Signal and Noise in Perceptual Learning

by

Jason M. Gold

A thesis submitted in conformity with the requirements

for the degree of Doctor of Philosophy

Graduate Department of Psychology

University of Toronto

For Andrea

Signal and Noise in Perceptual Learning

Jason M. Gold, Doctor of Philosophy, 2001

Graduate Department of Psychology, University of Toronto

Abstract

Performance in perceptual tasks often improves with practice. This effect is known as 'perceptual learning', and it has been the source of great interest and debate over the course of the last century. Although much is known about how perceptual learning changes the properties of cortical circuitry, little is known about how these changes manifest themselves at the level of behavior. In this thesis, the behavioral effects of perceptual learning are considered within the context of signal detection theory. Within a signal detection framework, an observer's sensitivity in a perceptual task is defined by the ratio of signal-to-noise within the system. Thus, according to signal detection theory, the improvements that take place with perceptual learning can be due to increases in *internal signal strength* or decreases in *internal noise*. These quantities cannot be measured directly. Instead, psychophysical techniques must be used to infer their magnitudes. In this thesis, two psychophysical techniques were used to discriminate between the effects of signal and noise as observers learned to identify sets of unfamiliar visual patterns. *Noise masking* was used to measure observers' *equivalent input noise* and

calculation efficiency, quantities that correspond to internal noise and internal signal strength, respectively, within the context of a simple black-box model of the visual system. Equivalent input noise only reflects the effects of an internal noise whose magnitude is independent of the magnitude of the stimulus. *Response consistency* was used to estimate the effect of learning on internal noise that depends on the magnitude of the stimulus. Calculation efficiency improved by as much as a factor of four across learning sessions for two very different pattern identification tasks (face and texture identification). However, neither form of internal noise changed significantly with learning. These results were used to test the prediction that an observer's calculation should become more similar to the calculation of an ideal discriminator with practice. *Response classification* was used to estimate observers' linear templates as learning took place, and showed observer's calculations became significantly more correlated with the ideal template with practice. Taken together, these results place new theoretical constraints on models of perceptual learning.

Acknowledgements

First and foremost, I would like to thank Patrick J. Bennett and Allison B. Sekuler for their patience, support and guidance during my years as their student. I would also like to thank Richard Murray for both his friendship and the countless number of important discussions we have had over the past several years; all of the observers that participated in my experiments; Bruce Schneider, Jim Dannemiller and Stan Hamstra for generously providing their time and expertise as members of my thesis committee; and all of the members of the University of Toronto Vision Lab for providing a stimulating and pleasant research environment. Finally, I would like to thank the University of Toronto Department of Psychology, the Government of Canada, and the Vision Science Research Program for their financial support during my graduate studies. This research was funded by National Sciences and Engineering Research Council grants OGP105494 and OGP0042133 to Patrick. J. Bennett and Allison B. Sekuler, 1998-1999 and 1999-2000 Ontario Graduate Scholarships to Jason M. Gold, and a 2000-2001 Vision Science Research Program grant to Jason M. Gold.

CONTENTS

LIST OF TABLES

Chapter 3

LIST OF EQUATIONS

Chapter 1

Chapter 2

Chapter 3

Appendix

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

1 Signal and Noise in Perceptual Learning: General Introduction

This chapter serves as a general introduction to the thesis. It will begin with a brief discussion of the general phenomena associated with perceptual learning. Next, these effects are considered within the context of signal detection theory, and related to the concepts of internal signal enhancement and internal noise reduction. Then, two techniques will be discussed -- *noise masking* and *response consistency* -- that can be used to estimate the magnitude of these quantities. Next, a technique called *response classification* will be discussed that can be used to estimate changes in the classification rules used by an observer as learning takes place in a discrimination task. Finally, an overview of the rationale and conclusions for each of the experiments in the subsequent chapters of the thesis will be provided.

Perceptual Learning

The observation that performance in perceptual tasks can improve with practice has been documented for over a century (Gibson, 1969; Gilbert, 1994). This improvement in perceptual discriminations with training is referred to as *perceptual learning*, and it has at least four characteristic attributes that, taken together, are generally thought to differentiate it from other forms of learning. First, perceptual learning can occur over a wide range of time scales, from within as little as 100 trials (Fahle, Edelman, & Poggio, 1995; Poggio, Fahle, & Edelman, 1992) to as long as several weeks (Fiorentini & Berardi, 1997; Karni et al., 1998; Karni & Sagi, 1993; Sathian & Zangaladze, 1998;

Schoups, Vogels, & Orban, 1995). Second, perceptual learning occurs for a wide variety of perceptual tasks, including very simple sensory discriminations such as visual and tactile acuity tasks (Fahle et al., 1995; Fahle & Morgan, 1996; Poggio et al., 1992; Sathian & Zangaladze, 1998), orientation discrimination (Matthews, Liu, Geesaman, & Qian, 1999; Schiltz et al., 1999; Schoups et al., 1995), motion discrimination (Ball & Sekuler, 1987; Matthews et al., 1999), texture discrimination (Fine & Jacobs, 2000; Karni & Sagi, 1991) and auditory pitch discrimination (Demany, 1985; Recanzone, Schreiner, & Merzenich, 1993). Third, perceptual learning is typically restricted to the exact specifications of the stimuli and task where training has occurred (Ahissar & Hochstein, 1997; Ball & Sekuler, 1987; Crist, Kapadia, Westheimer, & Gilbert, 1997; Fahle & Morgan, 1996; Fiorentini & Berardi, 1980). That is, the learning often does not transfer to other tasks, stimuli, or sensory locations. Finally, observers often do not require feedback in order to exhibit the learning effects described above (Ball & Sekuler, 1987; Fahle et al., 1995; Herzog & Fahle, 1997, 1999).

These last three findings (learning for simple stimuli, stimulus specificity, and implicit learning) have been taken as evidence that perceptual learning occurs at relatively early stages of sensory processing (Gilbert, 1994). As a result, much of the recent psychophysical and physiological work on this topic has been directed toward localizing the neural substrates that mediate perceptual learning in different tasks and modalities. Most of the evidence from these experiments suggests that perceptual

learning for many tasks takes place at or before the level of primary sensory cortex. For example, several studies have found partial or no inter-ocular transfer of learning for simple visual discrimination tasks (Ball & Sekuler, 1987; Fahle et al., 1995), suggesting some of the effects of learning for these tasks occur before cortical processing. Similarly, physiological studies have found that practice changes the response properties of neurons in primary cortical areas for simple discrimination tasks, such as visual orientation discrimination (Schiltz et al., 1999) and auditory frequency discrimination (Recanzone et al., 1993). Other physiological studies have investigated the topographic changes that take place in sensory cortical maps with practice (Buonomano & Merzenich, 1998; Recanzone et al., 1993). These studies have found sensory cortex to be highly plastic, with striking amounts of cortical reorganization and reallocation taking place as a result of extensive training or restricted experience. However, there is also evidence that suggests higher order mechanisms, such as those found in the prefrontal cortex, can also change with perceptual learning (Asaad, Rainer, & Miller, 2000).

Signal and Noise

But what is it about these mechanisms that changes with learning? One way to approach this problem is to consider the effects of learning within the context of *signal detection theory* (Green & Swets, 1966). Signal detection theory is a general framework designed to characterize and quantify an observer's decision processes and sensitivity in a

task. Although it was first developed within the context of problems in radar, it was soon realized that it also could be applied to problems related to human signal detection and recognition (Green & Swets, 1966; Peterson, Birdsall, & Fox, 1954). Unlike other theories (e.g., high threshold theory), signal detection theory assumes an observer's internal responses are probabilistic, so that a particular stimulus has only some probability of eliciting a particular internal response (or, conversely, an internal response only has a certain probability of originating from a particular stimulus). The theory assumes an observer makes decisions relative to some internal response criterion. One of the main assets of this model of the observer is the development of a bias-free estimate of sensitivity (d') that is independent of the observer's choice of criterion. This model of human decision processes makes explicit predictions about the relationship between task performance and criterion, as well as the shapes of the underlying probability distributions. Using these predictions, signal detection theory has been successfully applied to human performance in a wide range of psychophysical tasks (see Green & Swets (1988) for an exhaustive bibliography).

Beyond its incorporation of response bias and internal criterion, a second major feature of signal detection theory is that it can be used to quantify optimal decision processes, which can in turn be used to estimate the amount of information used by a non-optimal (e.g., a human) observer in a given task (Geisler, 1989; Tanner & Birdsall, 1958). Specifically, the decision rule used by an observer that makes statistically optimal

use of all available information can be derived through Bayesian inference (Tjan, 1996). The performance of this decision rule can be used to obtain a measure of ideal sensitivity in a given task. Ideal sensitivity reflects the performance of an observer with no internal constraints, and is thus only constrained by the physical availability of information in a given task. Human sensitivity can then be compared to ideal sensitivity to determine how much of the available information the human observer has used in the task. This method of quantifying the amount of information used by a non-ideal observer is referred to as *ideal observer analysis*, and a more formal treatment of it will be provided in Chapter 2 and in the Appendix.

To understand how these concepts relate to a human observer's behavior in a typical psychophysical experiment, consider the simple task of discriminating between two signals, S_0 and S_1. In a typical discrimination task, an observer might be shown either S_0 or S_1 in a brief interval. The observer's task is to decide which signal had appeared in the interval. Psychologically, the stimulus will produce an internal response within the observer, and this internal response must be used to make a decision as to which signal was shown. An ideal observer will always perform this task perfectly. However, a human observer is not ideal. One way that human and ideal observers differ is in terms of *internal noise*. If an ideal observer is shown the same exact stimulus several times, it will always make exactly the same decision. Unlike an ideal observer, human observers have internal variability or 'noise' (Barlow, 1956, 1957; Green, 1964).

This trial-by-trial variability is thought to originate from a variety of sources, ranging from the stochastic properties of sensory neurons (Croner, Purpura, & Kaplan, 1993; Tolhurst, Movshon, & Dean, 1983) to random fluctuations in strategy (Burgess, 1990; Raghavan, 1989). As a result, the same stimulus will not produce the same internal response on every presentation. Instead, they will produce a distribution of responses across identical presentations, and the variance of this distribution will be determined by the magnitude of the internal noise. In the case of the discrimination task described above, a human observer will have two distributions of internal responses, one for S_0 and one for S_1. Thus, for an observer limited by noise, the discrimination task described above becomes one of determining from which distribution the internal response originated.

A second way that human and ideal observers differ is in their ability to use available information. An ideal observer uses a computation that is guaranteed to make optimal use of all of the information available in a given task. Unlike an ideal observer, human observers perform sub-optimal computations[1]. These inefficiencies can arise from many sources, ranging from sub-optimal encoding by sensory organs (Banks, Geisler, & Bennett, 1987; Geisler, 1989) to the comparison of the sensory representation to a sub-optimal receptive field or template (Legge, Kersten, & Burgess, 1987).

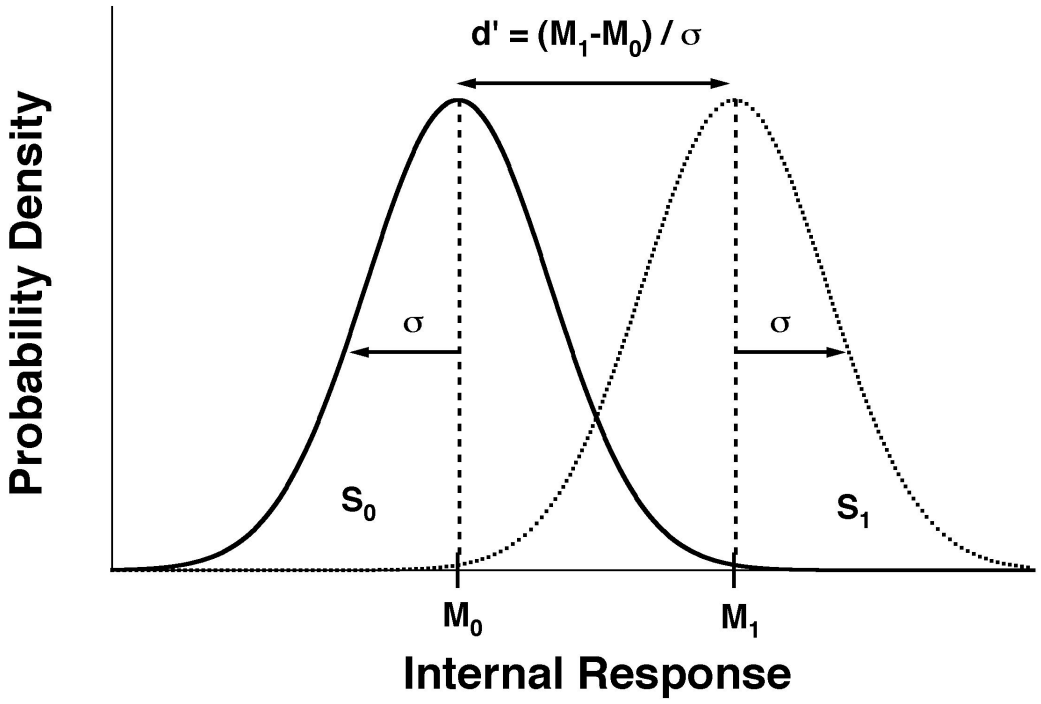


Figure 1.1 Hypothetical internal response distributions for a discrimination task. The leftmost distribution corresponds to signal S_0 and the rightmost to signal S_1. M_0 and M_1 are the means of the S_0 and S_1 distributions, respectively. The distributions are assumed to have equal variances (σ^2). Under these conditions, an observer's sensitivity (d') is equal to the difference between the means of the two distributions normalized by their common standard deviation.

This classical problem of signal-detection is illustrated schematically in Figure 1.1. Figure 1.1 depicts a pair of hypothetical internal response distributions produced by a human observer in the discrimination task described above. Each distribution is centered about a mean response *M* and has a variance of σ^2. A common assumption of signal detection models is that the underlying distributions are Gaussian distributed with

[1] Here, the discussion is restricted to non-stochastic (i.e., deterministic) inefficiencies in computation. Stochastic inefficiencies are more appropriately described as internal noise.

equal variances (Green & Swets, 1966)[2]. Under these conditions, an observer's sensitivity (d') is defined as the distance between the two means, normalized by the variance of the distributions. In this abstraction, one way that an observer's sensitivity can change is by altering the variance of the underlying distributions. This would occur if there were an increase or decrease in internal variability. Notice that sensitivity would also shift if the distance between the means of the internal distributions were to change. Such a shift would not correspond to a change in the stochastic aspects of an observer's computations. Instead, it would reflect a change in the efficiency of the observer's computations. In terms of signal detection theory, a shift in the mean strengths of the internal responses associated with each signal translates into a shift in the strength of the internal signal. Thus, in relation to perceptual learning, the changes in sensitivity that take place with training could be due to a decrease in internal noise, an increase in internal signal strength, or some combination of the two.

Measuring the Strength of Signal and Noise

Until recently, there were no techniques available to discriminate between changes due to internal noise or signal strength. Both kinds of changes predict changes in

[2] The assumption that the internal distributions are Gaussian with equal variance may be tested for a given task by comparing an observer's hits and false alarms for different criteria (Green & Swets, 1966).

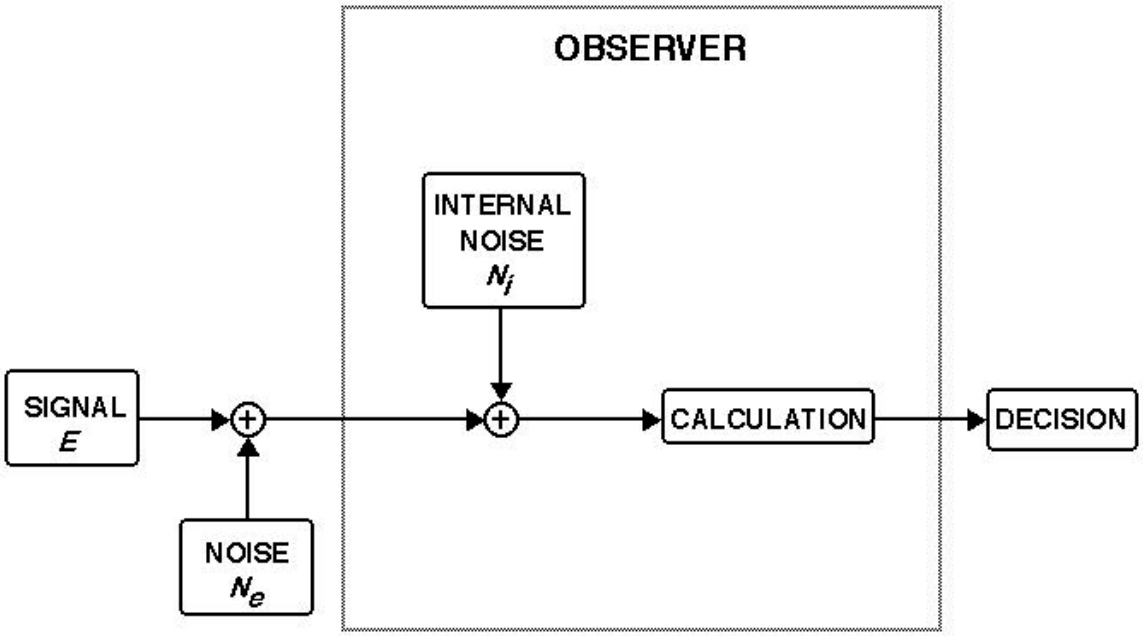


Figure 1.2 A black-box model of a human observer in a perceptual discrimination task (adapted from Pelli, 1981; 1990). The observer is treated like a black-box that receives a noisy external stimulus (E + N_e), introduces a fixed amount of variability to the stimulus (N_i), performs a calculation that is reduced to an internal response, and makes a decision based on the magnitude of the internal response.

sensitivity with learning. However, psychophysical techniques have been developed in recent years that, when used in combination, allow the effects of these quantities to be disambiguated. The techniques are called *noise masking* and *response consistency*.

Noise Masking

Referring the intrinsic noise of an electronic device to an external noise introduced into the system is a standard technique used by electrical engineers (Mumford & Schelbe, 1968). However, Pelli (1981) was among the first to apply a variant of this technique to human information processing. Pelli developed measures that he called an observer's *equivalent input noise* and *calculation efficiency*--quantities that correspond to

an observer's internal noise and internal signal strength, respectively, within the context of his signal detection model. To fully understand these quantities and how they are measured, it is useful to first consider Pelli's abstraction of the internal transformations performed by an observer in a detection or discrimination task. Pelli's 'black-box' model of a linear observer is illustrated schematically in Figure 1.2. In this model, the observer receives a physical stimulus (in this case, a signal corrupted by an externally added noise). The stimulus is converted into an internal representation, where an internal noise of fixed variance is introduced and a calculation is performed on the representation. A decision is then made based on the resulting internal response. Notice that the model characterizes the internal variability and calculation performed by the observer in purely abstract terms: the effects of all sources of variability are treated as arising from a single internal noise, and the effects of all computations are treated as arising from a single computation[3]. The model also assumes that the internal noise is *added* to the representation, that the observer only performs linear transformations on the stimulus until the stage of the decision, and that both the calculation and the variance of the internal noise are invariant with respect to the magnitude of the stimulus. In the

[3] A linear observer with multiple additive noise sources occurring before multiple calculations is mathematically equivalent to an observer that has a single additive noise source equal to the combined variances of the individual noise sources and performs a single combined calculation. The issues involved in the assumptions of linearity and early noise are discussed in Chapter 3.

experiments reported in this thesis, the stimuli are visual images varying in contrast across space. Accordingly, these quantities will be referred to as *contrast-invariant noise* and *contrast-invariant calculations*.

Given this framework, the observer's threshold will be linearly related to the magnitude of the external noise (Legge et al., 1987). This may be formalized as

$$E = k(N_e+N_i) \quad (1.1)$$

where E is the energy of the signal at threshold, N_e is the external noise power spectral density, and k and N_i are free parameters (see Chapter 2 for formal definitions of energy and noise power spectral density). Often, this equation is log-transformed to obtain

$$log(E) = log(k) + log(N_e+N_i). \quad (1.2)$$

The parameter N_i is referred to as the observer's equivalent input noise, and it is equal to the amount of external noise that must be added to the display in order to double the observer's noise-free threshold. The equivalent input noise is expressed in the same units as the external noise. The parameter k is inversely proportional to the goodness or 'efficiency' of the observer's calculation. An observer's *calculation efficiency* is computed

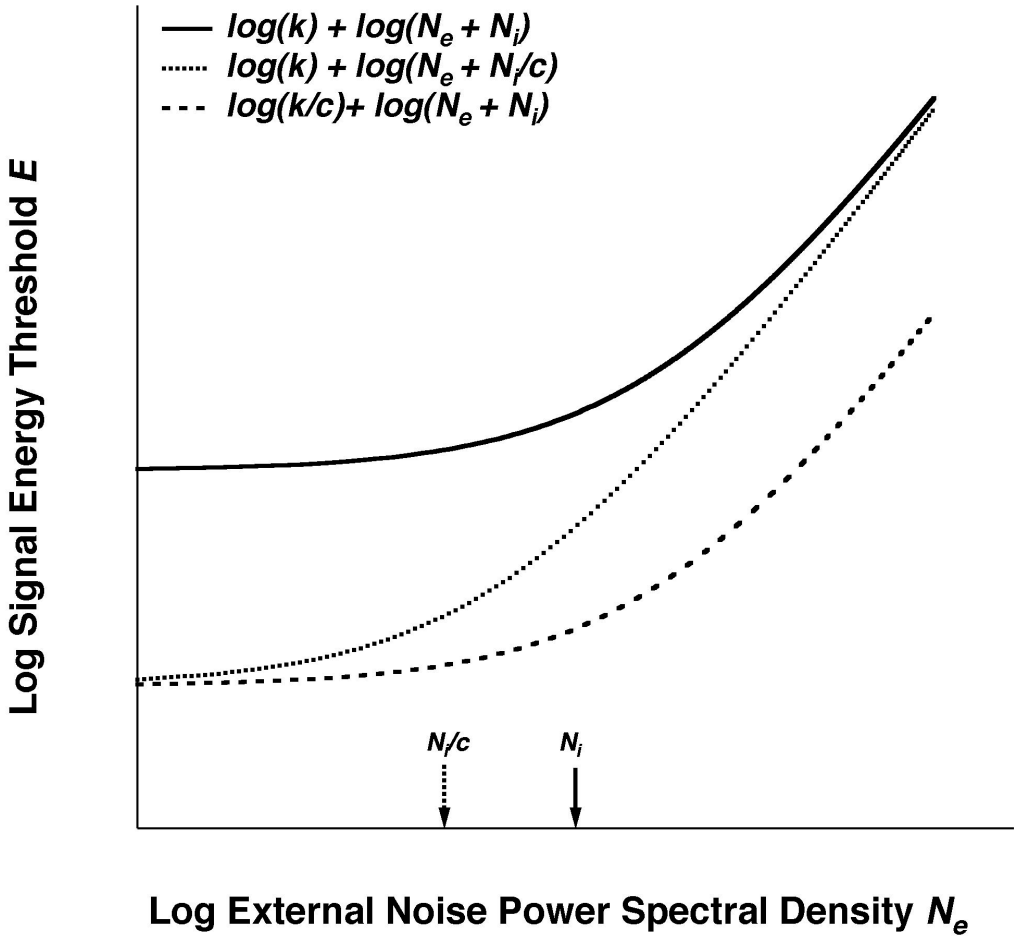


Figure 1.3 Hypothetical noise-masking functions for a human observer. Log of signal energy threshold (E) is plotted as a function of external noise power spectral density (N_e). The finely dashed line depicts a reduction in equivalent input noise N_i by a constant factor c relative to the solid line. The coarsely dashed line depicts an increase in calculation efficiency (indexed by k) by a constant factor c relative to the solid line.

by comparing k for a human observer to that an ideal observer, and is an index of the proportion of the available information used by the observer.

In the context of the signal-detection problem outlined above, equivalent input noise corresponds to the contrast-invariant internal stochastic constraints on performance (i.e., internal noise) and calculation efficiency corresponds to the internal deterministic

constraints on performance (i.e., internal signal strength). Given this model of the observer, it is possible to estimate the magnitude of an observer's internal noise and the efficiency of the observer's calculations by measuring thresholds in various amounts of externally added noise. First, consider the effects of internal noise. In Equation 1.2, it can be seen that, at very low external noise levels, altering the magnitude of the external noise will have little effect on an observer's threshold. It is only once the magnitude of the external noise begins to exceed the magnitude of the internal noise that thresholds will begin to rise. This relationship is illustrated in Figure 1.3. The solid line in Figure 1.3 depicts a hypothetical noise-masking function for a human observer, where log threshold energy (E) is plotted as a function of log external noise power spectral density (N_e). N_i corresponds to the amount of external noise that must be added to the signal in order to double an observer's zero-noise threshold (solid arrow in Figure 1.3). Thus, an observer's internal noise in this model is estimated by finding the magnitude of external noise that is equivalent to adding a fixed (i.e., contrast-invariant) level of noise to the external stimulus (hence the name 'equivalent input noise'). The finely dashed line in Figure 1.3 shows the effects of reducing this contrast-invariant noise N_i by a constant c. Changing N_i in this fashion reduces thresholds at only low external noise levels, shifting the 'kink point' of the noise-masking function to a lower value (the dashed arrow in Figure 1.3). Now consider the effects of calculation efficiency. In Equation 1.2, it can be seen that the index of calculation efficiency k will not have differential effects across

external noise levels. Instead, *k* will determine the overall 'height' of the noise masking function in log-log coordinates. The coarsely dashed line in Figure 1.3 shows the effects of reducing *k* by the constant *c*. Changing *k* in this fashion will shift the noise-masking function down uniformly at all levels of external noise in a log-log plot.

Equivalent input noise and calculation efficiency have been measured for a wide variety of tasks, including grating detection (Pelli, 1981), contrast discrimination (Legge et al., 1987); letter discrimination (Pelli & Farell, 1999; Raghavan, 1989; Tjan, Braje, Legge, & Kersten, 1995), object recognition (Tjan et al., 1995), divided attention (Dosher & Lu, 2000; Lu & Dosher, 1998), and motion discrimination (Lu, Liu, & Dosher, 2000). With one possible exception (Lu & Dosher, 1999), the form of the noise masking functions have conformed well to the model, showing a linear relationship between energy thresholds and external noise power spectral density.

However, there may be other sources of noise in the sensory systems that are not invariant with respect to stimulus magnitude (i.e., the combined magnitude of the signal and external noise). The effects of such a *contrast-dependent* internal noise in Pelli's black-box model can be seen by including a second independent noise source in Equation 1.1:

$$E = k(N_e+N_i + m(N_e+N_i+E)^P). \quad (1.3)$$

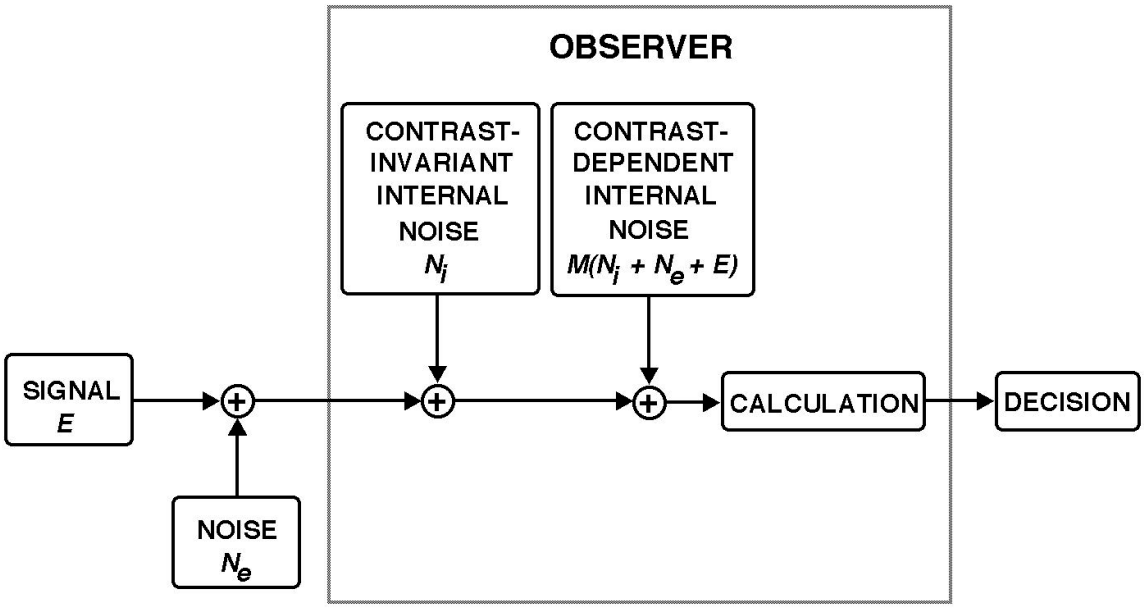


Figure 1.4 Revised black-box model of a human observer in a perceptual discrimination task. The model is identical to Figure 1.2, except for the inclusion of a contrast-dependent internal noise that depends upon the combined magnitude of the external stimulus and the contrast-invariant internal noise.

where the proportionality constant *m* and the exponent *P* determine the magnitude of the contrast-dependent internal noise. However, there is both physiological (Tolhurst, Movshon, & Dean, 1983) and psychophysical (Burgess & Colborne, 1988) evidence that the contrast-dependent noise is directly proportional to stimulus magnitude (i.e., the exponent *P* in Equation 1.3 is equal to unity). The effects of a proportional noise in Pelli's black-box model can be seen by setting *P* to unity in Equation 1.3:

$$E = k(N_e+N_i + m(N_e+N_i+E))$$

$$= (k(1+m)/(1-km))(N_e+N_i)$$

$$= k'(N_e+N_i)$$

$$log(E) = log(k') + log(N_e+N_i) \quad (1.4)$$

where k' is a constant equal to $k(1+m)/(1-km)$. Notice that k' is affected by changes in both proportional noise (m) and calculation efficiency (k), making these two factors confounded in the context of Pelli's black-box model. Pelli (1990) is explicit about this aspect of the model, and assumes that any proportional noise stems from the stochastic properties of the contrast-invariant calculation (i.e., random changes in the calculation across trials). However, there may be sources of proportional noise other than a noisy calculation (Lillywhite, 1981). Thus, contrast-dependent noise is included as a separate source of internal noise in the revised black-box model shown in Figure 1.4.

Response Consistency

Green (1964) and others (Burgess & Colborne, 1988; Spiegel & Green, 1981) have devised a method of measuring internal noise that is independent of the deterministic operations of the observer. The technique is called *response consistency*, and it takes advantage of the fact that internal noise will cause trial-by-trial variability in an observer's internal responses to identical stimuli. Consider again the task outlined above, where observers must identify a signal presented in external noise. If the signal and noise shown on every trial of the experiment were recorded and then the exact same trial-by-trial sequence was shown a second time, the task would be physically identical in

both passes through the experiment. The responses of a noiseless observer would also be identical on each corresponding trial in the sequence. However, for an observer with internal noise, there would be response inconsistency between corresponding identical trials in the two passes. The ratio of internal to external noise would determine the degree of inconsistency between the two passes. The contrast-dependent component of an observer's internal noise can be estimated by measuring response consistency under conditions of high external noise (where the contribution of the contrast-invariant internal noise will be relatively negligible). Similarly, the contrast-invariant component of an observer's internal noise can be estimated by measuring response consistency under conditions of low external noise (where the contrast-invariant internal noise dominates)[4]. In the context of perceptual learning, response consistency offers a way of measuring the ratio of internal to external noise as a function of learning independently of changes in an observer's calculation efficiency.

Measuring the Calculation

As described below in the overview, the results of the first series of experiments in the thesis show that calculation efficiency changes with perceptual learning and the

[4] Note that measures of response consistency include the effects of photon noise, which cannot be reproduced on corresponding trials. Photon noise is often included in estimates of internal/external noise ratios and equivalent input noise because of its relatively negligible effects.

amount of both contrast-invariant and contrast-dependent internal noise remains fixed. But what is it about observers' calculations that changes with practice? The measure of calculation efficiency only provides a gross index of the goodness of the calculation. It does not specify the nature of the transformations performed by the observer.

A technique developed by Ahumada and his colleagues (Ahumada & Lovell, 1971; Beard & Ahumada, 1998; Watson & Rosenholtz, 1997) called *response classification* offers a way of addressing this problem. Consider once again the identification task outlined above. On some trials, an observer will incorrectly classify the stimulus. For example, on some trials the observer will respond that the signal was S_1 when, in fact, the signal shown was actually S_0. If the signal was embedded in a large amount of external noise, there are two possible reasons for this mistake. One possibility is that internal contrast-dependent noise was high, causing the observer to misclassify the stimulus. A second possibility is that the external noise was distributed in such a way to make the stimulus look more like S_1 than S_0. As long as the internal contrast-dependent noise is not excessively high, the external noise will affect an observer's classifications in this fashion on many of the trials. The noise fields shown on each trial can be recorded and classified according to the identity of the signal shown and the response of the observer. After many trials, these noise fields can be averaged in each signal-response category and summed across categories in such a fashion as to produce a *classification image* (see Chapter 2 for a more formal treatment of computing classification images).

The classification image is a map that shows the locations in the stimulus waveform that have affected an observer's responses during the experiment. More specifically, it shows the correlation between the noise magnitude at each location in the waveform and an observer's responses throughout the experiment. In effect, it is an estimate of the linear classification rule used by the observer.

The response classification technique has been applied to a variety of tasks, including auditory (Ahumada & Lovell, 1971; Ahumada, Marken, & Sandusky, 1975) and visual (Abbey, Eckstein, & Bochud, 1999; Ahumada & Beard, 1999) detection, vernier acuity (Ahumada, 1996; Ahumada & Beard, 1998), letter discrimination (Watson, 1998; Watson & Rosenholtz, 1997), stereo vision (Neri, Parker & Blakemore, 1999) and visual completion (Gold, Murray, Bennett, & Sekuler, 2000). In the context of perceptual learning, the response classification technique offers a way of specifying the nature of the changes that occur in an observer's calculation over the course of training.

Overview of the Thesis

This section provides an overview of the main experimental results and conclusions of the thesis. The perceptual tasks in the thesis will be restricted to the recognition of 2-dimensional visual patterns. The general goal of the thesis is to specify the changes that take place with perceptual learning within the context of the signal

detection framework described above. The overall conclusion is that it is only calculation efficiency that changes with perceptual learning.

Chapter 2 covers most of the methods, terminology, and units used in the thesis, although a few details have been left for Chapters 3 and 4. The first part of Chapter 3 concerns the measurement of equivalent input noise and calculation efficiency as a function of learning in two visual pattern identification tasks. This first set of experiments trace the changes in both of these quantities as observers learned to identify 10 unfamiliar human faces or 10 unfamiliar abstract textures (band-pass filtered noise fields). The results of these experiments show over a factor of 4 increase in calculation efficiency for the face stimuli and over a factor of 2 increase for the texture with no corresponding changes in equivalent input noise. In addition, several other aspects of the data are explored, including shifts in linearity and uncertainty across learning sessions, within-session learning effects, and analysis of the face identification data according to gender rather than individual items.

The second part of Chapter 3 addresses the possibility that contrast-dependent internal noise contributed to the changes in calculation efficiency observed in the first set of experiments. Response consistency is used to measure the effects of both contrast-invariant and contrast-dependent internal noise, as a function of learning. Specifically, response consistency is measured for both the face and texture stimuli, in both high and low external noise. The results show response consistency does not change with learning

in the presence of either high or low external noise, even though performance improves in the same fashion as in the first experiments. These results rule out the possibility that either contrast-invariant or contrast-dependent internal noise changes with learning in these tasks, implying the perceptual learning is due only to changes in calculation efficiency. In addition, the predictions of a 'late' noise model (i.e., a model that assumes all of the internal noise occurs after the calculation) are considered within the context of these results.

The fourth and final chapter involves using the response classification technique to explore the changes that take place in observers' calculations as a function of learning. Classification images are measured across a series of learning sessions with two new faces and two new band-pass textures[5]. The main statistical test involves correlating a human observer's classification image with the classification image of an ideal observer. An ideal observer's classification image reveals the relative informativeness of each pixel in the identification task. Correlating the human observer's classification image with the ideal classification image offers an index of how the observer's calculation changes with

[5] The reason for reducing the number of stimuli in each set is to increase the statistical power of the analyses. The response classification technique requires a large number of trials to produce a clear image of an observer's classification rules. Reducing the number of items to be learned increases the power of the response classification technique, because the noise fields are sorted into a smaller number of stimulus-response categories. Despite this improvement in power, learning takes place over a relatively small number of trials, so the analyses in Chapter 4 often rely upon statistical tests rather than visual inspection of the classification images (see Chapter 2 for details on computing classification images).

learning. The results of the first two experiments show that the correlation with the ideal template increases with learning, indicating human classification rules become more ideal with training. A second experiment measures noise masking functions and response consistency with the new face and texture stimuli, and shows that the changes in signal and noise found in the previous experiment with 10 stimuli extend to a task with only two stimuli.

2 Methods

This chapter describes the general experimental methods used throughout the thesis. Some of the methodological details specific to particular experiments are discussed within their respective chapters.

Apparatus

Stimuli were displayed using a Macintosh G3 computer on a 13" Apple high resolution RGB color monitor. The monitor displayed 640 x 480 pixels, which subtended a visual angle of 12.9 x 9.6 degrees from the viewing distance of 100 cm, at a frame rate of 67 Hz (non-interlaced). Luminance calibrations were performed with a Hagner Optikon universal spot photometer, and the calibration data were used to build a linearized 1779-element look-up table (Tyler, Chan, Liu, McBride, & Kontsevich, 1992). The experiment was conducted in the MATLAB programming environment (version 5.1), using in-house software and the extensions provided by the Psychophysics Toolbox (Brainard, 1997) and the Video Toolbox (Pelli, 1997). When constructing the stimuli used on each trial, the computer software selected appropriate luminance values from the calibrated look-up table and stored them in the 8-bit hardware look-up table of the display. Luminance on the display ranged between 0.3 and 80.2 cd/m^2. Average luminance was defined as the luminance produced by the RGB combination (160,160,160), which was 28.8 cd/m^2. Pixel contrast (as defined by equation 1 below) on the display could be varied between -1.0 and 1.8.

Signals

There were two classes of signals used in the experiments. The first class of signal were digital images (256 x 256 pixels in size) of human faces that were constructed using Adobe Photoshop (version 3.0) and MATLAB. The second class of signals were randomly generated band-pass filtered Gaussian noise fields (also 256 x 256 pixels in size) generated using MATLAB (see below for details about the face and texture images). All of the images were created prior to the experiment and stored on disk. The values in each image represented the contrast (c_i) at pixel location *i*, defined by Equation 2.1:

$$c_i = \frac{l_i - L}{L} \tag{2.1}$$

where L is average luminance and l_i is the luminance of the i^{th} pixel. Each image file was normalized so that root-mean-square (RMS) contrast of the image equaled 1. RMS contrast is defined as

$$c_{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} c_i^2} \tag{2.2}$$

where *n* is the number of pixels in the image. Prior to each experimental session, the appropriate image set was read into memory. On each trial, the contrast of the image to be displayed was set to the desired value by multiplying the image data by an appropriate constant, and the contrast values were converted to luminance values. These luminance values were used to construct a linear 8-bit look-up table for the display. Finally, the image luminance values were mapped onto the values stored in the look-up table.

Faces

There were two sets of faces used in the experiments reported in this thesis. The first set, used in the noise masking and response consistency experiments reported in Chapter 3, consisted of 5 male and 5 female Caucasian faces. The second set, used in the response classification experiments in Chapter 4, consisted of two Caucasian male faces. All face models were members of the University of Toronto Department of Psychology. Each face was photographed in front of a uniform black field. Glasses, makeup, and any other non-facial cues were removed from the models' faces before being photographed. Each model's hair was held away from the face and forehead by a small head cap. None of the models had facial hair. All models were asked to look directly at the camera with a neutral facial expression. The film was developed directly to photographic CD-ROM, and each picture was digitally converted to grayscale and cropped to show only the inner portion of the face, eliminating non-facial cues such as hair and ears. The shape of the

visible region of each face was elliptical, and the size and height:width ratio were constant across all stimuli (198 pixels:140 pixels; 4.0° x 2.9°). The faces were centered within a 256 x 256 pixel (5.25° x 5.25°) background of zero contrast (i.e., average luminance).

The contrast values for each face were first linearly transformed so that they ranged from -1 to 1 and the background was set to zero. Next, differences in the amplitude spectra of the faces within each set were eliminated in the following way: First, the Fourier transform of each face was computed, and the modulus at each spatial frequency and orientation was averaged across all faces. After averaging, the DC component was set to zero. Finally, the amplitude spectrum for each face was replaced by the average amplitude spectrum, and the inverse Fourier transforms were computed. For each set of faces, the result of this process was a collection of faces with identical amplitude spectra (Figure 2.1)[6,7].

[6] The reasons for eliminating amplitude differences across faces are twofold. First, these faces had been generated in previous work to remove differences in the relative amplitude across frequency bands (Gold, Bennett & Sekuler, 1999). Second, amplitude differences across faces could lead to differences in stimulus detectability, which could be used as cues to identify the faces.

[7] The amplitude spectra of the original faces differed only slightly from each other, and so the appearance of faces in Figure 2.1 did not differ significantly from the original items.

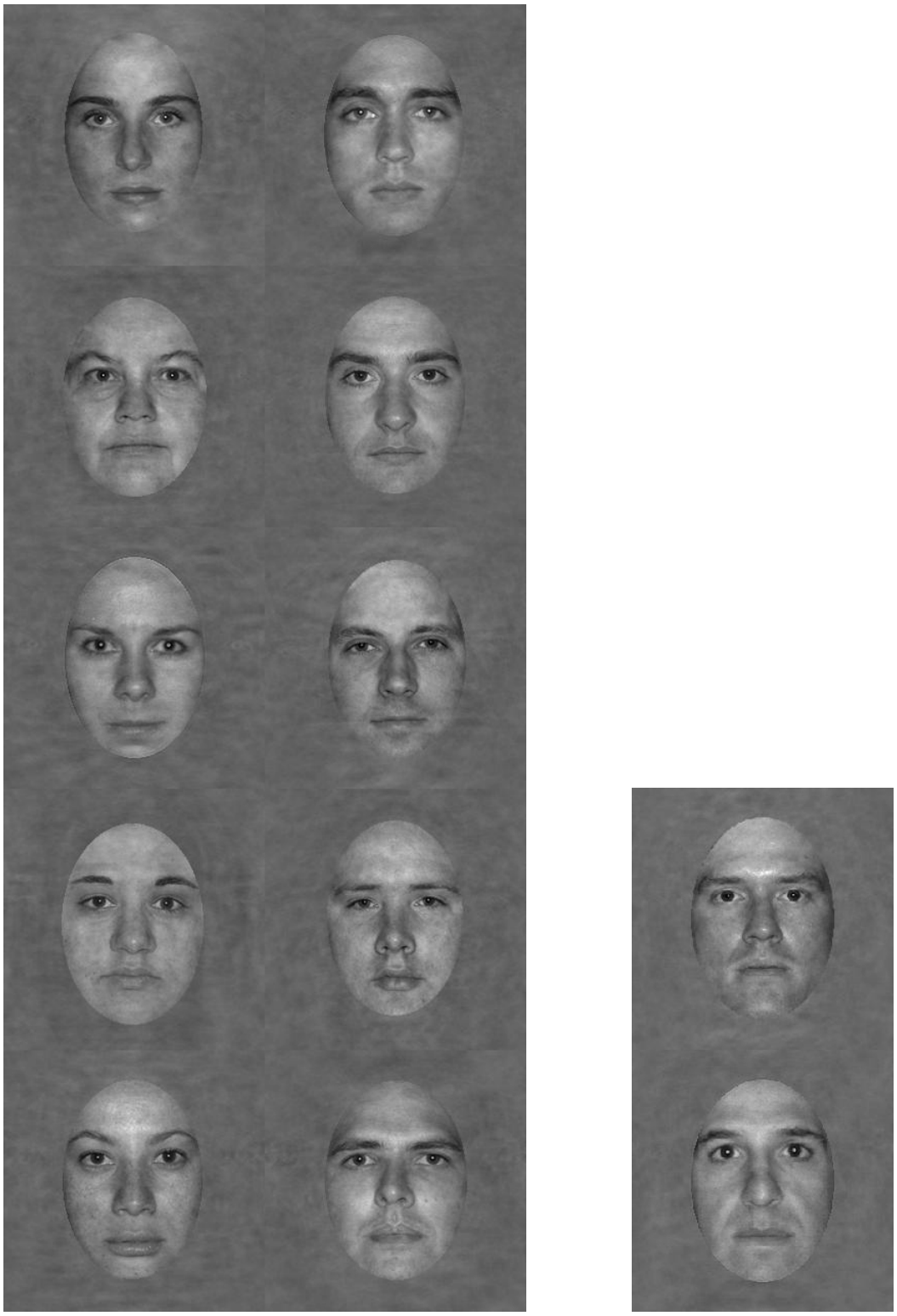


Figure 2.1 The two sets of face stimuli used in the experiments. The 10 faces on the left all share a common Fourier amplitude spectrum, as do the two faces on the right (see text for details).

Textures

Texture patterns were produced by randomly generating 256 x 256 pixel (5.25° x 5.25°) Gaussian white noise fields in MATLAB (see Display Noise below for details). The noise fields were converted into the spatial frequency domain, and filtered with a 2-4 cycle per image (c/image) ideal filter. The amplitude of all frequencies outside of the 2-4 c/image pass-band (including the DC component) were set to zero, and the amplitude within the pass-band remained unchanged. The images were then converted back into the spatial domain[8]. This low-frequency filtering produced the sets of blob-like textures, shown in Figure 2.2. As with the face stimuli, there were two sets of textures used in the experiments reported in this thesis. The first set, used in the noise masking and response consistency experiments reported in Chapter 3, consisted of 10 unique textures. The second set, used in the response classification experiments in Chapter 4, consisted of two unique textures.

Display Noise

The techniques used in this thesis rely heavily on corrupting signals with externally added noise. In all of the experiments reported in this thesis, the external noise

[8] Unlike the face stimuli, the amplitude spectra of the textures were not averaged. This is because Gaussian noise has equal power (on average) at all frequencies and the same filter was applied to each noise field. This made the differences in power across the texture patterns negligible.

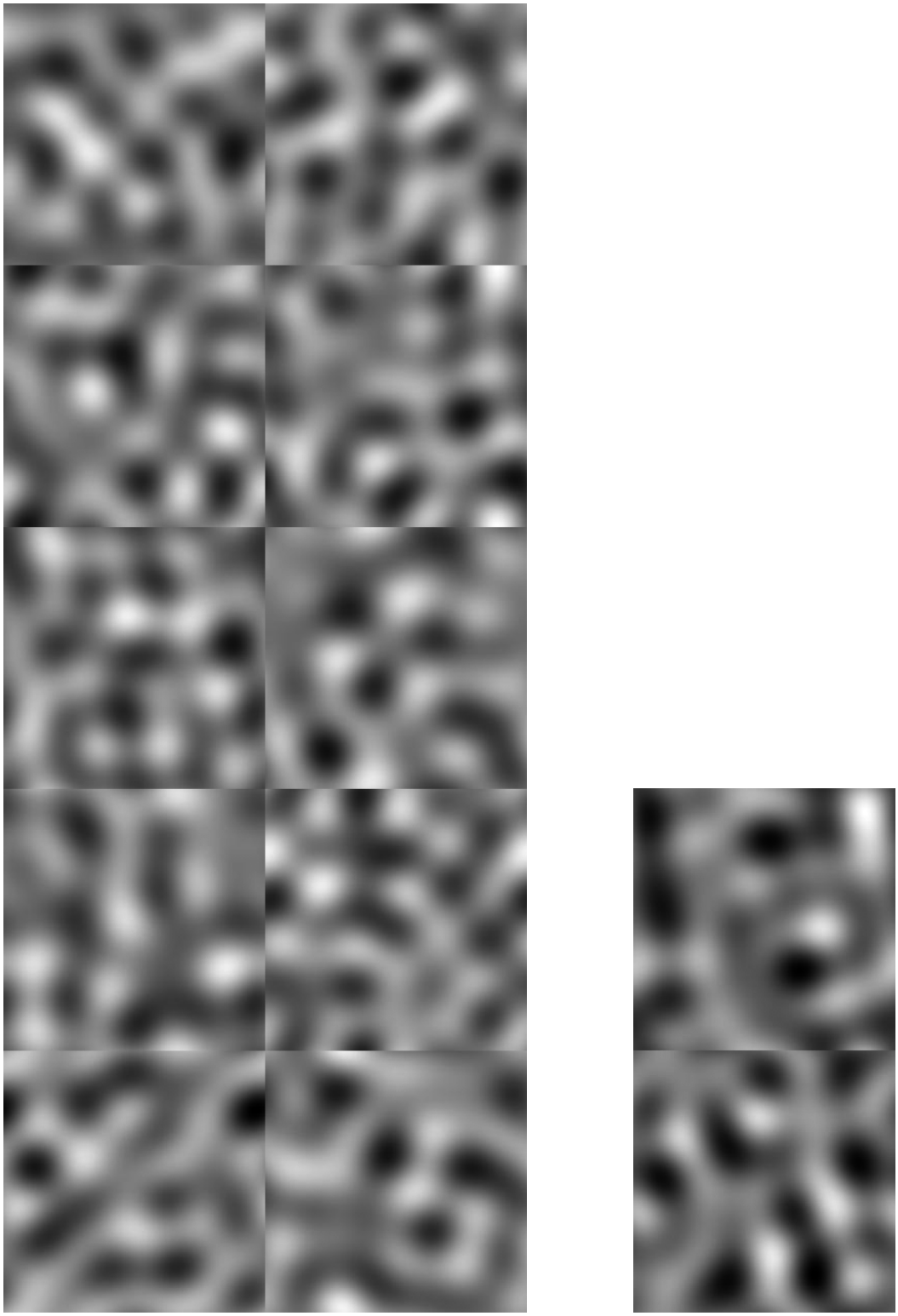


Figure 2.2 The two sets of texture stimuli used in the experiments. Each texture is a randomly generated Gaussian noise field, filtered by a 2-4 cycle per image rectangular frequency filter (see text for details).

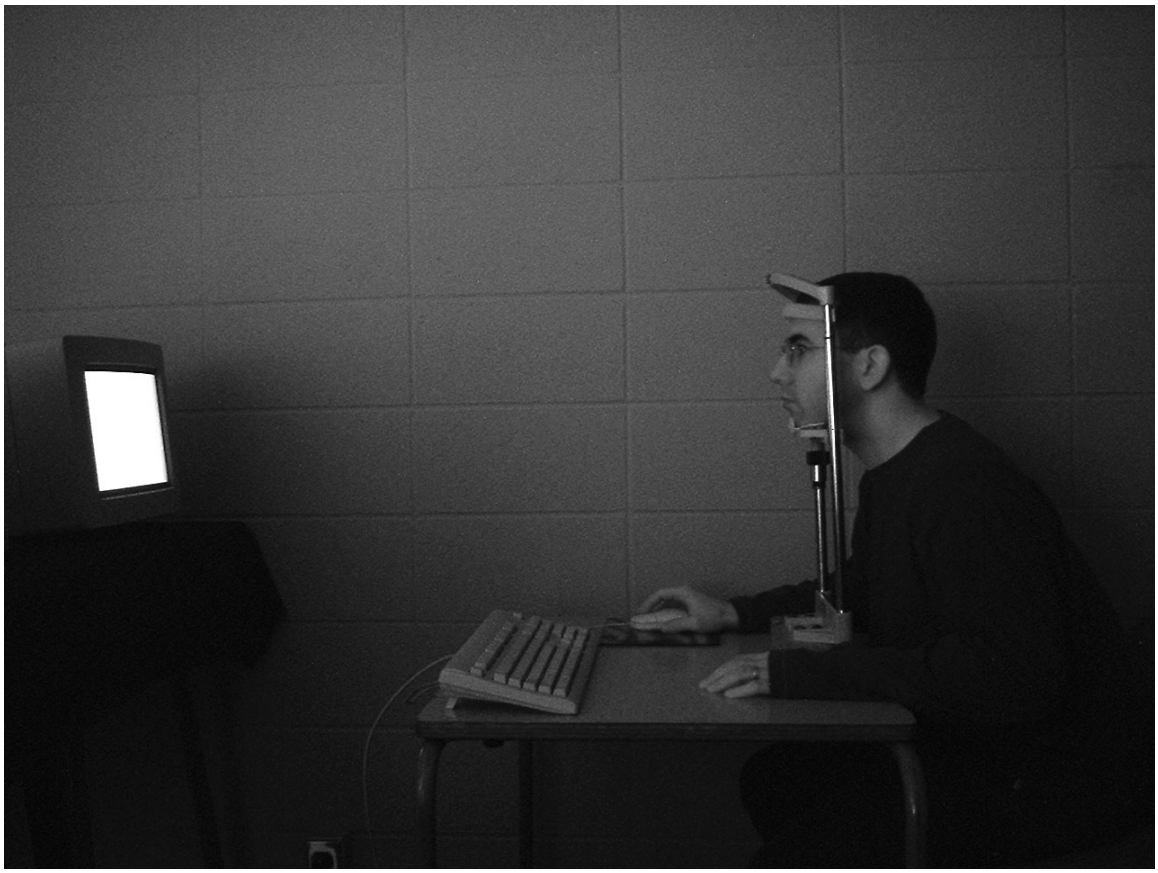


Figure 2.3 Viewing conditions during the experiments. The observer was positioned at a distance of 100 cm from the CRT, with his or her head stabilized by a chin rest.

added to the signal at each pixel was created by drawing a random sample from an independent Gaussian distribution of contrast values, with a mean of 0 and a variance as required by the condition. The values were produced in MATLAB with a pseudo-random number generator. The noise on each trial was static (i.e., did not temporally vary during the course of a trial), white (i.e., the power spectral density did not vary with frequency) and was matched to the size of the signal (256 x 256 pixels). Values beyond +/-2 standard deviations from the mean were discarded and replaced by random samples from the remaining contrast values.

Viewing Conditions

Viewing was binocular through natural pupils, and a headrest/chinrest stabilized the observer's head throughout the session (see Figure 2.3). The monitor supplied the only source of illumination during the experiment.

Human Observers

All participants had normal or corrected-to-normal visual acuity (self-reported). Participants ranged from 17 to 35 years of age, with a mean age of 24 years.

Tasks and Procedure

All experiments involved single-interval, 1-of-*m* identification tasks, where *m* corresponds to the number of signals in the set. The signal energy and power spectral density of the noise were varied according to the procedures detailed in the threshold estimation section below. Observers were familiarized briefly with high contrast versions of the stimuli before the beginning of each experiment. At the start of each trial, a fixation point appeared at the center of the screen (3 x 3 pixels in size), and a brief tone indicated a trial could commence with a mouse click. After the mouse was clicked, the stimulus (signal + noise combination) appeared for 34 frames (approximately 500 ms). Next, the display was set to average luminance, and after a brief 100 ms pause, 100 x 100 pixel high contrast thumbnail versions of the *m* possible signals appeared on the screen

surrounding the region where the stimulus had been displayed. Observers identified the stimulus by clicking the mouse on the appropriate image. Once an image was chosen, the displays were cleared and set to average luminance. Auditory feedback after each trial indicated the accuracy of the response.

Threshold Estimation

In all of the experiments reported in this thesis, identification thresholds at each level of external noise were measured by varying signal energy across trials. There were two different psychophysical methods used to estimate observer thresholds.

In the noise-masking experiments reported in Chapter 3, signal energy was manipulated according to the method of constant stimuli. Pilot studies identified several signal energy levels that spanned the threshold range for a typical unpracticed observer in each level of external noise power spectral density. Signal energy E is defined as

$$E = (c_{RMS})^2(n)(a) \tag{2.3}$$

where n is the number of image pixels and a is the area of a single pixel, in degrees of visual angle squared. Noise power spectral density N is defined as

$$N = \sigma^2 a \tag{2.4}$$

where a is as defined above and σ is the standard deviation of the noise, expressed in values of contrast. Signal energy levels were adjusted after each session for each observer as required by their rate of learning. Before each trial, a signal was chosen randomly to appear within the stimulus interval. There were the same number of trials at each signal energy level. Trials within each session were completely randomized with respect to signal energy and noise power.

In the response consistency experiments reported in Chapter 3 and the response classification experiments reported in Chapter 4, signal energy was manipulated using a UDTR ('up-down-transformed-response') adaptive staircase procedure. In the case of the response consistency experiments, two interleaved staircases were used to obtain measurements that spanned the range of the psychometric function. In the case of the response classification experiments, a single staircase was used to maintain a relatively constant level of performance throughout the experiment and to avoid presenting trials that fell largely outside of the threshold range. Signal energy levels were chosen that coarsely sampled a range of several log units. The staircase shifted through these levels according to the accuracy of the observer's responses. In the response consistency experiments, the first staircase used a 1-up-1-down rule (i.e., when the observer made an incorrect response, the signal energy level was increased by 1 sample; when the observer made a correct response, the signal energy level was reduced by 1 sample). The second staircase used a 1-up-2-down rule (i.e., when the observer made an incorrect response,

the signal energy level was increased by 1 sample; when the observer made two consecutive correct responses, the signal energy level was reduced by 1 sample). The 1-up-2-down rule was also used in the response classification experiment. The staircases maintained this process throughout each experimental session.

For all of the experiments, psychometric functions were estimated by maximum-likelihood fits to the data. The fitting function was of the form

$$p = 1 - (1 - \gamma)e^{\left(\frac{E}{\alpha}\right)^{\beta}} \tag{2.5}$$

where p is percent correct, E is signal energy, γ is the guessing rate, and α and β are free parameters. Threshold was defined as the signal energy yielding either 50% correct responses (in the case of 10 possible signals) or 71% correct responses (in the case of two possible signals). Confidence intervals for the fitted parameters and threshold estimates were calculated by bootstrap simulations. Each simulation consisted of at least 500 simulated data sets.

Ideal Observer

The ideal decision rule for the tasks and stimuli reported in this thesis is

$$\max_{j=1..m}\{\sum_{i=1}^{n} T_{ij}R_i\} \tag{2.6}$$

where m is the number of possible signals, n is the number of pixels in each signal, T_{ij} is the i^{th} pixel in the j^{th} noise-free normalized signal, and R_i is the i^{th} pixel in the noisy stimulus. This rule amounts to maximizing the cross-correlation between the stimulus (i.e., signal + noise combination) and each of the m possible signal matrices (templates). This rule is proven in the Appendix. Ideal observer thresholds were obtained in all conditions through Monte Carlo simulations, in which each template was compared to the stimulus at a range of signal energy values for each corresponding noise level tested with human observers. For each trial, the ideal observer simply chose the template that yielded the highest cross-correlation with the stimulus. Ideal thresholds were estimated from psychometric functions that were fit to the data (using the procedure described above in the section on threshold estimation) from at least 10,000 simulated trials.

Equivalent Input Noise and Calculation Efficiency

Recall from Chapter 1 that an observer's equivalent input noise and calculation efficiency are estimated by measuring signal identification energy thresholds across a range of external noise power spectral density levels. Equation 1.1 is fit to the thresholds, with the negative x-intercept N_i as the estimate of contrast-invariant internal noise and the slope k as an index of efficiency. As with the external noise, N_i is expressed in units of power spectral density.

It can be shown (Tjan et al., 1995) that the ideal observer's signal identification energy threshold is a linear function of noise power spectral density, i.e.

$$E_{ideal} = k_{ideal} N_e \tag{2.7}$$

where N_e is the power spectral density of the external noise. The slope parameter k_{ideal} varies with the set of signals and is directly related to the intrinsic difficulty of the task (i.e., the similarity of the templates). The human observer's calculation efficiency η is defined as

$$\eta = k_{ideal}/k. \tag{2.8}$$

Fits to both ideal and human thresholds were estimated by maximum-likelihood minimization, and bootstrap simulations provided confidence intervals for the fitted parameters (minimum 500 simulated experiments).

Internal / External Noise Ratio

This section describes the methods used in this thesis to estimate an observer's internal/external noise ratio (*I/E*). Recall from Chapter 1 that a difficulty associated with interpreting the estimates of calculation efficiency obtained using the methods above is that the parameter *k* is influenced by both the efficiency of internal calculations and any noise in the system that grows in proportion to the magnitude of the stimulus. Thus, changes in *k* can be due to changes in calculation efficiency, contrast-dependent internal noise, or both. However, response consistency can be used to tease these influences apart. Specifically, response consistency can be used to estimate an observer's *I/E* at high external noise levels (where internal contrast-dependent noise dominates) and low external noise levels (where internal contrast-invariant noise dominates). These measures reveal to what degree internal contrast-dependent noise contributes to the magnitude of *k* (assuming it is proportional to stimulus magnitude) and whether internal contrast-dependent noise changes as learning occurs.

The response consistency experiments reported in this thesis involved having observers make two consecutive passes through identical experiments within a given

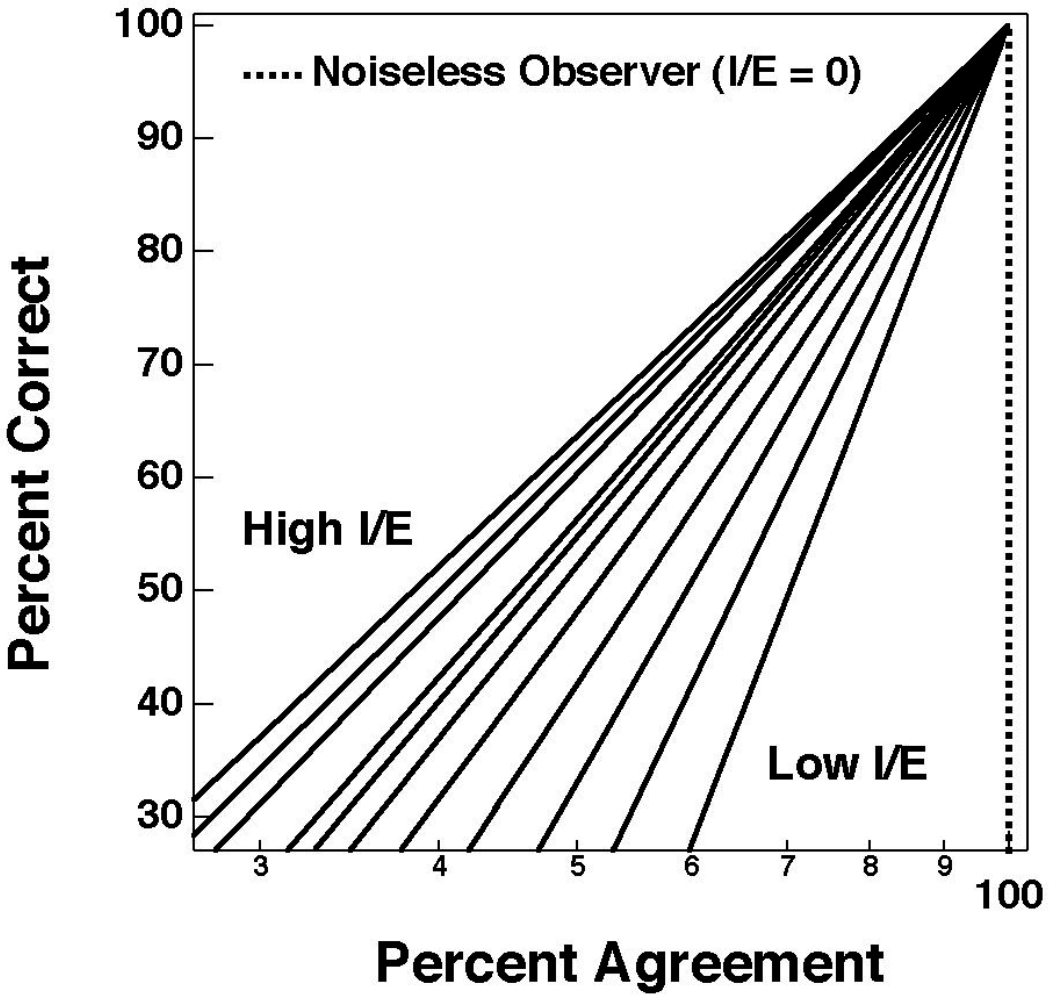


Figure 2.4 Hypothetical curves demonstrating the relationship between percent correct and percent agreement in a double-pass response consistency experiment for an observer with various internal/external noise ratios. Each line is described by Equation 2.9. The rightmost dashed line corresponds to a noiseless observer. Each solid line corresponds to the performance of a noisy observer with a particular internal/external noise ratio. The slope of the line is inversely related to the internal/external noise ratio.

experimental session. Specifically, observers were shown a particular sequence of stimuli in the first half of the session (according to one of the threshold estimation procedures discussed above). In the second half of the session, the exact same stimuli were reproduced in exactly the same sequence. An observer with no internal noise would be perfectly consistent between corresponding trials of the two passes through the

experiment. If we were to plot the percentage of agreement of responses between the two sessions as a function of percent correct responses for such a noiseless observer, it would look like the dashed line shown to the right in Figure 2.4. However, if this observer were to have internal noise, percent agreement would decrease for each level of percent correct (Burgess & Colborne, 1988). The performance of a noisy observer would be well approximated by one of the solid lines shown to the left in Figure 2.4. These lines follow the form

$$pc = k_{I/E} log_{10}(pa/100) + 100 \quad (2.9)$$

where *pc* is percent correct performance at a given level of signal energy, *pa* is the percent agreement between the two runs, and $k_{I/E}$ is a free parameter. Each line corresponds to the performance of an observer with a particular *I/E*. Larger values of *I/E* produce lower values of $k_{I/E}$. Thus, a particular *I/E* will coincide with an observed level of inconsistency. The exact relationship between *I/E* and $k_{I/E}$ is task dependent, but in the tasks reported here, it is has been fit by a function of the form

$$I/E = \alpha + \gamma_1 e^{-\beta_1 k_{I/E}} + \gamma_2 e^{-\beta_2 k_{I/E}} \quad (2.10)$$

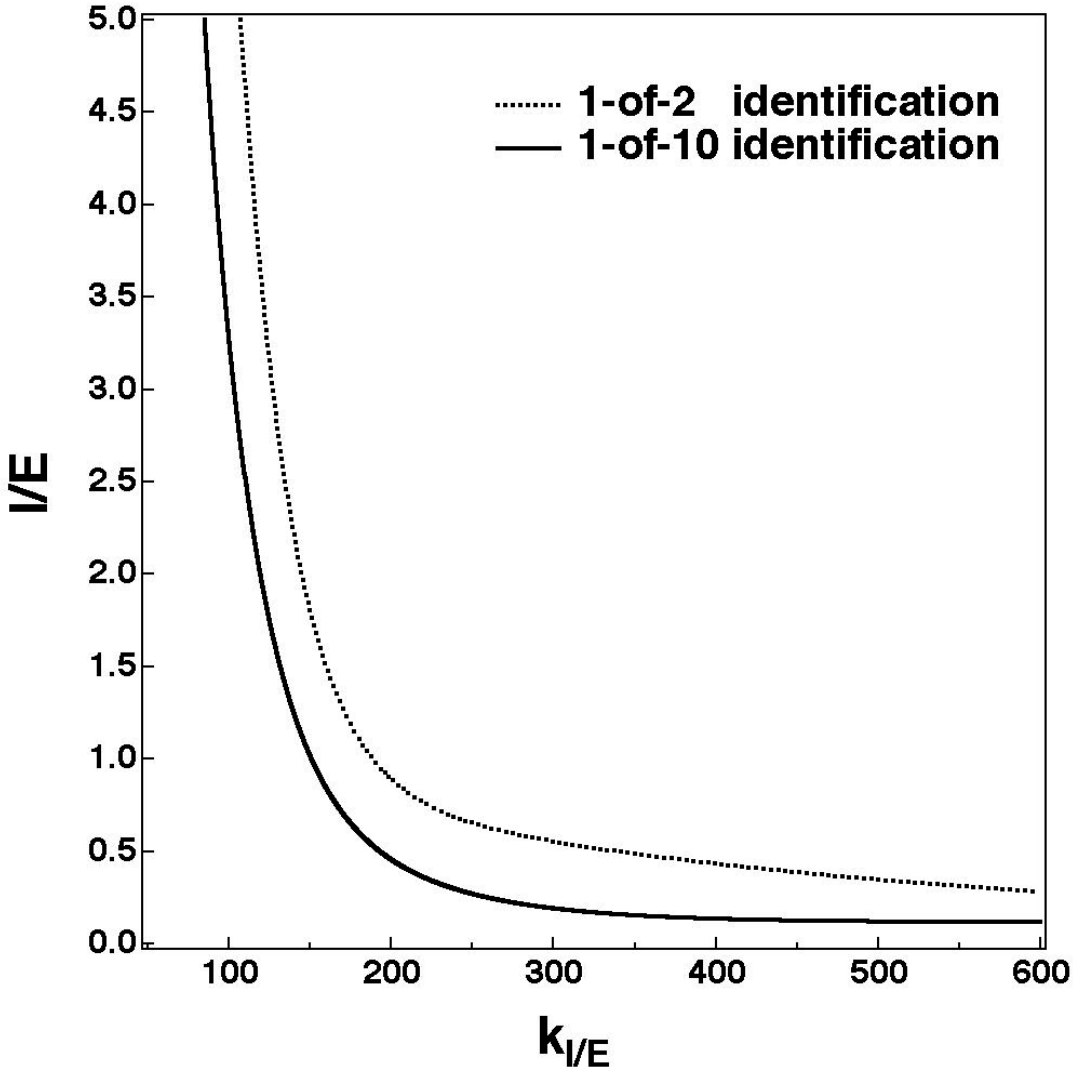


Figure 2.5 Relationship between internal/external noise ratios and slopes of functions of the form described in Figure 2.4 for 1-of-2 and 1-of-10 identification tasks. Each curve is the fit of Equation 2.10 to the simulated data of a noisy cross-correlator.

where α, γ_1, γ_2, β_1 and β_2 are free parameters. Simulated observers with a range of *I/E*'s were implemented in MATLAB to determine the parameters for equation 2.10 for the 1-of-10 and 1-of-2 identification tasks reported in this thesis. Figure 2.5 shows the relationship between *I/E* and $k_{I/E}$ for the simulated observers in these tasks. Maximum-likelihood minimization was used to fit equation 2.9 to the human data, and bootstrap

simulations were used to produce confidence intervals for the parameters (minimum 500 simulated experiments). The fitted parameters for Equation 2.10 for the appropriate task were then used to estimate *I/E*'s for each human observer[9].

Classification Images

This final section describes the methods used in this thesis for computing classification images. Recall from Chapter 1 that a classification image reveals the stimulus locations that have influenced an observer's responses over the course of an experiment. This is achieved by corrupting signals with external noise and correlating the noise contrast at each location with an observer's responses across trials. In order to insure the noise is having an effect on an observer's responses, a high level of external noise is used in conjunction with an adaptive staircase to maintain threshold performance.

After the data are collected, the first step in computing the classification images for a task is to average the noise matrices point-by-point according to each signal-response combination. In the case of only two signals, there will be four signal-response

[9] The fitted parameters for the 1-of-10 identification task are: $\alpha = 0.1148$, $\gamma_1 = 4.0700$, $\gamma_2 = 75.9068$, $\beta_1 = 0.0134$, $\beta_2 = 0.0358$. The fitted parameters for the 1-of-2 identification task are: $\alpha = 0.0727$, $\gamma_1 = 1.0755$, $\gamma_2 = 141.5219$, $\beta_1 = 0.0027$, $\beta_2 = 0.0329$.

combinations: *S1R1*, *S1R2*,*S2R1* and *S2R2*[10]. These four images may be combined to form a single image *C* as follows:

$$C = (S1R2 + S2R2) - (S1R1 + S2R1). \quad (2.11)$$

C is the observer's raw classification image. Unfortunately, it is often difficult to detect features in the raw classification image by visual inspection[11]. There are two major ways to address this issue. The first way is to smooth the classification images to remove some of the extraneous noise. The second way is to rely on statistical tests. Both of these methods are used to analyze the classification images reported in Chapter 4.

Frequency Filtering

In the case of the texture identification task, the signal is highly localized in the frequency domain (2-4 c/image). This aspect of the stimulus suggests a convenient filter to apply to the classification images, namely one that is matched to the stimulus (i.e., a 2-4 c/image ideal filter). However, the power spectrum of the faces is not highly localized in frequency space: like most natural images, it falls off in inverse proportion to spatial

[10] The response classification experiments in this thesis were restricted to two possible signals, so only this special case is considered here. However, see Watson (1998) for an example of a 1-of-3 identification task.
[11] This is presumably due to masking by high spatial frequencies in the noise.

frequency. Previous data (Gold et al., 1999a; Nasanen, 1999) have shown observers make especially efficient use of frequencies within a 2-octave wide band centered around 6 c/face, with efficiency gradually declining above and below the center frequency. This suggests an appropriate filter to apply to the face classification images would be a 2-octave wide frequency filter centered at 6 c/face. Unlike the filter for the texture classification images, the filter used in the face identification experiments reported in Chapter 4 was not ideal; rather, it was Gaussian in shape in log-frequency space, falling to half-height at 1 octave above and below 6 c/face.

Statistics

The most straightforward statistical test is to correlate the unfiltered classification image obtained from a human observer with the classification image of an ideal decision rule. For a 1-of-2 identification task, the ideal observer's classification image (template) is simply the difference between the two possible signals, i.e.

$$C_{ideal} = T_2 - T_1. \quad (2.12)$$

This rule is proven in the Appendix. The 'goodness' of a human observer's classification image can be computed by measuring it's cross correlation with the ideal template, i.e.

$$G = \sum_{i=1}^{n} C_i C_{i\,ideal} \,. \tag{2.13}$$

When C is normalized to unit variance and C_{ideal} is normalized to unit energy, the significance of the correlation can be computed by a z-test on the quantity G.

A second statistical test can be performed to remove the pixels from the classification that fall outside of some criterion level of significance. In the case of the raw classification image, the expected variance of each pixel is equal to the sum of the expected variances from each signal-response category. All of the pixels falling within some criterion number of standard deviations can then be replaced with zero contrast pixels. Both of the above tests are used to evaluate the classification images in the experiments reported in Chapter 4.

3 Internal Noise and Calculation Efficiency in Pattern Identification Learning

Introduction

This chapter presents the data from the experiments investigating the effects of perceptual learning on internal noise and calculation efficiency. It involves two main sets of experiments. The first set of experiments involve measurement of equivalent input noise (i.e., contrast-invariant internal noise) and calculation efficiency as observers learn to identify unfamiliar patterns. As discussed in Chapter 1, measurement of individual thresholds does not allow us to discriminate between the impact of contrast-invariant internal noise and calculation efficiency on performance. However, these factors can be teased apart by measuring two or more thresholds in sufficiently different magnitudes of external noise. Thus, the first set of experiments in this chapter consist of measuring signal identification energy thresholds in a range of external noise power spectral densities, across a series of learning sessions. Measures of equivalent input noise and calculation efficiency are then derived from these data, allowing us to trace any changes in these quantities as learning takes place.

The second set of experiments involve measurement of the contribution of internal contrast-dependent noise to the effects of learning found in the first set of experiments. As discussed in Chapter 1, changes in calculation efficiency and changes in proportional contrast-dependent noise have identical effects on thresholds across external noise levels (within the context of the black-box model described in Chapter 1). In the absence of an independent measure of contrast-dependent internal noise, there is no way

to disentangle the effects of these two factors. Thus, the second set of experiments use a different method -- response consistency -- to quantify the amount of contrast-dependent internal noise within the observer as learning takes place.

Two different kinds of patterns -- human faces and abstract textures -- were used as signals in the experiments reported in this Chapter. The rationale for using these two different kinds of patterns stems from recent debate regarding the mechanisms that mediate the perception of human faces. There is some evidence that there are specialized cortical mechanisms devoted to face perception (Kanwisher, McDermott, & Chun, 1997; Perrett, Hietanen, Oram, & Benson, 1992), while other evidence suggests the apparent special status of faces is rooted in expertise (Gauthier, Skudlarski, Gore, & Anderson, 2000; Gauthier & Tarr, 1997; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999). Faces were used in the current experiments because they reflect a complex, real-world perceptual learning problem that the visual system must solve on a relatively continual basis. However, if learning to recognize novel faces is mediated by face-specific mechanisms, the effects of learning may not apply to other kinds of patterns. The texture identification task was used to address this issue. The textures were used because they are spatially dissimilar to faces but are similar in complexity.

Equivalent Input Noise and Calculation Efficiency

This section consists of two main experiments and several post-hoc analyses. As mentioned above, the experiments in this section involve the measurement of equivalent input noise and calculation efficiency for the identification of faces (Experiment 3.1) and textures (Experiment 3.2) as perceptual learning takes place. Both experiments rely on the theoretical framework described in Chapter 1 and many of the general methods described in Chapter 2. Only the aspects of the methodology specific to the current set of experiments are presented here.

Methods

Stimuli**.** The stimuli used were the sets of 10 faces and 10 textures shown in Figures 2.1 and 2.2. In each task, signal energy identification thresholds were measured in five different levels of external noise power spectral density. For the face task, the external noise levels were: 0.04, 0.20, 1.02, 5.11 and 25.55 x 10^{-6} deg^2. Pilot studies suggested that equivalent noise was higher for the texture identification task, so the lowest external noise level was removed and replaced by a higher noise level of 51.10 x 10^{-6} deg^2. A unique noise field was generated on every trial.

Procedure. Signal energy thresholds were estimated according to the method of constant stimuli. Pilot studies identified five initial signal energy levels that spanned the threshold range in each level of external noise. The signal energy levels were adjusted

after each session as necessary (given the rate of learning). There were 31 trials per stimulus energy level within each session, yielding a total of 775 trials (31 trials x 5 signal levels x 5 noise levels). The level of external noise, signal energy and the identity of the signal were chosen randomly on each trial. Each session was completed without breaks and lasted about one hour. Only one session was completed each day. Observers in the face identification task completed six sessions within ten days. Observers in the texture identification task completed four sessions within seven days.

Observers. There were two observers that participated in each task, with one observer (AMC) participating in both tasks. One observer was the author (JMG) and the remaining two observers were naive to the purposes of the experiment.

Experiment 3.1: Face Identification

Psychometric functions for one observer (CGB) in the face identification task are shown in Figure 3.1 Each panel corresponds to a different level of external noise within a particular session. Each row corresponds to a different session, progressing from the first to the last, from top to bottom. Each column corresponds to a different external noise level, progressing from lowest to highest, from left to right. The symbols in each plot correspond to the empirically measured level of percent correct performance at a given level of signal energy. The smooth curves are maximum-likelihood fits to Equation 2.5.

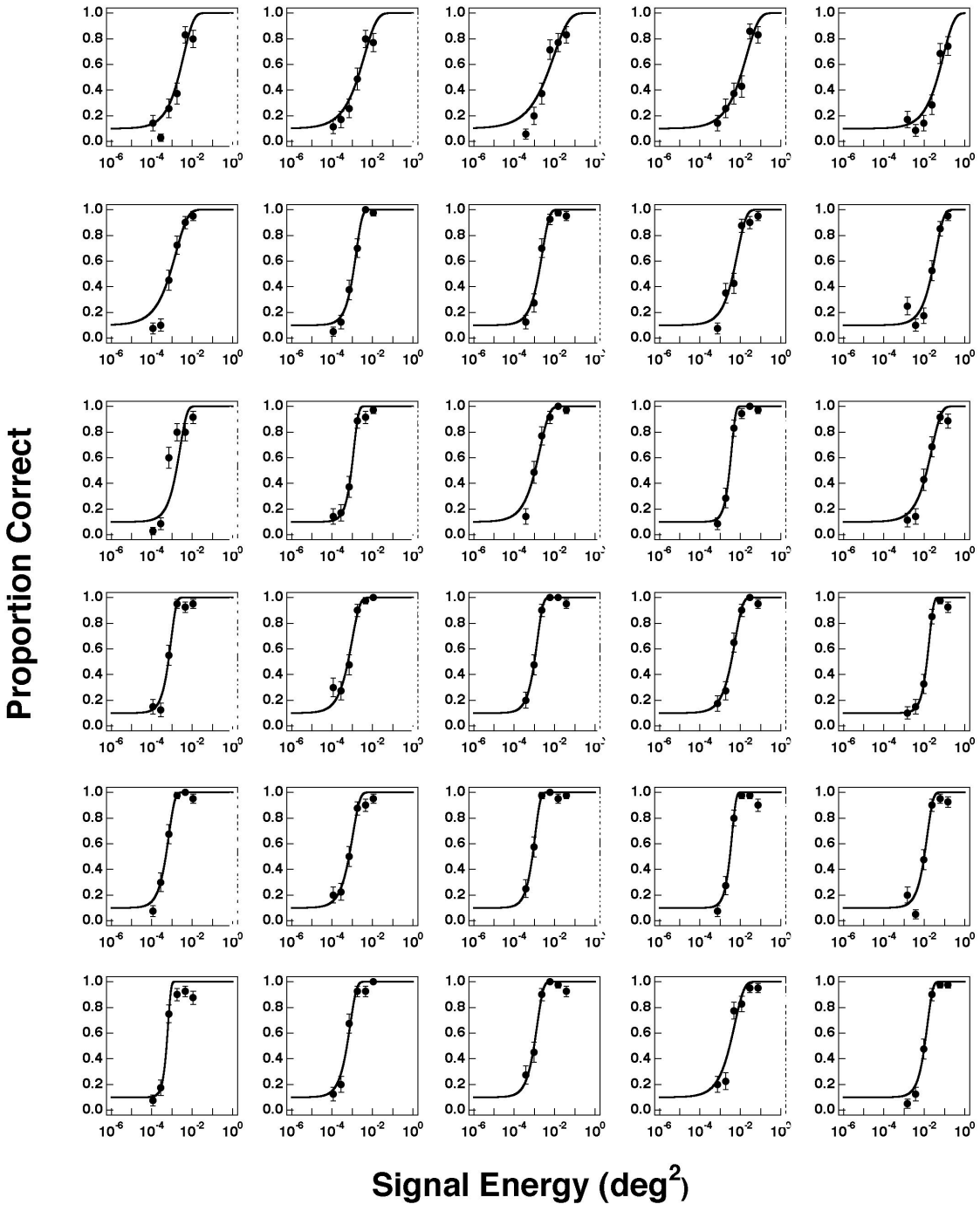


Figure 3.1 Psychometric functions plotting proportion correct as a function of signal energy for one observer (CMG) in the 1-of-10 face identification experiment. Each panel corresponds to performance at a particular level of external noise during a given session. Each row corresponds to a different session, with the sessions proceeding in sequence from top (first session) to bottom (last session). Each column corresponds to a different level of external noise power spectral density, with the external noise levels proceeding in sequence from the left (lowest noise level) to the right (highest noise level). The data points in each plot are the empirical measures of percent correct for each corresponding signal level, and the solid lines are the maximum likelihood fits to Equation 2.5. Error bars on each symbol correspond to ± 1 standard deviation.

	AMC			CGB		
Session	**r^2**	**F**	**p**	**r^2**	**F**	**p**
1	0.9681	91.07	3.92e-05	0.9988	2500.41	1.06e-08
2	0.9968	951.75	1.18e-07	0.9990	3043.44	6.46e-09
3	0.9969	976.75	1.10e-07	0.9962	789.40	1.86e-07
4	0.9955	664.08	2.88e-07	0.9981	1649.09	2.98e-08
5	0.9999	23762.60	3.80e-11	0.9973	1110.79	8.00e-08
6	0.9957	703.09	2.50e-07	0.9926	407.19	9.74e-07
Mean	*0.9922*	*4524.89*	*6.66e-06*	*0.9970*	*1583.39*	*2.15e-07*

Table 3.1 Statistics for linear threshold fits across sessions in the face identification task (see text for details).

Careful inspection of the plots reveal shifts to higher signal energy levels from left to right in each row. This shift corresponds to an increase in threshold with increasing external noise magnitude beyond the level of contrast-invariant internal noise. The curves also shift to lower signal energy levels from top to bottom across sessions in each column. This shift is due to the effects of learning.

Both of these effects are summarized in Figure 3.2, which shows the 50% correct threshold signal energy levels as a function of external noise power spectral density for two observers in the face identification task across all six sessions. Each symbol corresponds to a single threshold. The filled symbols correspond to the first three sessions, the open symbols the last three sessions. The smooth curves are maximum-likelihood fits to Equation 1.1. The goodness of the fits can be expressed in terms of r^2, which indicates the proportion of variability in the thresholds across external noise levels that is accounted for by the parameters of Equation 1.1. The r^2 values for both observers

in each session are summarized in Table 3.1, along with *F* statistics that indicate the significance of the fits. The r^2 values for both observers in almost every session were above 0.99 (with the exception of 0.97 for observer AMC in the first session), indicating the effect of noise on thresholds in this task is well characterized by a linear function.

Inspection of Figure 3.2 reveals a clear trend across sessions for both observers. Namely, the height of the function shifts down uniformly across external noise levels with practice. However, the kink point in the functions remains constant across sessions. Recall from Chapter 1 that this kind of shift in the noise masking function (in log-log axes) is consistent with the effects of increased internal signal strength with no corresponding decrease in the strength of internal contrast-invariant noise. These effects can be seen more clearly in Figure 3.3, which plots calculation efficiency and equivalent input noise for both observers, as a function of practice. Recall from the previous chapters that calculation efficiency corresponds to the height of the noise masking function (normalized by the performance of the ideal observer) and equivalent input noise corresponds to the kink point of the noise masking function.

There are several interesting aspects to note about these data. First, there is a striking increase in calculation efficiency across sessions but almost no change in internal noise: calculation efficiency increased by a factor of 4.4 for observer AMC and 4.3 for observer CGB with practice. Second, the effects of learning occurred mostly in the first three experimental sessions. These sharp increases in efficiency at the beginning of

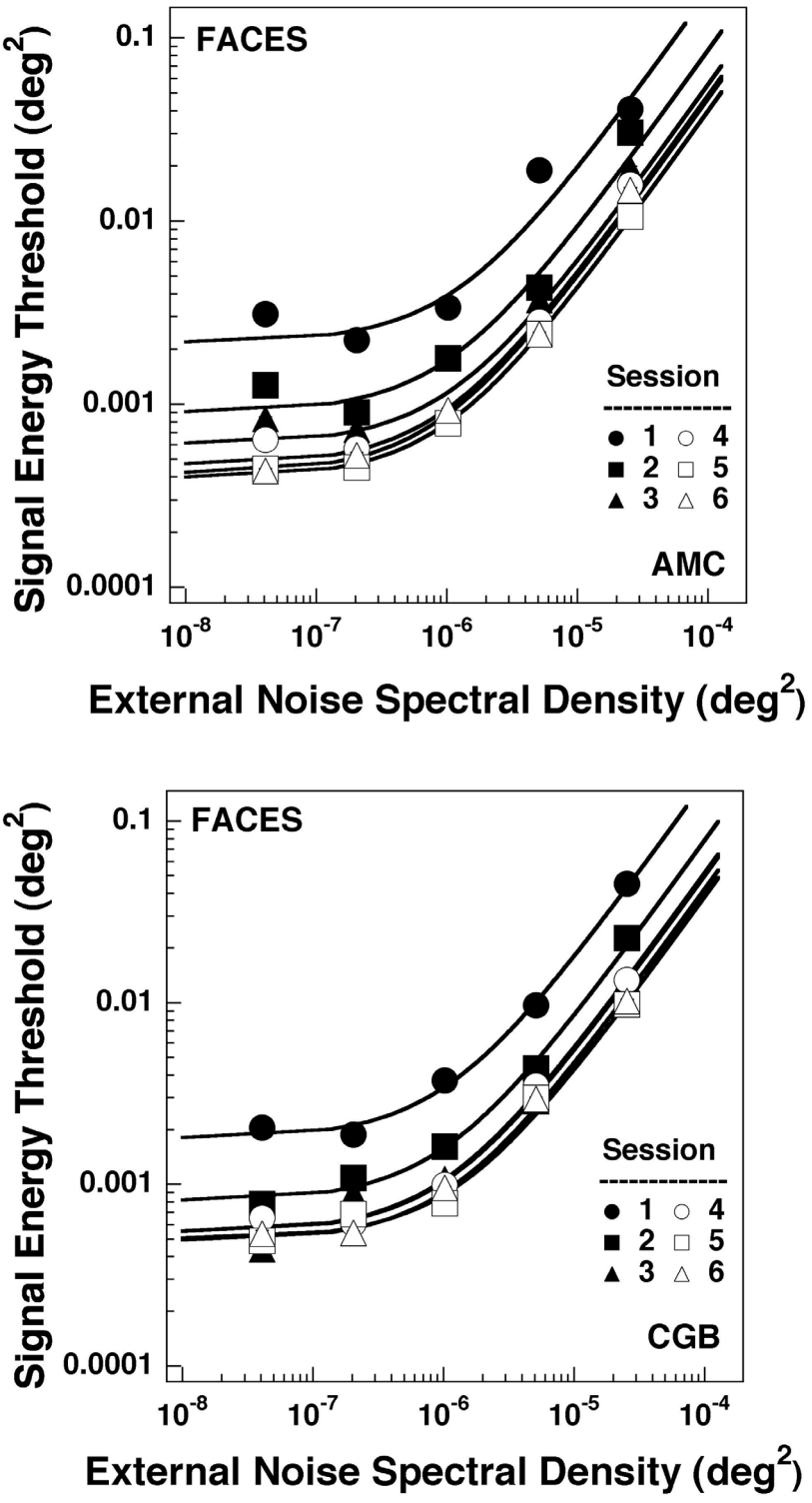


Figure 3.2 Noise masking functions for the two observers in the 1-of-10 face identification task across six sessions. Each panel plots the observer's signal energy thresholds (defined as 50% correct) as a function of external noise power spectral density. The filled symbols correspond to the first three sessions, the open symbols the last three sessions (see legend). Solid lines correspond to maximum-likelihood fits to Equation 1.1. Error bars on each symbol correspond to ± 1 standard deviation. In this figure, the error bars are smaller than the symbols.

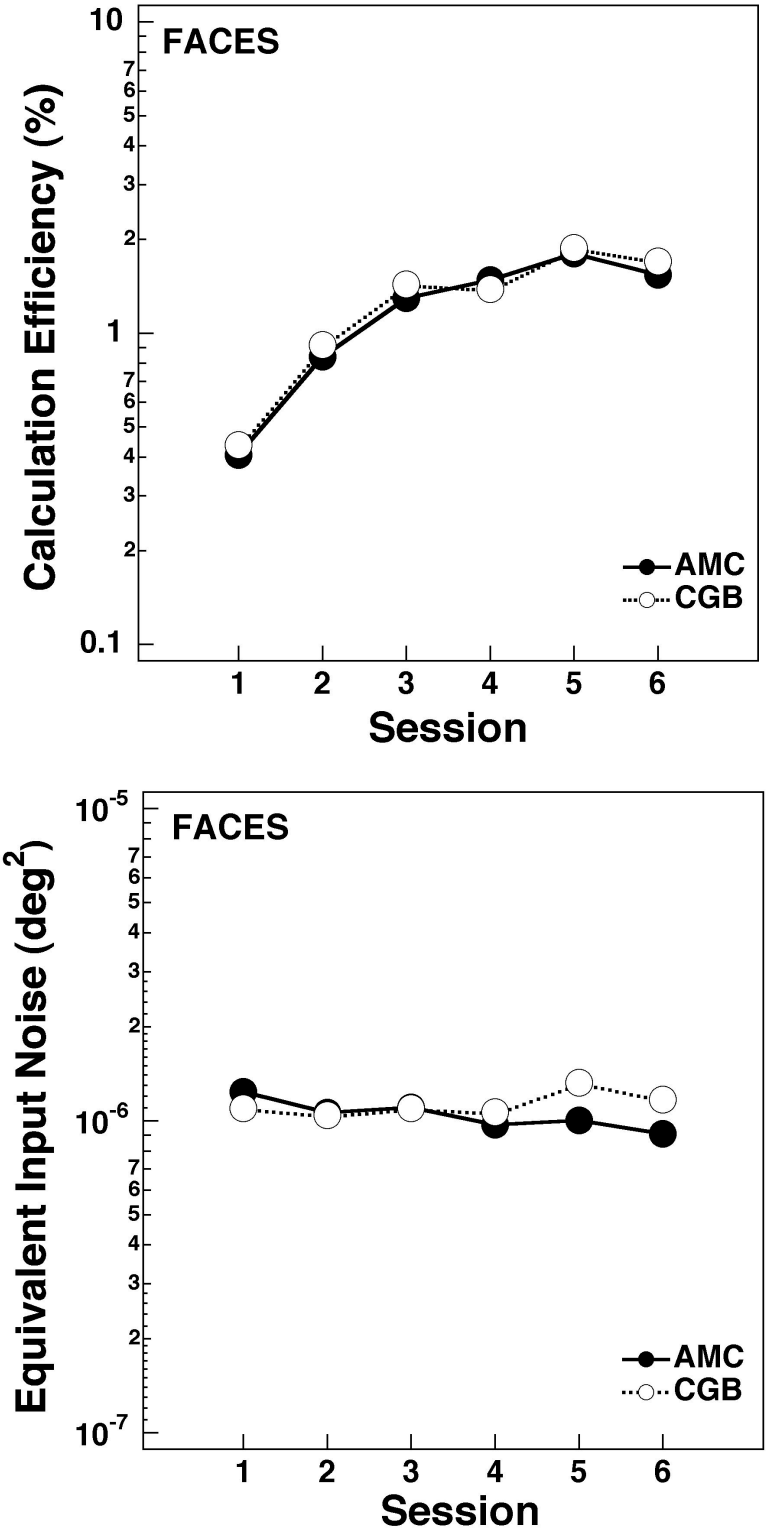
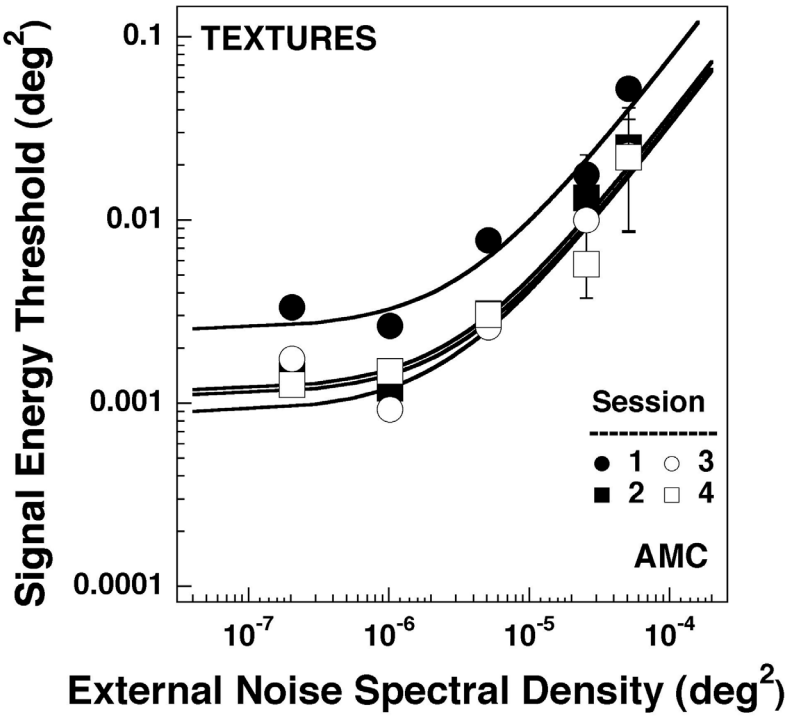


Figure 3.3 Calculation efficiency (top panel) and equivalent input noise (bottom panel) as a function of experimental session for the two observers in the 1-of-10 face identification task. Error bars on each symbol correspond to ± 1 standard deviation. In this figure, the error bars are smaller than the symbols.

training suggest there could be learning taking place within individual sessions (Fahle et al., 1995; Gilbert, 1994). This possibility is considered in the Discussion section below. Third, the estimates of calculation efficiency after learning are consistent with previous measures of efficiency for familiar face identification (Gold et al., 1999a). The absolute efficiencies in both cases are only about 1-2%, significantly lower than for other complex patterns (see Gold et al. (1999) for a summary). However, as Burgess (1990) points out, these measures of efficiency incorporate limitations imposed by contrast-dependent internal noise, and therefore should be considered lower bounds. The effects of contrast-dependent noise on face and texture identification efficiency will be considered below in the second half of this chapter.

Experiment 3.2: Texture Identification

As mentioned in the introduction, the recent controversy over the nature of face perception makes it necessary to test the generality of the effects found above in the face task. This issue was addressed by having observers learn to identify the set of 10 abstract texture stimuli described in Chapter 2. The noise masking functions for two observers in the texture identification task are shown in Figure 3.4. One of the observers (AMC) also participated in the face identification experiment. The $r^2$and F statistics for the fits are shown in Table 3.2. Figure 3.5 traces the changes in calculation efficiency and equivalent input noise across sessions for both observers.



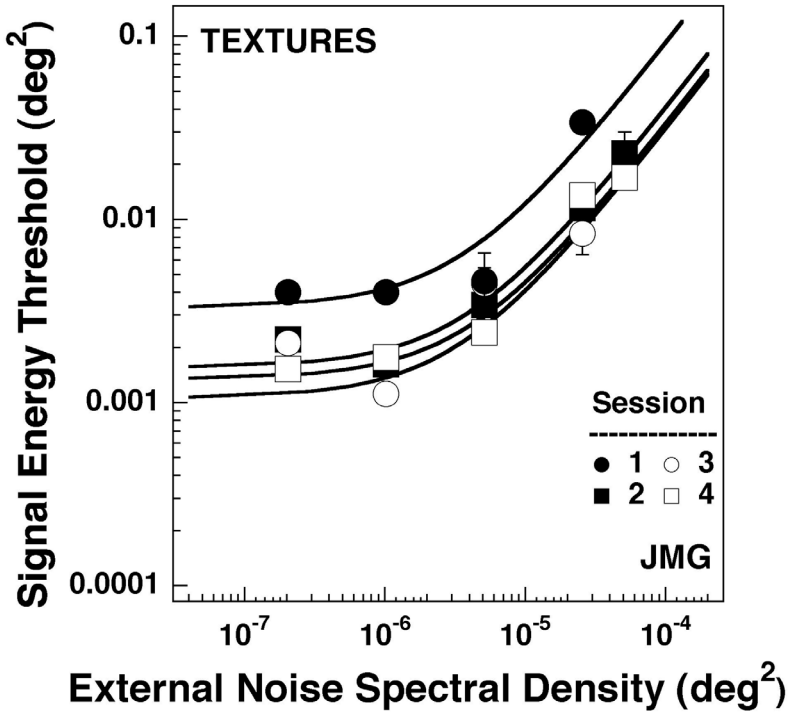


Figure 3.4 Noise masking functions for the two observers in the 1-of-10 texture identification task across four sessions. Each panel plots the observer's signal energy thresholds (defined as 50% correct) as a function of external noise power spectral density. The filled symbols correspond to the first two sessions, the open symbols the last two sessions (see legend). Solid lines correspond to maximum-likelihood fits to Equation 1.1. Error bars on each symbol correspond to ± 1 standard deviation. Often, the error bars are smaller than the symbols.

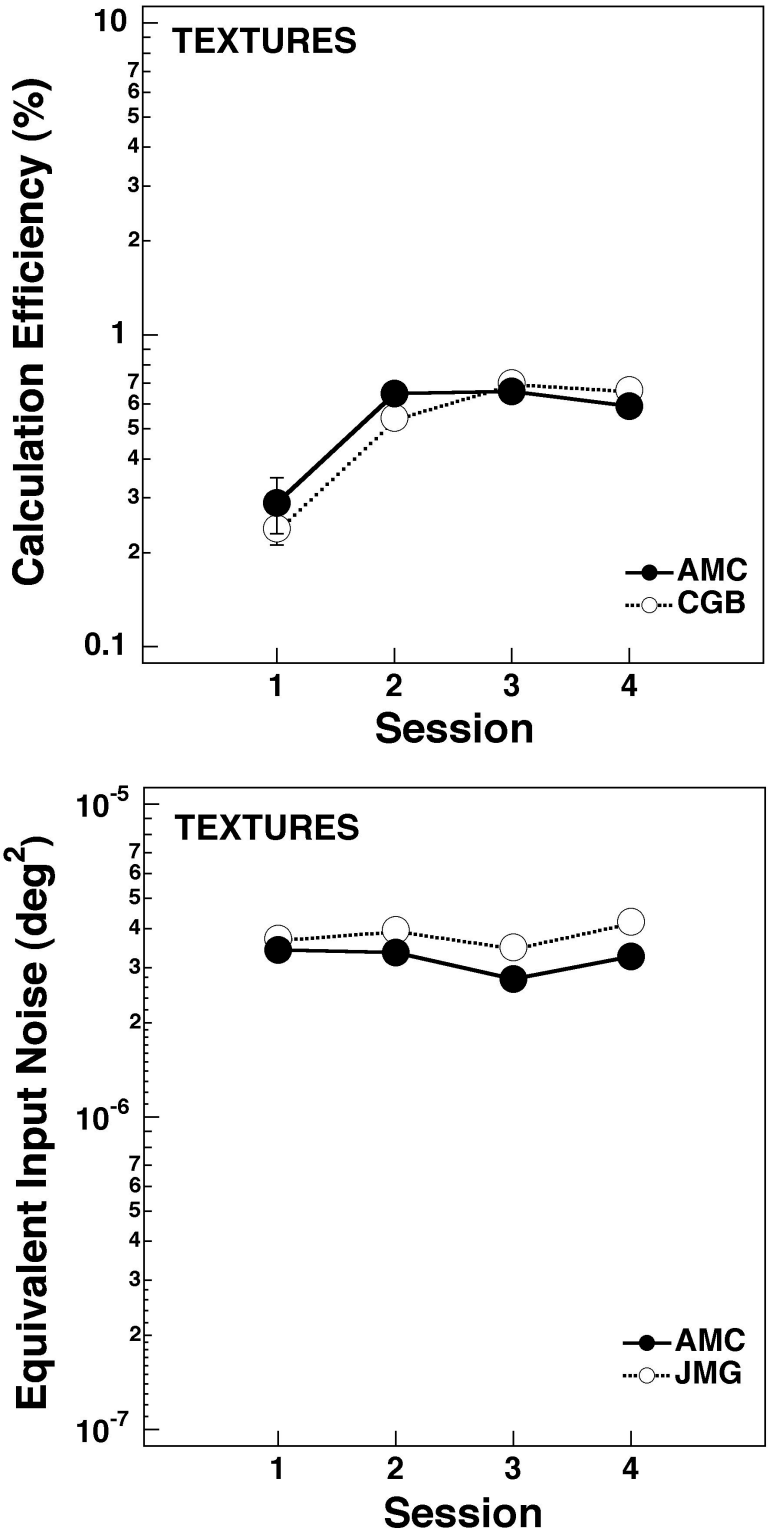


Figure 3.5 Calculation efficiency (top panel) and equivalent input noise (bottom panel) as a function of experimental session for the two observers in the1-of-10 texture identification task. Error bars on each symbol correspond to ± 1 standard deviation. Often, the error bars are smaller than the symbols.

It is clear from these data that practice produced a uniform downward shift in thresholds across external noise levels in the texture task just as it did in the face task. These results suggest it is unlikely that the learning effects found with faces are mediated by face-specific mechanisms. Instead, they most likely reflect the more general effects of learning in a complex pattern identification task. However, there are also some interesting differences between the results from the two tasks. First, learning reached a plateau earlier for the textures than for the faces. Also, the effect of learning was significantly lower in magnitude for the textures than for the faces: AMC improved by a factor of 2.3 and JMG by a factor of 2.9, nearly a factor of 2 less than the observers in the face identification task. Second, equivalent input noise is higher for textures than for faces. This result is surprising, because it is often thought that equivalent input noise reflects the effects of contrast-invariant internal noise occurring relatively early in the visual system (Pelli, 1981, 1990). If so, changing the task should have little or no effect on estimates of equivalent input noise (unless the noise depends on certain aspects of the stimuli, such as local contrast). The fact that equivalent noise was found to be significantly higher for texture identification is consistent with the existence of a central or 'late' contrast-invariant noise (see Raghavan (1989) for an in-depth treatment of localizing the sources of equivalent input noise).

	AMC			JMG		
Session	**r^2**	**F**	**p**	**r^2**	**F**	**p**
1	0.9874	234.58	3.82e-06	0.9726	71.050	1.90e-4
2	0.9994	4913.76	1.95e-09	0.9969	954.28	1.17e-07
3	0.9976	1248.09	5.98e-08	0.9716	106.57	2.67e-05
4	0.9273	38.28	3.14e-4	0.8976	26.30	7.51e-4
Mean	*0.9779*	*1608.68*	*7.94e-05*	*0.9599*	*289.55*	*2.42e-4*

Table 3.2 Statistics for linear threshold fits across sessions in the texture identification task (see text for details).

Discussion

Taken together, the results of the above experiments offer compelling evidence that perceptual learning increases internal signal strength but has no discernable effect on the strength of internal contrast-invariant noise. This interpretation rests upon a particular model of the human observer. In its simplest form, the observer is modeled as constrained by contrast-invariant internal noise and the efficiency of internal calculations. However, as discussed in Chapter 1, this simple model may be elaborated to included other plausible constraints. Most notably, the model does not allow us to discriminate between the effects of changes in calculation efficiency and changes in an internal noise that grows in proportion to stimulus energy. Both have the effect of uniformly reducing the height of the noise masking function on log-log axes. Within the context of the present experiments, the effects of learning could be due to changes in calculation efficiency or changes in internal proportional contrast-dependent noise (or both). This

issue is the subject of the next set of experiments, where response consistency is used to disentangle the effects of calculation efficiency and contrast-dependent internal noise.

However, before turning to the consistency measures, it is useful at this point to step back and explore some additional aspects of the noise masking data, including: (1) the assumptions of linearity and the effects of stimulus uncertainty in relation to the shapes of the psychometric functions and the effects of learning; (2) within-session learning effects; and (3) learning of gender in the face identification experiment.

Linearity, Uncertainty and the Psychometric Function. One of the implicit assumptions of the model outlined in Chapter 1 is that the observer only performs linear transformations upon the stimulus. The model also assumes that the observer has no intrinsic uncertainty about various aspects of the stimulus, most notably its location. However, several recent models of pattern detection and discrimination incorporate the effects of point-wise non-linearities (Lu & Dosher, 1999), stimulus uncertainty (Eckstein, Ahumada, & Watson, 1997; Eckstein, Whiting, & Thomas, 1996; Manjeshwar & Wilson, 2001a, 2001b) or both (Solomon, Lavie, & Morgan, 1997; Watson & Solomon, 1997). Each of these factors will affect the shape of an observer's psychometric functions. Specifically, the psychometric function in Gaussian external noise for a linear observer with no stimulus uncertainty will be linear, when expressed in terms of d' (instead of percent correct) and signal contrast (instead of energy). This is illustrated by the diagonal line in Figure 3.6. However, this relationship can break down at low signal contrasts in

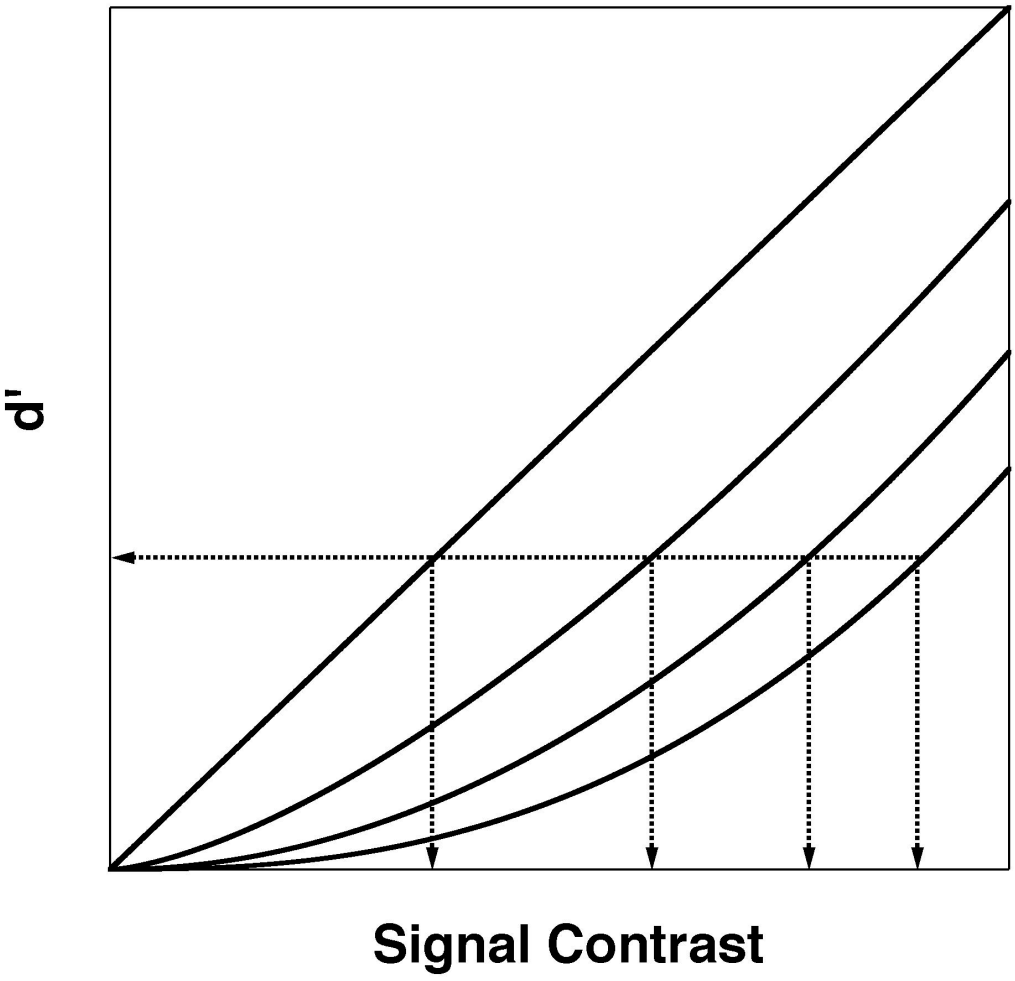


Figure 3.6 Demonstration of the effects of either stimulus uncertainty or an accelerating non-linearity on the shape of an observer's psychometric function. A linear observer with no uncertainty will have a linear psychometric function (diagonal line) when plotted in terms of *d'* and signal contrast. Both uncertainty and an accelerating non-linearity will have the effect of making the psychometric function non-linear. The curved lines show psychometric functions for observers with increasing uncertainty and/or non-linearity from left to right. The largest effects are at low signal contrasts. The dashed lines show how the estimate of contrast threshold for a given level of *d'* depends on uncertainty and the degree of non-linearity.

the presence of either an accelerating non-linearity or stimulus uncertainty. This effect is illustrated by the bottom three curves in Figure 3.6. The amount of uncertainty and the degree of the non-linearity manifest themselves in the form of a rightward shift in the

psychometric function at low contrasts: the greater the shift, the more uncertainty/non-linearity.

One way to test for the presence of either of these effects is to compare *d'* ratios and contrast ratios from a given psychometric function. For a linear observer with no stimulus uncertainty, the ratio of two different values of *d'* equals the ratio of the two corresponding signal contrasts on a given psychometric function (Lu & Dosher, 1999). However, this relationship breaks down in the presence of non-linearities and/or stimulus uncertainty. Specifically, decreasing either uncertainty or the degree of an accelerating non-linearity will serve to decrease the estimate of threshold for a given level of *d'*. This is illustrated by the dashed lines in Figure 3.6, which show the corresponding contrast threshold estimates for a given *d'* with various degrees of uncertainty and/or non-linearity.

Our analysis of the face and texture data above ignored the possible effects of non-linearities and stimulus uncertainty. However, given the above argument, it is possible that some or all of the effects of learning found in the first set of experiments were due to decreases in either uncertainty or non-linearity across sessions. That is, decreases in the contributions of these two factors with learning would predict a systematic decrease in threshold estimates for a given criterion level of percent correct (or *d'*) with practice. This possibility was tested by making use of the equal *d'* and contrast threshold ratio prediction of a linear model (Lu & Dosher, 1999). The

psychometric functions for all observers across all external noise levels and learning sessions in the face and texture tasks were analyzed using this technique. Two levels of *d'* (1.0 and 2.0) were chosen for comparison. These two levels correspond to percent corrects of 0.34 and 0.67, respectively, in a 1-of-10 identification task (Macmillan & Creelman, 1991). These two values were chosen because they span a wide range within the psychometric function, offering a strong test of the model. For each psychometric function and criterion percent correct, the corresponding signal energy level for each percent correct was determined by computing the inverse of Equation 2.5, and signal energy was then converted to contrast by computing the inverse of Equation 2.3. The prediction of a *d'* ratio of 0.5 for a linear observer was verified through computer simulations.

The results of this analysis are shown in Figure 3.7. The top panel corresponds to the results from the face identification task, the bottom panel the texture identification task. Each panel shows the mean contrast ratios (averaged across external noise levels) for each observer, as a function of experimental session. The error bars on each symbol indicate ± 1 standard deviation of the ratio estimates across external noise levels. The dashed line in each panel shows the ratio predicted by the linear model. Although the data are a bit noisy in the texture identification task, there is a clear shift away from linearity as learning takes place. Also note that, especially in the face identification task, observers are close to the predictions of a linear model before learning has taken place.

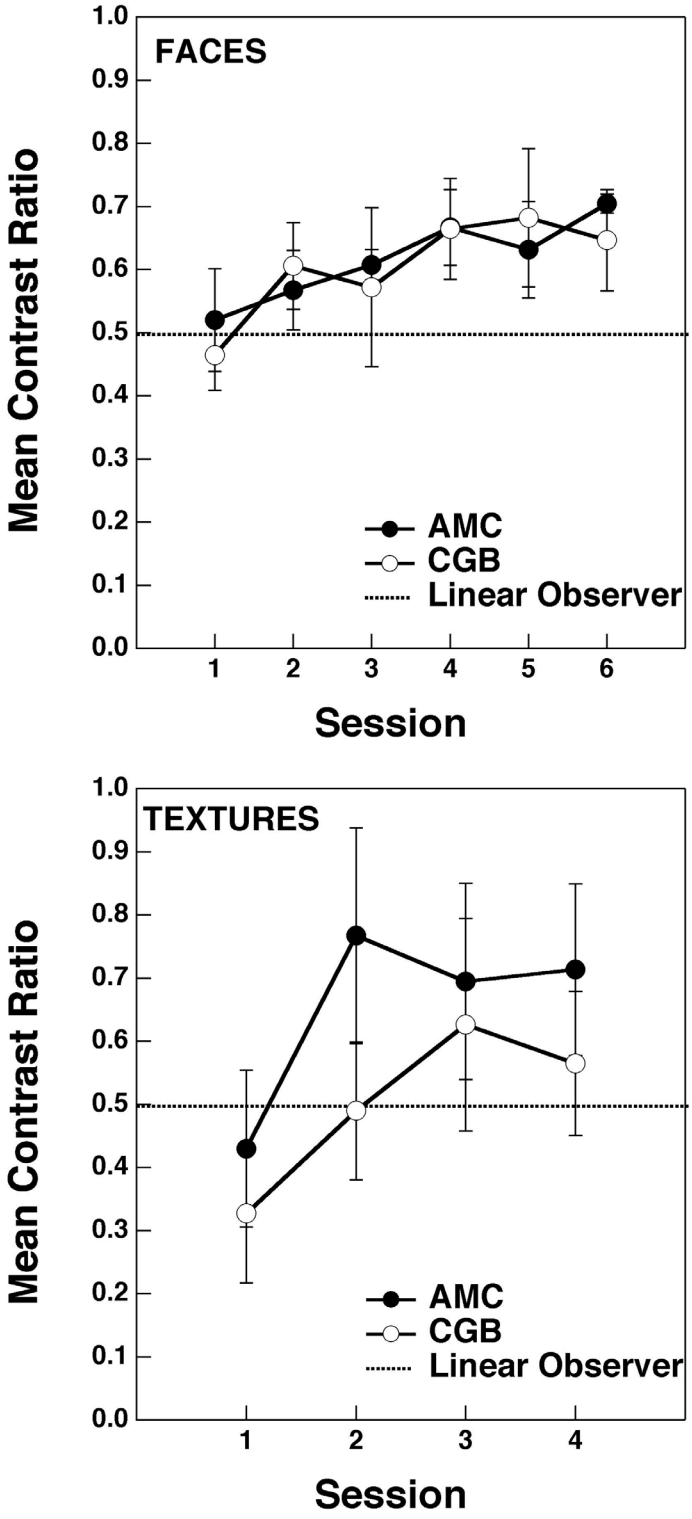


Figure 3.7 Ratios of the contrasts corresponding to two levels of *d'* (1.0 and 2.0) from the psychometric functions of the two observers in 1-of-10 face (top panel) and texture (bottom panel) identification tasks. The dashed line in each panel shows the predictions for a linear observer. Each data point is the average contrast ratio across external noise levels. Error bars on each symbol correspond to ± 1 standard deviation.

These results are inconsistent with the idea that the increases in calculation efficiency found in the first set of experiments were due to decreases in either uncertainty or the degree of point-wise non-linearities. Instead, they show that learning produces the non-intuitive result of *increasing* one or both of these effects. What kinds of changes in the visual system might predict these kinds of results? One possibility is that learning served to increase the number of templates used by observers to perform the task. For example, observers may have used only a few templates to make gross categorical discriminations during the first session (e.g., one for males and one for females). Then, as learning took place, observers could have introduced additional templates, eventually building up a much more elaborate set. Such changes could produce improvements in calculation efficiency (for example, if the templates became better matched to the stimuli) but could also produce increases in uncertainty as the number of templates grows (Eckstein, Ahumada, & Watson, 1997). Comparison of Figures 3.3, 3.5 and 3.7 shows that the patterns of across-session changes in calculation efficiency (Figures 3.3 and 3.5) and non-linearity (Figure 3.7) are quite similar, which is consistent with the idea that the number of templates used by observers increases with learning, producing parallel increases in both calculation efficiency and uncertainty with practice[12].

[12] These findings are also consistent with a recent model of template learning proposed by Beard and Ahumada (1999). Their model involves internal template refinement with practice, and predicts an increase in the non-linearity of the psychometric function as learning takes place.

Within-Session Learning Effects. Figures 3.3 and 3.5 show that the majority of learning in both the face and texture identification tasks took place within the first several sessions. However, these data do not reveal whether the learning took place within each experimental session or required a period of consolidation, in the form of a delay between sessions. There is recent evidence for both fast, within-session learning (Fahle et al., 1995; Gilbert, 1994) and slow, between session learning that requires either rapid-eye-movement (REM) sleep or slow-wave sleep (SWS) (Stickgold, James, & Hobson, 2000; Stickgold, Whidbee, Schirmer, Patel, & Hobson, 2000). Thus, an interesting question is whether the improvements in our face and texture identification experiments took place within a particular session or if the learning required consolidation periods between sessions. Also, there may have been within-session changes in internal noise and calculation efficiency that were obscured by the coarseness of the previous analysis.

To address these questions, the data from each session of the face and texture identification tasks for each observer were subdivided into three successive blocks of trials. The first and second blocks consisted of the first and second sets of 258 trials, respectively, and the third block consisted of the last 259 trials. Each block was then analyzed using the same methods described in Experiments 3.1 and 3.2. The results of this analysis are shown in Figures 3.8 and 3.9. Each figure shows estimates of calculation efficiency (top panels) and equivalent input noise (bottom panels) as a function of trial block. The vertical dashed lines indicate the beginning of each session.

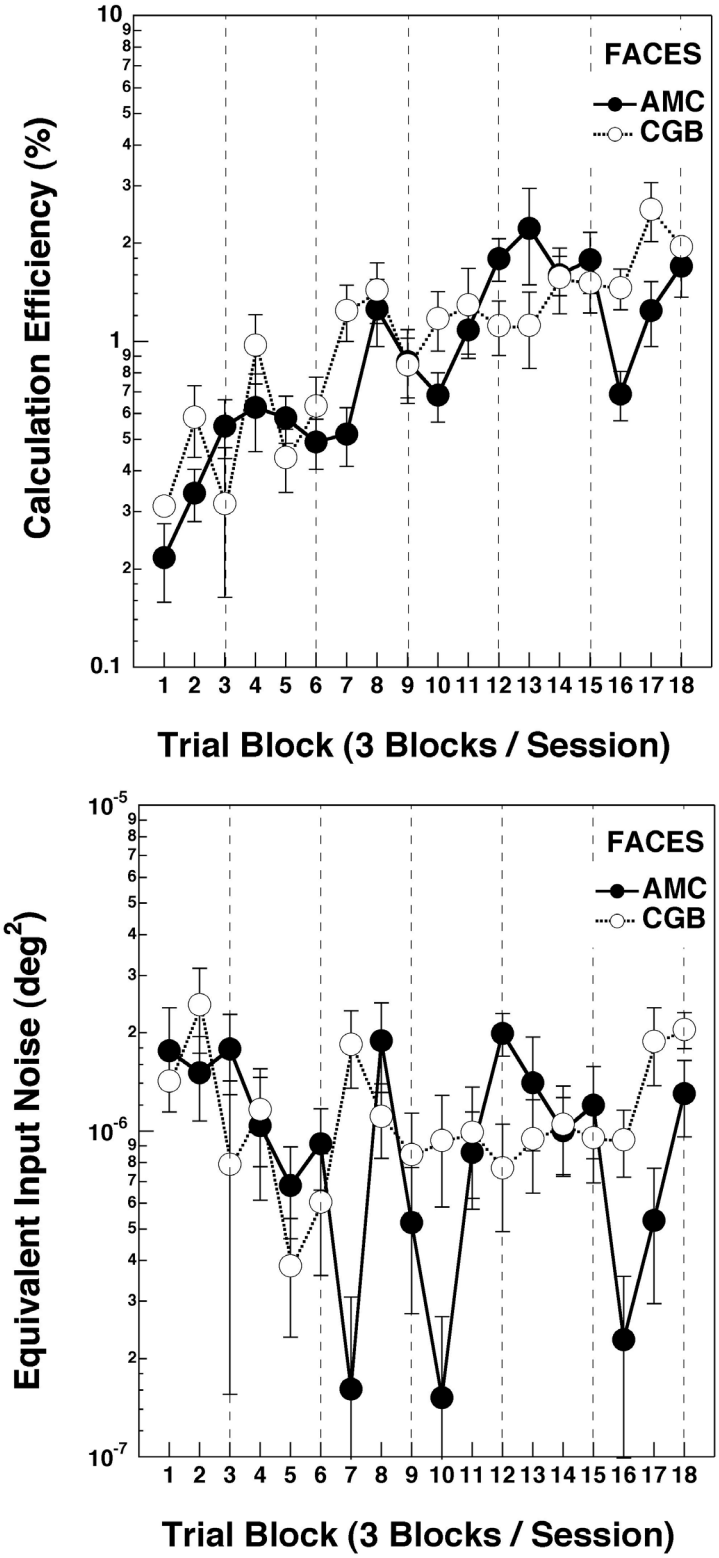


Figure 3.8 Within-session analysis of calculation efficiency (top panel) and equivalent input noise (bottom panel) for the two observers in the1-of-10 face identification task. Each session was separated into three blocks and analyzed as in the original set of experiments. Error bars on each symbol correspond to ± 1 standard deviation.

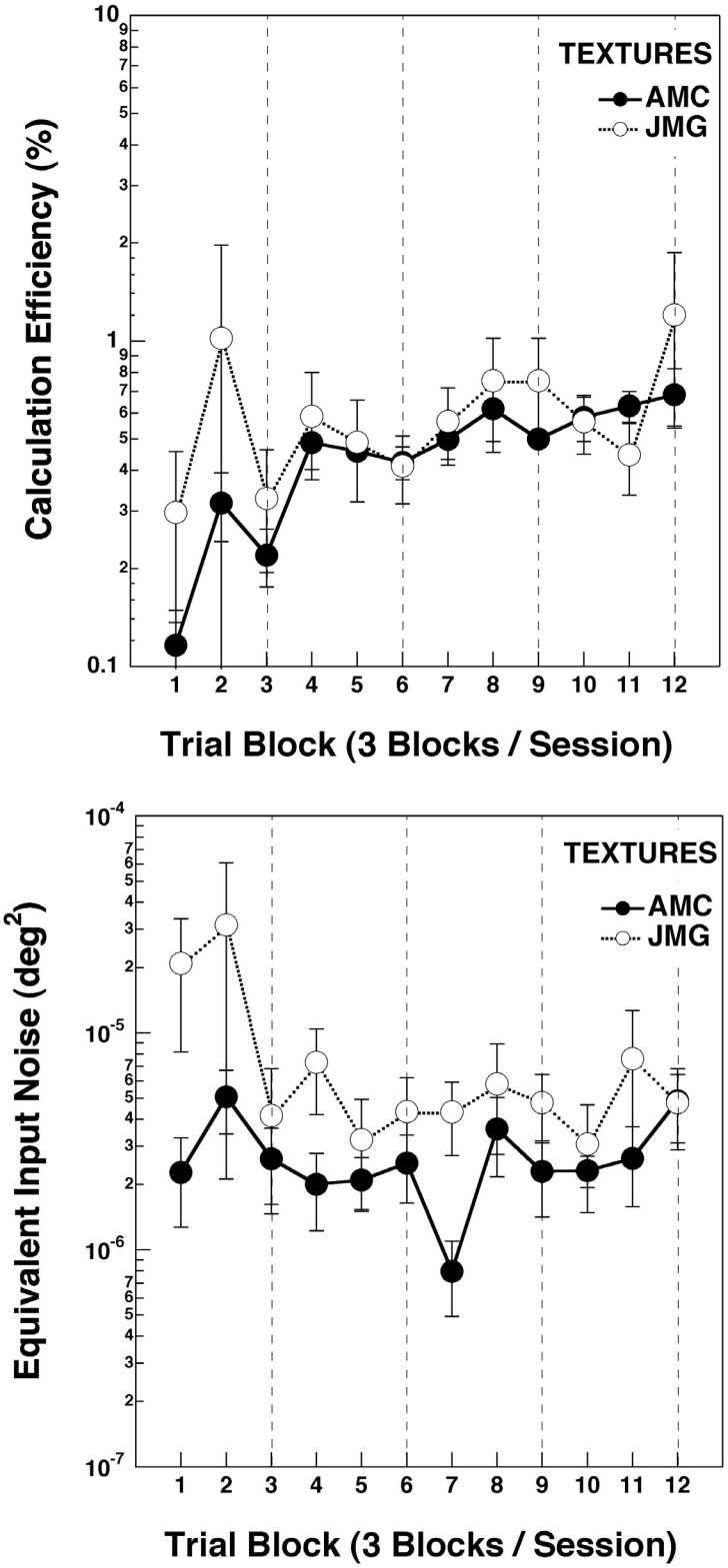


Figure 3.9 Within-session analysis of calculation efficiency (top panel) and equivalent input noise (bottom panel) for the two observers in the 1-of-10 texture identification task. Each session was separated into three blocks and analyzed as in the original set of experiments. Error bars on each symbol correspond to ± 1 standard deviation.

There are several interesting aspects to these data. First, notice that these data are far noisier that those obtained from the session-wise analysis. The noisiness is due to the fact that the psychometric fits were calculated from five stimulus levels that received only 10 or 11 trials. Second, with the possible exception of one observer in the texture identification task, we find the same general effect of learning found previously: an increase in calculation efficiency but no consistent change in equivalent input noise. The one exception is observer JMG in the texture identification task (open symbols in Figure 3.9), who shows a decrease in internal noise within the first session. However, notice that the data for these blocks are particularly noisy for this observer. Third, there is some weak evidence of within-session learning, most notably for observer AMC in the first and third sessions of the face identification task. However, it is difficult to draw very strong conclusions given the noisiness of the data[13]. Thus, this within-session analysis supports the conclusions of the previous session-wise analysis of a change in calculation efficiency but not internal noise with learning.

Gender Learning. In the face identification task, observers were asked to identify the specific face shown on each trial. At the beginning of learning, the observers were completely unfamiliar with the individual faces. However, they *were* implicitly familiar

[13] A more reliable way to test for within-session effects might be to use an adaptive psychometric procedure. With a staircase it is possible to obtain a reliable threshold estimate with fewer trials (with the consequence of a less reliable estimate of the shape of the entire psychometric function).

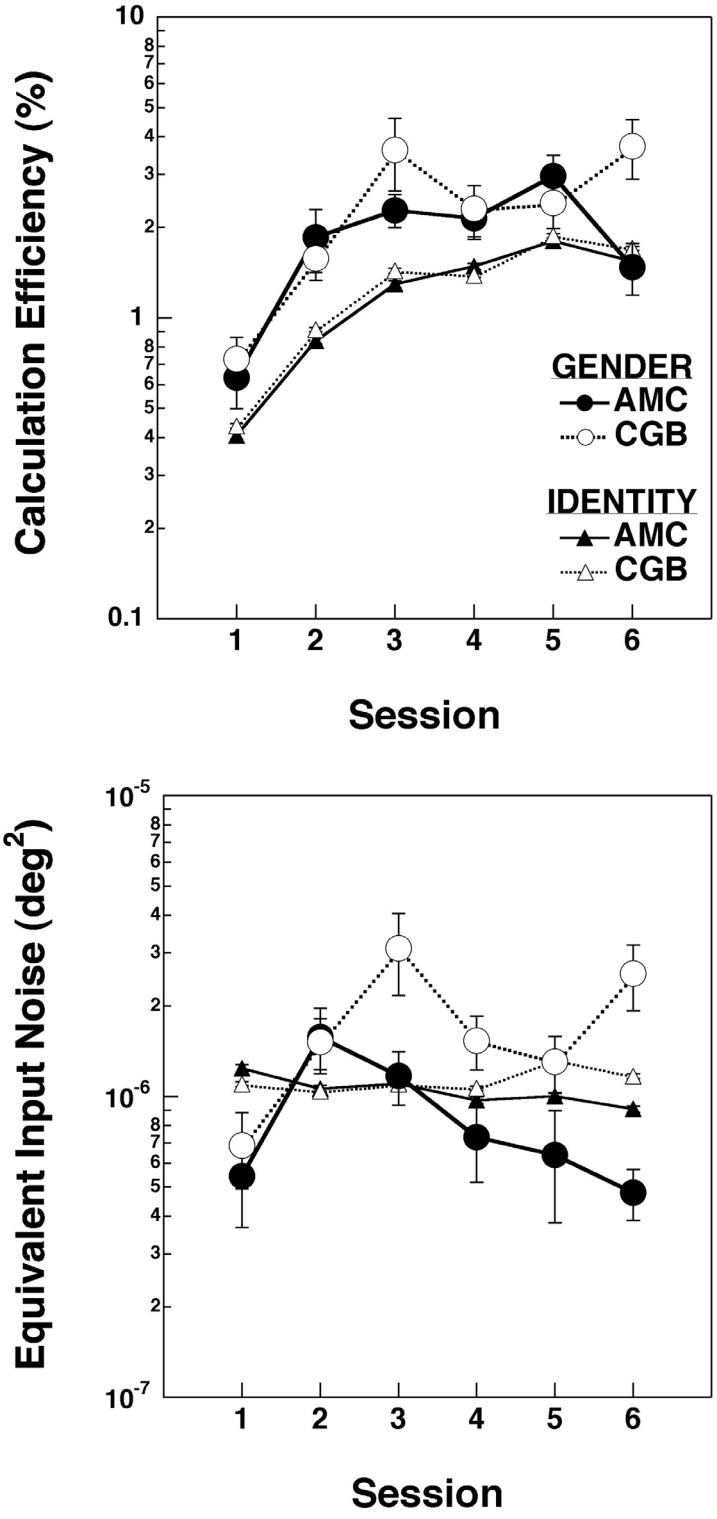


Figure 3.10 Gender-wise analysis of calculation efficiency (top panel) and equivalent input noise (bottom panel) as a function of experimental session for the two observers in the1-of-10 face identification task. The data from the original experiment were analyzed according to the correctness of responses with respect to the gender of the face rather than identify (circular symbols). The data from the original analysis according to identity are also plotted for comparison (triangular symbols). Error bars on each symbol correspond to ± 1 standard deviation.

with the differences between genders (recall that there were five male and five female faces in the set). Because of this ability, observers may have defaulted to focusing their efforts on making gender discriminations during the initial stages of learning. This idea is similar to Gibson's (1969) notion of differentiation, where initially indiscriminable percepts progressively become distinct with practice. This idea is also in accord with more recent empirical work by Gauthier and Tarr (1997), where human observers were able to learn basic level discriminations (e.g., gender) more quickly than subordinate level discriminations (e.g., identity) when trained to recognize unfamiliar objects (‘greebles’). Further, the non-linearity analyses above showed that part of the learning process involved increases in uncertainty, and these results are consistent with a model where the observer increases the number of templates used to perform the task with practice (i.e., a shift from gender-based to identity-based discriminations). Thus, a prediction of this model is that gender learning should occur primarily during the earliest stages of training. To explore this possibility, the data from the face identification experiment were re-analyzed according to the correctness of the observer's response with respect to the gender of the stimulus. That is, the observer's responses were scored as 'correct' if their response matched the gender of the stimulus, regardless of whether the response matched the stimulus identity. There were five male and five female faces in the set, and each was equally likely to appear on a given trial. As a result, there was no need to correct for unbalanced frequencies of occurrence. Psychometric functions were

fit to the data in each condition (with threshold defined as 71% correct). All of the data from both subjects and the ideal observer in the original face identification experiment were re-analyzed in this fashion, allowing us to compute calculation efficiency and equivalent input noise estimates for observers' ability to discriminate between male and female faces[14].

The results of this analysis is shown in Figure 3.10. The top panel corresponds to calculation efficiency estimates and the bottom panel equivalent input noise estimates for the original two observers in the face identification task. The data from the gender analysis (large circular symbols) have been plotted along with the data from the original identity analysis (small triangles) for reference. Although the reanalyzed data are much noisier than the original data[15], they show the familiar trend of an increase in calculation efficiency with no consistent change in equivalent input noise across learning sessions. Interestingly, gender discrimination efficiency was consistently higher than identity discrimination efficiency for both observers across sessions. This is particularly surprising, given that the task did not explicitly require observers to make discriminations

[14] Strictly speaking, the ideal decision rule in a categorical discrimination task (with more that one member within each category) is no longer the simple cross-correlation rule described in the Appendix. The ideal observer's identification data was nonetheless re-analyzed in the same fashion as the human observers' identification data because the human observers did not actually perform a gender discrimination task.

[15] The noisiness of the data is due to the fact that the spacing of the stimulus levels was not tailored for a gender discrimination task. Specifically, the stimulus levels tended to be shifted to higher locations on the psychometric functions (due to higher percent corrects), thus reducing the number of constraints placed on the psychometric fits.

according to gender. Finally, learning did not appear to plateau significantly earlier for gender discrimination that for identity discrimination. However, it is important to note that these two variables are highly confounded: if identity learning occurs, performance for *both* kinds of discriminations will improve. As a result, the increase in gender discrimination efficiency with learning may be an artifact of identity discrimination learning. Further experiments are necessary to tease these two factors apart.

Contrast-Dependent Internal Noise

The results of the Experiments 3.1 and 3.2 indicate that perceptual learning improves calculation efficiency but has no effect on internal contrast-invariant noise. However, as described in Chapter 1, the quantity we defined as 'calculation efficiency' confounds the effects of the goodness of internal calculations with the effects of any contrast-dependent internal noise that is proportional to the strength of the stimulus. Thus, the above experiments do not allow us to distinguish between increases in calculation efficiency and decreases in proportional contrast-dependent internal noise. Double-pass response consistency in high external noise (where contrast-dependent internal noise dominates) offers a way to discriminate between the impact of these two factors. Specifically, if decreases in proportional contrast-dependent internal noise contributed to the increases in the estimates of calculation efficiency found in Experiments 3.1 and 3.2, observers should become more consistent as learning takes

place in large amounts of external noise. This will show up as an increase in responses consistency between passes through identical sets of stimuli in high external noise.

Thus, the experiments described in this section involved the measurement of changes in contrast-dependent internal noise using the double-pass response consistency technique described in Chapter 1. As in the previous experiments, this technique is applied to the identification of both faces (Experiment 3.3) and textures (Experiment 3.4) as perceptual learning takes place. Again, the reader is referred to the theoretical framework described in Chapter 1 and the methods described in Chapter 2 for more details. Only the aspects of the methodology specific to the current set of experiments are presented here.

Methods

Stimuli**.** The stimuli used were the same sets of 10 faces and 10 textures used in the previous experiments. In each task, signal energy identification thresholds were measured in the highest levels of external noise power spectral density used in the previous experiments (faces: 25.55×10^{-6} deg^2; textures: 51.10×10^{-6} deg^2). *N/2* unique noise fields were generated for each experimental session, where *N* is the number of trials within a given session. The sequence of signal identities, signal energy levels, and seeds used to generate the noise fields during the first half of the experiment were saved before

every trial to allow for the exact reproduction of the same sequence of stimuli during the second half of the experiment.

Procedure. Signal energy was varied across trials during the first half of each session according to two interleaved UDTR staircases (see Chapter 2). Together, these two staircases offered a wide sample of the observer's psychometric function. This first half of the session consisted of 200 trials per staircase (400 trials total). The second half of the session consisted of an exact replication of the first half of the session (i.e., an exact pixel-by-pixel reproduction of the sequence of trials shown during the first half of the session was shown again during the second half of the session). Thus, each session consisted of a total of 800 trials. Each session was completed without breaks and lasted about one hour. Only one session was completed each day. Each observer completed six sessions within ten days. All of the observers were unaware that the first and second halves of each session were identical.

Observers. Two observers participated in the face identification task and two in the texture identification task. All of the observers were naive to the purposes of the experiment.

Experiment 3.3: Face Identification

Signal energy thresholds in the face identification task are plotted as a function of session for both observers in Figure 3.11. This figure shows a clear improvement across

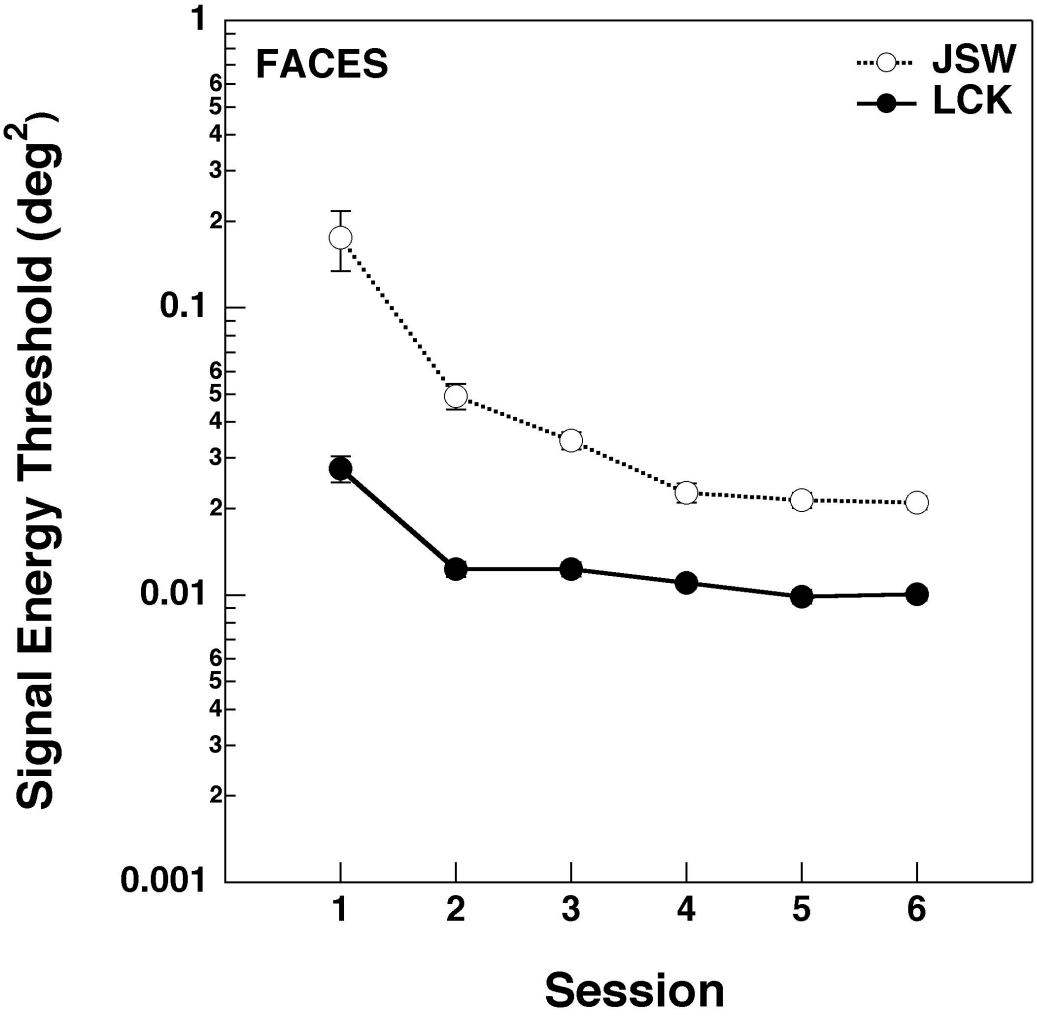


Figure 3.11 Signal energy thresholds in high external noise plotted as a function of session for the two observers in the 1-of-10 face identification task. The external noise power spectral density was set to the highest level used in Experiment 3.1 (25.55×10^{-6} deg^2). Error bars on each symbol correspond to ± 1 standard deviation.

sessions for both observers: one observer (LCK) exhibited an improvement across sessions similar to those found in Experiment 3.1 (about a factor of 2.8); the second observer (JSW) exhibited a markedly higher degree of improvement across sessions (about a factor of 8.2). However, note that this second observer's initial threshold was much higher than the other observers' initial thresholds (both from this experiment and

Experiment 3.1). Regardless of individual differences, these results demonstrate a clear learning effect for both observers.

Figure 3.12 shows the results of the corresponding consistency analysis for each observer. Each panel plots percent correct as a function of percent agreement for one of the observers. Each symbol corresponds to a single stimulus level within a given session. Recall that an adaptive staircase varied the contrast during the first half of the session, so not all of the data points received the same number of trials. Instead, the data points surrounding 50% and 79% correct received the most trials (due to the nature of the staircase rules chosen)[16]. In both panels, the closed symbols correspond to the first three sessions, the open symbols the last three sessions. If a decrease in contrast-dependent internal noise was responsible for some or all of the decrease in thresholds with learning, percent agreement should increase across sessions. That is, if contrast-dependent internal noise decreased with learning, we would expect to see a rightward shift in the data across sessions. However, there appear to be no systematic changes in percent agreement across sessions. Instead, for each observer the data from all of the sessions appear to fall around a single line. In each panel, the solid line corresponds to the predictions of an observer

[16] Percent correct is based on the percentage of correct responses from the entire session. Percent agreement is computed by comparing the responses on corresponding trials in the two passes through the stimulus set and calculating the percentage of trials where the responses were the same (regardless of accuracy). Notice that percent correct and percent agreement are highly correlated: this is because percent agreement will naturally increase with increasing percent correct.

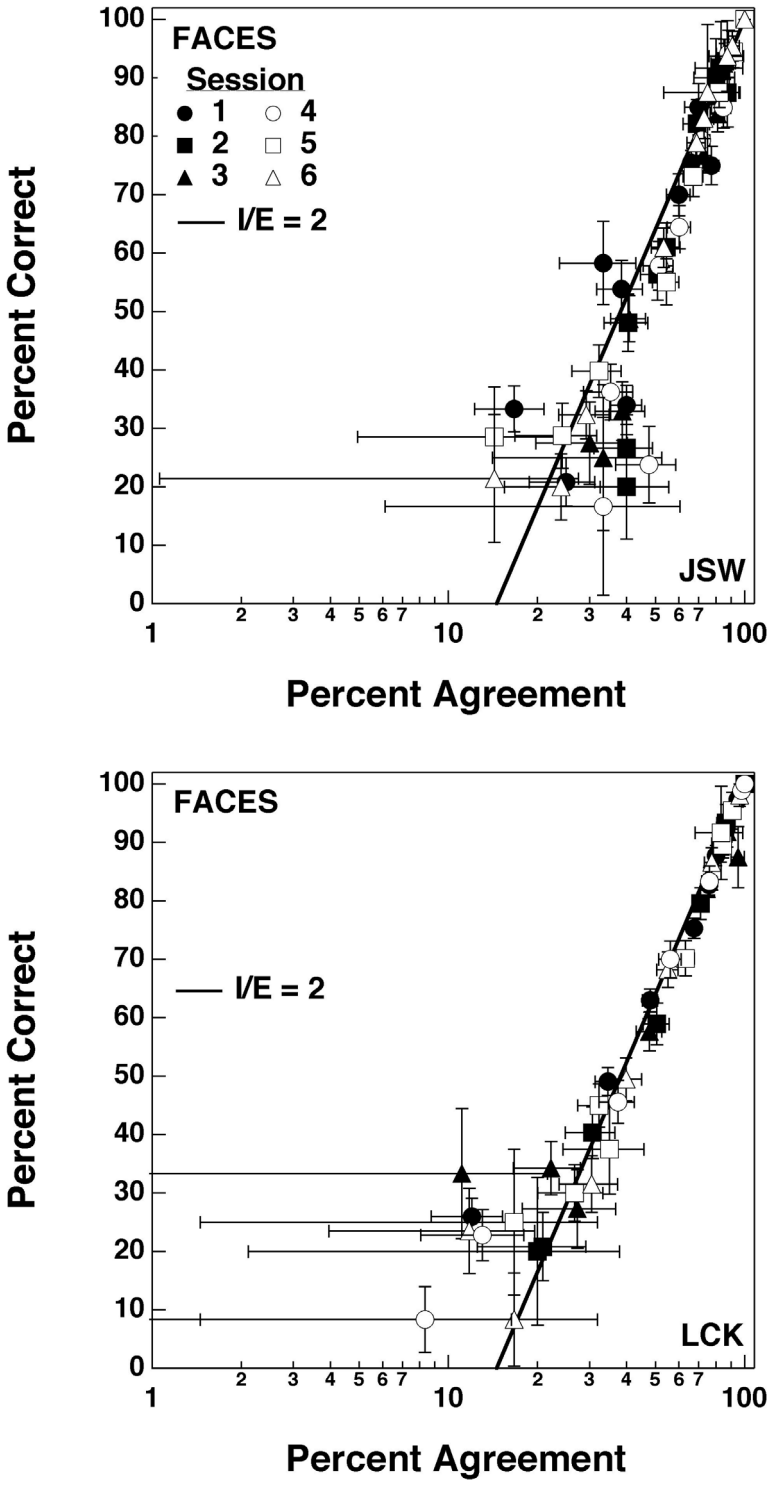


Figure 3.12 Response consistency plots for the two observers in the 1-of-10 face identification task across six sessions. Each panel plots percent correct as a function of percent agreement for each stimulus level tested. The filled symbols correspond to the first three sessions, the open symbols the last three sessions (see legend). Error bars on each symbol correspond to ± 1 standard deviation. The large variation in error bar magnitude is due to the unequal number of trials at each data point (a result of the use of a staircase procedure). Solid lines in each plot correspond to the performance of a simulated observer with an internal/external noise ratio approximately equal to the average of the estimated internal/external noise ratios across sessions (see Figure 3.13).

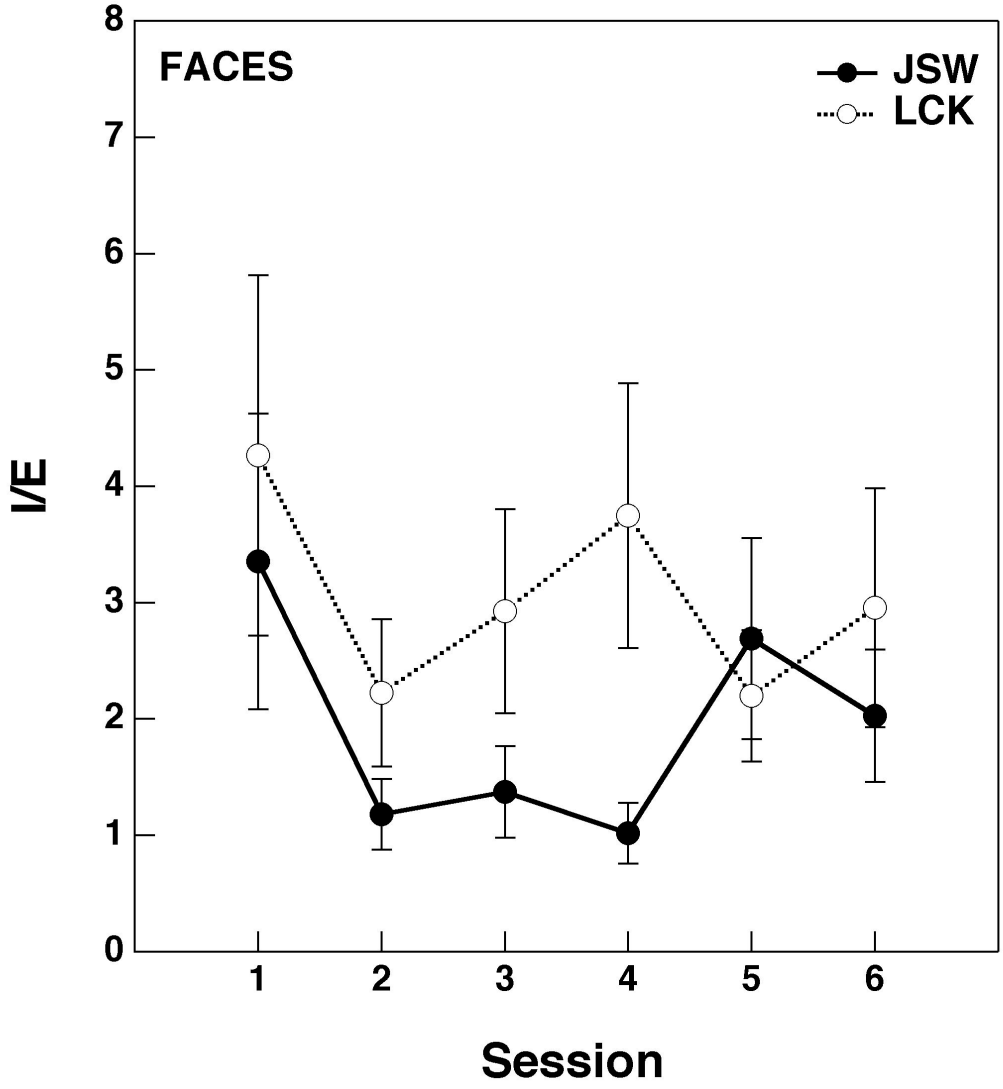


Figure 3.13 Internal/external noise ratio estimates for the two observers in the 1-of-10 face identification task across six sessions. Error bars on each symbol correspond to ± 1 standard deviation.

with an internal/external noise ratio of 2.0.

This trend is shown quantitatively in Figure 3.13, where the internal/external noise ratios for each observer are plotted as function of practice. The data for each observer from each session were fit to Equation 2.9, and Equation 2.10 was used to compute estimates of the internal/external noise ratios. Although there is a large degree of variability across sessions, there is no systematic decrease in the estimated

internal/external noise ratio with practice, implying contrast-dependent internal noise did not decrease significantly with learning. In addition, the average internal/external noise ratio for each observer is between 2.0 and 3.0, a range significantly higher than the estimates of 0.8-1.0 reported in previous experiments (Burgess & Colborne, 1988; Green, 1964). The presence of such a large contrast-dependent internal noise indicates the previous estimates of the efficiency of internal calculations were significantly underestimated by the analysis in Experiment 3.1 [17].

Experiment 3.4: Texture Identification

The results of Experiment 3.4 are shown in Figures 3.14-3.16. Figure 3.14 shows signal energy thresholds for both observers, as a function of session. As found the in previous experiments, the performance of both observers improved with practice. However, the magnitude of improvement greatly exceeded that found in the previous

[17] The internal/external noise ratio can be used to factor out the contribution of proportional internal noise to the previous estimates of calculation efficiency. Recall that the internal/external noise ratio is expressed in terms of noise standard deviation, and that the human and ideal noise-masking functions are linear when expressed in terms of signal energy and noise power spectral density. Noise power spectral density is proportional to the variance of the noise(squared standard deviation). Thus, the observer's calculation efficiency in the absence of proportional internal noise may be estimated by adding the uncorrected calculation efficiency to the product of the squared internal/external noise ratio and the uncorrected calculation efficiency. If we estimate the internal/external noise standard deviation ratio to be 2 in the face identification experiment, this would yield an internal/external noise variance ratio of about 4. This value would increase the calculation efficiency estimates from Experiment 3.1 from about 2% to about 10%. In the texture identification experiment reported below, the average estimated internal/external noise standard deviation ratio was about 1.3, yielding an internal/external noise variance ratio of about 1.7. This value would increase the calculation efficiency estimates from Experiment 3.2 from about 2% to about 5.5%.

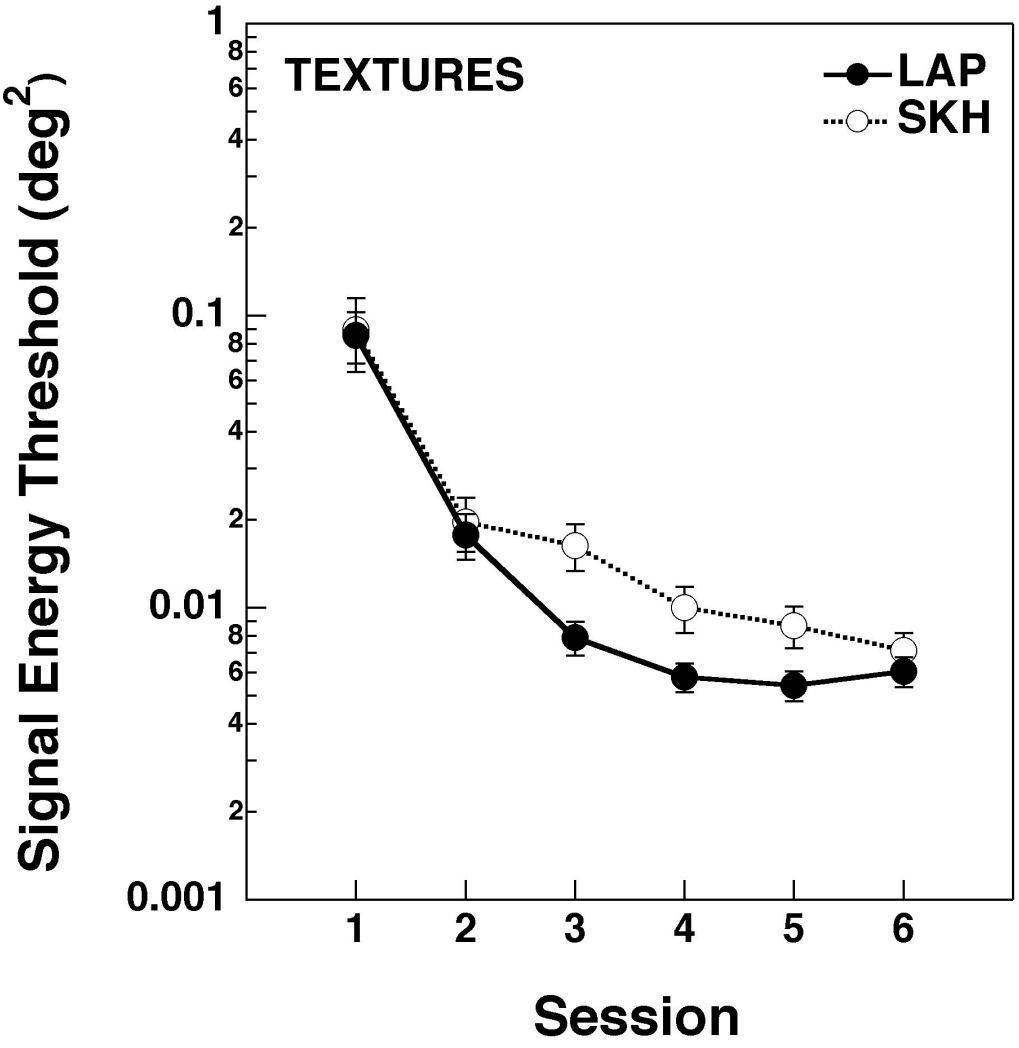


Figure 3.14 Signal energy thresholds in high external noise plotted as a function of session for the two observers in the 1-of-10 texture identification task. The external noise power spectral density was set to the highest level used in Experiment 3.2 (51.10×10^{-6} deg^2). Plotting conventions are the same as in Figure 3.11.

texture identification experiment: observer LAP improved by a factor of 14 and observer SKH by a factor of 12.5 Unless this discrepancy is simply due to individual differences, it must be a result of using an adaptive threshold estimation procedure and/or restricting the power of the external noise to a single level. Regardless, both observers showed a clear effect of learning in the task.

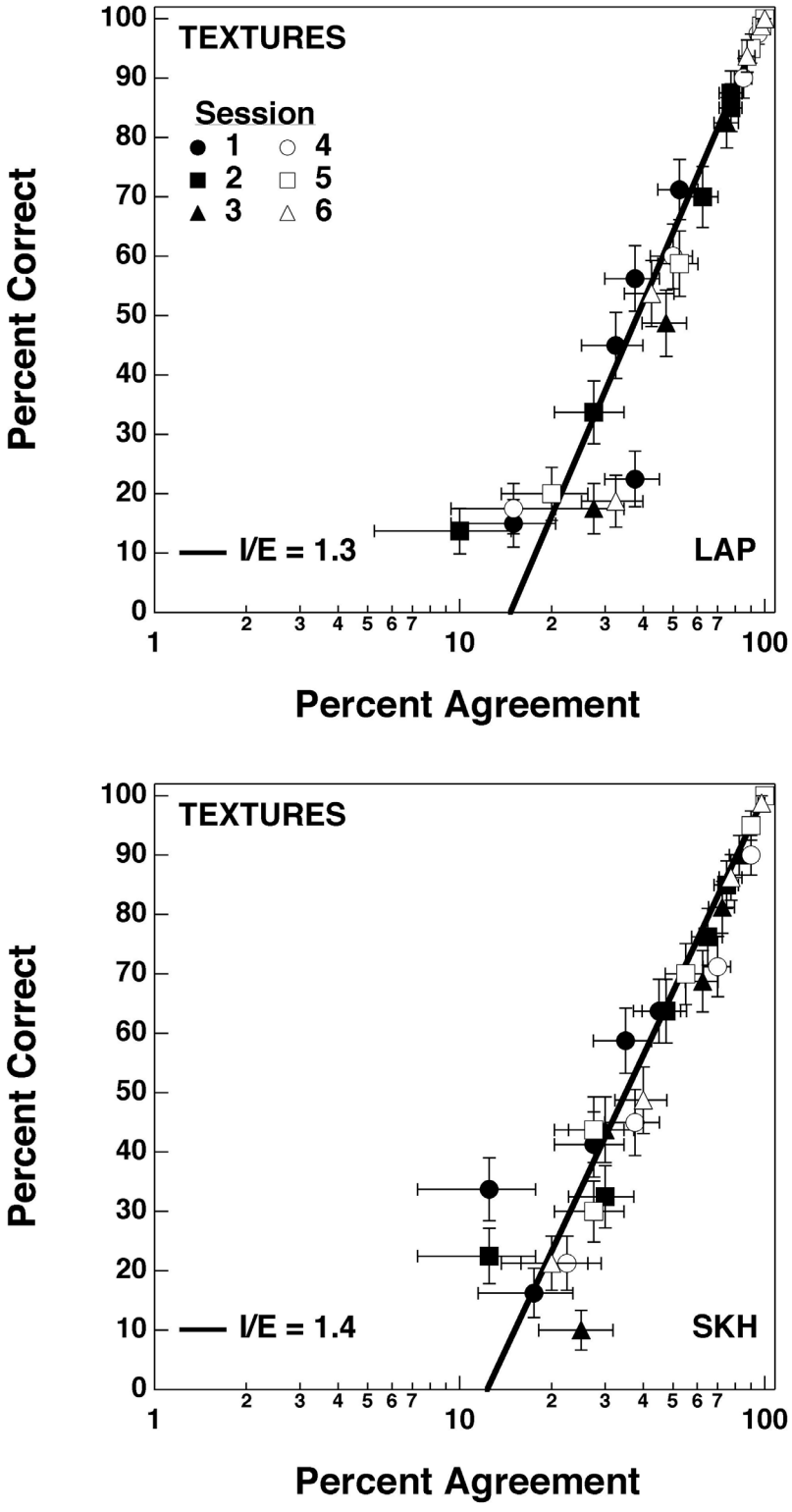


Figure 3.15 Response consistency plots for the two observers in the 1-of-10 texture identification task across six sessions. Plotting conventions are the same as in Figure 3.12.

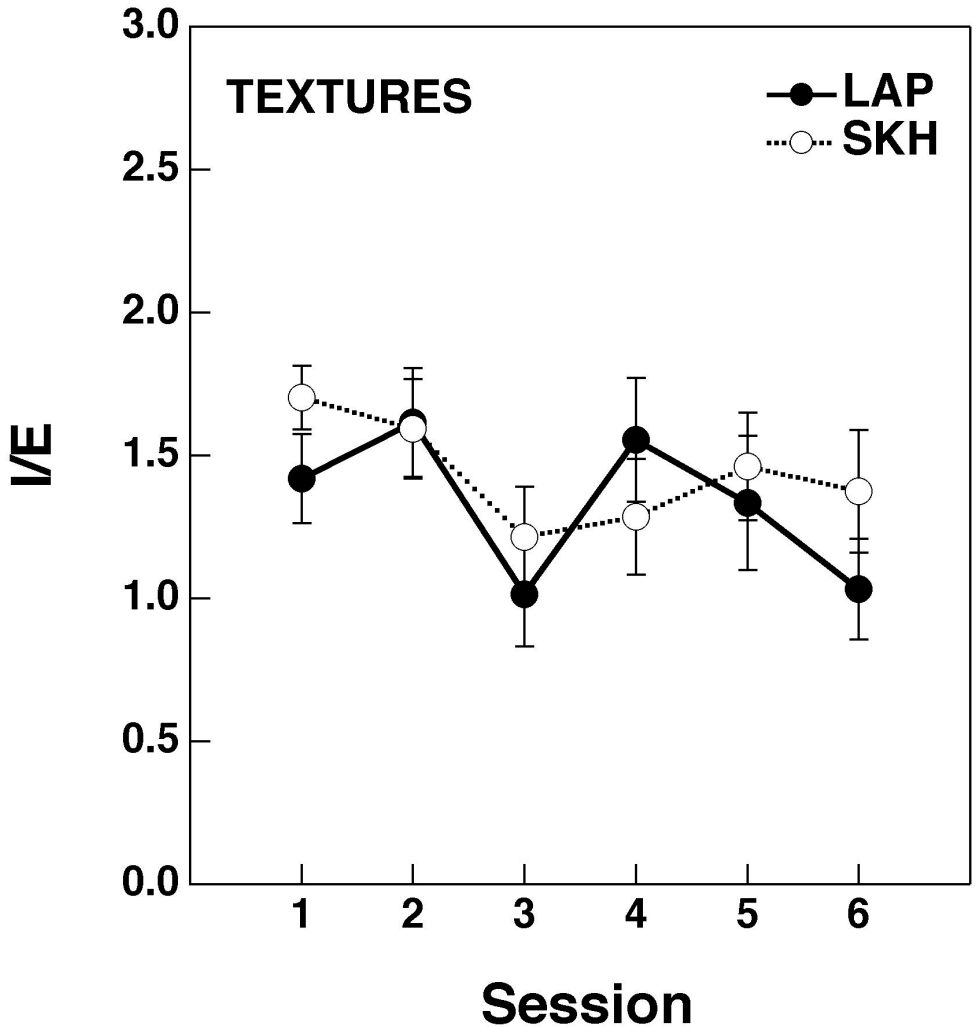


Figure 3.16 Internal/external noise ratio estimates for the two observers in the 1-of-10 texture identification task across six sessions. Plotting conventions are the same as in Figure 3.13.

Figure 3.15 shows the results of the corresponding consistency analysis for each observer. The plotting conventions are the same as in Figure 3.12. As in the face identification task, for each observer the data from all of the sessions fall around a single line. In each panel, the solid line corresponds to the predictions of an observer with a particular internal/external noise ratio (LAP: 1.3; SKH: 1.4). This trend is reflected in Figure 3.16, which shows the corresponding internal/external noise ratio estimates for

each observer across sessions. As in the face identification experiment, these data show no systematic changes in the internal/external noise ratio estimates across sessions, indicating contrast-dependent noise did not contribute to the reduction in thresholds with practice. However, unlike the face identification experiment, the average internal/external noise ratio for each observer falls between around 1.3, a value only marginally higher than the estimates of 0.8-1.0 reported in previous experiments.

Discussion

The results from the response consistency analyses of Experiments 3.3 and 3.4 show that contrast-dependent internal noise does not decrease significantly with learning in both the face and texture recognition tasks. These results, combined with the noise masking results from Experiments 3.1 and 3.2, are consistent with the idea that the only significant change that occurs with perceptual learning is an increase in the efficiency of internal calculations. In addition, the fact that the effects of learning were the same for faces and abstract textures strongly suggests that this effect is a general property of perceptual learning in pattern identification tasks. However, as alluded to in the introduction, one of the implicit assumptions of the model outlined in Chapter 1 is that all of the internal noise is introduced *before* the calculations have taken place (see Figure 1.4). In contrast, Dosher and Lu's Perceptual Template Model (Dosher & Lu, 1998; Dosher & Lu, 1999, 2000; Lu & Dosher, 1998, 1999; Lu et al., 2000) assumes all of the

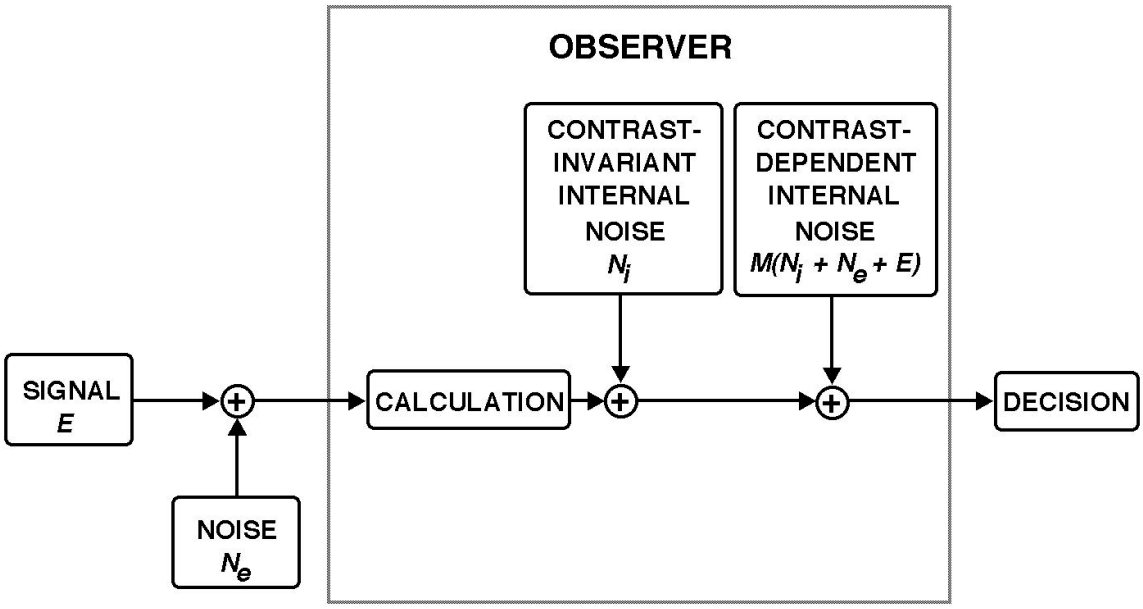


Figure 3.17 'Late' noise version of the black-box model illustrated in Figure 1.4. Unlike the model described in Figure 1.4, the internal noise is introduced after the internal calculations have taken place.

internal noise is introduced after the calculations have taken place (Figure 3.17). Such a 'late noise' model makes very different predictions about how an observer's noise masking function will be affected by changes in the efficiency of internal calculations. This can be seen by considering Equation 1.1 with a contrast-invariant noise whose effect on performance is independent of the calculation:

$$E = kN_e + N_i \tag{3.1}$$

This may be expanded to include a contrast-dependent internal noise that is some proportion m of the sum of E, N_e, and N_i:

$$E = kN_e + N_i + m(E + kN_e + N_i)$$

$$= (kN_e + N_i)((1 + m)/(1 - m))$$

$$E = k'(kN_e + N_i)$$

$$\log(E) = \log(k') + \log(kN_e + N_i) \quad (3.2)$$

where k' is equal to $(1+m)/(1-m)$. Comparison of Equations 1.3 (Chapter 1) and 3.2 reveals both early and late noise models make the same predictions about the effects of changes in both contrast-invariant and contrast-dependent internal noise. However, notice that the efficiency parameter k is not incorporated within the constant k'. Instead, it only affects the internal representation of the magnitude of the external noise. Accordingly, reducing k only affects thresholds at high external noise levels. This property is illustrated in Figure 3.18, which demonstrates the effects of changes in internal noise and calculation efficiency on the noise masking function of an observer with only late internal noise.

How does a late noise model account for our results? Experiments 3.1-3.2 showed a change in the noise masking function that is consistent with a change in proportional noise in a late noise model. However, Experiments 3.3-3.4 showed that contrast-dependent internal noise does not change with learning. Thus, in the absence of a change in contrast-dependent internal noise, the only way a late noise model is able

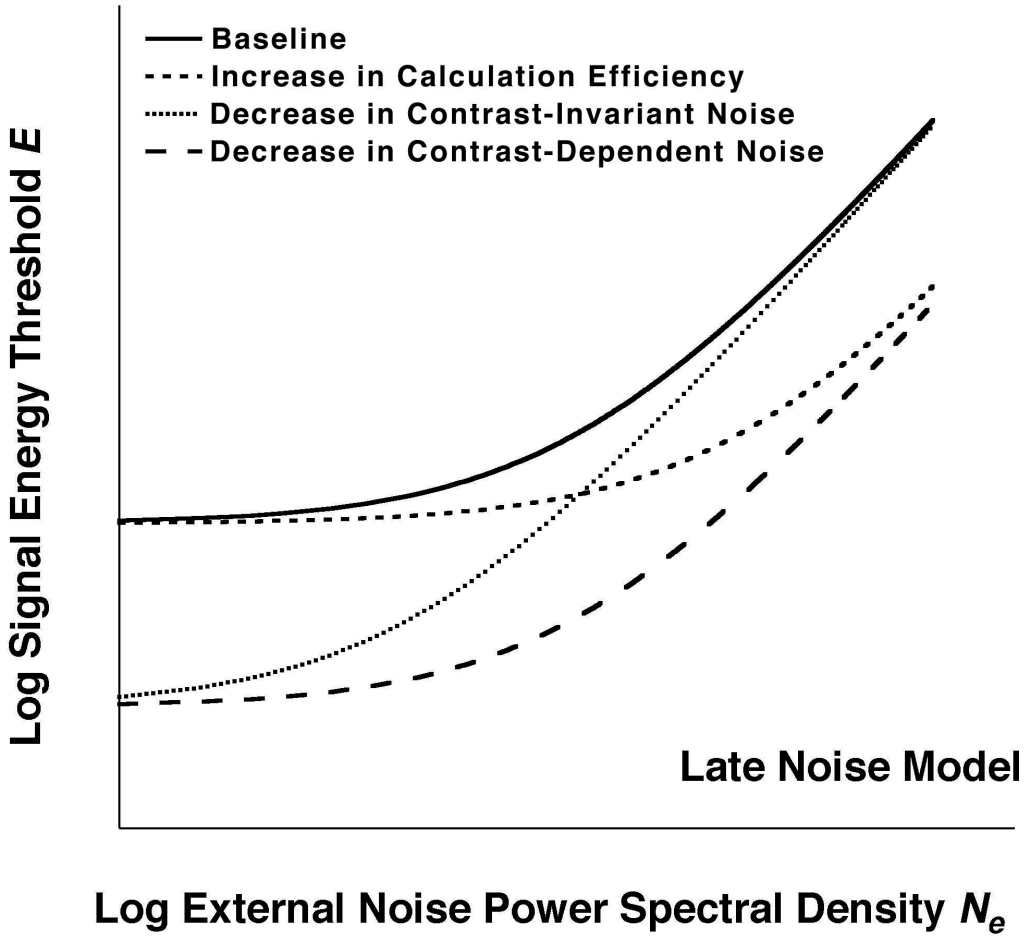


Figure 3.18 Hypothetical noise-masking functions for an observer with 'late' internal noise. Log of signal energy threshold (E) is plotted as a function of external noise power spectral density (N_e). The solid line depicts the possible performance of an observer that performs contrast-invariant calculations and has both contrast-invariant and contrast-dependent internal noise. The medium dashed line depicts the changes in the noise-masking function predicted by an increase in the efficiency of internal calculations. The finely dashed line depicts the changes in the noise-masking function predicted by decrease in contrast-invariant internal noise. The coarsely dashed line depicts the changes in the noise-masking function predicted by a decrease in contrast-dependent internal noise that is proportional to stimulus magnitude.

account for our results is by positing identical, concurrent reductions in both contrast-invariant internal noise and calculation efficiency as learning takes place. This prediction of the late noise model can be tested by measuring response consistency in a low level of external noise. Under low noise conditions, the dominant internal noise will be the

contrast-invariant component. Thus, a late noise model predicts that the observer's internal/external noise ratio should decline as learning takes place in low external noise.

Four observers performed the same face and texture identification tasks as in Experiment 3.3 and 3.4, with the exception that the external noise power spectral density was set to the lowest values used in Experiments 3.1 and 3.2 (faces:0.04×10^{-6} deg^2; textures: 0.20×10^{-6} deg^2). In addition, a set of 10 new textures were generated, allowing the two observers from Experiment 3.4 to participate in the low external noise condition (to allow for comparison of results across external noise conditions). Two new observers were recruited for the face identification task. All observers participated in six sessions within eight days, with the exception of one observer (RRB) who was unable to complete the final session due to illness.

The results are shown in Figures 3.19-3.21. Figure 3.19 shows thresholds for each observer in the face (top panel) and texture (bottom panel) identification tasks, as a function of session. Figure 3.20 shows the corresponding consistency plots and Figure 3.21 shows the estimated internal/external noise ratios for each observer. As in the previous experiments, thresholds declined significantly with practice in both tasks, but none of the observers showed a corresponding increase in consistency. These results validate our previous conclusion (based on the measures of equivalent input noise from Experiments 3.1 and 3.2) that internal contrast-invariant noise is not affected by learning, and are inconsistent with the predictions of a late noise model. However, we also found

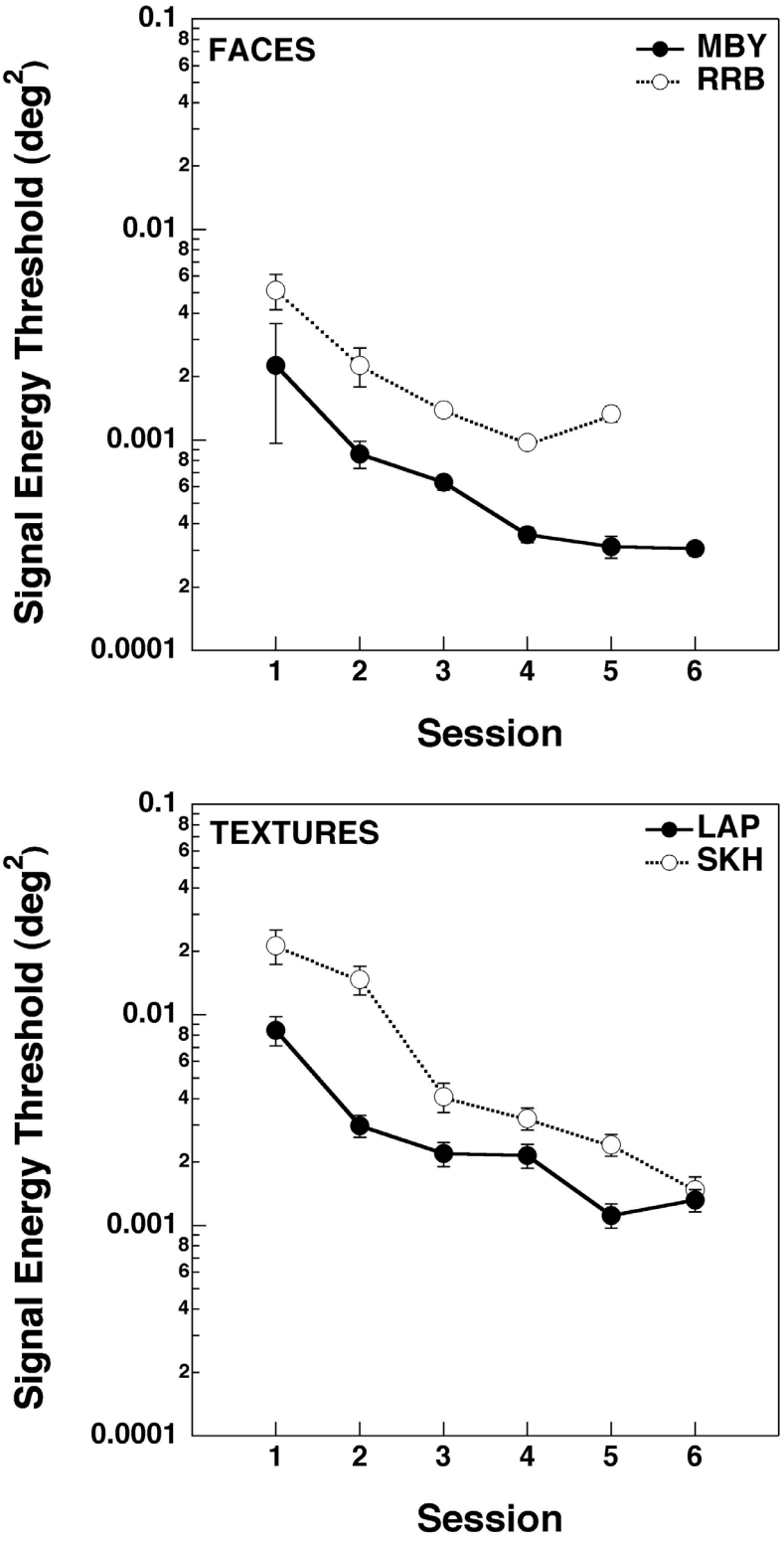


Figure 3.19 Signal energy thresholds in low external noise plotted as a function of session for observers in the 1-of-10 face (top panel) and texture (lower panel) identification tasks. The external noise power spectral density was set to the lowest level used in either Experiment 3.1 or 3.2 (faces:0.04 x 10^{-6} deg^2; textures: 0.20 x 10^{-6} deg^2). Plotting conventions in each panel are the same as in Figure 3.11.

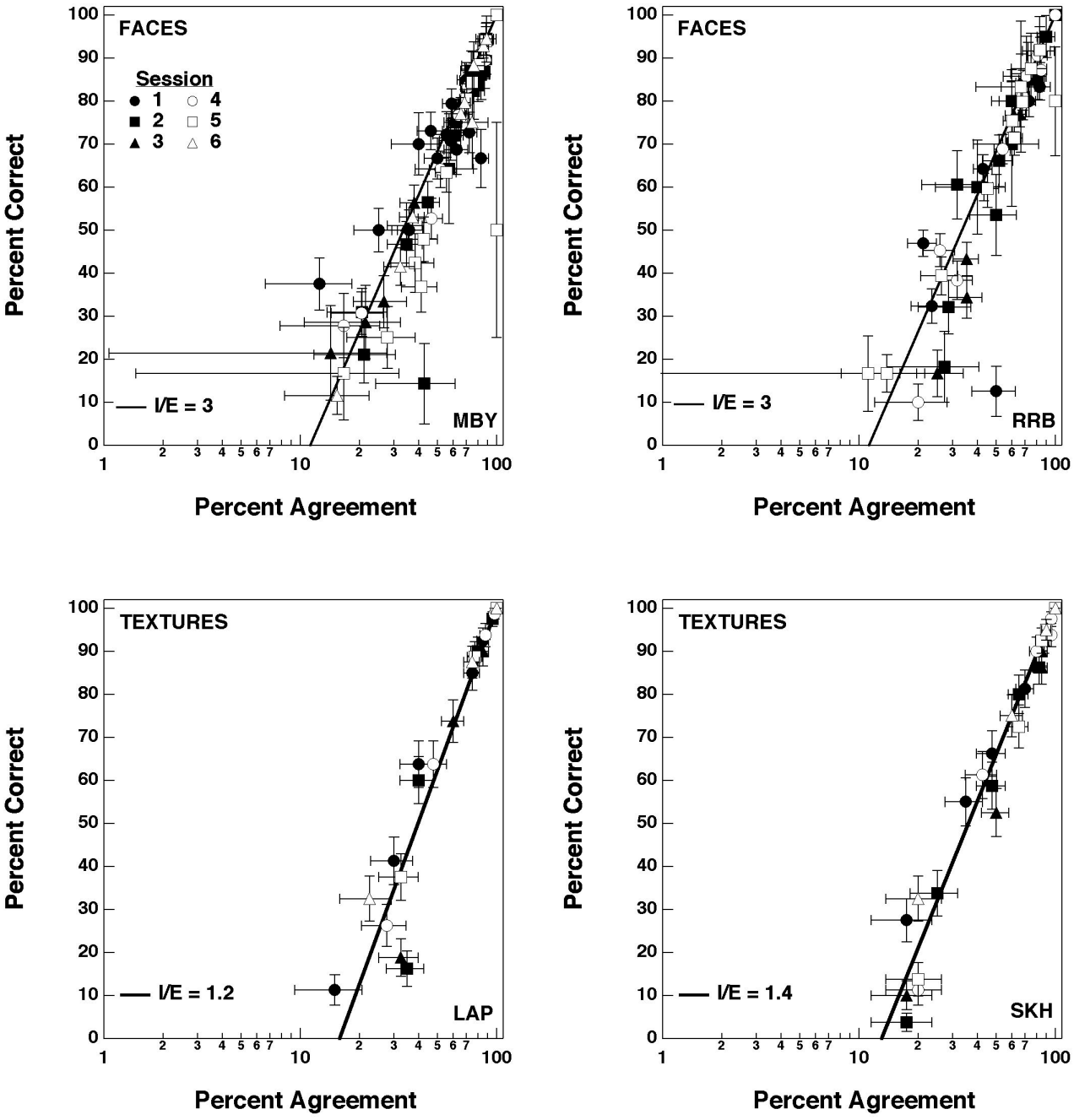


Figure 3.20 Response consistency plots for the observers in the 1-of-10 face (top two panels) and texture (bottom two panels) identification tasks in low external noise across sessions. Plotting conventions for each panel are the same as in Figure 3.12.

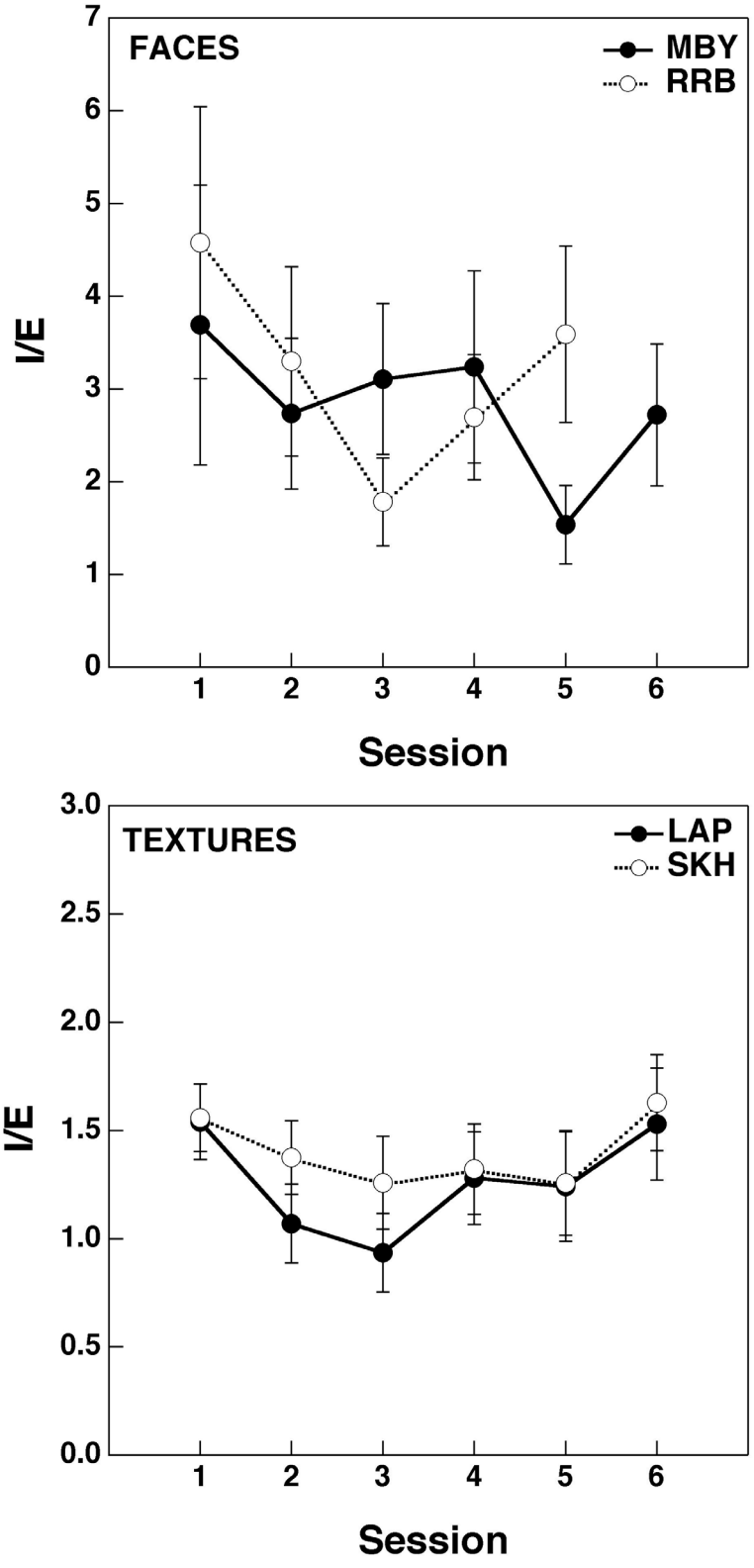


Figure 3.21 Internal/external noise estimates for the observers in the 1-of-10 face (top panel) and texture (bottom panel) identification tasks in low external noise across sessions. Plotting conventions are the same as in Figure 3.13.

in Experiments 3.1 and 3.2 that equivalent noise differed for faces and textures, a result that is consistent with the existence of a late internal noise. One possible explanation for this apparent discrepancy is that observers may have made use of the outputs of very different early mechanisms in the face and texture identification tasks (perhaps due to their extremely different bandwidths) and that these early mechanisms differ greatly in their stochastic properties.

Conclusions

The results of the experiments in this chapter offer compelling evidence that it is only the efficiency of internal calculations that changes with perceptual learning in visual pattern recognition tasks. In all of the experiments reported thus far, identification thresholds have been used to make inferences about the changes that take place with perceptual mechanisms as learning occurs. However, one shortcoming of the noise masking and response consistency techniques is that they do not allow us to specify the particular kinds of changes that occur within an observer's internal representations as learning takes place. The fact that calculation efficiency improves with learning implies that internal representations (i.e., templates) become more ideal with practice. This prediction of the model is the topic of Chapter 4, where the response classification technique described in Chapters 1 and 2 is used to measure observers' calculations as perceptual learning takes place.

4 Changes in Observer Calculations with Learning

Introduction

The purpose of this chapter is to extend the results of the previous set of experiments to include a more detailed description of the changes that take place with perceptual learning. The experiments in Chapter 3 showed that it is only the efficiency of internal calculations that change as learning occurs. However, the exact nature of those changes remains unspecified. That is, what is it about the calculations that changes with practice, exactly? This chapter will attempt to address this question by using a technique called *response classification*. Recall from Chapter 1 that the response classification technique involves randomly perturbing a signal to recover the aspects of the stimulus used by an observer in a task. This is achieved by correlating the random perturbations with the observer's responses across trials. In a visual pattern recognition task, this amounts to adding noise to a signal and correlating the contrast of the noise at each stimulus location with the observer's responses across trials. The end result is a correlation map called a *classification image* that shows the relative influence of each pixel on the observer's responses (see Chapter 1 for a more detailed description of response classification and Chapter 2 for details about how classification images are computed). For a linear observer, the classification image will be proportional to an observer's calculation (Abbey, Eckstein, & Bochud, 1999). In the context of perceptual learning, the response classification technique can be used to trace the changes in an observer's calculations as learning takes place. In addition, human and ideal

classification images can be compared to test the prediction that an observer's calculations become more efficient (and thus more ideal) as learning occurs.

Thus, the experiments in this chapter involve the application of the response classification technique to the identification of novel faces (Experiment 4.1) and textures (Experiment 4.2). As mentioned in Chapter 1, the stimulus sets were reduced to only two signals per set for in order to expedite the data collection and simplify the analyses. As in Chapter 3, the experiments rely heavily on the theory described in Chapter 1 and the methods described in Chapter 2. Only the aspects of the methodology specific to the present experiments are described here.

Methods

Stimuli**.** The stimuli used were the sets of two faces and two textures shown in Figure 2.1 and 2.2. In each task, signal energy was manipulated across trials according to a single UDTR staircase to maintain 71% correct performance throughout the session. The staircase was reset to its initial state at the beginning of each session. The signals were embedded in the highest level of external noise power spectral density used in the previous experiments (faces: 25.55×10^{-6} deg^2; textures: 51.10×10^{-6} deg^2). A unique noise field was generated before each trial. The sequence of signal identities, observer responses and seeds used to generate the noise fields were saved after every trial to allow for subsequent computation of the classification images (see Chapter 2 for details).

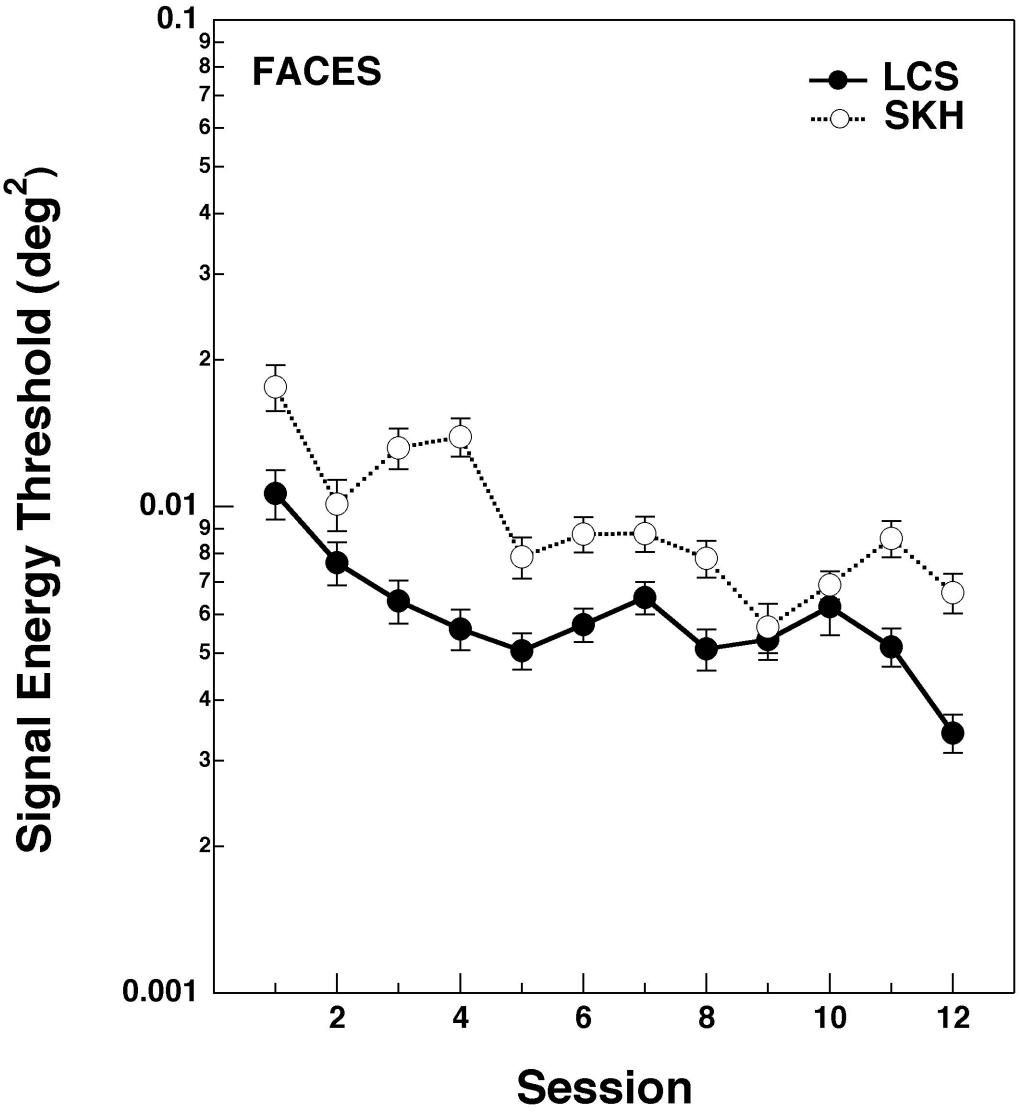


Figure 4.1 Signal energy thresholds plotted as a function of session for the two observers in the face identification task. The external noise power spectral density was set to the highest level used in Experiment 3.1 (25.55×10^{-6} deg^2). Error bars correspond to ± 1 standard error.

Procedure. Each session consisted of 800 trials that were completed without breaks over the course of about one hour. Only one session was completed each day. Each observer participated in a total of 12 sessions over the course of 16 days.

Observers. Two observers participated in the face identification task and two in the texture identification task. All of the observers were naive to the purposes of the experiment.

Experiment 4.1: Face Identification

Signal energy thresholds for the two observers in the face identification task are plotted as a function of session in Figure 4.1. Threshold was defined as the signal energy level yielding 71% correct performance, and was estimated by a maximum-likelihood fit of the staircase data to Equation 2.5. Figure 4.1 shows a clear effect of learning, with thresholds declining by about a factor of 3.0 for both observers over the course of the 12 sessions. Similar to the previous experiments, the majority of learning for both observers occurred within the first 4-6 sessions.

The corresponding classification images for each session are shown in Figure 4.2 for one observer (LCS). Sessions 1-3 are shown in the top row (from left to right), sessions 4-6 are shown in the second row; sessions 7-9 in the third row, and sessions 10-12 in the bottom row. Recall that the computation of the classification images involves sorting and averaging the noise according to the stimulus-response combinations and combining these averages to form a single classification image (see Equation 2.11). The contrast of a pixel in the classification image corresponds to the correlation between the contrast of the noise at that pixel across trials and the observer's responses (in this case,

the responses are 'face #1' or 'face #2'). In Figure 4.2, pixels that are brighter than mean gray indicate a positive correlation between the noise contrast at that pixel and the response 'face #2'. Alternatively, pixels that are darker than mean gray indicate a negative correlation between the noise contrast at that pixel and the response 'face #2'[18].

Inspection of Figure 4.2 reveals no obvious features emerging from the classification images. Unfortunately, these images are far too noisy to produce visible representations of the locations used by observers during each session. One reason for this noisy appearance is the relatively small number of trials used to compute each classification image. Typically, at least several thousand trials are needed to produce highly visible image features in the raw classification images (Abbey et al., 1999; Beard & Ahumada, 1998; Gold et al., 2000). A second reason for the noisy appearance most likely stems from visual masking by spatial frequencies outside of the range of frequencies used by the observer to perform the task (De Valois & De Valois, 1990). One way to address this issue is to filter the classification images according to the regions of the frequency spectrum that are thought to contribute to the observer's decisions. As mentioned in Chapter 2, previous results suggest observers make use of frequencies within a 2-octave wide band centered around 6 c/image for face identification. Figure 4.3

[18] The choice to correlate responses with 'face #2' is arbitrary. It results from choosing the second face in the set to correspond to *S2* and *R2* in Equation 2.11.

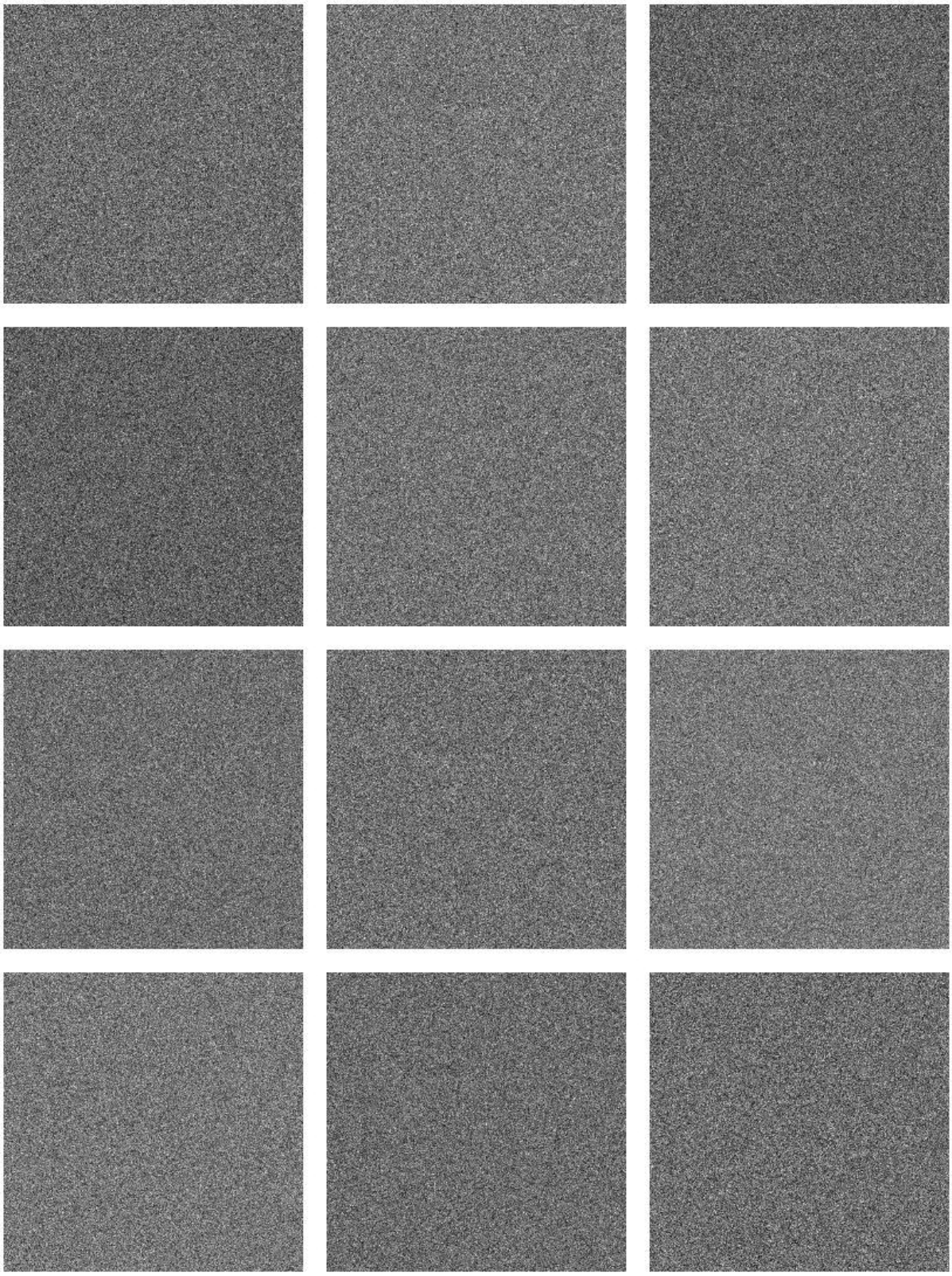


Figure 4.2 Raw classification images for observer LCS in the face identification task. Each panel corresponds to a single session. The top row corresponds to sessions 1-3 (from left to right), the second row to sessions 4-6, the third row to sessions 7-9, and the bottom row to sessions 10-12.

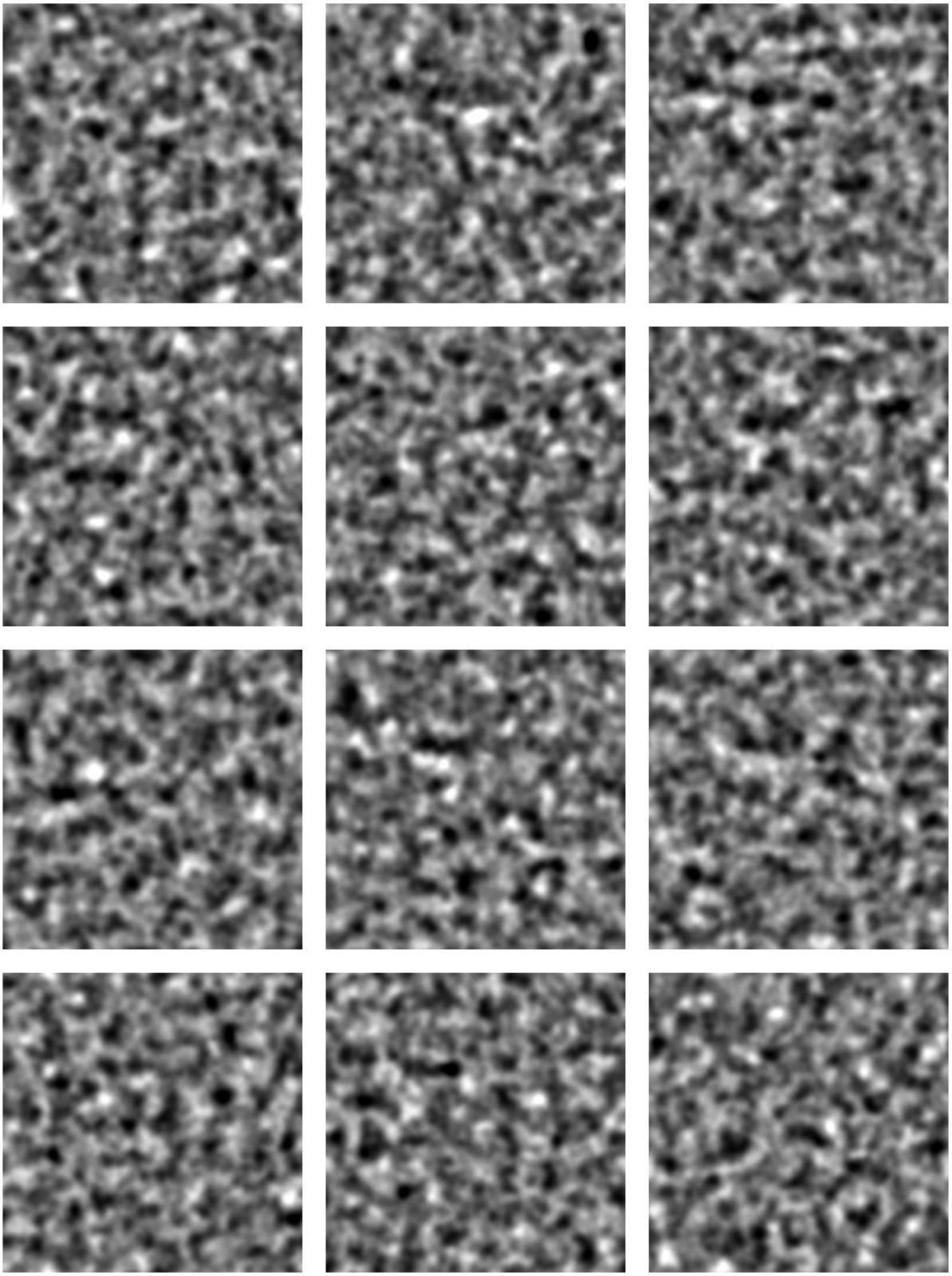


Figure 4.3 Smoothed classification images for observer LCS in the face identification task. Each panel corresponds to a single session. The top row corresponds to sessions 1-3 (from left to right), the second row to sessions 4-6, the third row to sessions 7-9, and the bottom row to sessions 10-12.

Figure 4.4 The ideal observer's classification image (template) in the face identification task (see Appendix for details).

shows the raw classification images from Figure 4.2 filtered with a log-Gaussian spatial frequency filter centered at 6 c/image with a bandwidth of 2 octaves (± 1 octave at half-height). Unfortunately, there still appear to be no discernible features within the classification images, presumably due to the small number of trials within each session.

Although it is not possible to visualize features in these data, statistical analyses can be used to test for global changes in the classification images. Specifically, the results from Chapter 3 imply that the observer's calculations become more efficient with learning. A strong prediction of these results is that the observer's classification images should become more similar to the ideal observer's classification image as learning takes

place[19]. It is shown in the Appendix that the ideal classification image or 'template' in a 1-of-2 identification task is the difference between the two possible signals (Equations 2.12 and A.5). The ideal template for the 1-of-2 face identification task is shown in Figure 4.4. The ideal template reveals that most of the information for this task and pair of signals is located around the eyes and eyebrows -- the places where the two faces differ the most. Figure 4.5 plots the normalized cross-correlation with the ideal template for the two observers in the 1-of-2 face identification task, as a function of session[20]. Each symbol corresponds to the correlation with the ideal template for one of the observers within a given session. If learning improves calculation efficiency, we would expect a significant increase in the correlation between the human and ideal templates across sessions. Although the data are noisy, there appears to be a gradual increase in correlation with practice. This trend was tested by computing least-squares linear fits to the data for each observer (the solid and dashed lines passing through the plots). The amount of variability in the data accounted for by a linear fit can be estimated by computing r^2, and the probability p that this value could have been obtained in the

[19] The ideal observer's classification image can be thought of as an 'information map', with the contrast at each pixel corresponding to its relative informativeness. This is a result of the fact that ideal observer is only constrained by the physical availability of information. If learning improves the efficiency of an observer's calculations, it follows that the calculations are becoming more similar to those used by the ideal observer. This should result in an increase in the similarity between the human and ideal classification images.

[20] The normalized cross-correlation ranges between -1.0 and 1.0, and is computed according to Equation 2.13.

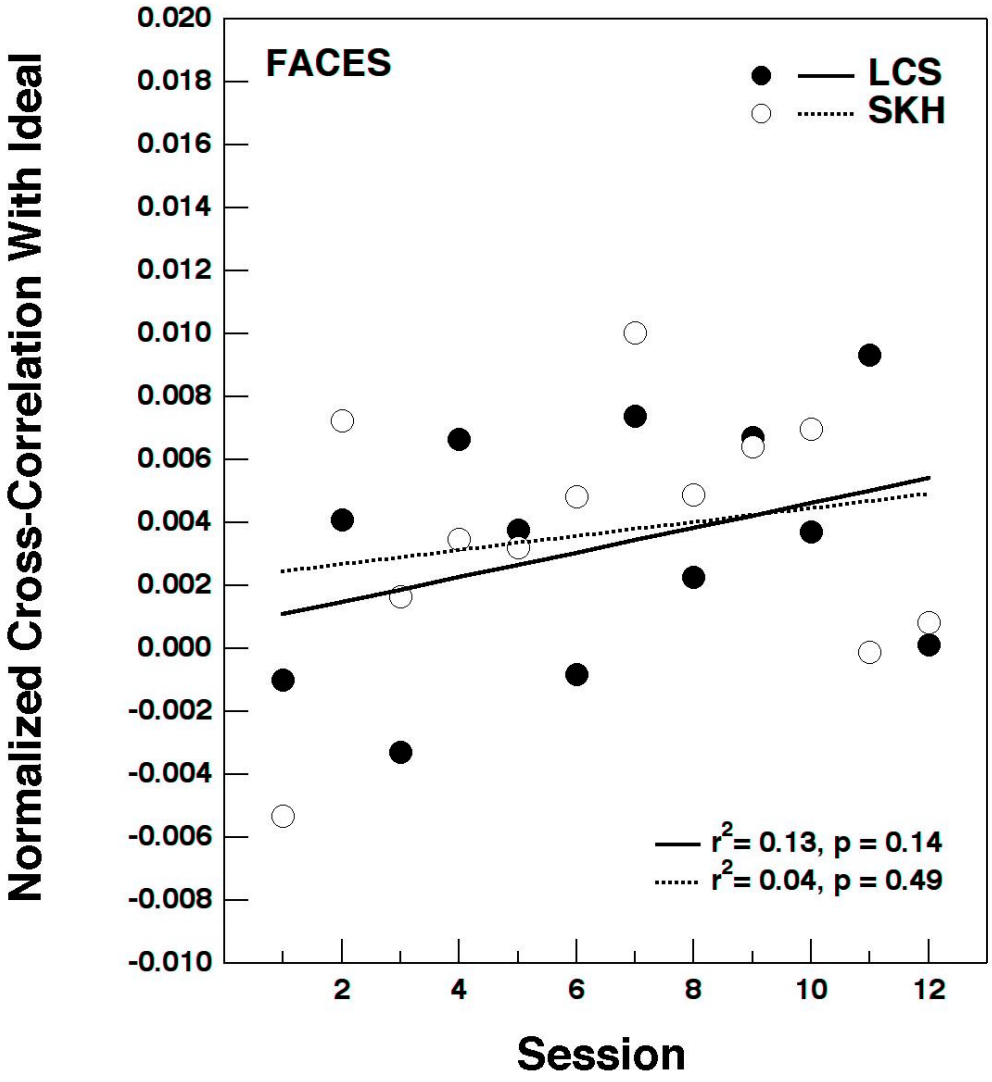


Figure 4.5 Normalized cross-correlations between the two observers' classification images and the ideal observer's template in the face identification task, plotted as a function of session. The lines passing through the plot correspond to least-squares linear fits to the data for each observer. The corresponding r^2 and p values for each fit are shown in the lower right corner.

absence of a real correlation can be estimated by using a t-test (McCall, 1986). The r^2, and p values for both observers are reported in the lower right corner of Figure 4.5. Despite the appearance of a gradual trend, these values show that there is not a significant linear increase in correlation with the ideal template across sessions for both observers.

The lack of a significant increase in correlation with the ideal template across sessions is not entirely surprising, given the relatively small number of trials (800) used to compute each classification image. If we are to assume that an observer uses some relatively small subset of pixels to perform the task, the noise falling on the vast majority of pixels will have no influence upon the observer's responses. As a result, the pixels used by the observer must be highly correlated with the ideal template in order to overcome the degrading effects of the remaining noisy pixels. One way this can be achieved is by increasing the number of trials in the classification image. Inspection of Figure 4.1 shows the majority of learning occurs within the first six sessions. This suggests that we may be able to increase the signal-to-noise ratio by collapsing across sessions within the first and second halves of the experiment. Figures 4.6 and 4.7 show the raw (top row) and smoothed (middle row) classification images for observers LCS and SKH, respectively, pooled across either the first (sessions 1-6; left column) or last (sessions 7-12; right column) half of the experiment. In addition, the bottom row in each figure shows the smoothed images after being submitted to a statistical test. In these images, all of the pixels that are not statistically significant ($p<0.01$) are set to mean gray. The remaining pixels have been set to unity and multiplied by the corresponding pixels in the ideal template. Inspection of both of these figures shows two major differences between the first and last halves of the experiment. First, both observers appear to concentrate more on the top left eyebrow region in the second half of the experiment.

LCS

Sessions 1-6 Sessions 7-12

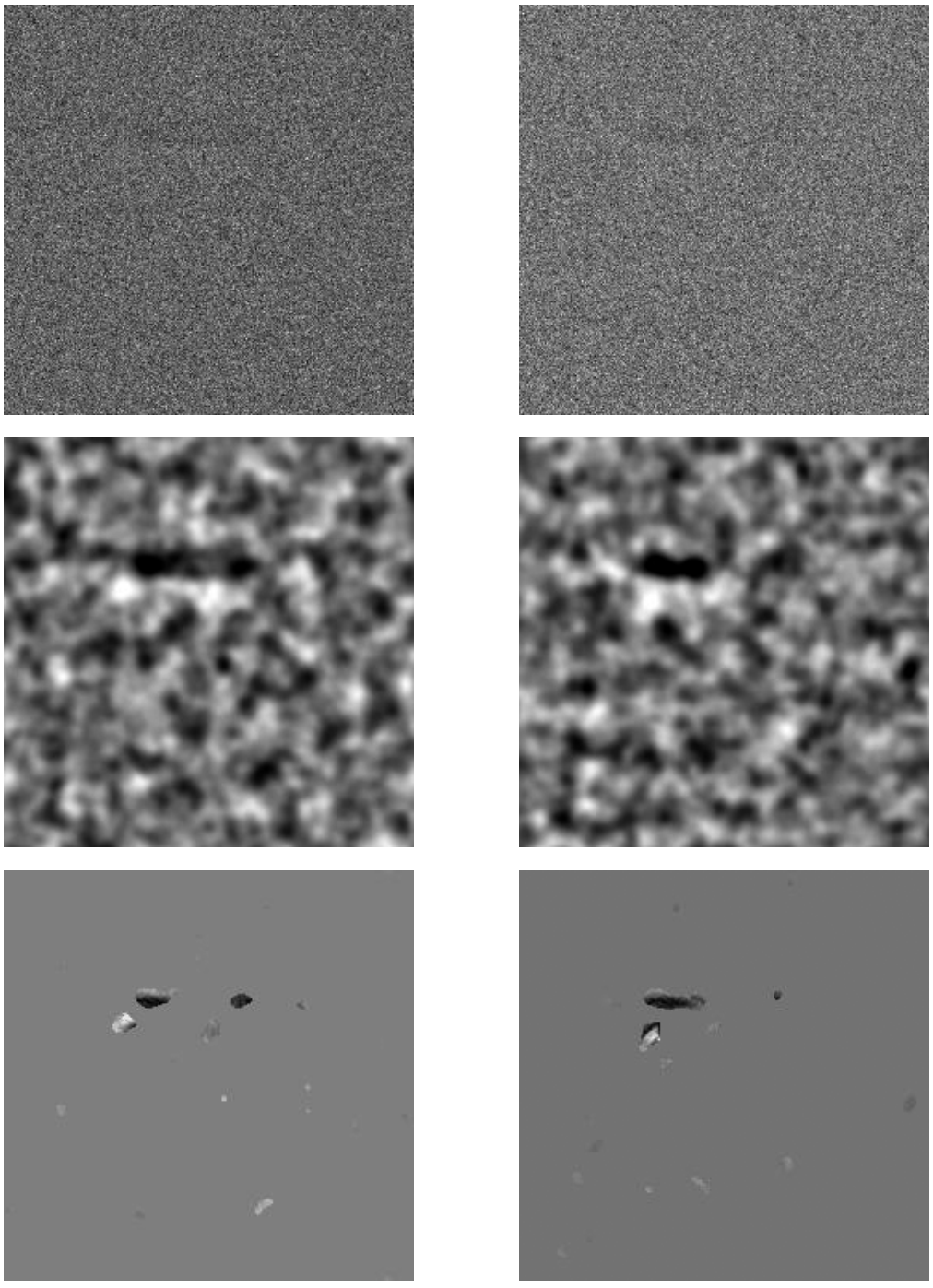


Figure 4.6 The raw (top row), smoothed (middle row) and statistical (bottom row) classification images for observer LCS in the face identification task, pooled across either the first (sessions 1-6; left column) or last (sessions 7-12; right column) half of the experiment. In these statistical images, all of the pixels that are not statistically significant ($p<0.01$) are set to mean gray. The remaining pixels have been set to unity and multiplied by the corresponding pixels in the ideal template.

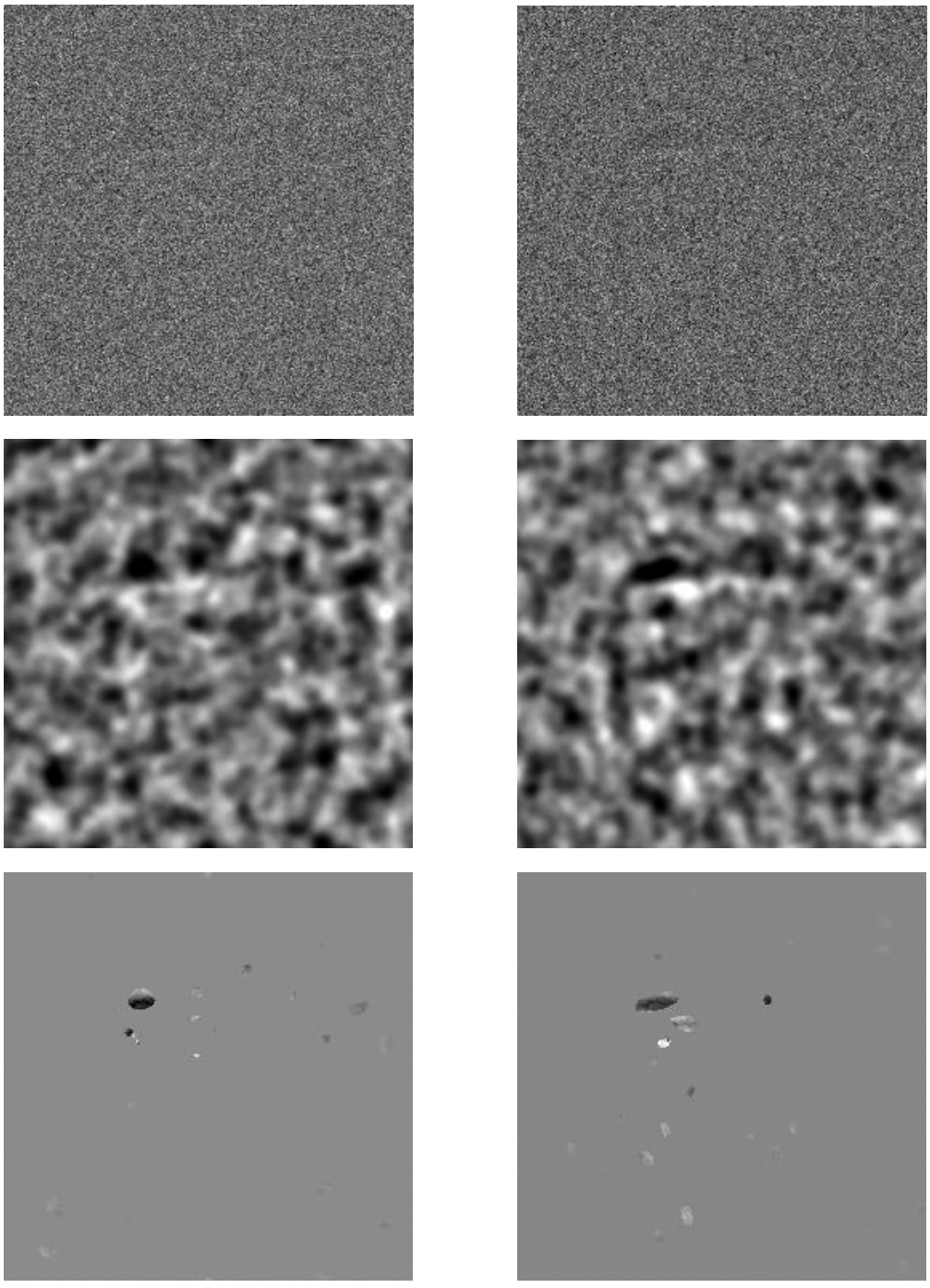


Figure 4.7 The raw (top row), smoothed (middle row) and statistical (bottom row) classification images for observer SKH in the face identification task, pooled across either the first (sessions 1-6; left column) or last (sessions 7-12; right column) half of the experiment.

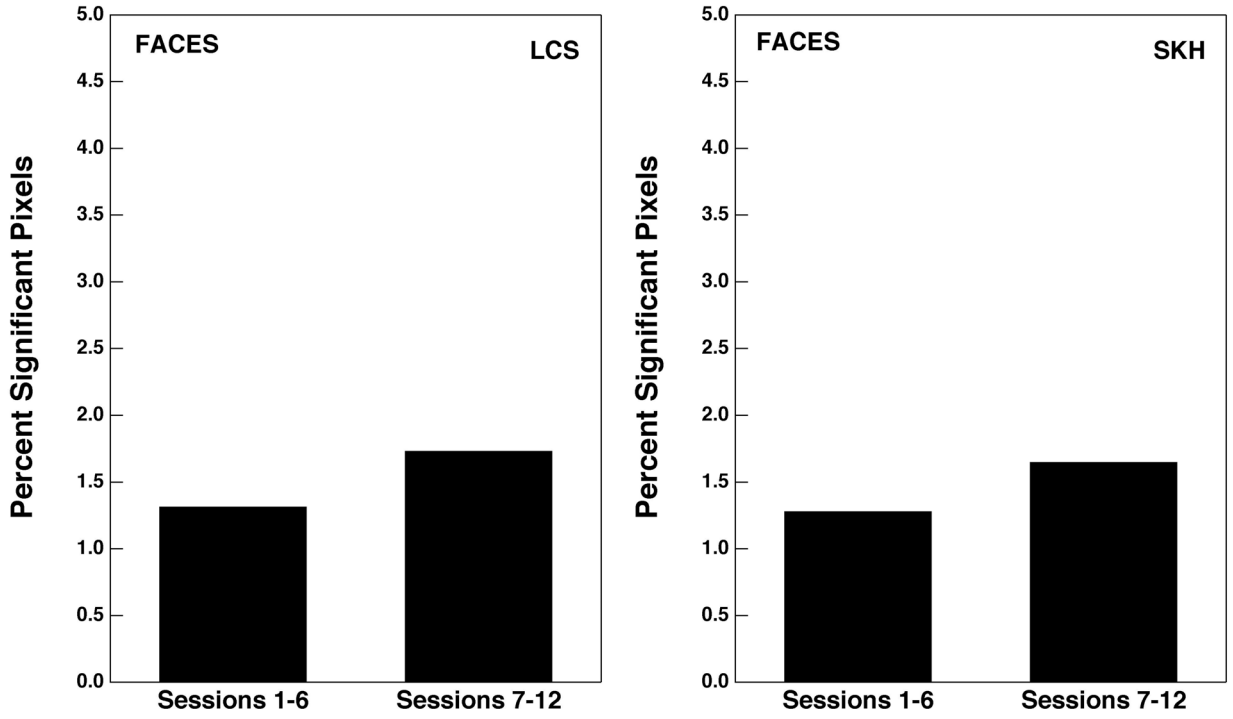


Figure 4.8 The percentage of classification image pixels that reached statistical significance in the first and last halves of the face identification experiment.

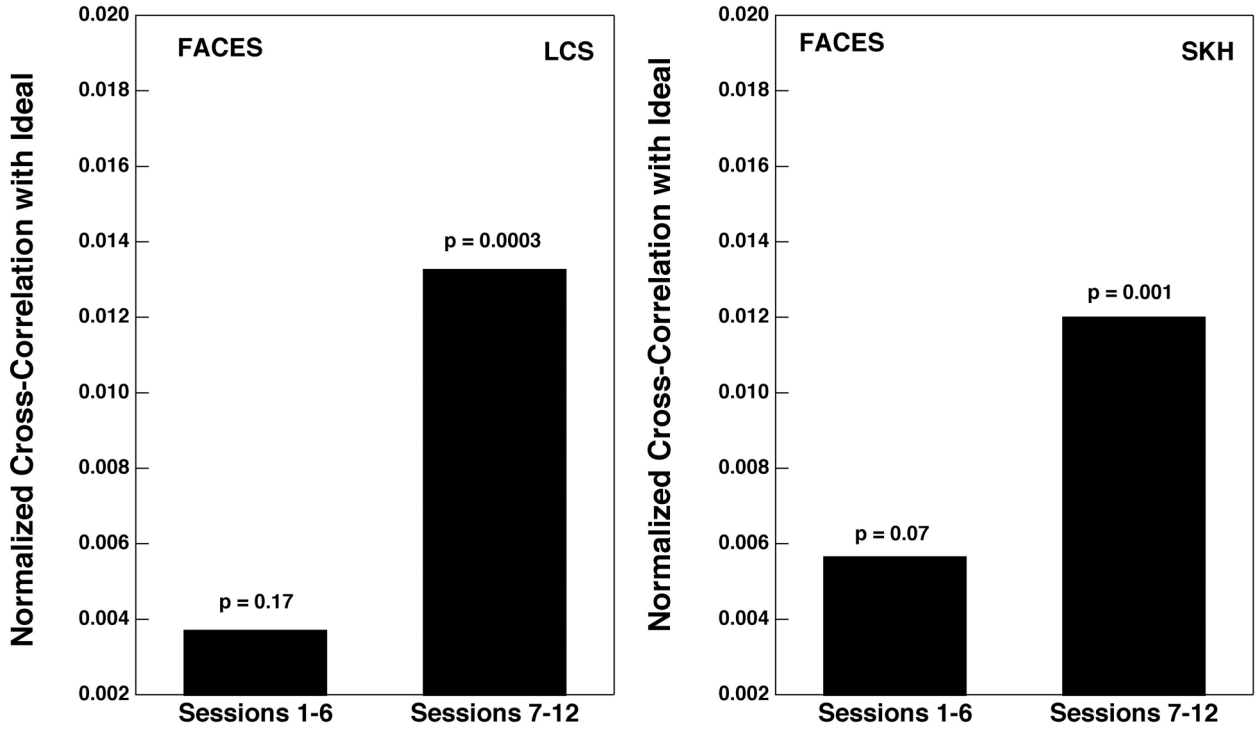


Figure 4.9 Normalized cross-correlations between the two observers' classification images and the ideal observer's template in the first and last halves of the face identification experiment.

Second, the size of the region appears to become larger in the second half of the experiment. This change in region size is reflected in Figure 4.8, which shows the percentage of image pixels that reached statistical significance in the two halves of the experiment. These data show that both observers did in fact use more pixels in the second half of the experiment[21]. Figure 4.9 shows that both of these changes dramatically improved the correlation with the ideal observer's template. This figure plots the normalized cross-correlation between the human classification images and the ideal template for the first and last halves of the experiment. The statistical significance of the correlation is shown above each bar in the histogram. These data show that the correlation was not statistically significant for both observers in the first half of the experiment ($p>0.05$), but was highly significant for both observers in the second half of the experiment ($p<0.001$).

Experiment 4.2: Texture Identification

Signal energy thresholds for the two observers in the texture identification task are plotted as a function of session in Figure 4.10. Thresholds were calculated in the same fashion as Experiment 4.1. As in all of the previous experiments, there was a large

[21] However, the use of more pixels does not necessarily insure that the calculation will be more efficient. For example, an observer could use twice as many pixels in the second half of the experiment, but if none of the pixels were informative, the calculation would clearly be less efficient.

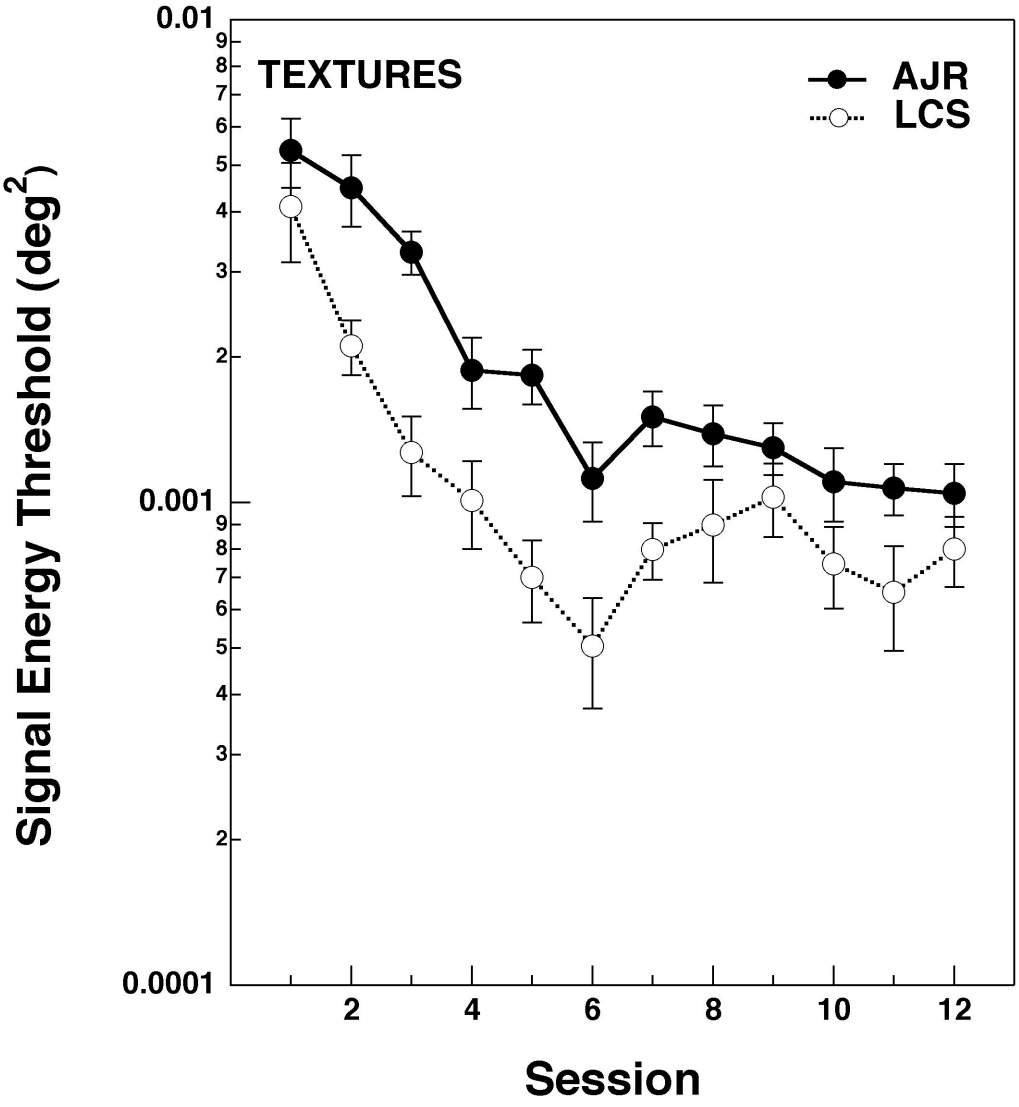


Figure 4.10 Signal energy thresholds plotted as a function of session for the two observers in the texture identification task. The external noise power spectral density was set to the highest level used in Experiment 3.1 (51.10 x 10^{-6} deg^2). Error bars correspond to ± 1 standard error.

effect of learning: thresholds declining by about a factor of 5.0 for both observers over the course of the 12 sessions, with the majority of learning for both observers occurring within the first six sessions.

The corresponding classification images for each session are shown in Figure 4.11 for one observer (LCS). Plotting conventions are the same as in Figure 4.2, with the

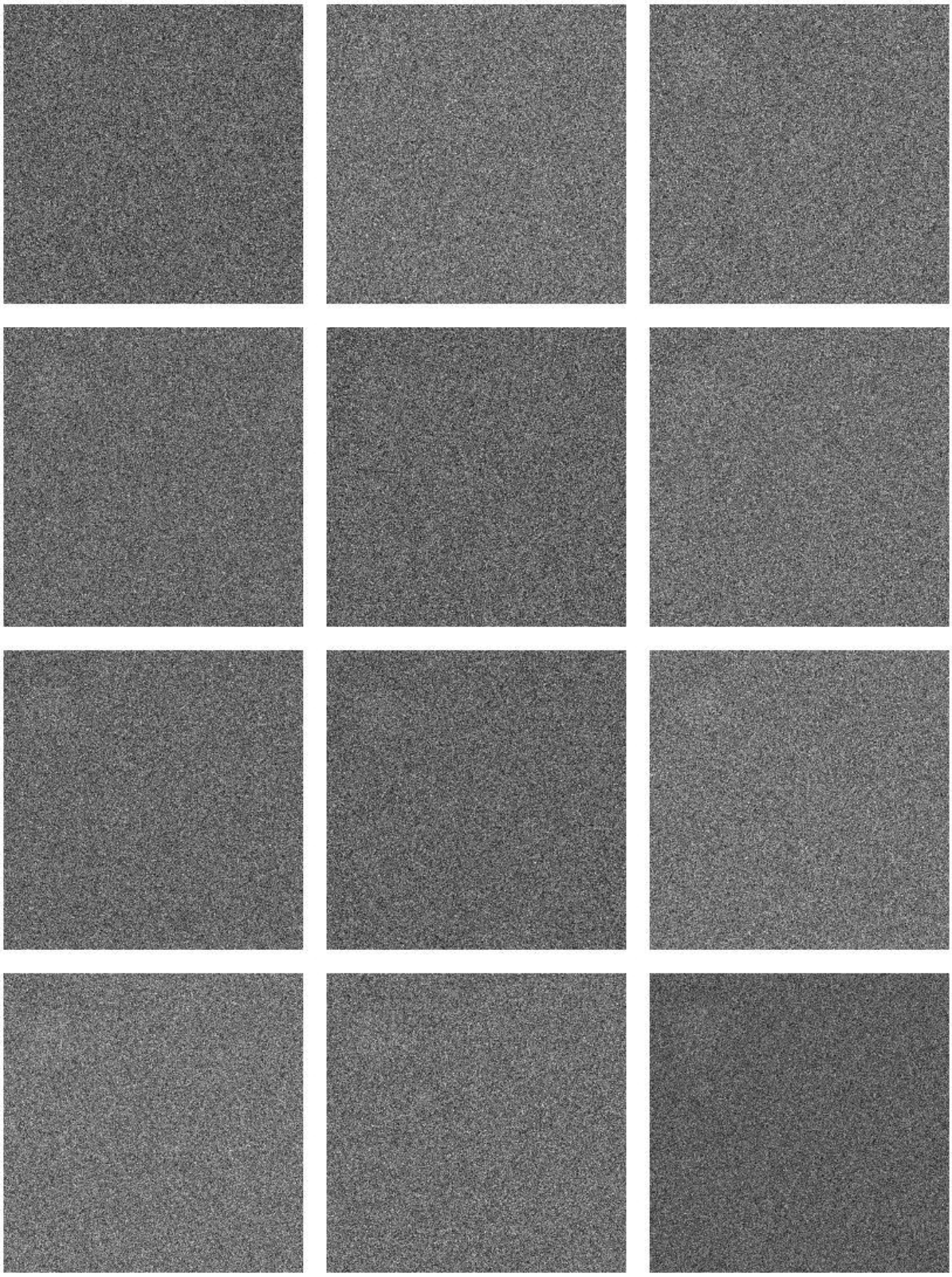


Figure 4.11 Raw classification images for observer LCS in the texture identification task. Each panel corresponds to a single session. The top row corresponds to sessions 1-3 (from left to right), the second row to sessions 4-6, the third row to sessions 7-9, and the bottom row to sessions 10-12.

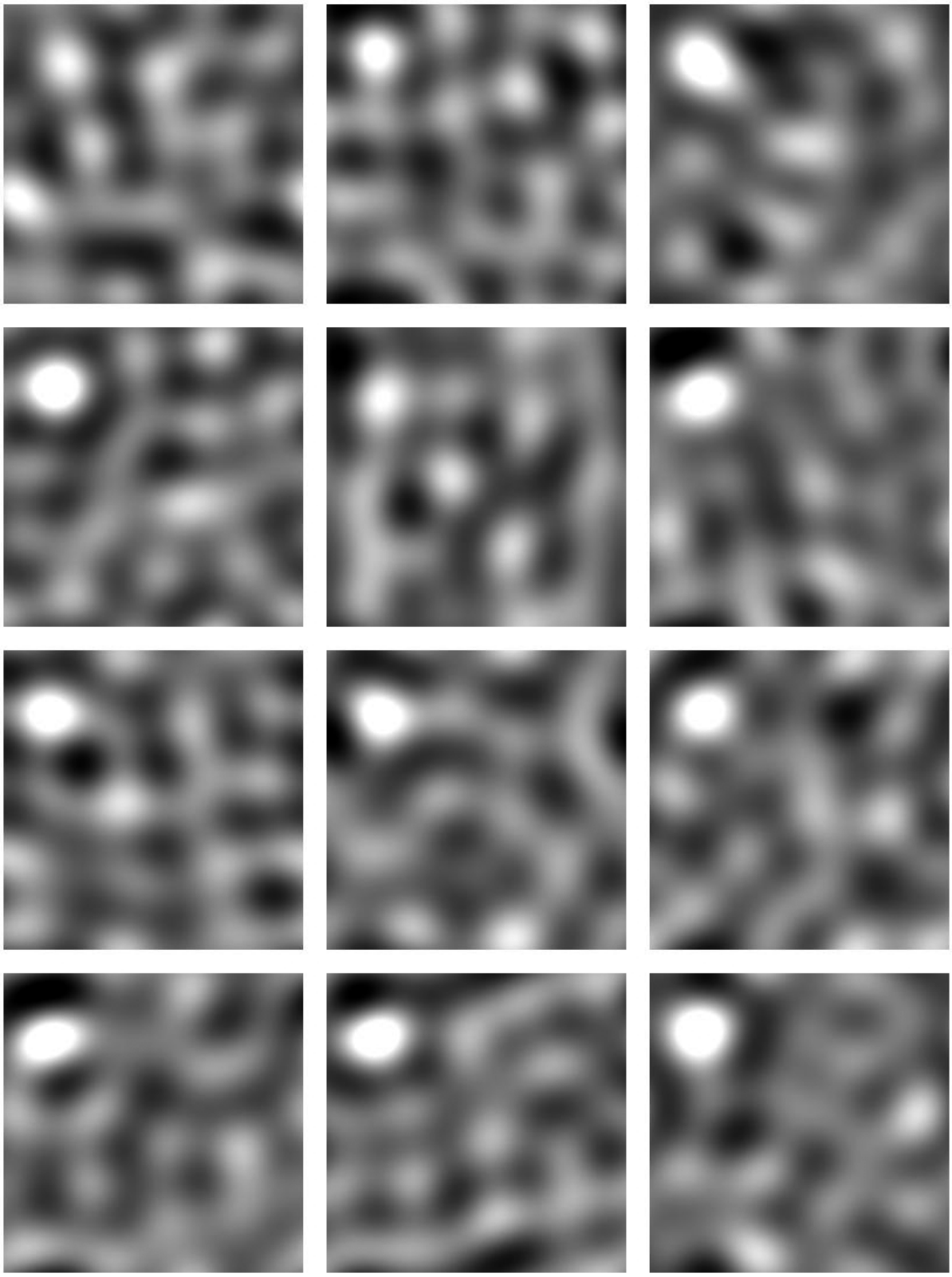


Figure 4.12 Smoothed classification images for observer LCS in the texture identification task. Each panel corresponds to a single session. The top row corresponds to sessions 1-3 (from left to right), the second row to sessions 4-6, the third row to sessions 7-9, and the bottom row to sessions 10-12.

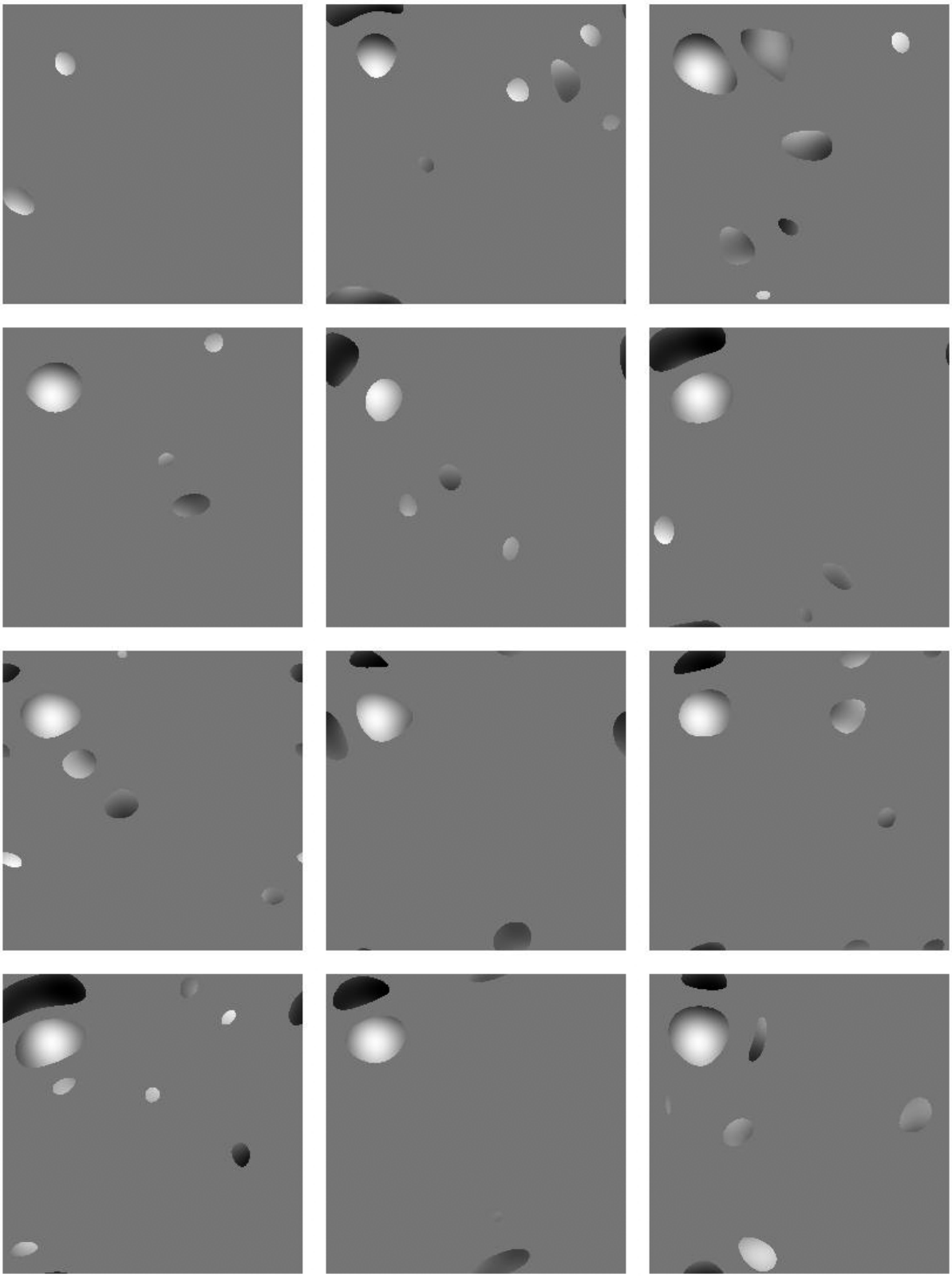


Figure 4.13 Statistical classification images for observer LCS in the texture identification task. Each panel corresponds to a single session. The top row corresponds to sessions 1-3 (from left to right), the second row to sessions 4-6, the third row to sessions 7-9, and the bottom row to sessions 10-12.

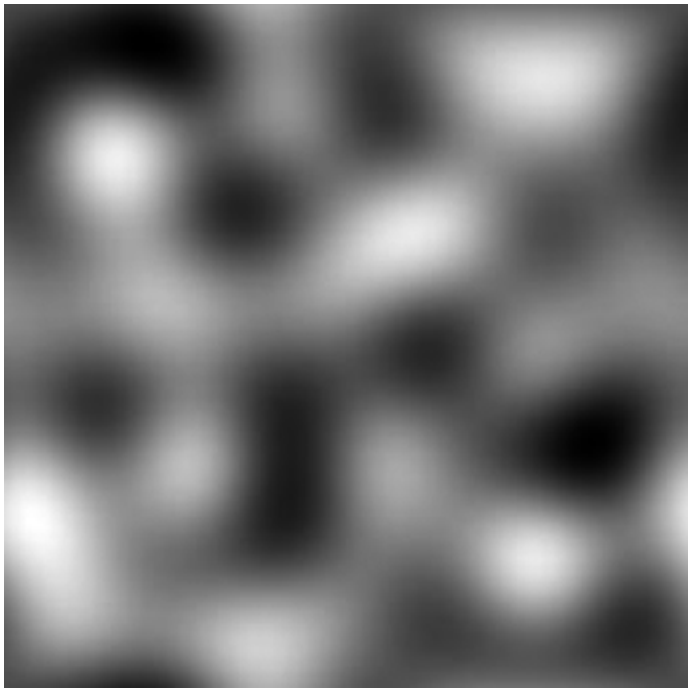


Figure 4.14 The ideal observer's classification image (template) in the texture identification task (see Appendix for details).

exception that pixels that are brighter than mean gray indicate a positive correlation between the noise contrast at that pixel and the response 'texture #2' and pixels that are darker than mean gray indicate a negative correlation between the noise contrast at that pixel and the response 'texture #2'. As with the face identification task, there are no obvious features that emerge from the raw classification images. However, unlike the face stimuli, the textures are highly localized in the frequency domain[22]. This property suggests that a filter matched to the bandwidth of the signals (2-4 c/image) could be highly successful at removing noise from the classification images. Figure 4.12 shows

[22] The textures are designed to have energy only within a narrow band of frequencies (2-4 c/image), whereas energy is inversely related to frequency in the faces.

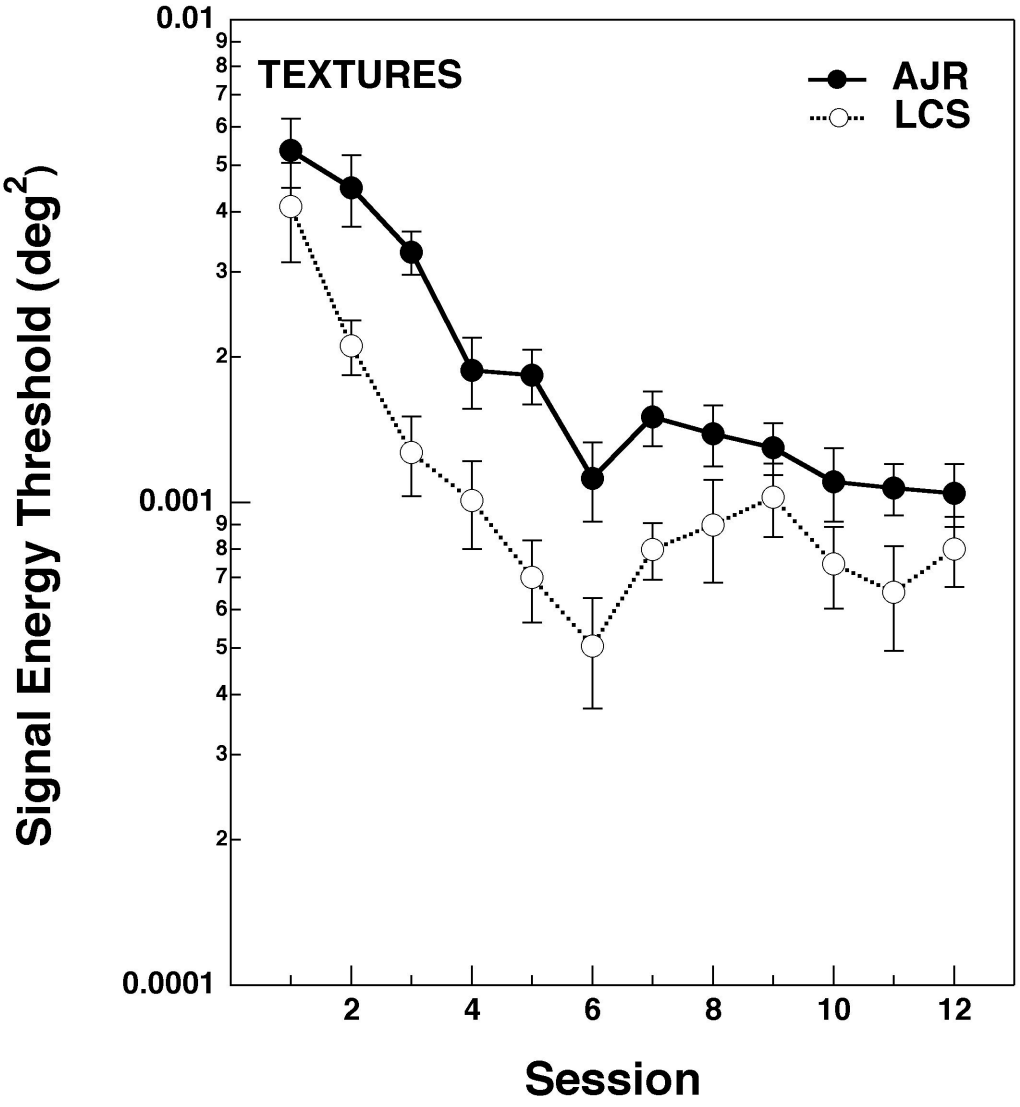


Figure 4.10 Signal energy thresholds plotted as a function of session for the two observers in the texture identification task. The external noise power spectral density was set to the highest level used in Experiment 3.1 (51.10 x 10^{-6} deg^2). Error bars correspond to ± 1 standard error.

the raw classification images from Figure 4.11, smoothed with the same filter used to create the texture patterns. Figure 4.13 shows the corresponding statistical classification images, computed as described in Experiment 4.1. Unlike the filtered face classification images, it is possible to see changes occurring in the classification images across sessions. Specifically, this observer appears to be making greater use of a region in the

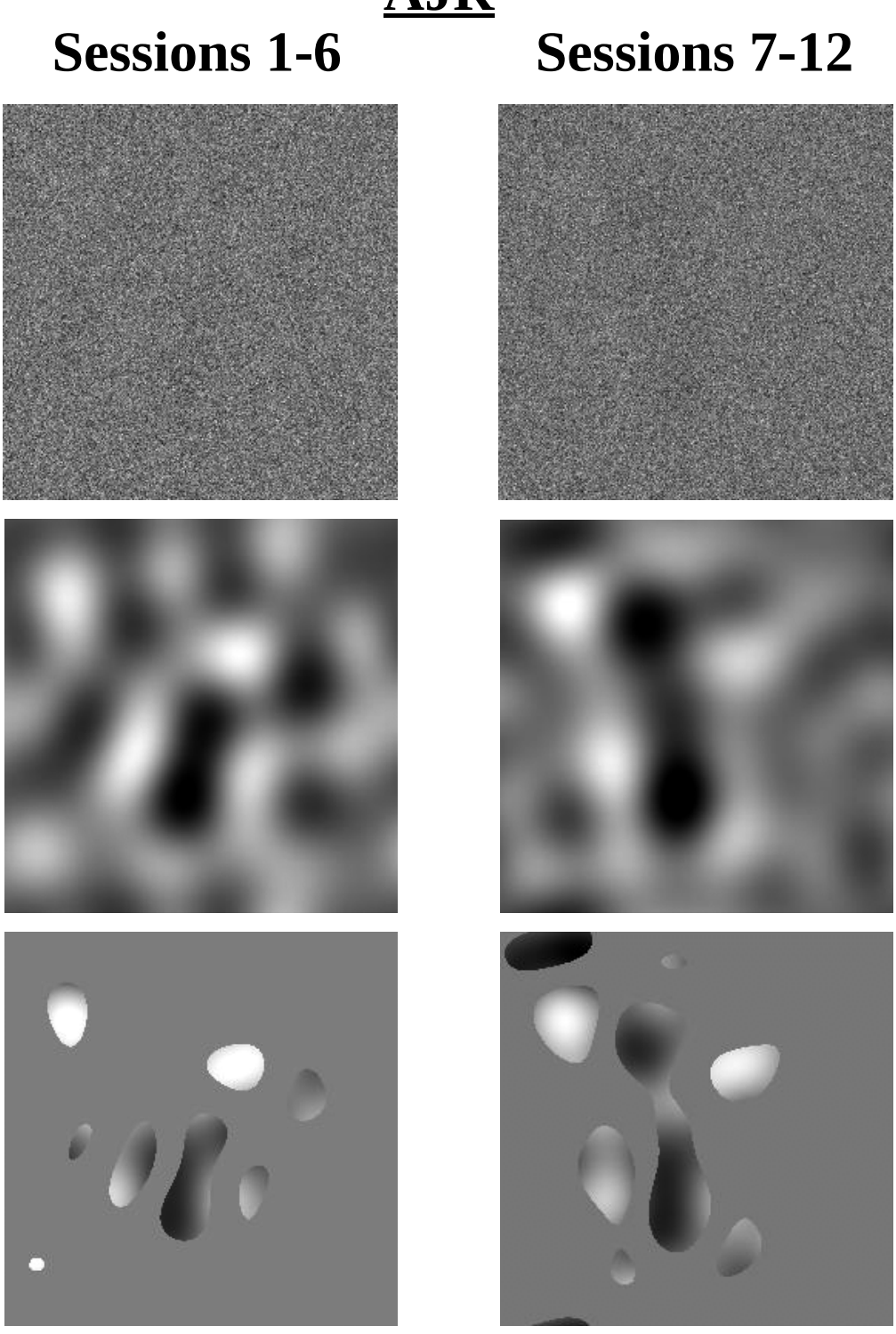


Figure 4.16 The raw (top row), smoothed (middle row) and statistical (bottom row) classification images for observer AJR in the texture identification task, pooled across either the first (sessions 1-6; left column) or last (sessions 7-12; right column) half of the experiment.

LCS

Sessions 1-6 Sessions 7-12

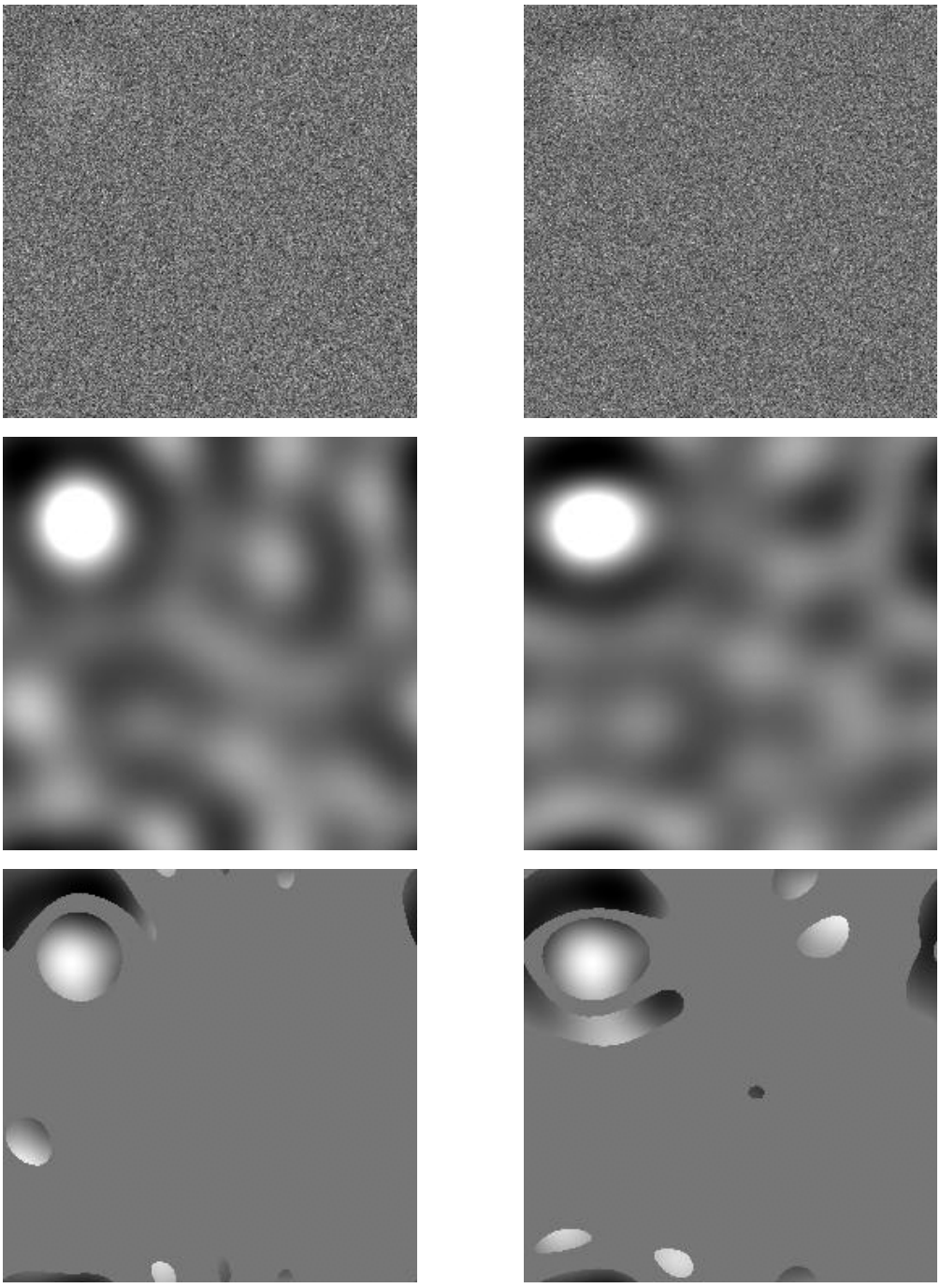


Figure 4.17 The raw (top row), smoothed (middle row) and statistical (bottom row) classification images for observer LCS in the texture identification task, pooled across either the first (sessions 1-6; left column) or last (sessions 7-12; right column) half of the experiment.

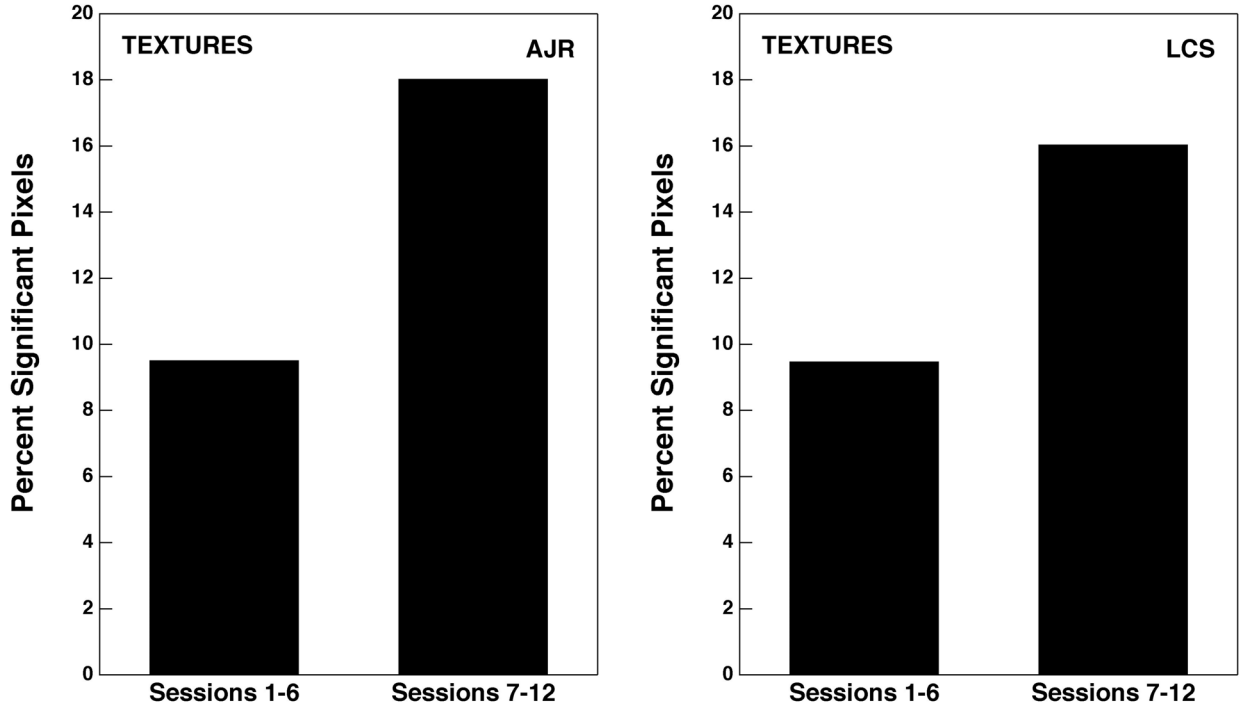


Figure 4.18 The percentage of classification image pixels that reached statistical significance in the first and last halves of the texture identification experiment.

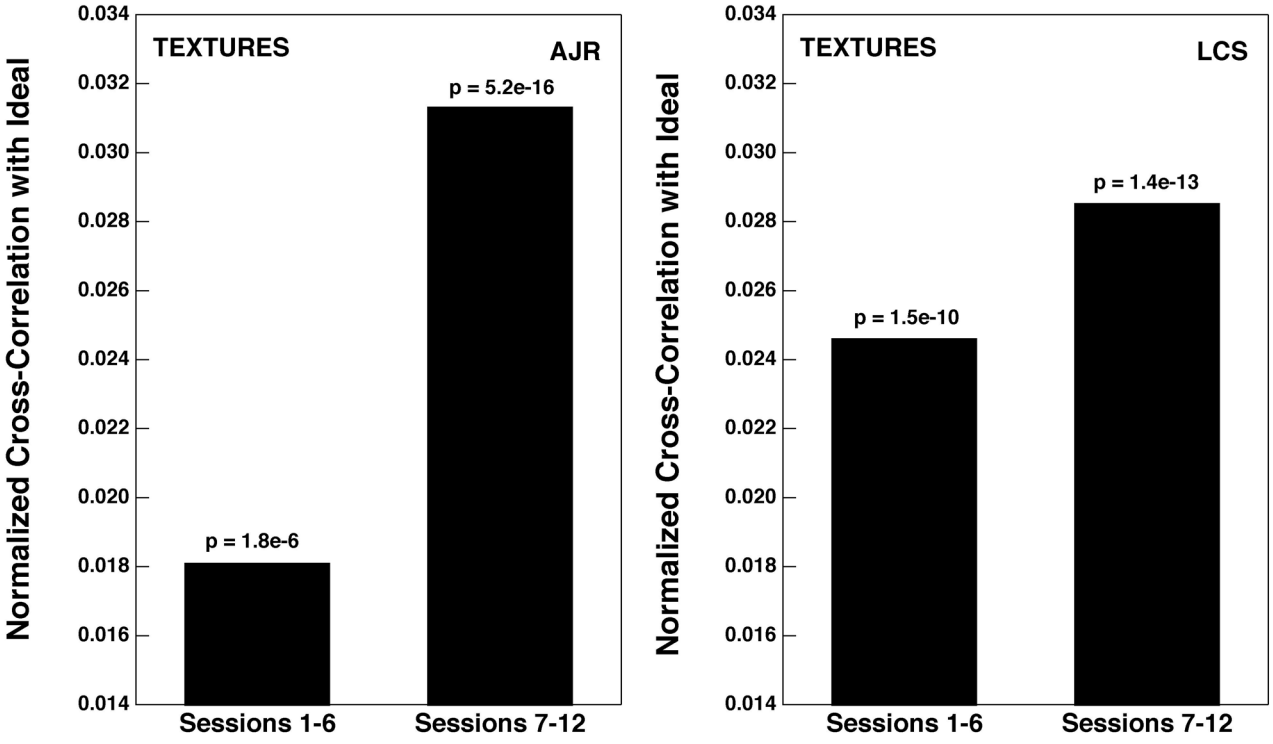


Figure 4.19 Normalized cross-correlations between the two observers' classification images and the ideal observer's template in the first and last halves of the texture identification experiment.

upper left corner as learning takes place. Comparison with the ideal observer's template (Figure 4.14) shows this region to be highly informative, suggesting calculation efficiency is improving with practice. This observation is borne out in Figure 4.15, which shows the normalized cross-correlation between the human classification images and the ideal template, as a function of session. r^2 values and p values for the linear fits are reported in the lower right corner of the figure. These data show a highly significant increase in the correlation with the ideal template with practice.

An even more robust effect of learning can be seen when the data are collapsed as in Experiment 4.1 across the first and last halves of the experiment. Figures 4.16 and 4.17 show the raw (top row), smoothed (middle row) and statistical (bottom row) classification images for observers AJR and LCS, respectively, pooled across either the first (sessions 1-6; left column) or last (sessions 7-12; right column) half of the experiment. Here, features can even be seen in the raw classification images, especially for observer LCS (see the top right corner of the images). However, the smoothed and statistical images are more effective in revealing the large changes that take place with learning. Specifically, observer AJR uses a localized region in the center of the images in throughout the first half of the experiment, and then greatly expands this region in the second half of the experiment to connect two disparate localized regions. Comparison with the ideal template (Figure 4.14) shows that both of these locations are highly informative. In contrast, observer LCS uses a localized region in the top left corner

throughout the experiment. However, this region expands in the second half of the experiment to include informative regions directly adjacent to it. Despite the fact that the two observers used very different strategies, each would predict a large increase in both the number of pixels used and the cross-correlation with the ideal template in the second half of the experiment. Figure 4.18 shows that the proportion of pixels that reached statistical significance is in fact far greater for both observers in the second half of the experiment, and Figure 4.19 shows that the normalized cross-correlation between the classification images and the ideal template is also significantly greater for both observers in the second half of the experiment.

Discussion

The above data has offered a direct view of the changes that take place in two observers' strategies as they learned to identify sets of unfamiliar faces and textures. Visual inspection of the classification images shows clear changes in both the locations and the sizes of the regions that they use to identify the patterns. Statistical analyses coincide with this impression, showing that learning both increases the similarity between the observers' classification images and the ideal template and increases the number of pixels observers use to perform the tasks. Both of these results are consistent with the experiments in Chapter 3 that showed observers' calculations become more efficient as learning takes place.

Although the experiments in Chapter 3 ruled out the possibility that internal noise decreases with learning in a 1-of-10 identification task, it is possible that the effects of learning in a 1-of-2 identification task are different. The importance of this possibility is made more apparent by noting that the changes in the classification images observed in Experiments 4.1 and 4.2 are also consistent with the effects of a decrease in internal contrast-dependent noise. This is because a reduction in contrast-dependent internal noise would increase the signal-to-noise ratio in the classification image, causing it to converge more quickly and thus increase the correlation with the ideal template and the number of significant pixels.

This possibility was tested by measuring calculation efficiency and internal noise (both contrast-invariant and contrast-dependent) for the 1-of-2 identification tasks used in Experiments 4.1 and 4.2. The experiments in Chapter 3 and nearly all previous studies have found a linear relationship between signal energy threshold and external noise power spectral density (Pelli & Farell, 1999), so two-point noise-masking functions were used to estimate calculation efficiency and equivalent input noise[23]. Signal energy thresholds were measured in the highest and lowest external noise levels used in the 1-of-10 identification tasks. In addition, the first and last halves of each session were identical

[23] A two-point noise masking function assumes the noise masking function is linear because the relationship between any two points is perfectly characterized by a single straight line.

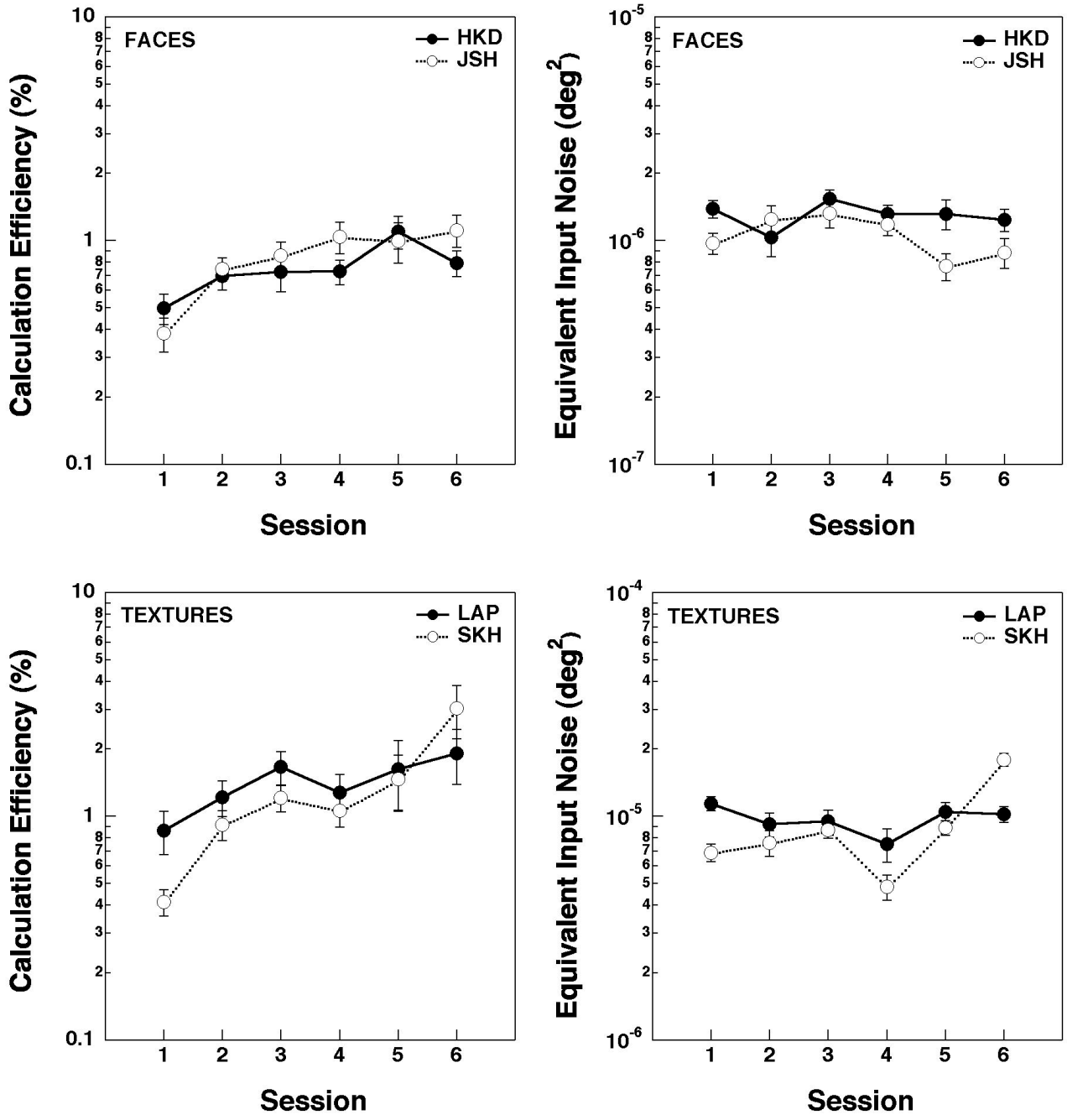


Figure 4.20 Calculation efficiency (left column) and equivalent input noise (right column) as a function of experimental session for the 1-of-2 face (top row) and texture (bottom row) identification tasks. Error bars on each symbol correspond to ± 1 standard deviation.

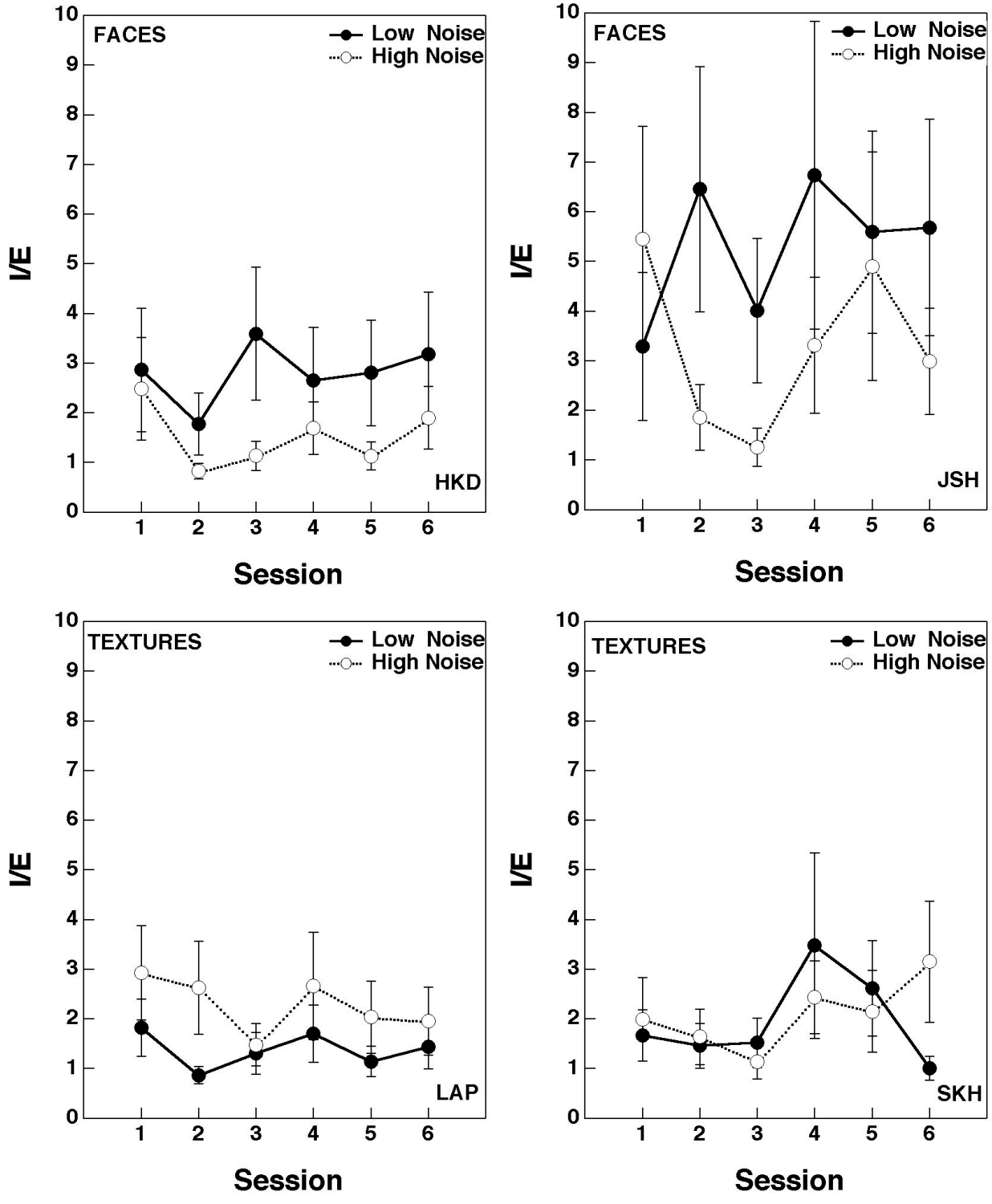


Figure 4.21 Internal/external noise ratio estimates in low (closed symbols) and high (open symbols) external noise as a function of experimental session for the 1-of-2 face (top row) and texture (bottom row) identification tasks. Error bars on each symbol correspond to ± 1 standard deviation.

to allow for the measurement of response consistency. Each session consisted of 200 repeated trials per external noise level, for a total of 800 trials (200 trials x 2 external noise levels x 2 passes). Two interleaved staircases (converging on 71% correct and 84% correct) were used to adjust the signal energy at each level of external noise. Two observers participated in the face discrimination task and two in the texture discrimination task (all experimentally naive). All other aspects of the experiment were the same as described in the experiments in Chapter 3.

The results are shown in Figures 4.20 and 4.21. Figure 4.20 shows the estimates of calculation efficiency (left column) and equivalent input noise (right column) for both the faces (top row) and textures (bottom row), as a function of session. Similar to the results with the 1-of-10 identification tasks, these data show learning served to increase calculation efficiency for both faces and textures, but had little or no effect on equivalent input noise[24]. Figure 4.21 shows the corresponding internal/external noise estimates from the response consistency analysis. Each panel plots the internal/external noise estimates for one observer in either the face (top row) or texture (bottom row) identification task. In each panel, the closed symbols correspond to the internal/external noise estimates in low external noise and the open symbols to high external noise. Although the data are noisy, there is no consistent change in the internal/external noise estimates across

sessions, indicating that neither contrast-invariant nor contrast-dependent noise changed significantly with practice. Taken together, these results suggest that an increase in calculation efficiency mediated the changes in observer classification images found in Experiments 4.1 and 4.2, and that the qualitative effects of learning do not depend upon set size.

Overall Conclusions

The experiments reported in this thesis were designed to explore the effects of perceptual learning under a signal detection framework. The experiments in Chapter 3 allowed us to discriminate between the effects of calculation efficiency and internal noise as observers learned to identify unfamiliar patterns. They showed that it is only calculation efficiency that changes with learning in these pattern recognition tasks. The experiments in the present chapter were designed to explore further the changes that take place in an observer's calculations as perceptual learning occurs. Classification images were used to estimate the computations employed by observers as they learned to recognize unfamiliar patterns. These experiments served two major purposes. First, they tested a strong prediction of the results from Chapter 3 -- namely, that an observer's

[24] The one exception to this trend was observer SKH in the texture identification condition, who showed an *increase* in equivalent noise in the last session.

calculations should become more efficient with perceptual learning. This prediction was verified by showing observers' classification images become more similar to the ideal template as learning takes place. Second, they extended the notion of an improvement in calculation efficiency with learning by offering a direct view of the particular calculations used by observers in the face and texture discrimination tasks, as well as how the calculations changed as observers learned to recognize the stimuli.

Taken together, the results of the experiments in this thesis place important new theoretical constraints upon models of perceptual learning. Beyond these theoretical implications, there may also be important practical applications for these techniques with respect to various phenomena associated with perceptual learning, such as visual learning disorders (e.g., dyslexia) or the development of visual expertise (e.g., surgery).

Appendix: Ideal Observer Analysis

The proof for the ideal 1-of-*m* identification decision rule reported in Chapter 2 is provided below, along with the derivation of the ideal template for a 1-of-2 identification task. Recall that the tasks we are interested in involves presenting one of *m* possible signals embedded in static Gaussian white noise of some power spectral density. A signal is chosen randomly on each trial, and the observer must indicate which signal been shown embedded in the noise during the trial. Signal energy is systematically varied across trials to obtain an identification threshold of some criterion percent correct.

The Ideal Decision Rule

Given the task and the statistics of the noise, the ideal decision rule can be derived from first principles. We shall assume that the ideal observer has access to the set of possible stimuli that may appear in the noise. In this sense, it can be thought of as a *stimulus-known-exactly* observer (Green & Swets, 1966). The stimuli can be thought of as templates, which the ideal observer can use in the task. The templates are defined as

T_j for $j = 1$ to m

where T_j is the j^{th} template and m is the number of templates (signals). The noise added to the stimulus is Gaussian and white, with mean 0 and variance σ^2. Also,

n = the number of pixels in each image

R = the signal+noise combination shown on a given trial.

We can use Bayes rule to derive the a posteriori probability that each template appeared in the noise on a given trial (Duda & Hart, 1973). Bayes rule is expressed as

$$P(T/R) = \frac{P(R|T) * P(T)}{P(R)} \approx P(R|T) * P(T) \quad \text{(A.1)}$$

P(R) falls out of the equation because it is a constant that scales the likelihoods produced but does not change their relative order. For our identification task, we apply the rule to each template, and choose the template that has the maximum likelihood of containing the signal. That is,

$$\max_{j=1..m}\{P(T) * P(R|T)\} \quad \text{(A.2)}$$

Because the order of the probabilities is what we are interested in, and the probability of seeing each signal is the same, *P(T)* falls out. Thus, we are left with maximizing *P(R|T)*. Because identification is maximized when the difference between the data and the

templates is minimized (i.e., the template that is most similar to the data produces the highest probability) and the noise is Gaussian, the quantity may be expressed as

$$\min_{j=1..m}\{\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(R_i-T_{ij})^2}\}$$

$$\min_{j=1..m}\{e^{n\ln(\frac{1}{\sqrt{2\pi\sigma^2}})+\sum_{i=1}^{n}-\frac{1}{2\sigma^2}(R_i-T_{ij})^2}\}$$

$$\min_{j=1..m}\{e^{n\ln(\frac{1}{\sqrt{2\pi\sigma^2}})}e^{\sum_{i=1}^{n}-\frac{1}{2\sigma^2}(R_i-T_{ij})^2}\}$$

Taking advantage of the fact that we are only interested in relative order, the first term falls out and we are left with:

$$\min_{j=1..m}\{e^{\sum_{i=1}^{n}-\frac{1}{2\sigma^2}(R_i-T_{ij})^2}\}$$

However, exponentials increase monotonically, so the exponentiation can be removed as well, leaving:

$$\min_{j=1..m}\{\sum_{i=1}^{n}-\frac{1}{2\sigma^2}(R_i-T_{ij})^2\}$$

$$\min_{j=1..m}\{\sum_{i=1}^{n}(R_i-T_{ij})^2\}$$

$$\min_{j=1..m}\{\sum_{i=1}^{n}R_i^2-2T_{ij}R_i+T_{ij}^2\} \quad \text{(A.3)}$$

If all of the templates have the same energy (which they do in the experiments reported in this thesis), the rule may be simplified further. Specifically, $\sum_{i=1}^{n} R_i^2$ and $\sum_{i=1}^{n} T_{ij}^2$ are the same for all of the images, so they fall out, leaving:

$$\min_{j=1..m}\{\sum_{i=1}^{n} -2T_{ij}R_i\}$$

The scalar constant can also be removed, and because maximization is equivalent to the negative of minimization, the final rule can be expressed as:

$$\max_{j=1..m}\{\sum_{i=1}^{n} T_{ij}R_i\} \qquad \text{(A.4)}$$

Thus, we have shown that the ideal decision rule for a 1-of-*m* identification task is to choose the template that produces the highest cross correlation with the stimulus.

The Ideal Template for 1-of-2 Identification

A special case arises when there are only two templates (signals), T_1 and T_2. First, the decision rule may be expressed as a difference between two cross-correlations. Second, the two-signal case allows us to derive a single template that the ideal observer

can use to perform the identification task. This template will represent the relative informativeness at each pixel *i* in our task, and it is identical to the ideal observer's classification image measured with an infinite number of trials according to the methods described in Chapter 2. Assuming the stimuli have equal probability of occurrence, equation A.4 for the 1-of-2 identification task is

$$\max\{\sum_{i=1}^{n} T_{i1}R_i, \sum_{i=1}^{n} T_{i2}R_i\}.$$

Because there are only two quantities, the decision rule may be thought of as a function of their difference:

$$P = \sum_{i=1}^{n} T_{i2}R_i - \sum_{i=1}^{n} T_{i1}R_i .$$

If $P > 0$, choose T_2; otherwise, choose T_1. These terms may be combined, and the ideal template may be expressed as the difference between the two templates:

$$\begin{aligned} P &= \sum_{i=1}^{n} T_{i2}R_i - \sum_{i=1}^{n} T_{i1}R_i \\ &= \sum_{i=1}^{n} (T_{i2}R_i - T_{i1}R_i) \end{aligned}$$

$$P = \sum_{i=1}^{n} R_i (T_{i2} - T_{i1}) \tag{A.5}$$

Thus, we have shown that the quantity T_2-T_1 (i.e., the difference between the two possible signals) can serve as a single template for the ideal observer to perform the identification task.

Bibliography

Abbey, C. K., Eckstein, M. P., & Bochud, F. O. (1999). *Estimation of human-observer templates in two-alternative forced-choice experiments.* Paper presented at the Proceedings of SPIE, San Diego, CA.

Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature, 387*(6631), 401-406.

Ahumada, A. J. (1996). Perceptual Classification images from vernier acuity masked by noise. *Perception, 25*, 18.

Ahumada, A. J., & Beard, B. L. (1998). Response classification images in vernier acuity. *IOVS, 39*(4), 1109.

Ahumada, A. J., & Beard, B. L. (1999). Classification images for detection. *IOVS, 40*(4), 3015.

Ahumada, A. J., & Lovell, J. (1971). Stimulus features in signal detection. *JASA, 49*(6-2), 1751-1756.

Ahumada, A. J., Marken, R., & Sandusky, A. (1975). Time and frequency analyses of auditory signal detection. *JASA, 57*(2), 385-390.

Asaad, W. F., Rainer, G., & Miller, E. K. (2000). Task-specific neural activity in the primate prefrontal cortex. *J Neurophysiol, 84*(1), 451-459.

Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Res, 27*(6), 953-965.

Banks, M. S., Geisler, W. S., & Bennett, P. J. (1987). The physical lmits of grating visibility. *Vis. Res., 27*, 1915-1924.

Barlow, H. B. (1956). Retinal noise and absolute threshold. *JOSA, 46*, 634-639.

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

Beard, B. L., & Ahumada, A. J. (1998). *Technique to extract relevant image features for visual tasks.* Paper presented at the SPIE, San Jose, CA.

Beard, B. L., & Ahumada, A. J. (1999). Detection in fixed and random noise in foveal and parafoveal vision explained by template learning. *J Opt Soc Am A Opt Image Sci Vis, 16*(3), 755-763.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 443-446.

Buonomano, D. V., & Merzenich, M. M. (1998). Cortical plasticity: from synapses to maps. *Annu Rev Neurosci, 21*, 149-186.

Burgess, A. E., & Colborne, B. (1988). Visual signal detection. IV. Observer inconsistency. *J. Opt. Soc. Am. A, 5*(4), 617-627.

Burgess, D. G. (1990). High level decision efficiencies. In C. Blakemore (Ed.), *Vision: coding and efficiency* (pp. 431-440). Cambridge, MA: Cambridge University Press.

Crist, R. E., Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (1997). Perceptual learning of spatial localization: specificity for orientation, position, and context. *J Neurophysiol, 78*(6), 2889-2894.

Croner, L. J., Purpura, K., & Kaplan, E. (1993). Response variability in retinal ganglion cells of primates. *Proc Natl Acad Sci U S A, 90*(17), 8128-8130.

De Valois, R. L., & De Valois, K. K. (1990). *Spatial Vision*. Oxford: Oxford University Press.

Demany, L. (1985). Perceptual learning in frequency discrimination. *J Acoust Soc Am, 78*(3), 1118-1120.

Dosher, B. A., & Lu, Z. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proc. Nat. Acad. Sci., 95*, 13988-13993.

Dosher, B. A., & Lu, Z. L. (1999). Mechanisms of perceptual learning. *Vision Res, 39*(19), 3197-3221.

Dosher, B. A., & Lu, Z. L. (2000). Mechanisms of perceptual attention in precuing of location. *Vision Res, 40*(10-12), 1269-1292.

Duda, R. O., & Hart, P. E. (1973). *Pattern classification and scene analysis*. New York: John Wiley & Sons.

Eckstein, M. P., Ahumada, A. J., & Watson, A. B. (1997). Visual signal detection in structured backgrounds. II. Effects of contrast gain control, background variations, and white noise. *J Opt Soc Am A, 14*(9), 2406-2419.

Eckstein, M. P., Whiting, J. S., & Thomas, J. P. (1996). Role of knowledge in human visual temporal integration in spatiotemporal noise. *J Opt Soc Am A, 13*(10), 1960-1968.

Fahle, M., Edelman, S., & Poggio, T. (1995). Fast perceptual learning in hyperacuity. *Vision Res, 35*(21), 3003-3013.

Fahle, M., & Morgan, M. (1996). No transfer of perceptual learning between similar stimuli in the same retinal position. *Curr. Biol., 6*(3), 292-297.

Fine, I., & Jacobs, R. A. (2000). Perceptual learning for a pattern discrimination task. *Vision Res, 40*(23), 3209-3230.

Fiorentini, A., & Berardi, N. (1980). Perceptual learning specific for orientation and spatial frequency. *Nature, 287*(5777), 43-44.

Fiorentini, A., & Berardi, N. (1997). Visual perceptual learning: a sign of neural plasticity at early stages of visual processing. *Arch Ital Biol, 135*(2), 157-167.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nat Neurosci, 3*(2), 191-197.

Gauthier, I., & Tarr, M. J. (1997). Becoming a 'Greeble' expert: exploring mechanisms for face recognition. *Vision Res, 37*(12), 1673-1682.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nat Neurosci, 2*(6), 568-573.

Geisler, W. S. (1989). Sequential ideal observer analysis of visual discriminations. *Psych. Rev., 96*(2), 267-314.

Gibson, E. J. (1969). *Principles of perceptual learning and development*. New York: Appleton.

Gilbert, C. D. (1994). Early perceptual learning [comment]. *Proc Natl Acad Sci U S A, 91*(4), 1195-1197.

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999a). Identification of band-pass filtered letters and faces by human and ideal observers. *Vis. Res.*(39), 3537-3560.

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999b). Signal but not noise changes with perceptual learning. *Nature, 402*(6758), 176-178.

Gold, J. M., Bennett, P. J., & Sekuler, A. B. (1999c). Learning improves calculation efficiency for complex pattern identification. *IOVS, 40*(4), 3080.

Gold, J. M., Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2000). Deriving behavioral receptive fields for visually completed contours. *Current Biology, 10,* 663-666.

Green, D. M. (1964). Consistency of auditory detection judgments. *Psych. Rev., 71*(5), 392-407.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: John Wiley and Sons.

Green, D. M., & Swets, J. A. (1988). *Signal Detection Theory and Psychophysics* (Reprint ed.). Los Altos, CA: Peninsula Publishing.

Herzog, M. H., & Fahle, M. (1997). The role of feedback in learning a vernier discrimination task. *Vision Res, 37*(15), 2133-2141.

Herzog, M. H., & Fahle, M. (1999). Effects of biased feedback on learning and deciding in a vernier discrimination task. *Vision Res, 39*(25), 4232-4243.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci, 17*(11), 4302-4311.

Karni, A., Meyer, G., Rey-Hipolito, C., Jezzard, P., Adams, M. M., Turner, R., & Ungerleider, L. G. (1998). The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. *Proc Natl Acad Sci U S A, 95*(3), 861-868.

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortext plasticity. *Proc Natl Acad Sci U S A, 88*, 4966-4970.

Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature, 350*, 250-252.

Legge, G., Kersten, D., & Burgess, A. E. (1987). Contrast discrimination in noise. *J. Opt. Soc. Am. A, 4*(2), 391-406.

Lillywhite, P. G. (1981). Multiplicative intrinsic noise and the limits to visual performance. *Vis. Res., 21*, 291-296.

Lu, Z. L., & Dosher, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Res, 38*(9), 1183-1198.

Lu, Z. L., & Dosher, B. A. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *J Opt Soc Am A Opt Image Sci Vis, 16*(3), 764-778.

Lu, Z. L., Liu, C. Q., & Dosher, B. A. (2000). Attention mechanisms for multi-location first- and second-order motion perception. *Vision Res, 40*(2), 173-186.

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

Manjeshwar, R. M., & Wilson, D. L. (2001a). Effect of inherent location uncertainty on detection of stationary targets in noisy image sequences. *J Opt Soc Am A Opt Image Sci Vis, 18*(1), 78-85.

Manjeshwar, R. M., & Wilson, D. L. (2001b). Hyperefficient detection of targets in noisy images. *J Opt Soc Am A, 18*(3), 507-513.

Matthews, N., Liu, Z., Geesaman, B. J., & Qian, N. (1999). Perceptual learning on orientation and direction discrimination. *Vision Res, 39*(22), 3692-3701.

McCall, R. B. (1986). *Fundamental statistics for behavioral sciences* (Fourth ed.). San Diego.

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

Nasanen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Res, 39*(23), 3824-3833.

Neri, P., Parker, A.J. & Blakemore, C. (1999). Probing the human stereoscopic system with reverse correlation. *Nature 401*(6754), 695-698.

Pelli, D. G. (1981). *Effects of visual noise.* Unpublished Ph. D., University of Cambridge, Cambridge.

Pelli, D. G. (1990). The quantum efficiency of vision. In C. Blakemore (Ed.), *Vision: coding and efficiency* (pp. 3-24). Cambridge, MA: Cambridge University 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? *J Opt Soc Am A Opt Image Sci Vis, 16*(3), 647-653.

Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philos Trans R Soc Lond B Biol Sci, 335*(1273), 23-30.

Peterson, W. W., Birdsall, T. G., & Fox, W. C. (1954). The theory of signal detectability. *Transactions of the IRE Professional Group on Information Theory, PGIT-4*, 171-212.

Poggio, T., Fahle, M., & Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *Science, 256*(5059), 1018-1021.

Raghavan, M. (1989). *Sources of visual noise.* Unpublished Ph. D., Syracuse University, Syracuse.

Recanzone, G. H., Schreiner, C. E., & Merzenich, M. M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J. Neurosci., 13*(1), 87-103.

Sathian, K., & Zangaladze, A. (1998). Perceptual learning in tactile hyperacuity: complete intermanual transfer but limited retention. *Exp Brain Res, 118*(1), 131-134.

Schiltz, C., Bodart, J. M., Dubois, S., Dejardin, S., Michel, C., Roucoux, A., Crommelinck, M., & Orban, G. A. (1999). Neuronal mechanisms of perceptual learning: changes in human brain activity with training in orientation discrimination. *Neuroimage, 9*(1), 46-62.

Schoups, A. A., Vogels, R., & Orban, G. A. (1995). Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. *J Physiol (Lond), 483*(Pt 3), 797-810.

Solomon, J. A., Lavie, N., & Morgan, M. J. (1997). Contrast discrimination function: spatial cuing effects. *J Opt Soc Am A, 14*(9), 2443-2448.

Spiegel, M. F., & Green, D. M. (1981). Two procedures for estimating internal noise. *J. Acoust. Soc. Am., 70*(1), 69-73.

Stickgold, R., James, L., & Hobson, J. A. (2000). Visual discrimination learning requires sleep after training. *Nature, 3*(12), 1237-1238.

Stickgold, R., Whidbee, D., Schirmer, B., Patel, V., & Hobson, J. A. (2000). Visual discrimination task improvement: a multi-step process occurring during sleep. *Journal of Cognitive Neuroscience, 12*(2), 246-254.

Tanner, W. P., & Birdsall, T. G. (1958). Definitions of d' and η as psychophysical data. *JOSA, 30*(10), 922-928.

Tjan, B. S. (1996). *Ideal observer analysis of object recognition.* Unpublished Ph.D., University of Minnesota, Minnesota.

Tjan, B. S., Braje, W. L., Legge, G. E., & Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vis. Res., 35*(21), 3053-3069.

Tolhurst, D. J., Movshon, J. A., & Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey cortex. *Vis. Res., 23*(8), 775-785.

Tyler, C. W., Chan, H., Liu, L., McBride, B., & Kontsevich, L. (1992). Bit-stealing: How to get 1786 or more grey levels from an 8-bit color monitor. In B. E. Rogowitz (Ed.), *Human vision, visual processing, and digital display III* (pp. 351-364).

Watson, A. B. (1998). Multi-category classification: template models and classification images. *IOVS, 39*(4).

Watson, A. B., & Rosenholtz, R. (1997). A Rorschach test for visual classification strategies. *IOVS, 38*(4), 2.

Watson, A. B., & Solomon, J. A. (1997). Model of visual contrast gain control and pattern masking. *J Opt Soc Am A, 14*(9), 2379-2391.