties

Dependent measure	LAM	LAUM	INM	INUM	PTM
d'	$\frac{\beta c}{\sqrt{\sigma_{ext}^2 + \sigma_{add}^2}}$	$\frac{\beta c}{\sqrt{\sigma_{ext}^2 + \sigma_{add}^2}}$	$\frac{\beta c}{\sqrt{(1+N_{mul}^2)\sigma_{ext}^2 + \sigma_{add}^2}}$	$\frac{\beta c}{\sqrt{(1 + N_{mul}^2)\sigma_{ext}^2 + \sigma_{add}^2}}$	$\frac{(\beta c)^{\gamma_1}}{\sqrt{\sigma_{ext}^{2\gamma_1} + N_{mul}^2\left[\sigma_{ext}^{2\gamma_2} + \frac{(\beta_2 c)^{2\gamma_2}}{2}\right] + \sigma_{add}^2}}$
d'_{obs}	d'	$f(d', U)$	d'	$f(d', U)$	d'
c_τ	$\frac{d'}{\beta}\sqrt{\sigma_{ext}^2 + \sigma_{add}^2}$	$\frac{d'}{\beta}\sqrt{\sigma_{ext}^2 + \sigma_{add}^2}$	$\frac{d'}{\beta}\sqrt{(1 + N_{mul}^2)\sigma_{ext}^2 + \sigma_{add}^2}$	$\frac{d'}{\beta}\sqrt{(1 + N_{mul}^2)\sigma_{ext}^2 + \sigma_{add}^2}$	$\left\{\frac{d'^2[(1 + N_{mul}^2)\sigma_{ext}^{2\gamma} + \sigma_{add}^2]}{\beta^{2\gamma} - N_{mul}^2\beta_2^{2\gamma}d'^2/2}\right\}^{\frac{1}{2\gamma}}$
P_C	$\int_{-\infty}^{+\infty} g(x - d', 0, 1)G(x, 0, 1)dx$	$\int_{-\infty}^{+\infty}[g(x - d', 0, 1)G^{2U+1}(x, 0, 1) + Ug(x, 0, 1)G^{2U}(x, 0, 1) G(x - d', 0, 1)]dx$	$\int_{-\infty}^{+\infty} g(x - d', 0, 1)G(x, 0, 1)dx$	$\int_{-\infty}^{+\infty}[g(x - d', 0, 1)G^{2U+1}(x, 0, 1) + Ug(x, 0, 1)G^{2U}(x, 0, 1) G(x - d', 0, 1)]dx$	$\int_{-\infty}^{+\infty} g(x - \beta^{\gamma_1}c^{\gamma_1}, 0, \sqrt{\sigma_{ext}^{2\gamma_1} + N_{mul}^2[\sigma_{ext}^{2\gamma_2} + (\beta_2 c)^{2\gamma_2}] + \sigma_{add}^2)} G(x, 0, \sqrt{\sigma_{ext}^{2\gamma_1} + N_{mul}^2\sigma_{ext}^{2\gamma_2} + \sigma_{add}^2)dx$
P_A	$\int_{-\infty}^{+\infty} g(x - \beta c, 0, \sqrt{2}\sigma_{ext}) \times dx\{G^2(x, 0, \sqrt{2}\sigma_{add}) + [1 - G(x, 0, \sqrt{2}\sigma_{add})]^2\}$	Simulation	$\int_{-\infty}^{+\infty} g(x - \beta c, 0, \sqrt{2}\sigma_{ext})dx \{G^2(x, 0, \sqrt{2(\sigma_{add}^2 + N_{mul}^2\sigma_{ext}^2)} + [1 - G(x, 0, \sqrt{2(\sigma_{add}^2 + N_{mul}^2\sigma_{ext}^2)}]^2\}$	Simulation	$\int_{-\infty}^{+\infty} g(x - (\beta c)^{\gamma_1}, 0, \sqrt{2}\sigma_{ext}^{\gamma_1})dx \times \{G^2(x, 0, \sqrt{N_{mul}^2[2\sigma_{ext}^{2\gamma_2} + (\beta_2 c)^{2\gamma_1}] + 2\sigma_{add}^2}) + [1 - G(x, 0, \sqrt{N_{mul}^2[2\sigma_{ext}^{2\gamma_2} + (\beta_2 c)^{2\gamma_2}] + 2\sigma_{add}^1)]^2\}$
$\frac{c_{\tau_1}}{c_{\tau_2}}$	$\frac{d'_1}{d'_2}$	$\frac{d'_{obs1}}{d'_{obs2}}$	$\frac{d'_1}{d'_2}$	$\frac{d'_{obs1}}{d'_{obs2}}$	$\left[\frac{d'^2_2}{d'^2_1}\frac{\beta^{2\gamma} - N_{mul}^2\beta_2^{2\gamma}d'^2_1/2}{\beta^{2\gamma} - N_{mul}^2\beta_2^{2\gamma}d'^2_2/2}\right]^{\frac{1}{2\gamma}}$

Note. LAM = linear amplifier model; LAUM = linear amplifier model with uncertainty; INM = induced noise model; INUM = induced noise with uncertainty model; PTM = perceptual template model.

increases, the total amount of internal noise is increasingly dominated by multiplicative noise, and therefore, the family of probability correct (P_C) versus probability agreement (P_A) curves for all the external noise conditions in the double-pass procedure collapse into a single curve, with a constant internal-to-external-noise ratio. Models without multiplicative noise, that is, the LAM and LAUM, on the other hand, predict a much greater dispersion of the probability correct versus probability agreement functions.

Empirical Tests of the Observer Models

In this section, we evaluate the theoretical predictions of the five observer models described in the previous section against empirical data. We focus on results from the triple-TvC and double-pass experiments. These data provide evidence about threshold ratios and performance agreement functions. As shown in the previous section, all the models can provide a very good account of TvC functions at a single performance criterion level. The model predictions differ significantly in the triple-TvC or the double-pass consistency tests or in the joint application of both. We first present some critical results in the literature that discriminate these models. We then present a new data set that allowed us to statistically compare these models. This literature review focuses on a strong and interrelated set of data in the domain of visual tasks and visual processes.

Existing Evidence

Triple TvC. By measuring TvC functions at three criterion performance levels, the triple-TvC method allows us to derive two sets of threshold ratios. Originally, we found that, in two-interval forced-choice Gabor detection and 2AFC Gabor orientation identification, the threshold contrast ratios are indeed invariant to external noise level (Lu & Dosher, 1999); Also, the threshold ratios are significantly less than the corresponding d' ratios (Lu & Dosher, 1999). Since the publication of the original studies, we have conducted many more triple-TvC experiments using a wide range of perceptual tasks, including pseudo-character identification in peripheral vision with central cuing (Lu & Dosher, 2000) and peripheral cuing (Lu & Dosher, 2000), Gabor orientation identification in peripheral vision paired with central rapid serial visual presentation character identification (Dosher & Lu, 1999), first-order motion direction discrimination in peripheral vision (Lu et al., 2000), second-order motion direction discrimination in peripheral vision (Lu et al., 2000), and Gabor orientation identification in peripheral vision with both valid and invalid precuing (Dosher & Lu, 2000b). In all these cases, we found that (a) threshold contrast ratios are invariant across external noise levels, (b) the threshold contrast ratios are compressed (less than, and closer to, one another) relative to the corresponding d' ratios, and (c) in all the 2AFC tasks (see Figures 16a, 16b, 16c, and 16d) we have conducted so far, the threshold ratio between 75% correct and 65% correct is around 1.29, and the threshold ratio between 85% and 75% correct is around 1.22. Figure 16 presents a summary of the threshold ratios in some of the published data.

These results suggest that the threshold ratio between two performance criterion levels in a given external noise condition is a compressive nonlinear function of the corresponding d' ratio, invariant to the external noise contrast.[10] This is related to the observation that d' increases as a nonlinear function of signal contrast (Cohn et al., 1974; Foley & Legge, 1981; Leshowitz et al., 1968; Nachmias, 1981; Nachmias & Kocher, 1970; Nachmias & Sansbury, 1974; Stromeyer & Klein, 1974; Tanner, 1961). The important finding here is that the threshold ratios are invariant in many experiments. The invariance may place very strong constrains on the functional form of observer models (Iverson & Pavel, 1981), suggesting that uncertainty and/or a nonlinear transducer function are necessary components of observer models.[11]

Double-pass consistency. In a classic study, Burgess and Colborne (1988) applied the double-pass method to estimate the internal noise (σ_{int}) in a series of pattern discrimination or detection studies in a range of external noise conditions. The studies used two-interval forced-choice sine-wave amplitude discrimination and detection with observer-controlled viewing time. The signal stimuli were sine-wave gratings at two frequencies, 4.6 cycles/degree and 9.2 cycles/degree, in the two experiments. Probability correct versus probability agreement functions were measured over a range of external noise levels. As replotted in Figure 17, Burgess and Colborne found that most of the data points in the P_C versus P_A scatterplot were on or nearly on a single theoretical curve for a given ratio of internal noise and external noise standard deviation, although a wide range of external noise levels were used in the study. As shown in Figure 18, the standard deviation of internal noise increased linearly with that of the external noise σ_{ext} with a constant slope ($\alpha = 0.75 \pm 0.10$), independent of the external noise level once it exceeded a certain level. Burgess and Colborne concluded that internal noise has two components: one constant component (the additive noise in LAM) and the other induced component, with a standard deviation that is directly proportional to that of the external noise: $\sigma_{induced} = N_{mul}\sigma_{ext}$.

The double-pass method has been used to estimate internal noise in both the auditory and the visual modalities. All the studies found that the amplitude of (total) internal noise is proportional to that of external noise. The typical ratio of internal to external noise amplitudes is between 1.0 and 1.4 in auditory tests (Green, 1964; Swets, Shipley, McKey, & Green, 1959), and between 0.65 and 1.00 in visual tests (Burgess & Colborne, 1988; Chung et al., 2005; Gold et al., 1999; Levi & Klein, 2003). As discussed in Green (1964), these estimates are lower bounds of the internal-to-external-noise ratio because they have not explicitly considered

[10] In a contrast discrimination task in external noise, Legge et al. (1987) found that increment contrast threshold approached a linear function of d' in high external noise. The apparent discrepancy from the compressive d' ratios observed in detection and contrast discrimination is due to a task difference. In fact, all five observer models predict a linear threshold and d' relationship in high-external-noise conditions in a contrast discrimination task, consistent with the empirical results in Legge et al.: For a given observer model, we can compute the internal responses to stimuli with contrast $c + \Delta c$ and c (e.g., Equation 1) and their difference, involving essentially a computation of the derivative of the internal response function. Discrimination threshold Δc can be computed for a fixed difference of internal response, which gives rise to a particular performance level.

[11] The constancy of the ratios over drastic changes in accuracy of performance due to attention or learning (Dosher & Lu, 1998, 1999, 2000a; Lu & Dosher, 1999, 2000, 2004) indicates that these threshold ratios are a core property of the system. This suggests that nonlinearity is the more likely account. This point is treated in the discussion.

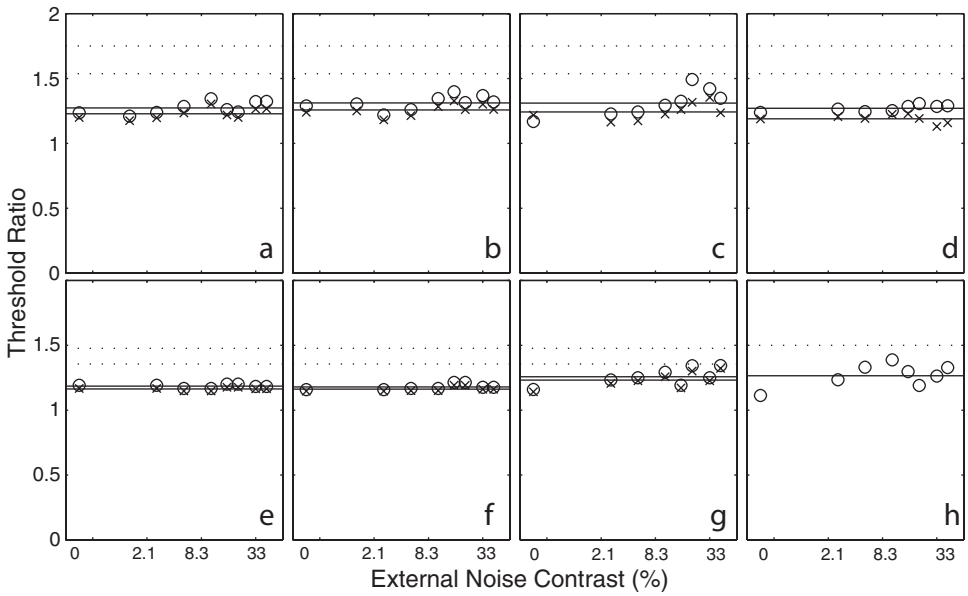


Figure 16. Threshold ratio versus external noise contrast functions from eight published experiments. The dotted lines indicate corresponding d' ratios. The symbols indicate measured threshold ratios. The solid lines indicate the mean of measured thresholds. All the panels except h contain two different symbols, indicating two sets of threshold ratios from triple-threshold-versus-external-noise-contrast (triple-TvC) measurements. Panel h contains data from a double-TvC measurement. All the ratios are geometric means across all the subjects. a: Gabor detection in fovea (Lu & Dosher, 1999). b: Gabor orientation identification in fovea (Lu & Dosher, 1999). c: First-order motion direction discrimination in peripheral vision, averaged across attended and unattended conditions (Lu et al., 2000). d: Second-order motion direction discrimination in peripheral vision, averaged across attended and unattended conditions (Lu et al., 2000). e: Pseudo–character identification in peripheral vision with central cuing, averaged across pre- and simultaneous cuing conditions (Lu & Dosher, 2000). f: Same as e except peripheral cues were used (Lu & Dosher, 2000). g: Gabor orientation identification in peripheral vision, averaged across valid and invalid precuing conditions (Dosher & Lu, 2000b). h: Gabor orientation identification in peripheral vision, averaged across the last 6 training days (Dosher & Lu, 1999).

response bias, which increases the agreement between two passes and therefore reduces the estimated internal-to-external-noise ratio. Two recent studies also showed that the ratio was invariant to significant performance improvements following perceptual learning (Chung et al., 2005; Gold et al., 1999). All these results imply that induced or multiplicative noise (or equivalently, contrast-gain control) is a necessary component of observer models.

Summary. Only two observer models, the INUM and the PTM, are qualitatively consistent with the combined results from the double-pass and triple-TvC procedures because, among the five observer models described in this article, only these two have both induced or multiplicative noise and decision uncertainty or a nonlinear transducer. The simple LAM is inadequate in that it has neither multiplicative noise nor a mechanism to generate nonlinearity in the thresholds. The INM has multiplicative noise but lacks a mechanism to generate nonlinearity. The LAUM uses uncertainty to accommodate nonlinearity but lacks multiplicative noise.

We (Lu & Dosher, 1999) fit the PTM and the INUM to the data from two experiments, one based on a two-interval forced-choice Gabor detection task and the other based on a 2AFC Gabor orientation identification task. In both data sets, full psychometric functions (performance accuracy vs. signal contrast) were collected over a range of external noise conditions. The double-pass procedure was not used in these studies. The analysis found the following: (a) The PTM and the INUM accounted for the data almost equally well. (b) The best fitting INUM, however, had very large *U*s, in the range of 20 to 200. (c) Not only were the estimated *U*s outside of the range considered by Eckstein et al. (1997; maximum $U = 3$), they also seem to be too large to be physiologically plausible. (d) The INUM did not sharply constrain the estimate of U in all but the smallest U values. The availability of the full psychometric functions in these experiments provided data on multiple criterion levels and therefore incorporated the threshold ratio constraints on the observer models. The analysis, consistent with the prior research on the properties of nonlinearity and multiplicative noise (identified separately), supports the models that incorporate both components.

A New Experiment

We conducted a new experiment to directly compare the five observer models using both triple-TvC and double-pass procedures. To our knowledge, this is the first use of data from both methods to jointly constrain and test observer models. Observers judged the orientation ($\pm 45°$) of a single Gabor patch in fovea. Full psychometric functions, measuring identification accuracy as a function of signal stimulus contrast, were collected, using the method of constant stimuli over a range of external noise condi-

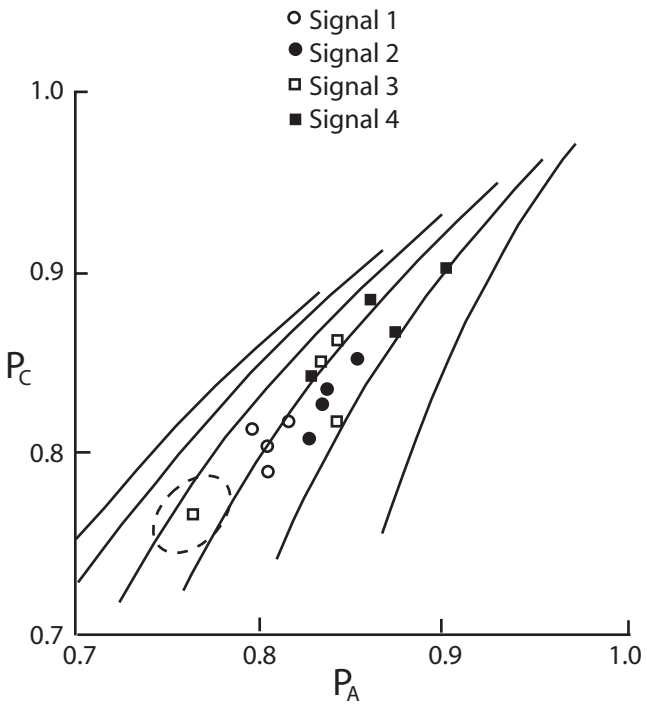


Figure 17. Covariation of the percentage of correct responses (P_C) and the percentage of agreement (P_A) of decisions made on two passes through a set of images. The solid lines are loci for observers with various ratios of internal noise/external standard deviation. Extremeness observers with no internal noise and overwhelmingly large internal noise would have P_A values of 1.0 and 0.5, respectively, for the two passes through the image set. The dotted ellipse around one data point represents the one-standard-deviation region. Replotted from "Visual Signal Detection: IV. Observer Inconsistency," by A. E. Burgess and B. Colborne, 1988, p. 620, Figure 2. Copyright 1988 by the Optical Society of America. Adapted with permission.

tions in a two-alternative forced-identification task. The entire trial–stimulus sequence was repeated using the double-pass procedure. We fit eight observer models to the full data set and statistically compared their ability to account for the data (see Appendix F for detailed experimental methods), including the five described in the previous section and three new models: a reduced PTM with $\beta_2 = \beta$ (rPTM), an altered PTM with a linear transducer and decision uncertainty (uPTM), and a fully saturated model that consisted of a PTM with decision uncertainty (fullM). The additional models were included to complete a model lattice and to test variants of the real contenders for a fully articulated observer model.

P_C versus P_A functions. Probability correct (P_C) is graphed as a function of probability agreement (P_A) for a range of external noise levels for three observers in Figure 19. The best fitting parameters of the eight observer models are listed in Table 4. The fullM, which combines a PTM and decision uncertainty, is the most saturated model. This includes all the major components of the observer models, including nonlinearity and decision uncertainty, as well as both forms of internal noise. χ^2 statistics (see Equation F3 in Appendix F) were used to compare quality of the fits of the seven reduced models with that of the most saturated model. The chi-squares and the corresponding degrees of freedom are listed in the last two columns of Table 4.

For all three observers, only the PTM provided equivalent fits to the data ($p \approx 1.0$) in comparison to the fullM. All the other models provided significantly inferior fits to the data ($p < 0.000001$). These include uPTM, which replaced the transducer function in the PTM with decision uncertainty, as well as the INUM, which had previously been shown to be equivalent to the PTM in accounting for multiple TvC functions (Lu & Dosher, 1999). The joint constraints of multiple criterion levels and the agreement data were, however, a challenge for the INUM. Rejecting the uPTM suggests that the inadequacy of the INUM is not solely due to the difference between the induced noise in the INUM and the multiplicative noise in the PTM. The predictions of the best fitting PTMs are plotted in Figure 19 as smooth curves. The slight misfits in the two lowest external noise conditions for observer SJ are due to effects of bias in his responses, which tended to increase P_A. The issue of

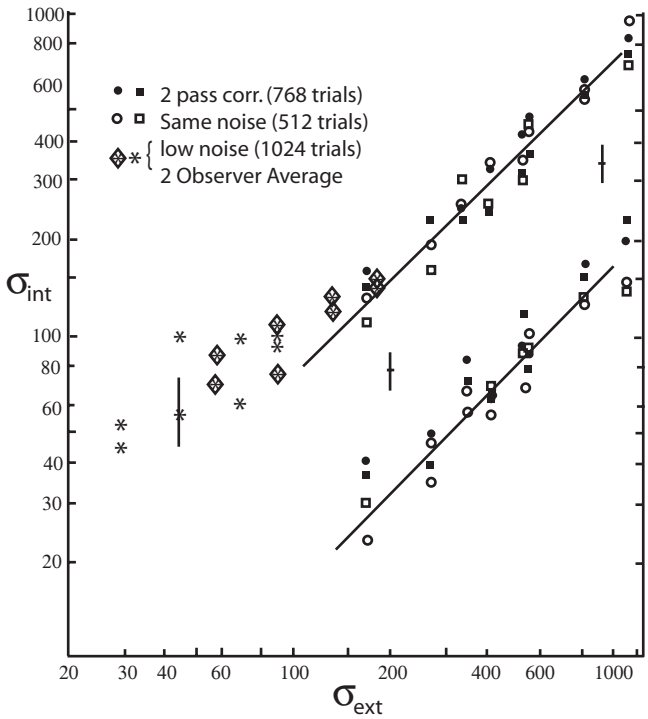


Figure 18. Observer internal noise standard deviations, σ_{int}, as a function of external (image) noise standard deviation, σ_{ext}, in relative units. The two-alternative forced-choice (2AFC) detection threshold ($d' = 1$) for the external noise is about 60 on this scale. The data are for a 2AFC amplitude discrimination of sine waves (4.6 and 9.2 cycles/degree). The circles represent data for noise fields equal to the signal size, and the squares stand for noise fields twice as large as the signal. The error bars represent ± 1 standard deviation. Two different measurement techniques have been used: two passes through stored image sets and 2AFC trials with identical noise fields. The points labeled low noise were done with low image-noise levels and hence have poor accuracy, even after averaging over two observers. The results for the two observers agree with experimental error. Replotted from "Visual Signal Detection: IV. Observer Inconsistency," by A. E. Burgess and B. Colborne, 1988, p. 621, Figure 3. Copyright 1988 by the Optical Society of America. Adapted with permission.

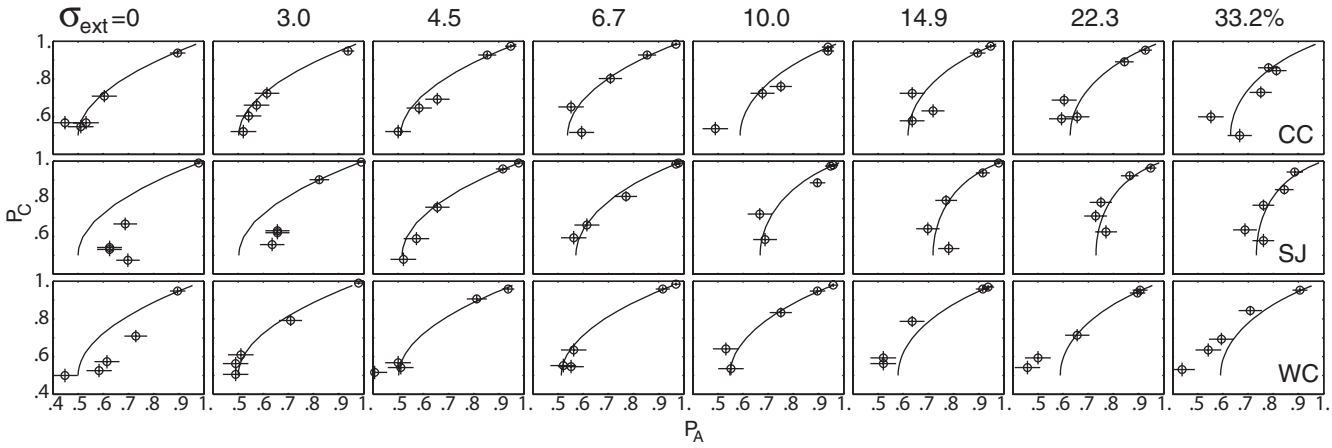


Figure 19. Probability correct (P_C) versus probability agreement (P_A) function for a range of external noise levels for three observers (CC, SJ, and WC). The smooth curves depict predictions of the best fitting perceptual template model. Error bars indicate one standard deviation of the corresponding probability.

bias in double-pass procedures was discussed in Green (1964) and treated in Lu and Dosher (2007).

To summarize the quality of the joint fit, we computed r^2s for the best fitting PTM separately for probability correct and probability agreement. For the three observers, r^2_C = 0.9364, 0.9599, 0.9753, and r^2_A = 0.9439, 0.9165, 0.9459. The PTM and rPTM (with $\beta_2 = \beta$) provided good accounts of probability correct as a function of contrast and external noise, with equivalent or nearly equivalent fits and parameter estimates. This explains why we had not previously required the full PTM with different β and β_2 (e.g., Lu & Dosher, 1999) when we considered only multiple-TvC functions without the double-pass procedure. The double-pass procedure, combined with the multiple-TvC data, can be used to provide stronger constraints that refine the exact form of multiplicative noise. That $\beta \neq \beta_2$ in the best fitting model implies that the perceptual template that controls the magnitude of multiplicative noise (or, equivalently, the relative quantity for divisive gain control) is different from the perceptual template in the signal

Table 4
Best Fitting Parameters of the Eight Observer Models

Obs	Model	*K*	N_{mul}	σ_{add}	β	β_2	γ	*U*	r^2_C	r^2_A	χ^2	*df*
CC	LAM	2	0.000	0.1188	2.020	0.000	1.000	0	0.8308	0.7463	380.87	4
	LAUM	3	0.000	0.0885	2.839	0.000	1.000	8	0.9127	0.8297	190.80	3
	INM	3	1.356	0.1368	2.942	0.000	1.000	0	0.8389	0.9284	128.52	3
	INUM	4	1.183	0.1269	4.025	0.000	1.000	4	0.9168	0.9280	62.69	2
	rPTM	4	0.552	0.0043	1.634	1.634	2.277	0	0.9321	0.8587	90.99	2
	PTM	5	1.228	0.0095	1.869	1.161	2.051	0	0.9364	0.9439	0.00	1
	uPTM	5	1.050	0.1197	4.031	0.640	1.000	5	0.9175	0.9317	49.65	1
	fullM	6	1.228	0.0095	1.869	1.161	2.051	0	0.9364	0.9439		
SJ	LAM	2	0.000	0.0929	2.167	0.000	1.000	0	0.8544	0.8238	318.55	4
	LAUM	3	0.000	0.0754	3.256	0.000	1.000	8	0.9393	0.8737	147.71	3
	INM	3	0.751	0.0942	2.520	0.000	1.000	0	0.8374	0.9187	184.94	3
	INUM	4	0.598	0.0873	3.406	0.000	1.000	4	0.9270	0.9040	104.35	2
	rPTM	4	0.456	0.0027	1.730	1.730	2.377	0	0.9605	0.8934	24.90	2
	PTM	5	0.704	0.0038	1.791	1.431	2.269	0	0.9599	0.9165	0.00	1
	uPTM	5	0.018	0.0807	3.604	50.441	1.000	7	0.9246	0.8967	96.57	1
	fullM	6	0.704	0.0038	1.791	1.431	2.269	0	0.9599	0.9165		
WC	LAM	2	0.000	0.1728	1.726	0.000	1.000	0	0.8876	0.7278	495.71	4
	LAUM	3	0.000	0.1234	2.335	0.000	1.000	8	0.9513	0.7928	377.23	3
	INM	3	1.817	0.1729	2.308	0.000	1.000	0	0.8995	0.9203	172.60	3
	INUM	4	1.502	0.1687	2.982	0.000	1.000	2	0.9442	0.9363	123.34	2
	rPTM	4	0.572	0.0052	1.320	1.320	2.451	0	0.9735	0.7833	215.16	2
	PTM	5	1.600	0.0078	1.485	0.908	2.357	0	0.9753	0.9459	0.16	1
	uPTM	5	0.694	0.1135	2.925	1.138	1.000	5	0.8977	0.8896	56.94	1
	fullM	6	1.482	0.0071	1.669	1.089	2.356	2	0.9680	0.9529		

Note. *K* refers to the number of parameters in the model. Obs = observer; LAM = linear amplifier model; LAUM = linear amplifier model with uncertainty; INM = induced noise model; INUM = induced noise with uncertainty model; rPTM = reduced perceptual template model; PTM = perceptual template model; uPTM = altered perceptual template model; fullM = fully saturated model.

pathway. The contributions to the contrast-gain pool are generally thought to be rather broadly tuned (Cannon & Fullenkamp, 1991; Solomon, Sperling, & Chubb, 1993).

TvC functions. The TvC functions of the three observers are shown in Figure 20, along with the ratios between thresholds at 85% and 75% correct and at 75% and 65% correct. The smooth curves in the upper row of Figure 20 depict the predictions of the best fitting PTMs to the probability correct (P_C) versus probability agreement (P_A) functions. The curves are not generated from independent fits to the TvC functions. As expected from the goodness of fit to the psychometric functions (which correspond to more than three accuracy levels), the PTM provided excellent fits to the TvC functions, accounting for 98.44%, 96.31%, and 97.70% of the variance for the three observers, respectively.

We also computed the ratios between thresholds at 75% and 65% correct and at 85% and 75% correct. The ratios are 1.34 ± 0.05, 1.33 ± 0.05, and 1.33 ± 0.06 between thresholds at 75% and 65% correct and 1.29 ± 0.04, 1.28 ± 0.04, and 1.28 ± 0.5 between thresholds at 85% and 75% correct. As can be seen in the lower row of Figure 20, the ratios are virtually the same across all the external noise levels. These values are also generally consistent with those observed in the many previous studies that have measured these ratios (see Figure 18).

Summary and discussion. The triple-TvC method emphasizes the range of external noise and performance levels. The double-pass method provides measures of the total amount of internal noise in each signal and external noise condition. Taken together, TvC functions across a range of performance levels (or equivalently, psychometric functions in a range of external noise conditions) along with the measure of total internal noise in all the stimulus conditions from the double-pass method jointly provide very strong constraints on observer models. Somewhat to our surprise, the INUM and the rPTM provided inferior fits to the data compared with the PTM, although we had previously shown that INUM and rPTM are statistically equivalent in fitting multiple-TvC functions alone. Decision uncertainty and transducer-based models are not fully equivalent when both multiple-TvC and double-pass agreement are jointly considered. The transducer function provides a better fit to the relationship between the many P_C versus P_A functions across different external noise conditions.

Importantly, it is necessary to apply the double-pass procedure in a wide range of external noise and performance levels to constrain the observer models. Applying the double-pass procedure in a narrow range of signal and/or external noise conditions, as is commonly done, is not sufficient to provide strong constrains on observer models. To obtain strong constraints on models, it is necessary to repeat triple-TvC measurements—or to jointly vary external noise and signal contrast across good ranges—using the double-pass procedure.

Discussion

Findings. The empirical study reported here used signal contrasts to span a wide range of performance across many levels of external noise, together with the double-pass test. From the joint constraints of this experiment, we have identified the transducer-

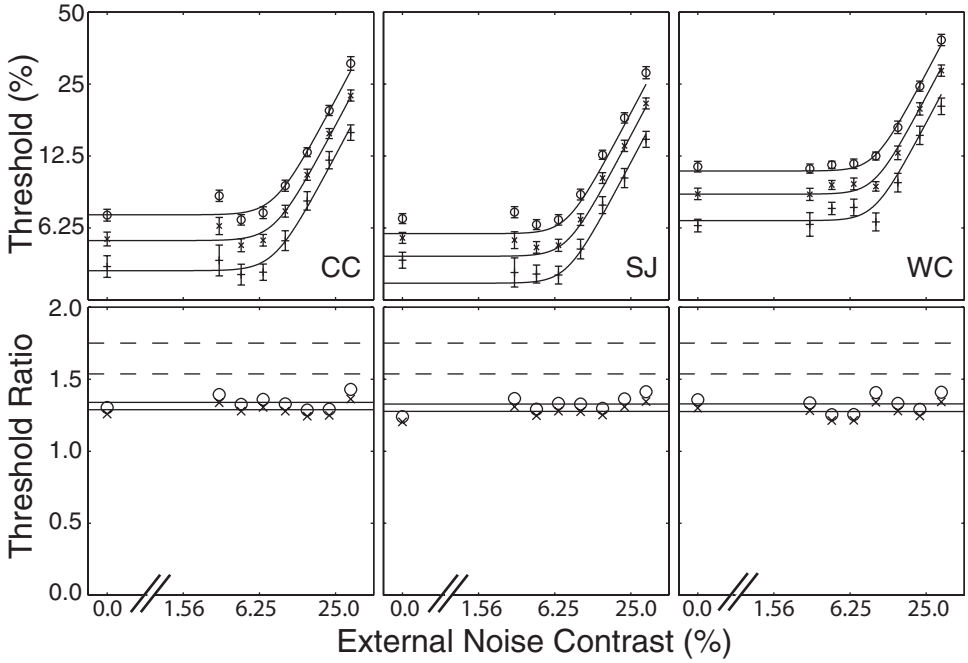


Figure 20. Upper row: Threshold-versus-external-noise-contrast functions at 65% (–s), 75% (×s), and 85% (circles) correct for three observers (CC, SJ, and WC). The smooth curves depict the predictions of the best fitting perceptual template models resulted from fitting the probability correct (P_C) versus probability agreement (P_A) functions. Lower row: Ratios of thresholds between 75% and 65% correct (circles) and 85% and 75% correct (×s). The dashed lines indicate the corresponding d' ratios. The solid lines represent the mean ratios across different external noise conditions.

based PTM as providing significantly better accounts of the data than the decision-uncertainty-based models. At the same time, adding decision uncertainty to the PTM does not improve the model fits. Although the INUM and the PTM are both able to account for the triple TvC (or, equivalently, full psychometric functions at a range of external noise levels), theoretical reasons of consistency with the effects in pattern masking lead us to prefer the nonlinearity/PTM form. The experiment adds direct evidence from converging methods to show that the nonlinear transducer does, in fact, provide a better account of the observed regularities. The uncertainty explanations are ruled out due to the overprediction of separation between agreement functions for different external noise levels, while the data and the PTM predict a convergence of agreement functions for higher external noise levels.

Additional theoretical issues bearing on the selection of the nonlinear form of the PTM are considered below.

Birdsall's theorem. According to Birdsall's theorem, when the variance of any source of noise prior to the nonlinear transducer is large enough so that other sources of noise in the experiment can be neglected, the resulting d' psychometric function will be linear (Lasley & Cohn, 1981). Because d' psychometric functions are generally nonlinear in high external noise, the theorem has been used to reject nonlinear transducer models in perception (Lasley & Cohn, 1981). However, Birdsall's theorem is based on three assumptions: (a) The transducer function is monotonic, (b) sources of noise other than the external noise can be neglected, and (c) observers try their best to maximize their performance accuracy. The second assumption is clearly violated in all experiments that are impacted by a significant amount of induced/multiplicative noise, as in Burgess and Colborne (1988), as well as in the experiment reported in this section. For many experimental situations with noticeable multiplicative noise, the theorem does not apply; for experimental situations with negligible multiplicative and additive noises, Birdsall's theorem may apply. We suggest that the former situation is more common in typical detection and discrimination tasks in the perceptual domain.

Nonlinear transducer. Nonlinear transducer function is one key component of many successful observer models in visual tasks. They are widely used in modeling pattern masking (Burbeck & Kelly, 1981; Burgess & Colborne, 1988; Foley, 1994; Klein & Levi, 1985; Legge & Foley, 1980; G. Sperling, 1989; Stromeyer & Klein, 1974; Watson & Solomon, 1997). The concept of a nonlinear transducer function is consistent with nonlinear properties of visual neurons (Albrecht & Geisler, 1991; Albrecht & Hamilton, 1982; Bonds, 1991; Derrington & Lennie, 1981; Heeger, 1993; Kaplan & Shapley, 1982; Ohzawa et al., 1982; Sclar et al., 1990). In stimulus identification by well-practiced observers, previous evidence suggests that stimulus uncertainty does not appear to play a major role. In a perceptual learning experiment studying Gabor orientation identification in peripheral vision, we (Dosher & Lu, 1999) demonstrated that the threshold ratio between two d' levels at all the external noise levels for each observer was constant across days, even though the thresholds themselves were improved by a factor of almost three. In the LAUM and INUM, the threshold ratio between two d' levels is a function of the number of hidden detectors (or degree of decision uncertainty). The result 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.

On the other hand, nonlinearity resulting from uncertainty has not shown strong constancies, and the degree of estimated uncertainty can be highly variable and large. For example, to account for their data, Eckstein et al. (1997) had to vary the degree of uncertainty for different external noise levels for the same observer in the same experiment in nonsystematic ways. We (Lu & Dosher, 1999) also showed that the INUM fits are essentially equivalent in a wide range of values of uncertainty, with 20 to 200 hidden detectors in the best fitting model. Current models of early visual system specify fewer visual channels.

More on uncertainty. Abbey and Eckstein (2006) considered early (pretemplate) and late (posttemplate) nonlinearities, as well as intrinsic target location uncertainty, in explaining the differences between classification images they obtained using stimuli with the same relevant spatial profile in detection, contrast discrimination, and identification tasks. Classification images estimate the relevant spatial features in display images by averaging the external noise samples from trials with a specific response since external noise with contrast patches consistent with the signal template will contribute to these responses. Abbey and Eckstein found that none of the models they considered fully explained the observed data and suggested a need for further investigations of the combined effects of these and other forms of nonlinearities on classification images.

The developments in this article, following most of the prior literature (e.g., Eckstein et al., 1997), assume stochastically independent detectors. Recently, uncertainty models that assume detectors with correlated responses have been developed (Abbey & Eckstein, 2006; Manjeshwar & Wilson, 2001; Zhang, Pham, & Eckstein, 2006). Perhaps future developments that incorporate uncertainty with correlated detectors might improve the uncertainty-based observer models.

Additional Theoretical Notes on the Perceptual Template Model

In this section, we summarize some additional technical and theoretical considerations of the PTM.

Equivalence Between Multiplicative Noise and Contrast-Gain Control

Multiplicative noise has been shown to be a necessary component in observer models both in the current analyses and in related earlier applications. Dao et al. (2006) showed that for TvC functions, the multiplicative noise formulation of the PTM is mathematically equivalent to a contrast-gain control formulation of the PTM. Although most of the existing psychophysical data do not distinguish these model forms, data in neurophysiology seem to favor the contrast-gain control form. Reformulated as a contrast-gain control model, the PTM is also completely consistent with the notion of a constant noise after nonlinear transduction of the input signal (Gorea & Sagi, 2001; Katkov, Tsodykd, & Sagi, 2006).

It is important to note that the mathematical equivalency between multiplicative noise and contrast-gain control versions of the PTM depends critically on the functional form of the multiplicative noise. In the PTM, the magnitude of multiplicative noise

is determined by the total amount of contrast energy in the input stimuli, including both signal and external noise. Including a signal contribution to the magnitude of multiplicative noise is necessary to make the model equivalent to contrast-gain control models and exhibit Weber's law-like behavior. In the INM and INUM, the amplitude of the induced noise is only related to external noise contrast. These two models cannot be reformulated into a contrast-gain control form and are not consistent with Weber's law.

In the uPTM, we replaced the induced noise in the INUM with the form of multiplicative noise used in the PTM. That the uPTM provides inferior fits to the data compared to the PTM indicates that the lack of a nonlinear transducer, rather than the form of induced noise, is the source of the problem for the INUM.

Relationship to Models in Pattern Masking

There has been a significant parallel development of observer models in visual pattern masking (Foley, 1994; Foley & Legge, 1981; Fredericksen & Hess, 1997; Gorea & Sagi, 2001; Klein & Levi, 1985; Kontsevich et al., 2002; Legge & Foley, 1980; Nachmias & Sansbury, 1974; Watson & Solomon, 1997). In pattern masking studies, instead of external noise, pattern masks (e.g., sine waves of the same or different frequencies, orientations, etc.) are used to probe the properties of the visual system. Pattern masking models usually describe the internal response as a function of the target and mask patterns (Foley, 1994):

$$R = \frac{\max\left(0, \sum_{ij} c_{ij} S_{\mathrm{E}ij}\right)^p}{\sum_j \left(\sum_i c_{ij} S_{\mathrm{I}ij}\right)^q + Z}, \tag{17}$$

where i and j index orientation and spatial frequency, c_{ij} denotes the contrast of grating ij, $S_{\mathrm{E}ij}$ and $S_{\mathrm{I}ij}$ denote the excitatory and inhibitory sensitivities of the pattern detector to grating ij, p and q are exponents of the nonlinear transducer functions in the excitatory and inhibitory pathways, and Z is a constant that is stimulus independent. In this model, the inhibitory terms corresponding to the same orientation i are summed prior to being raised to power q.

Although they are developed in rather different experimental domains with different focus on the properties of the visual system (nonlinearity vs. internal noise), the functional forms of the contrast-gain control formulation of the PTM and the pattern masking models are very similar (Dao et al., 2006). They should be consistent because they both describe the same visual system. The other four observer models are not consistent with the pattern masking models. The compatibility of the PTM and the pattern masking models lends further support to the PTM.

Class of Equivalent Models

The PTM in Figure 14 shows additive noise following multiplicative noise and nonlinearities. Yet some forms of noise considered by earlier investigators, such as photon noise or sampling noise (de Vries, 1943; Pelli, 1981; Rose, 1948), occur early in the visual system, possibly preceding the perceptual template or filter. Indeed, there are three locations in the model where additive noise might be introduced (see Figure 21): (a) prior to the perceptual template, (b) after the perceptual template but before nonlinearity and multiplicative noise, or (c) after nonlinearity and multiplicative noise. Although additive noise in these locations may be related to distinct physiological processes, we (Dosher & Lu, 1999) showed that a complex model with additive noises in all three locations, or any model with noise in any one or two of the locations, can be reexpressed in terms of a model with a single additive noise after multiplicative noise for the purpose of modeling behavioral choice. Conversely, any model with additive noise after multiplicative noise can also be reexpressed as a model with noise in all three locations, although there is no unique solution. The practical consequence is that it is impossible to rule out the model where all the additive noise sources are after the multiplicative noise (Location 3) and it is not possible on the basis of whole-system behavior to uniquely partition additive noise into these three sources, although certain patterns of condition differences may place constraints on the partition (e.g., Pelli, 1991). For example, a number of attention effects have been shown to be isolated to situations of high external noise and hence reflect changes in sensitivity to external noise due to attention; the fact that attention has no effect in the absence of external noise rules out the existence of significant amounts of internal noise prior to the template, which in turn restricts the plausible form to internal additive noise that occurs following the template (Dosher & Lu, 2000a).

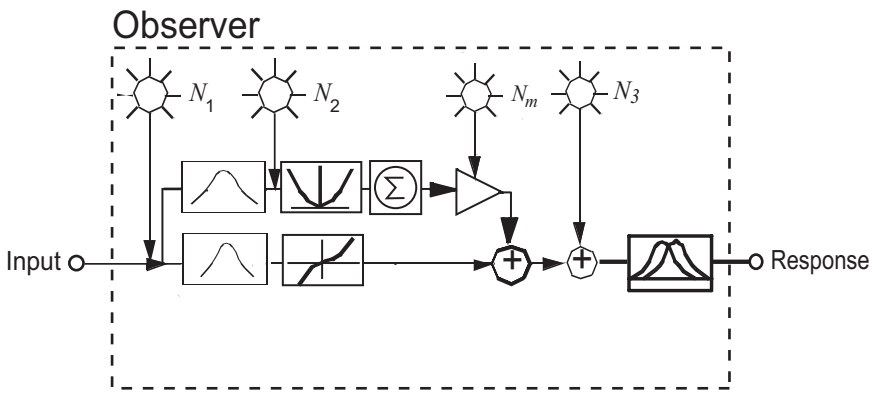


Figure 21. A perceptual template model with additive noise at three potential locations.

Cross Terms and Stochastic Simulations

All the observer models considered in this article are analytical simplifications of the more natural stochastic observer models. In approximating the stochastic models with the analytic models, we have made two simplifications: (a) using the expectations of the random variables in place of the random variables and (b) ignoring all the cross products. We (Dosher & Lu, 2000a) showed that the stochastic PTM exhibits all the key characteristics derived for the (analytic) PTM. In general, the analytic PTM is a close approximation to the stochastic PTM and provides a good approach to model testing: The (analytic) PTM fits all the data we have collected very well. In the special case where $\gamma = 1.0$, the (analytic) PTM is identical to the stochastic PTM. In the two extreme regions of the external noise manipulation, that is, when internal additive noise dominates or when external noise dominates, the (analytic) PTM approaches the stochastic model asymptotically (Dosher & Lu, 2000a). In the PTM development, we have assumed that the internal response distributions in the decision stage are Gaussian. If the noises are Gaussian, the assumption will be a good approximation for the range of nonlinearities ($\gamma = 2 \pm 1$) that we have encountered in all the empirical studies so far, so long as we restrict ourselves to performance ranges less than 95% correct.

Characterizing Perceptual Templates

In presenting the external noise methods, we have focused on the use of white external noise. The primary manipulations are the magnitude of the external noise and the contrast of the signal stimulus. A number of related techniques have been developed that involve manipulations of the characteristics of the external noise (e.g., spatial or temporal frequency, spatial or temporal extent, orientation, etc.) to estimate the corresponding characteristics, or sensitivity, of the perceptual template. For example, the profile of spatial frequency sensitivity of the perceptual template has been measured for visual tasks through the use of external noise systematically varying in its band-pass characteristics (Lu & Dosher, 2001; Solomon & Pelli, 1994; Talgar et al., 2004). The classification image method has also been widely used to infer the spatial form of the perceptual templates, for example, those pixels of a visual display that are the most influential in the selection of a response (Ahumada, 2002; Ahumada & Lovell, 1971; Eckstein, Shimozaki, & Abbey, 2002). Because all these methods depend on the use of observer models, a better understanding of the observer models is essential for the accurate and valid application of these methods. The results of this review suggest that some techniques based on the LAM need important modifications.

Extending Observer Models to Overlapping Perceptual Templates

All the observer models presented here have been formulated for experimental situations where any single signal stimulus plausibly activates only one perceptual template (e.g., Gabors of $\pm 45°$), that is, the gain of the template to the matched signal stimulus is β and to the nonmatched signal stimulus is 0. The observer models must be extended to handle cases in which close, more similar, to-be-discriminated stimuli that may activate more than one perceptual template (e.g., Gabors of $\pm 3°$) are tested. In two-alternative identification tasks, for example, the observer must identify a stimulus as one of two targets. Two templates are involved. The gain of the matching template with the stronger match to the stimulus is β, and the gain of the other template to the same, nonmatching stimulus is β'. The observer models developed here can be extended to situations where similar targets must be discriminated by considering overlapping templates. For signal stimuli that are quite distinct, $\beta' \approx 0$, and the response to external noise will be approximately independent, as assumed in the development here. For cases where overlap is significant, $\beta' > 0$, and correlated responses to external noise must be considered. The extended observer models can therefore provide an integrated framework within which to understand the performance limitations of the observer in two fundamental measurement regimes: contrast thresholds holding stimulus differences constant and/or feature thresholds holding contrast constant (Jeon, Lu, & Dosher, 2006).

General Discussion

In this article, we have reviewed three major external noise methods and five observer models. The work has translated previous theoretical analyses into a common formalism and derived systematic model predictions for a common set of behavioral tests based on the three major external noise methods. The theoretical development has enabled us to conduct a comprehensive evaluation of the existing observer models against empirical data.

We have found that five component processes, a perceptual template, a nonlinear transducer, both additive and multiplicative noises, and a decision stage, are necessary to simultaneously account for all the data from the three major external noise methods. As implemented in the PTM, these component processes map the physical properties of the input stimulus into internal perceptual representations, providing the necessary internal response distributions for the decision stage (see Figure 2). Although estimated from empirical results based on particular tasks and input stimuli, the components reflect the intrinsic characteristics and limitations of the perceptual system and are independent of the characteristics of the stimuli. Therefore, once all its components are specified, the observer model provides a principled way to predict observer performance in a range of related tasks from the results in a particular experiment.

These findings have major implications for the application of external noise methods in the literature. Often, these studies examine only single-TvC curves for assayed conditions. This in turn leads to the use of the LAM as a description. However, we now know that, as soon as multiple criteria or the slope of the psychometric function or the double-pass method is used, the LAM will be inconsistent with the larger ranging data set. This means that measurement of single TvCs and use of the LAM will fail in generalizing predictions to almost any other condition. Because the parameters of the LAM depend on the particular criterion performance level of the TvC function, the conclusions are at best performance-level dependent and at worst misleading.

One important conclusion of the current review is that it is necessary to measure TvC functions at multiple performance levels with double-pass agreement to more fully constrain observer models. Previous studies that measured only a single-TvC function at one criterion performance level with and without measures of

the double-pass agreement over a limited range of external noise conditions provide insufficient constrains on the observer models (e.g., Gold et al., 1999). If, however, a full analysis has specified the PTM as the correct model in a particular domain, then the use of triple-TvC or full psychometric functions over a range of external noise conditions will be sufficient to constrain the most important aspects of the model.

Key Findings in the Visual Domain

The empirical tests of the five observer models reported here focus on simple visual detection or discrimination tasks. Widely used in the study of human visual perceptual processes, these methods are important tools for quantifying, describing, and testing observer performance. Several quite consistent observations have been made over a range of tasks and studies in the visual domain: (a) The ratio of contrast thresholds at two performance criterion levels in a given external noise condition is invariant to the external noise condition; (b) in all the 2AFC tasks we have conducted so far, the threshold ratio between 75% correct and 65% correct is around 1.30, and the threshold ratio between 85% and 75% correct is around 1.23, which is equivalent to a transducer nonlinearity close to a power function, consistent with most of the observations in the visual domain; and (c) applications of the double-pass method in the visual domain have found that the amplitude ratio of internal and external noise is in the range of 0.65 to 1.00, suggesting 50% to 70% reliability in visual performance.

If observed more widely, these relatively consistent results across tasks and studies may shed light on some fundamental properties of the visual system. They may also place very strong constraints on the functional form of, as well as the type and degree of nonlinearity in, observer models (Iverson & Pavel, 1981). The internal-to-external-noise amplitude ratio also provides an upper bound on the desired level of performance of any model that attempts to simulate trial-to-trial behavior of human observers: With an internal-to-external-noise amplitude ratio of 1.0, an observer performs at a 70% consistency level, and therefore, the best a model can do is to make correct trial-by-trial predictions 70% of the time.

Extensions to Other Perceptual Domains

The external noise methods and the observer model approach have obvious extensions to the study of auditory processes. Indeed, the use of external noise or noise masking manipulations is widespread in auditory studies (e.g., Ahumada & Lovell, 1971; Bos & Deboer, 1966; Eijkman et al., 1966; Hartmann & Pumplin, 1988; Humes & Jesteadt, 1989; Moore, 1975; Osman, 1971; Richards et al., 1991). The double-pass method was originated in that context, where the internal-to-external-noise amplitude ratio was found to be between 1.0 and 1.4 (Green, 1964; Swets et al., 1959). Observer models in the auditory domain have been largely focused on theoretical questions other than the roles of nonlinearity, threshold ratios, and multiplicative noise, such as the nature and interaction of feature banks in the auditory system (Dau, Kollmeier, & Kohlrausch, 1997). However, extending the theoretical considerations of observer models and external noise methods developed in this article to the auditory domain is quite direct. Empirical methods and theoretical developments parallel to those in the visual domain could potentially have strong contributions to make in refining the existing auditory observer models. For example, applications of the external noise methods reviewed in this article could not only provide additional tests of the existing auditory models but also specify the various internal noises in those models.

In the tactile domain, the external noise methods and observer models are less developed, although internal noise has been used to explain human and animal behavior (Eijkman & Vendrick, 1963; Rollman, 1969; Wu et al., 1994). The systematic development of the external noise methods and the observer models in visual perception may serve as an example for applications in the tactile domain.

Efficient TvC Measurements

One important result emphasized here is that repeated measurements of TvC functions at multiple performance levels provide strong constraints on observer models. Reliable measurement of multiple TvC functions has been demanding in terms of the amount of data collection (often 2,000 trials). Recently, Lesmes, Jeon, Lu, and Dosher (2006) developed a novel Bayesian adaptive procedure (the *qTvC method*) to ease data collection. Exploiting the known regularities in empirical TvC functions, the qTvC method generalizes a strategy, previously used to estimate psychometric threshold and slope (Kontsevich & Tyler, 1999), to adaptively estimate three parameters: the threshold in low external noise c_0, the critical noise level N_c where external noise starts to dominate performance (the joint of the TvC function), and the common slope, η, of the psychometric functions across external noise conditions. Using one-step-ahead search, the qTvC selects the stimulus for each trial that minimizes the entropy of the three-dimensional posterior probability distribution, $p(N_c, c_0, \eta)$. Simulations showed that 300 trials were sufficient to reach TvC estimates at three widely separated performance levels with less than 1% bias and approximately 1 dB mean root-mean-square error. Using an orientation discrimination task, Lesmes et al. found excellent agreement between TvCs obtained with qTvC and the method of constant stimuli, although the qTvC estimates were based on only 12% of the data collection (240 vs. 1,920 trials).

The qTvC method can also be used in conjunction with the double-pass procedure. In a recent study, Jeon et al. (2006) recorded the trial sequence and stimulus samples in a qTvC experiment and asked the observers to rerun the experiment with the recorded stimuli. A maximum-likelihood procedure has been developed to analyze the measured TvC functions and the double-pass agreements (Jeon et al., 2006).

Implications for Mechanism Studies

Our original motivation for developing the PTM was to provide a theoretical framework to characterize the changes of intrinsic limitations of the perceptual system underlying apparent changes in human performance due to attention (Lu & Dosher, 1998) or perceptual learning (Dosher & Lu, 1998). In a typical study, TvC functions at multiple performance levels are measured under joint manipulations of external noise and observer state, such as attention or training. By analyzing how the intrinsic limitations of the perceptual system vary as a function of the observer state, the PTM provides a mathematical framework to distinguish three mecha-

nisms of attention/perceptual learning: stimulus enhancement, external noise exclusion, and reduction of multiplicative noise.

As reviewed in the introduction, the idea of using external noise to quantify changes of the limiting factors in perceptual sensitivity and therefore identify the mechanisms underlying changed perceptual performance has been extended and applied to studies of a wide range of cognitive, developmental, and disease processes. A partial list includes studies of attention (Dosher & Lu, 2000a, 2000b; Lu & Dosher, 1998, 2000; Lu et al., 2000; Talgar et al., 2004), perceptual learning (Chung et al., 2005; Dosher & Lu, 1998, 1999; Gold et al., 1999; R. W. Li, Levi, & Klein, 2003; Lu & Dosher, 2004; Lu et al., 2005), adaptation (Dao et al., 2006), amblyopia (Huang et al., 2007; Levi & Klein, 2003; Xu et al., 2006), perceptual interaction (Yu et al., 2001), dyslexia (A. Sperling et al., 2005), and visual memory (Gold et al., 2005).

All these mechanism studies used the equivalent input noise method, with measurements of TvC functions at single or multiple performance criterion levels. Some studies (Chung et al., 2005; Gold et al., 1999, 2005) measured a TvC at a single criterion and the double-pass agreement in a limited range of external noise conditions and therefore did not assess a wide range of external noise conditions, and many studies measured only TvC functions at a single criterion performance level. None of the prior studies jointly modeled the TvC functions at multiple criteria and the double-pass agreement data. As discussed earlier, it is necessary to measure TvC functions at multiple performance levels with double-pass agreement to more fully constrain observer models. However, many of these mechanism studies in the literature used the LAM as a default without serious considerations of the adequacy of the model. The current review of the literature shows that the LAM does not provide an adequate characterization of the perceptual system in a single state and therefore will be inadequate or misleading as a basis for interpreting mechanisms.

The double-pass method has become a popular method for measuring internal noise in different observer states. For example, it has been used to answer the question whether internal noise changes with perceptual learning (Gold et al., 1999). All these studies found that the P_C versus P_A functions did not change as a function of state. They concluded that the internal noise did not change. As detailed in Appendix B, the P_C versus P_A function is completely determined by the ratio of the standard deviation of internal noise to the standard deviation of external noise. However, what is widely unappreciated is that in these mechanism studies, the standard deviation of the effective external noise depends on whether the perceptual template changes as a function of state. If perceptual learning retunes the perceptual template, then the effective external noise in the system is reduced. If the double-pass agreement function and therefore the internal to external noise ratio did not change, then this implies that the internal noise was reduced approximately equivalently to the effective external noise. Without a full analysis using full TvCs at multiple performance levels (or full psychometric functions over a wide range of external noise conditions) to investigate potential changes of the template together with the P_C versus P_A functions, it is not possible to interpret the restricted-test double-pass results in those experiments. The conclusion that internal noise did not change should be reevaluated in those studies, as it assumed that the template did not change either—an assumption directly at odds with the primary conclusions of those studies.

The current analysis indicates—at least in the visual domain—that the PTM is the correct form of observer model and should be the basis of analysis in studies that determine the mechanisms of observer state changes. In those studies that measured TvC functions at multiple performance levels (e.g., Dosher & Lu, 1999), invariance across observer state changes of threshold ratio (between performance levels in a given external noise condition) implied that the nonlinearity and multiplicative noise in the observer model are invariant across different observer states. Observation of invariant threshold ratios reduced the importance of double-pass agreement in those studies, especially when the functional form of the observer model was known (e.g., the PTM). For studies in a new task domain where the functional form of the observer model is unknown or in investigations whose goal is the identification of the form of multiplicative noise, the full triple TvC with double-pass procedure is still recommended.

In addition to specifying the correct observer model for a particular domain, the interest of the external noise plus observer model approach is that it can provide new methods for understanding and classifying performance changes between observer states and/or different observer populations. Often, such as in the applications to attention, the external noise and observer model framework provides new insights into the nature of cognitive effects and can provide a means of classifying those effects. The results also provide estimates of some fundamental properties of the observer system and constraints on the representations. We suggest that these insights may help to understand which aspects of brain responses are most relevant to processing in the corresponding tasks.

Perturbation of High-Level Cognitive Processes

Application of the empirical and modeling approach outside the perceptual domains to the study of high-level perceptual and cognitive processes, though the implementation of external noise is less obvious, has the potential for addressing new questions in these domains. The external noise paradigms belong to the general class of perturbation methods that are widely used in many domains of science. Indeed, the external noise manipulations were directly inspired by methods in physics and engineering. As we have shown in this article, by perturbing the input signal stimulus with external noise and observing the behavior of the human observer under these variations in the perceptual stimulus, the external noise methods can generate very constraining data that reveal essential observer properties. One of the key strengths of the external noise methods is that internal properties are referenced to external stimulus manipulations with known physical properties and measures. The general perturbation approach may be extended to study higher-level cognitive processes with an appropriate construction of the dimensions of variation (see Ashby & O'Brien, 2005, for an example in category learning). The challenge is to design and quantify perturbations in a property of the stimulus representation relevant to the limiting processes or templates. Once a good perturbation method is found, the empirical methods and theoretical considerations reviewed in this article could inspire new developments in these domains.

References

Abbey, C. K., & Eckstein, M. P. (2006). Classification images for detection, contrast discrimination, and identification tasks with a common ideal observer. *Journal of Vision, 6,* 335–355.

Ahumada, A. J., Jr. (1967). *Detection of tones masked by noise: A comparison of human observers with digital-computer-simulated energy detectors of varying bandwidths.* Doctoral dissertation, University of California, Los Angeles.

Ahumada, A. J., Jr. (1987). Putting the visual system noise back in the picture. *Journal of the Optical Society of America, 4*(A), 2372–2378.

Ahumada, A. J., Jr. (2002). Classification image weights and internal noise level estimation. *Journal of Vision, 2,* 121–131.

Ahumada, A. J., Jr., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America, 49,* 1751–1756.

Ahumada, A. J., Jr., & Watson, A. B. (1985). Equivalent-noise model for contrast detection and discrimination. *Journal of the Optical Society of America, 2*(A), 1133–1139.

Albrecht, D. G., & Geisler, W. S. (1991). Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Visual Neuroscience, 7,* 531–546.

Albrecht, D. G., & Hamilton, D. B. (1982). Striate cortex of monkey and cat: Contrast response function. *Journal of Neurophysiology, 48,* 217–237.

Ashby, F. G. (1992). *Multidimensional models of perception and cognition.* Hillsdale, NJ: Erlbaum.

Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Sciences, 9,* 83–89.

Barlow, H. B. (1956). Retinal noise and absolute threshold. *Journal of the Optical Society of America, 46,* 634–639.

Barlow, H. B. (1957). Incremental thresholds at low intensities considered as signal/noise discrimination. *Journal of Physiology (London), 136,* 469–488.

Bonds, A. B. (1991). Temporal dynamics of contrast gain in single cells of the cat striate cortex. *Visual Neuroscience, 6,* 239–255.

Bos, C. E., & Deboer, E. (1966). Masking and discrimination. *Journal of the Acoustical Society of America, 39*(A), 708–715.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.

Burbeck, C. A., & Kelly, D. H. (1981). Contrast gain measurements and the transient/sustained dichotomy. *Journal of the Optical Society of America, 71,* 1335–1342.

Burgess, A. E. (1985). Visual signal detection: III. On Bayesian use of prior knowledge and cross correlation. *Journal of the Optical Society of America, 2*(A), 1498–1507.

Burgess, A. E., & Colborne, B. (1988). Visual signal detection: IV. Observer inconsistency. *Journal of the Optical Society of America, 2*(A), 617–627.

Burgess, A. E., Wagner, R. F., Jennings, R. J., & Barlow, H. B. (1981, October 2). Efficiency of human visual signal discrimination. *Science, 214,* 93–94.

Cannon, M. W., & Fullenkamp, S. C. (1991). Spatial interactions in apparent contrast-inhibitory effects among grating patterns of different spatial frequencies, spatial positions and orientations. *Vision Research, 31,* 1985–1998.

Chung, S. T. L., Levi, D. M., & Tjan, B. (2005). Learning letter identification in peripheral vision. *Vision Research, 45,* 1399–1412.

Cohn, T. E., Thibos, L. N., & Kleinstein, R. N. (1974). Detectability of a luminance increment. *Journal of the Optical Society of America, 64,* 1321–1327.

Dao, D. Y., Lu, Z.-L., & Dosher, B. A. (2006). Adaptation to sine-wave gratings selectively reduces the contrast gain of the adapted stimuli. *Journal of Vision, 6,* 739–759.

Dau, T., Kollmeier, B., & Kohlrausch, A. (1997). Modeling auditory processing of amplitude modulation: 1. Detection and masking with narrow-band carriers. *Journal of the Acoustical Society of America, 102,* 2892–2905.

Derrington, A. M., & Lennie, P. (1981). Spatial and temporal contrast sensitivities of neurons in lateral geniculate nucleus of macaque. *Journal of Physiology (London), 357,* 219–240.

de Vries, H. L. (1943). 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.

Dosher, B. A., & Lu, Z.-L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proceedings of the National Academy of Sciences, USA, 95,* 13988–13993.

Dosher, B. A., & Lu, Z.-L. (1999). Mechanisms of perceptual learning. *Vision Research, 39,* 3197–3221.

Dosher, B. A., & Lu, Z.-L. (2000a). Mechanisms of perceptual attention in precuing of location. *Vision Research, 40,* 1269–1292.

Dosher, B. A., & Lu, Z.-L. (2000b). Noise exclusion in spatial attention. *Psychological Science, 11,* 139–146.

D'Zmura, M., & Knoblauch, K. (1998). Spectral bandwidths for the detection of color. *Vision Research, 38,* 3117–3128.

Eckstein, M. P., Ahumada, A. J., Jr., & Watson, A. B. (1997). Visual signal detection in structured backgrounds: II. Effects of contrast gain control, background variations, and white noise. *Journal of the Optical Society of America, 14*(A), 2406–2419.

Eckstein, M. P., Shimozaki, S. S., & Abbey, C. K. (2002). The footprints of visual attention in the Posner cueing paradigm revealed by classification images. *Journal of Vision, 2,* 25–45.

Eijkman, E., Thijssen, J. M., & Vendrik, A. J. (1966). Weber's law, power law, and internal noise. *Journal of the Acoustical Society of America, 40,* 1164–1173.

Eijkman, E., & Vendrick, A. J. H. (1963). Detection theory applied to absolute sensitivity of sensory systems. *Biophysical Journal, 3,* 65–78.

Fletcher, H. (1940). Auditory patterns. *Review of Modern Physics, 12,* 47–65.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: Masking experiments require a new model. *Journal of the Optical Society of America, 11*(A), 1710–1719.

Foley, J. M., & Chen, C.-C. (1999). Pattern detection in the presence of maskers that differ in spatial phase and temporal offset: Threshold measurements and a model. *Vision Research, 39,* 3855–3872.

Foley, J. M., & Legge, G. E. (1981). Contrast detection and near-threshold discrimination in human vision. *Vision Research, 21,* 1041–1053.

Fredericksen, R. E., & Hess, R. F. (1997). Temporal detection in human vision: Dependence on stimulus energy. *Journal of the Optical Society of America, 14*(A), 2557–2569.

Friis, H. T. (1944). Noise figures of radio receivers. *Proceedings of the IRE, 32,* 419–422.

Gegenfurtner, K. R., & Kiper, D. C. (1992). Contrast detection in luminance and chromatic noise. *Journal of the Optical Society of America, 9*(A), 1880–1888.

Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review, 96,* 267–314.

Gilkey, R. H., Frank, A. S., & Robinson, D. E. (1978). Estimates of internal noise. *Journal of the Acoustical Society of America, 64,* S36(A).

Gilkey, R. H., Frank, A. S., & Robinson, D. E. (1981). Estimates of the ratio of external to internal noise obtained using repeatable samples of noise. *Journal of the Acoustical Society of America, 69,* S23(A).

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999, November 11). Signal but not noise changes with perceptual learning. *Nature, 402,* 176–178.

Gold, J., Murray, R., Sekuler, A. B., Bennett, P. J., & Sekuler, R. (2005). Visual memory decay is deterministic. *Psychological Science, 16,* 769–774.

Gorea, A., & Sagi, D. (2001). Disentangling signal from noise in visual contrast discrimination. *Nature Neuroscience, 4,* 1146–1150.

Graham, N. V. S. (1989). *Visual pattern analyzers.* New York: Oxford University Press.

Green, D. M. (1964). Consistency of auditory detection judgments. *Psychological Review, 71,* 392–407.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Hartmann, W. M., & Pumplin, J. (1988). Noise power fluctuations and the masking of sine signals. *Journal of the Acoustical Society of America, 83,* 2277–2289.

Hay, G. A., & Chesters, M. S. (1972). Signal-transfer functions in threshold and suprathreshold vision. *Journal of the Optical Society of America, 62,* 990–998.

Hays, W. L. (1981). *Statistics* (3rd ed.). New York: Holt, Rinehart & Winston.

Hays, W. L. (1988). *Statistics* (4th ed.). Fort Worth, TX: Holt, Rinehart & Winston.

Heeger, D. J. (1993). Modeling simple-cell direction selectivity with normalized, half-squared, linear operators. *Journal of Neurophysiology, 70,* 1885–1898.

Hood, D. C., & Finkelstein, M. A. (1986). Sensitivity to light. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 1. Sensory processes and perception* (pp. 5-1–5-66). New York: Wiley.

Huang, C. B., Tao, L. M., Zhou, Y. F., & Lu, Z.-L. (2007). Treated amblyopes remain deficient in spatial vision: A contrast sensitivity and external noise study. *Vision Research, 47,* 22–34.

Humes, L. E., & Jesteadt, W. (1989). Models of the additivity of masking. *Journal of the Acoustical Society of America, 85,* 1285–1294.

Iverson, G. J., & Pavel, M. (1981). On the functional form of partial masking functions in psychoacoustics. *Journal of Mathematical Psychology, 24,* 1–20.

Jeon, S.-T., Lu, Z.-L., & Dosher, B. (2006). Extending observer models for more difficult identification and discrimination. *Journal of Vision, 6,* 192.

Kaplan, E., & Shapley, R. M. (1982). X and Y cells in the lateral geniculate nucleus of macaque monkeys. *Journal of Physiology (London), 330,* 125–143.

Katkov, M., Tsodyks, M., & Sagi, D. (2006). Singularities in the inverse modeling of 2AFC contrast discrimination data. *Vision Research, 46,* 259–266.

Klein, S. A., & Levi, D. M. (1985). Hyperacuity thresholds of 1 sec: Theoretical predictions and empirical validation. *Journal of the Optical Society of America, 2*(A), 1170–1190.

Kontsevich, L. L., Chen, C. C., & Tyler, C. W. (2002). Separating the effects of response nonlinearity and internal noise psychophysically. *Vision Research, 42,* 1771–1784.

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research, 39,* 2729–2737.

Lasley, D. J., & Cohn, T. E. (1981). Why luminance discrimination may be better than detection. *Vision Research, 21,* 273–278.

Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America, 70,* 1458–1471.

Legge, G. E., Kersten, D., & Burgess, A. E. (1987). Contrast discrimination in noise. *Journal of the Optical Society of America, 4*(A), 391–404.

Leshowitz, B., Taub, H. B., & Raab, D. H. (1968). Visual detection of signals in the presence of continuous and pulsed backgrounds. *Perception & Psychophysics, 4,* 207–213.

Lesmes, L. A., Jeon, S.-T., Lu, Z.-L., & Dosher, B. A. (2006). Bayesian adaptive estimation of threshold versus contrast external noise functions: The quick TvC method. *Vision Research, 46,* 3160–3176.

Levi, D., & Klein, S. (2003). Noise provides some new signals about the spatial vision of amblyopes. *Journal of Neuroscience, 7,* 2522–2526.

Li, R. W., Levi, D. M., & Klein, S. A. (2003). Perceptual learning improves efficiency by re-tuning the "template" for position discrimination. *Nature Neuroscience, 7,* 178–183.

Li, X., Lu, Z.-L., Xu, P., Jin, J., & Zhou, Y. (2003). Generating high gray-level resolution monochrome displays with conventional computer graphics cards and color monitors. *Journal of Neuroscience Methods, 130,* 9–18.

Logan, G. D. (2004). Cumulative progress in formal theories of attention. *Annual Review of Psychology, 55,* 207–234.

Lu, Z.-L., Chu, W., Dosher, B. A., & Lee, S. (2005). Independent perceptual learning in monocular and binocular motion systems. *Proceedings of the National Academy of Sciences, USA, 102,* 5624–5629.

Lu, Z.-L., & Dosher, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Research, 38,* 1183–1198.

Lu, Z.-L., & Dosher, B. A. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *Journal of the Optical Society of America, 16*(A), 764–778.

Lu, Z.-L., & Dosher, B. A. (2000). Spatial attention: Different mechanisms for central and peripheral temporal precues? *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1534–1548.

Lu, Z.-L., & Dosher, B. A. (2001). Characterizing the spatial-frequency sensitivity of perceptual templates. *Journal of the Optical Society of America, 18*(A), 2041–2053.

Lu, Z.-L., & Dosher, B. A. (2004). Perceptual learning retunes the perceptual template in foveal orientation identification. *Journal of Vision, 4,* 44–56.

Lu, Z.-L., & Dosher, B. A. (2007). *Response bias in double-pass agreement versus percent correct functions.* Manuscript in preparation.

Lu, Z.-L., Liu, C. Q., & Dosher, B. A. (2000). Attention mechanisms for multi-location first- and second-order motion perception. *Vision Research, 40,* 173–186.

Lu, Z.-L., & Sperling, G. (1996). Contrast gain control in first- and second-order motion perception. *Journal of the Optical Society of America, 13*(A), 2305–2318.

Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and "nonparametric" indexes. *Psychological Bulletin, 107,* 401–413.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* New York: Cambridge University Press.

Manjeshwar, R. M., & Wilson, D. L. (2001). Effect of inherent location uncertainty on detection of stationary targets in noisy image sequences. *Journal of the Optical Society of America, 18*(A), 78–85.

Moore, B. C. J. (1975). Mechanisms of masking. *Journal of the Acoustical Society of America, 57,* 391–399.

Mumford, W. W., & Schelbe, E. H. (1968). *Noise performance factors in communication systems.* Dedham, MA: Horizon House-Microwave.

Nachmias, J. (1981). On the psychometric function for contrast detection. *Vision Research, 21,* 215–223.

Nachmias, J., & Kocher, E. C. (1970). Visual detection and discrimination of luminance increments. *Journal of the Optical Society of America, 60,* 382–389.

Nachmias, J., & Sansbury, R. V. (1974). Grating contrast: Discrimination may be better than detection. *Vision Research, 14,* 1039–1042.

Nagaraja, N. S. (1964). Effect of luminance noise on contrast thresholds. *Journal of the Optical Society of America, 54,* 950–955.

Nolte, L. W., & Jaarsma, D. (1967). More on detection of one of *M* orthogonal signals. *Journal of the Acoustical Society of America, 41,* 497–505.

North, D. O. (1942). The absolute sensitivity of radio receivers. *RCA Review, 6,* 332–344.

Ohzawa, I., Sclar, G., & Freeman, R. D. (1982, July 15). Contrast gain control in the cat visual cortex. *Nature, 298,* 266–268.

Osman, E. (1971). A correlation model of binaural masking level differences. *Journal of the Acoustical Society of America, 50,* 1494–1511.

Pelli, D. G. (1981). *Effects of visual noise.* Doctoral dissertation, Cambridge University, Cambridge, England.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America, 2*(A), 1508–1532.

Pelli, D. G. (1990). The quantum efficiency of vision. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 3–24). Cambridge, England: Cambridge University Press.

Pelli, D. G. (1991). Noise in the visual system may be early. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 147–151). Cambridge, MA: MIT Press.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Pelli, D. G., & Farell, B. (1999). Why use noise? *Journal of the Optical Society of America, 16*(A), 647–653.

Richards, V. M., Heller, L. M., & Green, D. M. (1991). The detection of a tone added to a narrow band of noise: The energy model revisited. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 43*(A), 481–501.

Rollman, G. B. (1969). Detection models: Experimental tests with electrocutaneous stimuli. *Perception & Psychophysics, 5,* 377–380.

Rose, A. (1948). The sensitivity performance of the human eye on an absolute scale. *Journal of the Optical Society of America, 38,* 196–208.

Sclar, G., Maunsell, J. H., & Lennie, P. (1990). Coding of image contrast in central visual pathways of the macaque monkey. *Vision Research, 30,* 1–10.

Solomon, J. A., & Pelli, D. G. (1994, June 2). The visual filter mediating letter identification. *Nature, 369,* 395–397.

Solomon, J. A., Sperling, G., & Chubb, C. (1993). The lateral inhibition of perceived contrast is indifferent to on-center off-center segregation, but specific to orientation. *Vision Research, 33,* 2671–2683.

Sperling, A., Lu, Z.-L., Manis, F. R., & Seidenberg, M. (2005). Deficits in perceptual noise exclusion in developmental dyslexia. *Nature Neuroscience, 8,* 862–863.

Sperling, G. (1989). Three stages and two systems of visual processing. *Spatial Vision, 4,* 183–207.

Sperling, G., & Dosher, B. A. (1986). Strategy and optimization in human information processing. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and performance* (Vol. 1, pp. 1–85). New York: Wiley.

Spiegel, M. F., & Green, D. M. (1981). Two procedures for estimating internal noise. *Journal of the Acoustical Society of America, 70,* 69–73.

Stromeyer, C. F., & Klein, S. (1974). Spatial frequency channels in human vision as asymmetric (edge) mechanisms. *Vision Research, 14,* 1409–1420.

Swets, J. A. (1996). *Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers.* Hillsdale, NJ: Erlbaum.

Swets, J. A., Shipley, E. F., McKey, M. J., & Green, D. M. (1959). Multiple observations of signals in noise. *Journal of the Acoustical Society of America, 31,* 514–521.

Talgar, C. P., Pelli, D. G., & Carrasco, M. (2004). Covert attention enhances letter identification without affecting channel tuning. *Journal of Vision, 4,* 22–31.

Tanner, W. P., Jr. (1961). Physiological implications of psychophysical data. *Annals of the New York Academy of Sciences, 89,* 752–765.

Tanner, W. P., Jr., & Birdsall, T. G. (1958). Definitions of d' and n as psychophysical measures. *Journal of the Acoustical Society of America, 30,* 922–928.

Tjan, B. S., Braje, W. L., Legge, G. E., & Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research, 35,* 3053–3069.

Tolhurst, D. J., Movshon, J. A., & Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research, 23,* 775–785.

Van Meeteren, A., & Barlow, H. B. (1981). The statistical efficiency for detecting sinusoidal modulation of average dot density in random figures. *Vision Research, 21,* 765–777.

Watson, A. B., & Solomon, J. A. (1997). Model of visual contrast gain control and pattern masking. *Journal of the Optical Society of America, 14,* 2379–2391.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling and goodness of fit. *Perception & Psychophysics, 63,* 1293–1313.

Wickelgren, W. A. (1968). Unidimensional strength theory and component analysis of noise in absolute and comparative judgments. *Journal of Mathematical Psychology, 5,* 102–122.

Wilkens, T. D. (2002). *Elementary signal detection theory.* New York: Oxford University Press.

Woodworth, R. S. (1938). *Experimental psychology.* New York: Holt.

Wu, J.-Y., Tsau, Y., Hopp, H.-P., Cohen, L. B., Tang, A. C., & Falk, C. X. (1994). Consistency in nervous systems: Trial-to-trial and animal-to-animal variations in the responses to repeated applications of a sensory stimulus in Aplysia. *Journal of Neuroscience, 14,* 1366–1384.

Xu, P., Lu, Z.-L., Qiu, Z., & Zhou, Y. (2006). Identify mechanisms of amblyopia in Gabor orientation identification with external noise. *Vision Research, 46,* 3748–3760.

Yu, C., Levi, D., & Klein, S. (2001). Surround modulation of perceived contrast and the role of brightness induction. *Journal of Vision, 1,* 18–31.

Zhang, Y., Pham, B. T., & Eckstein, M. P. (2006). The effect of nonlinear human visual system components on performance of a channelized Hotelling observer in structured backgrounds. *IEEE Transactions on Medical Imaging, 25,* 1348–1362.

Appendix A

Some Basic SDT Equations

We first provide a brief standard treatment of the signal detection theory (SDT) because many SDT concepts and equations are used in this article.

SDT in a Yes–No Task

There are two types of trials in a simple yes–no task, signal-present and signal-absent trials. The SDT postulates that, in every trial, the input stimulus generates an internal response x, which could have been generated by a signal stimulus with a probability density of $g(x, \mu_S, \sigma)$ or a signal-absent stimulus with a probability density of $g(x, \mu_N, \sigma)$, where $\mu_S > \mu_N$ (see Figure 1A in the main text). To decide whether the signal is present ("yes") or absent ("no") based on the single internal response x, the observer chooses a subjective criterion response C. If $x > C$, the observer responds with a "yes"; otherwise, she or he responds with a "no."

For a given C, we can compute the probability of all four possible outcomes of each trial, hit, miss, false alarm (FA), and correct rejection (CR):

$$P_{Hit}(C) = 1 - \int_{-\infty}^{C} g(x, \mu_S, \sigma)dx = \int_{-\infty}^{C} 1 - G(C, \mu_S, \mu)$$

$$= 1 - G\left(\frac{C - \mu_N}{\sigma}, \frac{\sigma_S - \mu_N}{\sigma}, 1\right), \quad \text{(A1)}$$

$$P_{Miss}(C) = 1 - P_{Hit} = G\left(\frac{C - \mu_N}{\sigma}, \frac{\mu_S - \mu_N}{\sigma}, 1\right), \quad \text{(A2)}$$

$$P_{FA}(C) = 1 - \int_{-\infty}^{C} g(x, \mu_N, \sigma)dx = 1 - G(C, \mu_N, \sigma)$$

$$= 1 - G\left(\frac{C - \mu_N}{\sigma}, 0, 1\right), \quad \text{(A3)}$$

and

$$P_{CR}(C) = 1 - P_{FA} = G\left(\frac{C - \mu_N}{\sigma}, 0, 1\right). \quad \text{(A4)}$$

One can solve Equation A3 to obtain the criterion response $\frac{C - \mu_N}{\sigma}$ as a function of the false-alarm probability:

$$\frac{C - \mu_N}{\sigma} = G^{-1}(1 - P_{FA}, 0, 1) . \quad \text{(A5)}$$

Substituting Equation A5 into Equation A1 results in the functional relationship between hit and false-alarm rates (the receiver operating characteristics [ROCs]):

$$P_{Hit} = 1 - G\left(G^{-1}(1 - P_{FA}, 0, 1), \frac{\mu_S - \mu_N}{\sigma}, 1\right). \quad \text{(A6)}$$

With the definition $d' = \frac{\mu_S - \mu_N}{\sigma}$, we can rewrite Equation A6 as

$$P_{Hit} = 1 - G(G^{-1}(1 - P_{FA}, 0, 1), d', 1). \quad \text{(A7)}$$

Therefore, the functional relationship between the hit and false-alarm rates (the ROC curve; see Figure 1B in the main text) is determined by the sensitivity of the observer, d', the signal-to-noise ratio in the internal response distribution. Conversely, the sensitivity of the observer, d', can be empirically observed by measuring the ROC curve.

SDT in a 2AFC Task

In a two-alternative forced-choice (2AFC) task, an observer is presented with two input stimuli, one from each of two stimulus categories. The SDT postulates that there are two internal responses, x_A and x_B, in each trial; the probability density that an internal response x is generated by a stimulus in Category 1 is $g(x, \mu_1, \sigma_1)$ and by a stimulus in Category 2 is $g(x, \mu_2, \sigma_2)$ where $\mu_2 > \mu_1$ (see Figure 1C in the main text). To decide whether x_A is generated by a stimulus in Category 1 (and therefore, x_B is generated by a stimulus in Category 2) or by a stimulus in Category 2 (and therefore, x_B is generated by a stimulus in Category 1), an unbiased observer compares x_A and x_B. If $x_A - x_B > 0$, she or he concludes that x_A is generated by a stimulus in Category 2; otherwise, she or he concludes that x_A is generated by a stimulus in Category 1.[A1]

The probability that the observer makes a correct response P_C can be computed in two different but equivalent ways. In the first way, P_C is the probability that $x_A - x_B > 0$, given that x_A is generated by a stimulus in Category 2 and x_B is generated by a stimulus in Category 1:

$$P_c = P(x_A - x_B > 0 \,|\, x_A \in C2 \;\&\; x_B \in C1)$$

$$= 1 - \int_{-\infty}^{0} g(x, \mu_2 - \mu_1, \sqrt{\sigma_1^2 + \sigma_2^2})dx$$

$$= I - G\left(0, \frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}, 1\right), \quad \text{(A8)}$$

[A1] Although some may favor the maximum-likelihood decision rule, it is not the optimal decision rule under some circumstances. In general, for unequal presentation probabilities or asymmetric payoffs, maximizing likelihood is not equivalent to maximizing a posterior probability of payoffs (Graham, 1989; Green & Swets, 1966). The maximum-likelihood decision rule is equivalent to the max rule and the difference rule when the signal and noise distributions have the same variance. On the other hand, in 2AFC or more general classification experiments, in which the stimuli are all simple stimuli, the max rule is plausible (Graham, 1989; Green & Swets, 1966; Macmillan & Creelman, 1991). The max rule is equivalent to the difference rule when the simple stimuli are equally detectable and far apart and each excites only one detector.

(*Appendixes continue*)

where $g(x, \mu_2 - \mu_1, \sqrt{\sigma_1^2 + \sigma_2^2})$ is the probability density function of the difference between two Gaussian random variables with probability density functions $g(x, \mu_1, \sigma_1)$ and $g(x, \mu_2, \sigma_2)$. The equation is illustrated by the shaded area in Figure 1D in the main text.

In the second way, P_C is derived from a different kind of reasoning. The probability density of obtaining an internal response x from a stimulus in Category 2 is $g(x, \mu_2, \sigma_2)$. The probability that the internal response x is greater than any random sample from distribution $g(x, \mu_1, \sigma_1)$ is $G(x, \mu_1, \sigma_1)$. The probability that all possible internal responses from stimuli in Category 2 are greater than those from stimuli in Category 1 (P_C) is the product of the two probability functions integrated over all the possible values of x:

$$P_c = \int_{-\infty}^{\infty} g(x, \mu_2, \sigma_2)G(x, \mu_1, \sigma_1)dx$$

$$= \int_{-\infty}^{\infty} g(x, \mu_2, \sigma_2)G(x, \mu_1, \sigma_1)dx. \quad \text{(A9)}$$

Equations A8 and A9 are mathematically equivalent. Although Equation A8 is more intuitive, Equation A9 is more readily extended to situations with decision uncertainty.

If $\sigma_2 = \sigma_1 = \sigma$, we can define $d' = \frac{\mu_2 - \mu_1}{\sigma}$ and simplify Equation A9:

$$P_c = \int_{-\infty}^{\infty} g(x - d', 0, 1)G(x, 0, 1)dx. \quad \text{(A10)}$$

SDT in 2AFC With Uncertainty

Although the observer is presented with two input stimuli, one from each of two stimulus categories, in a 2AFC task, it is possible that $(U + 1)$ independent detectors respond to each stimulus. Only one of those detectors is task relevant, but the observer cannot identify it and has to make a decision based on the internal responses of all the detectors. The observer therefore has to monitor a total of $2(U + 1)$ internal responses of which one is associated with the stimulus from Category 1, one is associated with the stimulus from Category 2, and $2U$ are associated with task-irrelevant detectors. Here, we consider the case in which (a) the internal responses of all the $2(U + 1)$ detectors are independent and Gaussian distributed, (b) the stimulus from Category 2 generates an internal response distribution with mean μ_S and standard deviation σ_S in the task-relevant detector, and (c) the distributions of the internal responses of all the other $2U + 1$ detectors have mean 0 and standard deviation σ_N. To decide which set of the $(U + 1)$ internal responses is generated by a stimulus in Category 2 (and therefore, the other set is generated by a stimulus in Category 1), the observer could potentially use several different decision rules.

With the optimal summation rule, the observer sums the $(U + 1)$ internal responses from each of the two stimuli and decides that the set with the larger sum is generated by a stimulus from Category 2. For the sum of the $(U + 1)$ random variables, the mean is equal to the sum of the means; the variance is the sum of the variances. Equation A9 therefore describes performance accuracy with the mean and variance of the sum of the $(U + 1)$ internal responses.

Often, a maximum rule, or max rule, is used instead of the summation rule. In many conditions, the max rule is a reasonable approximation (Nolte & Jaarsma, 1967) and approximates the optimal decision rule. With the maximum rule, the observer compares all the $2(U + 1)$ internal responses and labels the interval or sample corresponding to the maximum of all as generated by a stimulus from Category 2. The observer could make a correct response in two different ways: (a) The internal response of the task-relevant detector to a stimulus in Category 2, x, is greater than the other $2U + 1$ internal responses, or (b) the internal response of one of the task-irrelevant detectors to a stimulus in Category 2, x, is greater than the other $2U + 1$ internal responses, including the one from the task-relevant detector, although the observer generates the correct response for the wrong reason. The two possibilities are reflected in the two terms of the following equation:

$$P_c = \int_{-\infty}^{+\infty} [g(x - \mu_S, 0, \sigma_S)G^{2U+1}(x, 0, \sigma_N)$$

$$+ Ug(x, 0, \sigma_N)G^{2U}(x, 0, \sigma_N)G(x - \mu_S, 0, \sigma_S)]dx. \quad \text{(A11)}$$

The maximum rule can also be formulated in a different but equivalent way. The observer could first extract the maximum of the $(U + 1)$ internal responses to each stimulus in a trial and decide that the one that contains the greater maximum internal response is generated by the stimulus in Category 2. For the $(U + 1)$ internal responses generated by a stimulus in Category 1, all with the same probability density function $g(x, 0, \sigma_N)$, the probability density function of the maximum is

$$p_1(x|\max) = (U + 1)g(x, 0, \sigma_N)G^U(x, 0, \sigma_N). \quad \text{(A12)}$$

For the $(U + 1)$ internal response generated by a stimulus in Category 2, the probability density function of one of them is $g(x - \mu_S, 0, \sigma_S)$; the probability density function of the other U of them is $g(x, 0, \sigma_N)$. The probability density function of the maximum internal response is

$$p_2(x|\max) = g(x - \mu_S, 0, \sigma_S)G^U(x, 0, \sigma_N)$$

$$+ Ug(x, 0, \sigma_N)G^{U-1}(x, 0, \sigma_N)G(x - \mu_S, 0, \sigma_S). \quad \text{(A13)}$$

The probability of making a correct response is equal to the probability that samples from $p_2(x|\max)$ are greater than samples from $p_1(x|\max)$:

$$P_c = \int p_2(x|\max)P_1(x|\max)dx$$

$$= \int_{-\infty}^{+\infty} [g(x - \mu_S, 0, \sigma_S)G^{2U+1}(x, 0, \sigma_N)$$

$$+ Ug(x, 0, \sigma_N)G^{2U}(x, 0, \sigma_N)G(x - \mu_S, 0, \sigma_S)]dx. \quad \text{(A14)}$$

If $\sigma_S = \sigma_N = \sigma$, we can define $d' = \frac{\mu_S}{\sigma}$ and simplify Equation A11 to

$$P_c = \int_{-\infty}^{+\infty} [g(x - d', 0, 1)G^{2U+1}(x, 0, 1)$$

$$+ Ug(x, 0, 1)G^{2U}(x, 0, 1)G(x - d', 0, 1)dx. \quad \text{(A15)}$$

Appendix B

Double-Pass Consistency

We illustrate the mathematical basis of the double-pass procedure in a two-alternative forced-choice (2AFC) paradigm. Similar development can be found in Gilkey et al. (1978) and Burgess and Colborne (1988). We assume that (a) two independent perceptual detectors are used by the observer to perform the 2AFC task; (b) a given signal stimulus S generates a fixed response S in Detector 1 and response 0 in Detector 2; (c) a given external noise stimulus N_{ext} generates a response N_{ext1} in Detector 1, and an independent response N_{ext2} in Detector 2; (d) for a given pair of signal S and external noise N_{ext}, S, N_{ext1}, and N_{ext2} are invariant over time, that is, they stay the same in two passes of the same stimulus sequence; (e) the total internal noise of the observer in a given (stimulus and external noise magnitude) condition is a Gaussian random variable with mean 0 and standard deviation σ_{int}, the particular sample of internal noise has a value of N_{int1a} in Detector 1 and N_{int2a} in Detector 2 when the observer processes S and N_{ext} in the first pass, and the total amount of internal noise is N_{int1b} in Detector 1 and N_{int2b} in Detector 2 when the observer processes S and N_{ext} in the second pass; (f) the observer chooses response 1 if $(S + N_{ext1} + N_{int1}) > (N_{ext2} + N_{int2})$ and vice versa; and (g) N_{ext1}, N_{ext2}, N_{int1a}, N_{int2a}, N_{int1b}, and N_{int2b} are independent and normally distributed:

$$p(N_{ext1}) = g(N_{ext1}, 0, \sigma_{ext1}), \quad p(N_{ext2}) = g(N_{ext2}, 0, \sigma_{ext2}),$$

$$p(N_{int1a}) = g(N_{int1a}, 0, \sigma_{int1}), \quad p(N_{int2a}) = g(N_{int2a}, 0, \sigma_{int2}),$$

$$p(N_{int1b}) = g(N_{int1b}, 0, \sigma_{int1}), \quad p(N_{int2b}) = g(N_{int2b}, 0, \sigma_{int2}). \quad \text{(B1)}$$

The probability that the observer makes a correct response is

$$\begin{aligned} P_c &= \{P[(S + N_{ext1} + N_{int1a}) > (N_{ext2} + N_{int2a})] \\ &\quad + P[(S + N_{ext1} + N_{int1b}) > (N_{ext2} + N_{int2b})]\}/2.0 \\ &= \int g(x - S, 0, \sqrt{\sigma_{ext1}^2 + \sigma_{int1}^2}) G(x, 0, \sqrt{\sigma_{ext2}^2 + \sigma_{int2}^2}) dx, \end{aligned} \quad \text{(B2)}$$

where $G(x, 0, \sigma)$ is the cumulative distribution of a Gaussian random variable with mean 0 and standard deviation σ. If we assume that $\sigma_{ext1} = \sigma_{ext2} = \sigma_{ext}$, and $\sigma_{int1} = \sigma_{int2} = \sigma_{int} = \alpha\sigma_{ext}$, we have

$$P_c = \int g(x - S, 0, \sqrt{1 + \alpha^2}\sigma_{ext}) G(x, 0, \sqrt{1 + \alpha^2}\sigma_{ext}) dx. \quad \text{(B2a)}$$

Therefore, for a given stimulus condition, that is, S and σ_{ext}, P_C depends only on the ratio of standard deviation of the internal and external noise α.

When the same stimulus (signal + external noise) is passed to the observer twice, the probability that the two responses are consistent is

$$\begin{aligned} P_A &= P[(S + N_{ext1} + N_{int1a}) > (N_{ext2} + N_{int2a})]P[(S + N_{ext1} \\ &\quad + N_{int1b}) > (N_{ext2} + N_{int2b})] + P[(S + N_{ext1} + N_{int1a}) \\ &\quad < (N_{ext2} + N_{int2a})]P[(S + N_{ext1} + N_{int1b}) < (N_{ext2} + N_{int2b})] \\ &= P[(S + N_{ext1} - N_{ext2}) + (N_{int1a} - N_{int2a}) > 0] \\ &\quad \times P[(S + N_{ext1} - N_{ext2}) + (N_{int1b} - N_{int2b}) > 0] \\ &\quad + P[(S + N_{ext1} - N_{ext2}) + (N_{int1a} - N_{int2a}) < 0] \\ &\quad \times P[(S + N_{ext1} - N_{ext2}) + (N_{int1b} - N_{int2b}) < 0] \\ &= \int_{-\infty}^{+\infty} g(x - S, 0, \sqrt{\sigma_{ext1}^2 + \sigma_{ext2}^2}) \{G^2(x, 0, \sqrt{\sigma_{int1}^2 + \sigma_{int2}^2}) \\ &\quad + [1 - G(Zx, 0, \sqrt{\sigma_{int1}^2 + \sigma_{int2}^2})]^2\} dx. \end{aligned} \quad \text{(B3)}$$

Again, if we assume that $\sigma_{ext1} = \sigma_{ext2} = \sigma_{ext}$, and $\sigma_{int1} = \sigma_{int2} = \sigma_{int} = \alpha\sigma_{ext}$, we have

$$P_A = \int_{-\infty}^{+\infty} g(x - S, 0, \sqrt{2}\sigma_{ext}) \{G^2(x, 0, \sqrt{2}\alpha\sigma_{ext}) + [1 - G(x, 0, \sqrt{2}\alpha\sigma_{ext})]^2\} dx. \quad \text{(B3a)}$$

Similar to P_C, P_A depends on the ratio of standard deviation of the internal and external noise α for a given stimulus condition (S and σ_{ext}).

Appendix C

Linear Amplifier Model

For a signal stimulus with root-mean-square contrast c superimposed on white Gaussian noise images—images made of pixels whose contrasts are drawn from jointly independent, identically distributed Gaussian random variables with mean zero and standard deviation N_{ext}—the signal can be expressed as a function of space and time: $S(x, y, t) = cS_0(x, y, t)$ rescaled such that $\int\int\int S_0^2(x, y, t)dxdydt = 1.0$. The external noise can be expressed as $N(x, y, t) = \sigma_{ext}G(x, y, t)$ where the value of $G(x, y, t)$ at a particular point (x, y, t) is drawn from a Gaussian distribution with mean 0 and standard deviation 1.0. Again, we assume that two detectors, one matched to the signal and another orthogonal to the signal stimulus, are involved.

(*Appendixes continue*)

Matching the task-relevant template $T(x, y, t)$ to a signal-valued stimulus yields

$$T_S = \int\int\int T_1(x, y, t)S(x, y, t)dxdydt = c\int\int\int\int T_1(x, y, t)S_0(x, y, t)dxdydt; \quad \text{(C1)}$$

matching the template to the external noise yields

$$T_{N1} = \int\int\int T_1(x, y, t)N(x, y, t)dxdydt = \sigma_{ext} \times \int\int\int T_1(x, y, t)G(x, y, t)dxdydt, \quad \text{(C2a)}$$

and

$$T_{N2} = \int\int\int T_2(x, y, t)N(x, y, t)dxdydt = \sigma_{ext} \times \int\int\int T_2(x, y, t)G(x, y, t)dxdydt. \quad \text{(C2b)}$$

For a fixed template and a fixed signal stimulus, $T_{S_0} = \int\int\int T_1(x, y, t) S_0(x, y, t)dxdydt$ is a constant; $T_{G1} = \int\int\int T_1(x, y, t)G(x, y, t)dxdydt$ and $T_{G2} = \int\int\int\int T_2(x, y, t)G(x, y, t)\, dxdydt$ are Gaussian random variables with mean 0 and a fixed standard deviation σ_{T_G}. Because, mathematically, T_{S_0} and σ_{T_G} can only be known up to a constant, without losing any generality, we set σ_{T_G} to 1.0. This essentially sets the total gain of the perceptual template (integrated over space and time) to 1.0. The other way to state the normalization is that we set the gain of the perceptual template to $\beta = \frac{T_{S_0}}{\sigma_{T_G}}$. The outputs from template matching are

$$T_S = \beta c, \quad \text{(C3)}$$

$$T_{N1} = \sigma_{ext}\tilde{G}_1(0, 1), \quad \text{(C4a)}$$

and

$$T_{N2} = \sigma_{ext}\tilde{G}_2(0, 1), \quad \text{(C4b)}$$

where $\tilde{G}_1(0, 1)$ and $\tilde{G}_2(0, 1)$ are two samples from the standard normal distribution.

At the decision stage, the total variance of the external and internal (additive) noise is the sum of the variances of the external and the internal noise in the two detectors:

$$N^2_{total1} = \sigma^2_{ext} + \sigma^2_{add}, \quad \text{(C5a)}$$

and

$$N^2_{total2} = \sigma^2_{ext} + \sigma^2_{add}. \quad \text{(C5b)}$$

Signal discriminability, d', is determined by the signal-to-noise ratio:

$$d' = \frac{T_S}{N_{total1}} = \frac{T_S}{N_{total2}} = \frac{\beta c}{\sqrt{\sigma^2_{ext} + \sigma^2_{add}}}. \quad \text{(C6)}$$

For a two-alternative forced-choice task, probability correct can be expressed as a function of d' (Macmillan & Creelman, 1991):

$$P_c = \int_{-\infty}^{+\infty} g(x - \beta c, 0, \sqrt{\sigma^2_{ext} + \sigma^2_{add}})G(x, 0, \sqrt{\sigma^2_{ext} + \sigma^2_{add}})dx = \int_{-\infty}^{+\infty} g(x - d', 0, 1)G(x, 0, 1)dx, \quad \text{(C7)}$$

where $g(x, \mu, \sigma)$ and $G(x, \mu, \sigma)$ are the probability density, and cumulative density functions of a Gaussian distribution.

The probability that the observer responds to two passes of the same stimuli sequence consistently can be derived from Equation B3 (see Appendix B) by replacing S, σ_{ext}, and σ_{int} in the equation with the constructs from the linear amplifier model (LAM):

$$P_A = \int_{-\infty}^{+\infty} g(x - S, 0, \sqrt{\sigma^2_{ext1} + \sigma^2_{ext2}})\{G^2(x, 0, \sqrt{\sigma^2_{int1}+\sigma^2_{int2}}) + [1 - G(x, 0, \sqrt{\sigma^2_{int1} + \sigma^2_{int2}})]^2\}dx = \int_{-\infty}^{+\infty} g(x - \beta c, 0, \sqrt{2}\sigma_{ext})\{G^2(x, 0, \sqrt{2\sigma_{add}}) + [1 - G(x, 0, \sqrt{2\sigma_{add}})]^2\}dx. \quad \text{(C8)}$$

By inverting Equation C6, we can also express the threshold signal contrast energy c^2_τ required for the observer to maintain a given performance criterion level, that is, a fixed percent correct or fixed d', as a function of external noise contrast:

$$c^2_\tau = \left(\frac{d'}{\beta}\right)^2 [\sigma^2_{ext} + \sigma^2_{add}]. \quad \text{(C9a)}$$

Defining $k = \left(\frac{d'}{\beta}\right)^2$, we can rewrite Equation C9a as

$$c^2_\tau = k(\sigma^2_{ext} + \sigma^2_{add}). \quad \text{(C9b)}$$

This is the efficiency relation between threshold and external noise in the LAM. The parameter k is called *observer efficiency*. Because k is proportional to d'^2, it is obvious that it (and thus, observer efficiency) depends on the particular criterion performance level at which the threshold is defined. The more fundamental parameter in this model is the gain of the perceptual template to the signal stimulus β, which is independent of the performance criterion.

The LAM makes a very simple yet highly constraining prediction on the ratio between thresholds at different performance criteria. For any given external noise condition, σ_{ext}, the contrast threshold $c_{\tau 1}$ at performance criterion level d'_1 is

$$c_{\tau 1} = \frac{d'_1}{\beta}\sqrt{\sigma^2_{ext} + \sigma^2_{add}}. \quad \text{(C10)}$$

The contrast threshold $c_{\tau 2}$ at performance criterion level d'_2 is

$$c_{\tau 2} = \frac{d'_2}{\beta}\sqrt{\sigma^2_{ext} + \sigma^2_{add}}. \quad \text{(C11)}$$

The ratio between the two thresholds for the particular external noise condition is thus

$$\frac{c_{\tau 1}}{c_{\tau 2}} = \frac{d'_1}{d'_2}. \quad \text{(C12)}$$

Appendix D

MATLAB Code for Probability Correct and Agreement

```
function PC_PA= PC_PA(dS, N_ext, N_pre, N_abs, U, Trials)

%dS:     Magnitude of the signal at the decision stage
%N_ext:  Standard deviation of the external noise
%N_pre:  Standard deviation of the internal noise in signal
%          present detector
%N_abs:  Standard deviation of the internal noise in signal
%          absent detector
%U:      Number of hidden detectors
%Trials: Number of trials in each pass

if U==0
    N=sqrt(N_ext^2+(N_pre^2+N_abs^2)/2);
    t=dS+(-10000:10000)/100*N;
    PC=sum(0.01*(N)*normpdf(t,dS,N_pre).*normcdf(t,0,N_abs));

    sX=N_ext*sqrt(2);
    b=sqrt(N_pre^2+N_abs^2);
    t=dS+(-10000:10000)/100*sX;
    PA=sum(0.01*(sX)*normpdf(t,dS,sX)...
         .*(normdcdf(t,0,b).^2+(1-normcdf(t,0,b)).^2));
    PC_PA=[PC PA];
else
    PC=0;
    PA=0;

    for i=1:Trials
       n_ext1=randn(1,1+U)*N_ext;
       n_ext2=randn(1,1+U)*N_ext;
       n_int1=randn(1,1+U)*N_abs;
       n_int2=randn(1,1+U)*N_abs;
       n_int3=randn(1,1+U)*N_abs;
       n_int4=randn(1,1+U)*N_abs;
       n_ext1(1)=n_ext(1)+S;
       n_int1(1)=randn(1)*N_pre;
       n_int3(1)=randn(1)*N_pre;

   %first pass
      interval1=n_ext1+n_int1;
      interval2=n_ext2+n_int2;
      if (max(interval1) > max(interval2) )
          rsp1=1;
          PC=PC+1;
      else
          rsp1=2;
      end
   %second pass
      interval1=n_ext1+n_int3;
      interval2=n_ext2+n_int4;
      if (max(interval1) > max(interval2) )
          rsp2=1;
          PC=PC+1;
      else
          rsp2=2;
      end
      if (rsp1==rsp2)
              pA=pA+1;
         end
      end
      PC=PC/(2*Trials);
      PA=PA/Trials;
      PC_PA=[PC PA];
end
```

(*Appendixes continue*)

Appendix E

The Perceptual Template Model

In the perceptual template model (PTM; see Figure 15a in the main text), input stimuli are processed in two pathways. In the signal pathway, input stimuli pass through a perceptual template with certain selectivity for stimulus characteristics (e.g., color, spatial frequency, orientation, temporal/spatial windowing, etc.). As in the linear amplifier model, the gain of the template to white Gaussian noise is 1.0 because the total gain of the template is normalized to 1.0. The gain of the template to the signal stimulus is β (see Equation 10 in the main text). A template matching function might, however, be far more complex, for example, templates for objects, faces, and so on. It is related to the concept of a matched filter in prior investigations of identification performance (Burgess, 1985). The output of the perceptual template is then processed by an expansive nonlinear transducer function (Output = sign(Input)$|\text{Input}|^{\gamma_1}$), chosen from the pattern vision literature (Foley & Legge, 1981; Nachmias & Sansbury, 1974). After the nonlinear transducer, the expected magnitude of the response of the signal template to the signal stimulus is therefore

$$S_1 = \beta^{\gamma_1} c^{\gamma_1}. \tag{E1}$$

The expected standard deviations of external noise in the signal-present and signal-absent detectors are[E1]

$$\bar{\sigma}_{N1} = \sigma_{\text{ext}}^{\gamma_1}, \tag{E2a}$$

and

$$\bar{\sigma}_{N2} = \sigma_{\text{ext}}^{\gamma_1}. \tag{E2b}$$

In the multiplicative internal noise pathway, the input passes through a different perceptual template (gain to signal stimulus: β_2; gain to white external noise: 1.0) and a rectified, nonlinear transducer function (Output = $|\text{Input}|^{\gamma_2}$). In computing multiplicative noise, stimulus energy over a broad range of space, time, and features may be integrated.[E2] The variance of multiplicative noise is proportional to the total stimulus energy in each detector. In the signal-present detector,

$$\sigma_{\text{mul1}}^2 = N_{\text{mul}}^2[N_{\text{ext}}^{2\gamma_2} + (\beta_2 c)^{2\gamma_2}]. \tag{E3a}$$

In the detector not matched to the input signal,

$$\sigma_{\text{mu12}}^2 = N_{\text{mul}}^2 N_{\text{ext}}^{2\gamma_2}. \tag{E3b}$$

At the decision stage, the signal is combined with external noise from the signal path, the multiplicative noise, and the additive internal noise. The details of the decision process depend on the particular task, for example, detection versus identification. These have been modeled elsewhere (Macmillan & Creelman, 1990). Here, we summarize the total variance in the signal-present and signal-absent detectors:

$$\sigma_{\text{total1}}^2 = \sigma_{\text{ext}}^{2\gamma_1} + N_{\text{mul}}^2[\sigma_{\text{ext}}^{2\gamma_2} + (\beta_2 c)^{2\gamma_2}] + \sigma_{\text{add}}^2, \tag{E4a}$$

and

$$\sigma_{\text{total2}}^2 = \sigma_{\text{ext}}^{2\gamma_1} + N_{\text{mul}}^2 \sigma_{\text{ext}}^{2\gamma_2} + \sigma_{\text{add}}^2, \tag{E4b}$$

In the PTM, probability correct for a two-alternative forced-choice task is therefore

$$P_c = \int_{-\infty}^{+\infty} g(x - S_1, 0, \sigma_{\text{total1}}) G(x, 0, \sigma_{\text{total2}}) dx$$

$$= \int_{-\infty}^{+\infty} g(x - \beta^{\gamma_1} c^{\gamma_1}, 0, \sqrt{\sigma_{\text{ext}}^{2\gamma_1} + N_{\text{mul}}^2[\sigma_{\text{ext}}^{2\gamma_2} + (\beta_2 c)^{2\gamma_2}] + \sigma_{\text{add}}^2}$$

$$\times G(x, 0, \sqrt{\sigma_{\text{ext}}^{2\gamma_1} + N_{\text{mul}}^2 \sigma_{\text{ext}}^{2\gamma_2} + \sigma_{\text{add}}^2}) dx \tag{E5}$$

The average signal-to-noise ratio (d') in the PTM can be calculated:

$$d' = \frac{S_1}{\sqrt{(\sigma_{\text{total1}}^2 + \sigma_{\text{total2}}^2)/2}}$$

$$= \frac{(\beta c)^{\gamma_1}}{\sqrt{\sigma_{\text{ext}}^{2\gamma_1} + N_{\text{mul}}^2 \left[\sigma_{\text{ext}}^{2\gamma_2} + \frac{(\beta_2 c)^{2\gamma_2}}{2} \right] + \sigma_{\text{add}}^2}}. \tag{E6}$$

In the special case where $\gamma = \gamma_1 = \gamma_2$, corresponding to the situation where the rising portion of the threshold-versus-external-noise-contrast function has a slope of 1.0, we can solve Equation E6 to obtain threshold signal contrast c_τ as a function of external noise contrast σ_{ext} at a given performance criterion (i.e., d'):

$$c_\tau = \left\{ \frac{d'^2[(1 + N_{\text{mul}}^2)\sigma_{\text{ext}}^{2\gamma} + \sigma_{\text{add}}^2]}{\beta^{2\gamma} - N_{\text{mul}}^2 \beta_2^{2\gamma} d'^2/2} \right\}^{\frac{1}{2\gamma}}. \tag{E7}$$

When the same stimulus (signal + external noise) is passed to the PTM twice, the probability that the two responses are consistent can be derived from Equation B3 (see Appendix B):

$$P_A = \int_{-\infty}^{+\infty} g(x - S, 0, \sqrt{\sigma_{\text{ext1}}^2 + \sigma_{\text{ext2}}^2}) \{ G^2(x, 0, \sqrt{\sigma_{\text{int1}}^2 + \sigma_{\text{int2}}}$$

$$+ [1 - G(x, 0, \sqrt{\sigma_{\text{int1}}^2 + \sigma_{\text{int2}}^2})]^2 \} dx$$

[E1] In the PTM development, the external noise in the stimulus had a Gaussian distribution, corresponding to white external noise. After nonlinear transduction, the distribution of the external noise might deviate from the Gaussian distribution. However, spatial and temporal summation in the perceptual system should reduce this deviation. When 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.

[E2] The perceptual templates in the signal path and the gain-control path could be identical, a form that we used in a number of earlier studies (Dosher & Lu, 1998, 1999, 2000a, 2000b; Lu & Dosher, 1998, 1999, 2000, 2004), in which case $\beta = \beta_2$. This more general form allows for the possibility that the template for the signal may be relatively tightly tuned to the signal stimulus, while the gain control may be more broadly driven. This latter possibility relates to the observation from the physiology and psychophysics that the normalization pools are very broadly tuned.

$$= \int_{-\infty}^{+\infty} g(x - (\beta c)^{\gamma_1}, 0, \sqrt{2}\sigma_{ext}^{\gamma_1}) \times \{G^2(x, 0, \sqrt{N_{mul}^2[2\sigma_{ext}^{\gamma_2} + (\beta_2 c)^{2\gamma_2}] + 2\sigma_{add}^2}) + [1 - G(x, 0, \sqrt{N_{mul}^2[2\sigma_{ext}^{2\gamma_2} + (\beta_2 c)^{2\gamma_2}] + 2\sigma_{add}^2})]^2 dx. \quad \text{(E8)}$$

In all the applications of the PTM approach so far, we have found that the PTM with $\gamma = \gamma_1 = \gamma_2$ has provided adequate description of the empirical data. In the rest of this article, we will restrict our discussion to this reduced set of PTMs. The same logic could be followed to understand the properties of PTMs with $\gamma_1 \neq \gamma_2$. It follows directly from Equation E7 that, for any given external noise contrast $\forall N_{ext}$, the threshold signal contrast ratio between two performance criterion levels (corresponding to d_2' and d_1') is

$$\frac{c_{\tau_2}}{c_{\tau_1}} = \left[\frac{d_2'^2\beta^{2\gamma} - N_{mul}^2\beta_2^{2\gamma}d_1'^2/2}{d_1'^2\beta^{2\gamma} - N_{mul}^2\beta_2^{2\gamma}d_2'^2/2}\right]^{\frac{1}{2\gamma}}. \quad \text{(E9)}$$

Thus, the PTM predicts that threshold signal contrast ratio between two performance criterion levels for any given external noise contrast is a nonlinear function of the corresponding d's, independent of the particular external noise level.

Appendix F

Experimental Methods

Apparatus

The experiment was conducted on a Macintosh Power G4 computer running PsychToolbox extensions (Brainard, 1997; Pelli, 1997). The stimuli were presented on a Hewlett Packard hp91 color monitor with a 120-Hz refresh rate. A special circuit (X. Li, Lu, Xu, Jin, & Zhou, 2003) was used to display monochromatic images on the monitor with high grayscale resolution (>12.5 bits). A lookup table, obtained with a psychophysical procedure and photometric measurements, was used to linearize the luminance levels. Stimuli were viewed binocularly with natural pupils at a viewing distance of approximately 72 cm in dim light. Observers used a chinrest to maintain head position and fixation throughout the experiment.

Observers

Three observers, CC, SJ, and WC, participated in the experiment. All of them had corrected-to-normal vision and were experienced in psychophysical experiments but naïve to the purpose of the experiment.

Stimuli

The signal stimuli were Gaussian-windowed sinusoidal gratings, oriented $\theta = \pm 45°$ from vertical. The luminance profile of the Gabor stimulus is described by

$$L(x, y) = L_0\left\{1.0 + c\sin[2\pi f(x\cos\theta + y\sin\theta)]exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)\right\}, \quad \text{(F1)}$$

where c is the signal contrast, $\sigma = 0.57°$ is the standard deviation of the Gaussian window, and the background luminance L_0 was set in the middle of the dynamic range of the display ($L_{min} = 1$ cd/m^2; $L_{max} = 55$ cd/m^2).

The signal stimuli were rendered on a 64 × 64 pixel grid, extending 2.78° × 2.78° of visual angle. External noise images were constructed using 2 × 2 pixel elements (0.087° × 0.087°). Each noise element's contrast level was drawn independently from a Gaussian distribution with mean of 0 and standard deviation ranging from 0.0 to 0.33. Because the maximum achievable contrast is ±1.0 on the display, a noise sample with standard deviation of 0.33 conforms reasonably well to a Gaussian distribution. In a given trial, external noise images were made of elements with jointly independent, identically distributed contrasts. Eight external noise levels (0, 0.030, 0.045, 0.067, 0.100, 0.149, 0.223, and 0.332) were used in the experiment.

Design

The method of constant stimuli was used to measure psychometric functions in each of the eight external noise conditions. In each external noise condition, the psychometric function was sampled at five different signal stimulus contrast levels, specified for each observer based on pilot data to span the full range of performance levels.

Each observer completed 16 sessions of 480 trials. In each session, all external noise and signal contrast conditions were randomly mixed. Observers first ran four experimental sessions with different random external noise images and trial sequences. The same stimuli and trial sequences were used in the next four sessions (double-pass). New stimuli and trial sequences were used in Sessions 9 to 12. Sessions 13 to 16 repeated the stimuli from Sessions 9 to 12. An experimental session lasted about 15 to 20 min.

Procedure

In the beginning of each trial, a fixation cross was presented in the center of the screen for 250 ms. The subsequent stimulus sequence consisted of three 8.3-ms frames: a noise frame, a signal frame, and another (independent) noise frame. Observers were instructed to identify the orientation of the Gabor stimulus using the computer keyboard. A beep immediately followed each incorrect response. The next trial started half a second after the feedback.

Data Analysis and Statistical Tests

The measured probability correct (P_C) versus probability agreement (P_A) curves for each observer were fit with eight observer

(*Appendix continues*)

models using the maximum-likelihood procedure (Hays, 1988). The eight observer models included the five described in the section in the main text on observer models and three new models: a reduced PTM with $\beta_2 = \beta$, an altered PTM with a linear transducer and decision uncertainty, and a fully saturated model that consisted of a PTM with decision uncertainty. The additional models were included to complete a model lattice and to test variants of the models.

For each observer model, the probability correct (P_{Ci}) and probability agreement (P_{Ai}) were computed from the model for each signal contrast and external noise condition *i.* There were a total of 40 conditions. *Likelihood* is defined as a function of the total number of trials N_i, the number of correct trials K_{Ci}, and the number of pairs of trials with the same response in the two passes of the experiment K_{Ai} in each stimulus condition *i*:

$$likelihood = \prod_{i=1}^{40} \frac{N_i!}{K_{Ci}!(N_i - K_{Ci})!} P_{Ci}^{K_{Ci}}(1 - P_{Ci})^{N_i - K_{Ci}} \prod_{i=1}^{40} \times \frac{(N_i/2)!}{K_{Ai}!(N_i/2 - K_{Ai})!} P_{Ai}^{K_{Ai}}(1 - P_{Ai})^{N_i/2 - K_{Ai}}. \quad \text{(F2)}$$

A MATLAB function, *fminsearch,* was used to find the best fitting parameters for each observer model that maximized log(*likelihood*). Nested models were compared using a chi-square statistic:

$$\chi^2(df) = 2.0 \times \log\left(\frac{\max\ likelihood_{\text{full}}}{\max\ likelihood_{\text{reduced}}}\right), \quad \text{(F3)}$$

where $df = k_{\text{full}} - k_{\text{reduced}}$ is the difference between the number of parameters of the two models.

To derive threshold-versus-external-noise-contrast functions, a Weibull function,

$$P_c = [\max - (\max - 0.5) \times 2^{-\left(\frac{c}{\rho}\right)^{\eta}}] \times 100\%, \quad \text{(F4)}$$

was fit to the psychometric functions (Wichmann & Hill, 2001) using a maximum-likelihood procedure (Hays, 1981), where max, *c,* ρ, and η denote the maximum fraction correct, signal contrast, threshold, and the slope of the psychometric function, respectively. For each observer, we constrained max to be the same across all the external noise conditions. Statistical tests showed that imposing the constraint did not significantly reduce the quality of the fits. Threshold signal contrasts at 65%, 75%, and 85% correct were calculated from the best fitting Weibull functions.

Received March 9, 2007
Revision received October 9, 2007
Accepted October 10, 2007 ■

Correction to Stout and Miller (2007)

The article "Sometimes-Competing Retrieval (SOCR): A Formalization of the Comparator Hypothesis," by Steven C. Stout and Ralph R. Miller (*Psychological Review*, Vol. 114, No. 3, pp. 759–783) contained errors.

In the right column of Table 1, the first equation currently reads as follows:
$\Delta V_{X,O} = \alpha_X * \beta_O(\lambda_O \Sigma - V_j)$.
It should read as follows:
$\Delta V_{X,O} = \alpha_X * \beta_O(\lambda_O - \Sigma V_j)$.

In the right column of Table 1, the 11th equation currently reads as follows:
$\Delta Op_{X,j,O} = X * k3 * V_{X,j} * V_{j,O} * (1.0 - Op_{X,j,O})$, when $V_{X,O} = 0$.
It should read as follows:
$\Delta Op_{X,j,O} = \alpha_X * k3 * V_{X,j} * V_{j,O} * (1.0 - Op_{X,j,O})$, when $V_{X,O} = 0$.

Equation 6B currently reads as follows:
$\Delta Op_{X,j,O} = X * k3 * V_{X,j} * V_{j,O} * (1.0 - Op_{X,j,O})$, when $V_{X,O} = 0$.
It should read as follows:
$\Delta Op_{X,j,O} = \alpha_X * k3 * V_{X,j} * V_{j,O} * (1.0 - Op_{X,j,O})$, when $V_{X,O} = 0$.

DOI: 10.1037/0033-295X.115.1.