

# Evidence for criterion shifts in visual perceptual learning: Data and implications

**MICHAEL J. WENGER**

*Pennsylvania State University, University Park, Pennsylvania*

**ANGELINA M. COPELAND**

*University of Nevada, Las Vegas, Nevada*

**JENNIFER L. BITTNER**

*Pennsylvania State University, University Park, Pennsylvania*

AND

**ROBIN D. THOMAS**

*Miami University, Oxford, Ohio*

---

Work on visual perceptual learning for contrast detection has shown that reliable decreases in detection thresholds are accompanied by reliable increases in false alarm rates (Wenger & Rasche, 2006). The present study assesses the robustness and replicability of these changes, demonstrating that they are independent of a variety of task demands (i.e., the specific method used for perceptual practice and threshold estimation) and the presence or absence of trial-by-trial feedback and that the source of the increases can be found in shifts in changes in sensitivity and in bias for detection, identification, or both. Although the increase in false alarm rates suggests a strategic shift in response criteria for detection, we demonstrate that there are multiple potential explanations, including explanations that do not require strategic shifts on the part of the observer. The empirical evidence and analysis of alternative explanations reinforce the inference that visual perceptual learning may involve more than changes in perceptual sensitivity and that cortical circuits beyond the primary visual areas may be involved.

---

Human perceptual abilities are remarkably responsive to the effects of practice. A broad and accumulating literature on perceptual learning (for reviews, see Ahissar & Hochstein, 2004; Fahle & Poggio, 2002; Gibson, 1953, 1969; Gibson, Walk, Pick, & Tighe, 1958; Herzog & Fahle, 1998; Jacobs & Michaels, 2006; Sinha & Poggio, 2002) documents the range of tasks in which practice on elementary visual judgments can produce dramatic improvements in performance. This behavioral evidence has been used to motivate the development of a range of models for the mechanisms of perceptual learning (e.g., Adini & Sagi, 2001; Adini, Sagi, & Tsodyks, 2002; Gold, Bennett, & Sekuler, 1999; Gold, Sekuler, & Bennett, 2004; Petrov, Doshier, & Lu, 2005; Yu, Klein, & Levi, 2004), with the majority of these models focused on how perceptual sensitivity to either the presence of or a variation in a defining characteristic changes with experience. The work presented here explores the possibility that the changes observed in visual perceptual learning may also reflect changes in response (decisional) criteria for both detection and identification, in addition to or instead of changes in perceptual sensitivity.

In order to ask these questions, it is critical to consider the extent to which the behavioral evidence available in

any study of perceptual learning can speak in a reasonably restrictive manner to changes in perceptual sensitivity (Fahle & Poggio, 2002; Sagi & Tanne, 1994; Zenger & Sagi, 2002). If that evidence can reflect changes due to factors *other* than changes in sensitivity—such as changes in biases for detection, identification, response selection, or all three—inferences regarding the underlying mechanisms may be in error. Consider, in particular, the possibility that experience may produce changes in both sensitivity and bias. The latter could suggest important roles for selective attention, memory, or other influences, such as might be available by way of back-projections within the visual system (e.g., Di Lollo, Enns, & Rensink, 2000; Lamme & Roelfsema, 2000; Li, Piëch, & Gilbert, 2004; Ling & Carrasco, 2006; Sigman & Gilbert, 2000; Sigman et al., 2005). However, if the experimental approach does not provide the data necessary to measure both types of potential changes, or if it does not adequately control for the possible effects of different types of response bias, any inferences regarding any single basis for change is potentially suspect.

In a large number of the studies on perceptual learning, an experimental approach—an adaptive version of a

---

M. J. Wenger, [mjw19@psu.edu](mailto:mjw19@psu.edu)

---

two-alternative forced choice (2AFC) method—has been used that is widely assumed to minimize or obviate potential influences of response bias (see also Klein, 2001, for a much more complete methodological discussion of these issues). There are, however, at least four lines of evidence suggesting that this assumption may not be completely valid.

First, consider Klein's (2001) review of contemporary methods for threshold estimation. That review emphasizes, repeatedly, that the standard 2AFC adaptive procedure has a number of problems, including the fact that response bias may not be completely controlled (see Klein, 2001, pp. 1421, 1424, 1428, 1429, 1435, 1436, and 1448). Second, there are published reports that document the presence of a response bias for identification in 2AFC tasks (Green & Swets, 1966; Macmillan & Creelman, 2005; Nisbett & Wilson, 1977). Third, the claim of a lack of a response bias addresses only the identification component of performance on any task in which the target is always present, which has been the case in a large number of studies in which the adaptive 2AFC task has been used. Fourth, the claim either conflates or ignores the distinction between detection and identification, a distinction that we will argue is critical for understanding the types of changes involved in perceptual learning.

This evidence, along with data suggesting that perceptual practice can produce reliable liberal shifts in detection responses (Seitz, Nanez, Holloway, Koyama, & Watanabe, 2005; Wenger & Rasche, 2006; Wild & Busey, 2004), motivates the two central hypotheses of the present effort, which we tested in the context of perceptual learning for the detection of grayscale contrast. The first hypothesis is that practice can lead to appreciable liberal shifts in response bias for detection without any concomitant changes in response bias for identification. Evidence consistent with this hypothesis would support the general argument that the changes associated with perceptual learning reflect more than simple improvement in perceptual sensitivity and would provide additional motivation for considering theoretical models that address more than coding in the early portions of the visual system. The second hypothesis is that the liberal shift in response bias for detection can be obtained even in the absence of a requirement for explicit present/absent judgments. Evidence consistent with this hypothesis would advance the argument that the patterns we have observed may reflect a regularity of perceptual learning and are not just an artifact of our experimental approach.

### Empirical and Theoretical Motivations

The empirical motivation for the importance of distinguishing biases in detection from biases in identification and distinguishing both of these from changes in sensitivity comes from work by Copeland (2003; Copeland & Wenger, 2003) on perceptual learning for contrast detection and discrimination. That work documented that reliable reductions in detection and discrimination thresholds were accompanied by false alarm rates that either did not decrease or reliably increased. These findings were replicated and extended (Wenger & Rasche, 2006) using

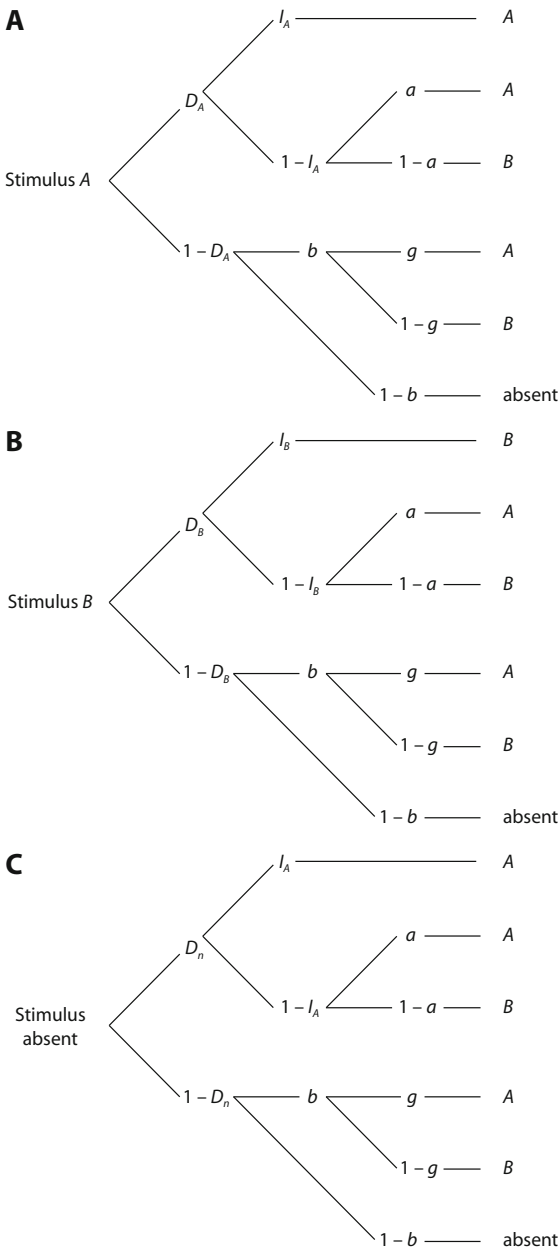
a modified version of the classic (Fechner, 1860/1966) method of constant stimuli (MCS), and similar results have been reported by other investigators (e.g., Seitz et al., 2005; Wild & Busey, 2004).

The theoretical motivation for this work has come from our assumptions regarding the roles of both detection and identification processes in perceptual learning, with a specific concern for tasks in which the stimuli are present at or near observers' thresholds for detection. In order to make explicit our assumptions regarding the role of potential biases in both the detection and the identification components of a perceptual learning task, we developed a simple multinomial model (Batchelder & Riefer, 1990, 1999; Hu & Batchelder, 1994; Riefer & Batchelder, 1988).<sup>1</sup> This model was intended mainly as a simple descriptive summary of our assumptions regarding the roles of detection and identification processes, including associated biases. This is an issue that has been considered and modeled in detail (outside of the context of perceptual learning) by others (see, e.g., Macmillan & Creelman, 2005; Olzak & Thomas, 1981; J. P. Thomas, 1985; J. P. Thomas, Gille, & Barker, 1982). We also will illustrate the basic viability of the model in our discussion of the data from Experiment 2.

Our strong assumption is that high levels of performance on any perceptual identification task in which the stimuli are present at or near observers' threshold for detection involves *both* detection and identification processes, even if the task nominally requires only an identification response. Examples would be tasks in which observers must make identification judgments (e.g., labeling based on orientation; Doshier & Lu, 1999) as an orthogonal dimension (e.g., contrast) is varied at physical levels that may be at or below detection thresholds. Consequently, changes in performance as a function of practice can reflect improvements in detection or identification abilities or both and, as we will argue, changes in bias in both detection and identification.

Assume, for the sake of example, a task in which the stimulus can be either present or absent and, when present, can be in one of two states, *A* or *B*. Assume that the observer's task is to identify the stimulus as being in one of these two states when the stimulus is perceived to be present and, otherwise, to indicate that the stimulus is absent. This is a variation on the standard 2AFC approach, used in earlier work (Kaernbach, 1990) and in Experiment 2 of the present study. The three models (for the three stimulus states: *A*, *B*, and absent) are presented in Figure 1, and the full set of equations for all of the response probabilities are presented in Table 1.

Consider, for example, a trial on which the stimulus is present in state *A*. Let  $D_A$  be the probability that stimulus *A* provides sufficient perceptual evidence for detection, noting that this probability can vary as a function of both experience and stimulus intensity. Let  $I_A$  be the probability that stimulus *A* provides sufficient evidence for identifying it as being in state *A*, given that the perceptual evidence is sufficient for detection, noting that this probability can vary with experience. Let *a* be the observer's bias for responding *A*, given that the perceptual



**Figure 1.** Multinomial models for an example task. (A) Model for processing when the stimulus is present in state A. (B) Model for processing when the stimulus is present in state B. (C) Model for processing when the stimulus is absent.  $D_A$ ,  $D_B$ , and  $D_n$ , probability that the perceptual evidence is sufficient to detect the stimulus when it is in state A or B or when it is absent, respectively;  $I_A$  and  $I_B$ , probability that the perceptual evidence is sufficient to identify the stimulus when it is in state A or B, respectively;  $a$ , probability of identifying a stimulus as being in state A given that the perceptual evidence is sufficient for detection but not identification;  $b$ , probability of judging a stimulus to be present given that the perceptual evidence is not sufficient for detection;  $g$ , probability of guessing that the stimulus is in state A given that the perceptual evidence is not sufficient for detection yet the observer judges the stimulus to be present.

evidence is sufficient for detection but not for identifying the stimulus as being in state A. Let  $b$  be the observer's bias for judging the stimulus to be present, given that the perceptual evidence is not sufficient for detection. Finally, let  $g$  be the probability that the observer guesses that the identity of the stimulus is A, given that the perceptual evidence is not sufficient for detection but considering the observers' bias for judging the stimulus to be present.

Given that the stimulus is present in state A, the observer can generate the correct response (A) in three different ways. First, the stimulus can provide sufficient evidence for both detection and correct identification, and this will happen with a probability of  $D_A \times I_A$ . Second, the stimulus can provide sufficient evidence for detection but not for identification. In this case, the observer can still respond A, with the probability of responding A in this case being  $D_A \times (1 - I_A) \times a$ . Third, if the stimulus does not provide sufficient evidence for detection, the observer can still choose to judge the stimulus as present and then guess at its identity. In this case, the probability of responding A will be  $(1 - D_A) \times b \times g$ . The overall probability of responding A, given that the stimulus is present in state A, is then

$$P(R = A | S = A) = D_A I_A + [D_A (1 - I_A) a] + [(1 - D_A) b g].$$

Table 1 gives the equations for each possible response in each of the three possible stimulus states.

This basic model can be used to illustrate how potential biases in detection and identification can affect threshold estimates, the standard form of evidence for perceptual learning. Consider four possibilities. First, perceptual practice can lead to an improvement in detection, so that for a given level of stimulus energy, the probability that the perceptual evidence will be sufficient for detection will increase as a function of practice. We can model this by increasing the value of  $D_A$  for each level of stimulus energy. Second, perceptual practice can produce an increase in the probability that, across all levels of stimulus energy, the stimulus will provide evidence sufficient for correct identification. We can model this by increasing the value of  $D_I$  across all values of stimulus energy. Third, perceptual practice can produce a liberal shift in the observer's bias for detection, increasing the probability that the observer will judge the stimulus to be present, independently of the stimulus state. We can model this by increasing the value of  $b$  across all values of stimulus energy. Fourth, perceptual practice can lead to an increase in the probability that the observer will choose to identify the stimulus as being in state A when the perceptual evidence is sufficient only for detection. We can model this by increasing the value of  $a$  across all values of stimulus energy.<sup>2</sup>

The behavioral outcomes of these possibilities are illustrated in the top four panels of Figure 2. (The specific parameter values used to produce these data are provided in Appendix A.) These four panels plot the psychometric functions (correct identifications as a function of stimulus strength, left ordinate) and the probability of two types of errors (interval errors [incorrect identifications] and false

**Table 1**  
**Equations for the Response Probabilities in an Example Perceptual Learning Task**  
**in Which the Stimulus Can Be in One of Three States (A, B, or Absent)**  
**and the Observer Must Give One of Three Responses (A, B, or Absent)**

Stimulus State	Observer's Response	Response Probability
A	A	$P(R = A   S = A) = D_A I_A + [D_A(1 - I_A)a] + [(1 - D_A)bg]$
	B	$P(R = B   S = A) = [D_A(1 - I_A)(1 - a)] + [(1 - D_A)b(1 - g)]$
	Absent	$P(R = \text{absent}   S = A) = (1 - D_A)(1 - b)$
B	A	$P(R = A   S = B) = [D_B(1 - I_B)a] + [(1 - D_B)bg]$
	B	$P(R = B   S = B) = D_B I_B + [D_B(1 - I_B)(1 - a)] + [(1 - D_B)b(1 - g)]$
	Absent	$P(R = \text{absent}   S = B) = (1 - D_B)(1 - b)$
Absent	A	$P(R = A   S = \text{absent}) = D_N I_A + [D_N(1 - I_A)a] + [(1 - D_N)bg]$
	B	$P(R = B   S = \text{absent}) = [D_N(1 - I_A)(1 - a)] + [(1 - D_N)b(1 - g)]$
	Absent	$P(R = \text{absent}   S = \text{absent}) = (1 - D_N)(1 - b)$

Note—See the caption to Figure 1 for an explanation of each of the variables.  $P(R = X | S = Y)$  is the probability that the observer's response is X, given stimulus Y.

alarms, right ordinate) at the beginning and end of a hypothetical perceptual learning experiment. A critical feature of these data is that, in all four cases, the change associated with learning results in a leftward shift of the psychometric function, meaning that in all four cases—including the two in which practice produces only a change in bias for detection or identification—threshold estimates decline. The explanation for this can be found in the equation for  $P(R = A | S = A)$  (Table 1), where it can be seen that increases in any of the four parameters under consideration will increase the value of at least one of the terms in that equation, thus increasing the probability of a correct response at any level of stimulus energy. Thus, in terms of the standard evidence for perceptual learning, the four types of changes will produce identical behavioral outcomes.

The four changes are, however, distinguishable at the level of the error rates. In the case in which practice improves detection (Figure 2A), the rate of misidentification declines slightly as a function of practice, and the rate of false alarms is unaffected. The first result is due to the fact that the two terms in the equations for misidentification [e.g.,  $P(R = B | S = A)$ ; Table 1] involve a term that would produce an increase and a term that would produce a decrease in the overall probability of an identification error. The second result is due to the fact that the probability of a false alarm can be expressed as the complement of a correct rejection (last equation in Table 1), which is not affected by changes in  $D_A$  or  $D_B$ .

In the case in which practice leads to an improvement in identification (Figure 2B), rate of misidentification is greatly reduced, and the false alarm rate is unaffected. The first result follows naturally from the effect an increase in  $I_A$  or  $I_B$  has on the value of the first term in the equations for identification errors. The second result follows from the fact that changing either of these parameters has no effect on the probability of a correct rejection and, thus, its complement.

In the case in which practice leads to a liberal shift in the bias for detection (Figure 2C), the rate of misidentification increases at the lowest levels of stimulus energy, and the rate of false alarms increases. The first result follows from the fact that increasing the value of  $b$  will increase the value of the second term in the equations for

identification errors, and the effect of that increase will be most pronounced when  $D_A$  and  $D_B$  are smallest (when  $1 - D_A$  and  $1 - D_B$  are consequently largest), which is the case for the lowest values of stimulus energy. The second result follows from the fact that increasing the value of  $b$  will produce fewer correct rejections (making  $1 - b$  smaller) and, thus, more false alarms.

In the case in which practice leads to a shift in the bias for giving one of the two identification responses (Figure 2D), the rate of misidentification increases with both more practice and increases in stimulus energy, and the false alarm rate is unaffected. The first result follows from the effect of increasing the value of  $a$  on the first term in the equation for an identification error when the stimulus is B. Here, an increase in  $a$  will interact with the increase in the value of the detection parameter as a function of stimulus energy. The second result follows from the fact that changing  $a$  has no effect on the probability of a correct rejection and, thus, false alarms.

This simple model thus suggests that shifts in bias for detection and for identification can influence estimates of threshold and that these influences will be unobservable without allowing errors in both identification and detection to be tracked. To better see how this would be the case for the widely used 2AFC task, note that the 2AFC task, in the context of a staircase procedure for threshold estimation, converges to an accuracy level of .79 (Klein, 2001; Leek, 2001; Levitt, 1971). Each point on the surfaces plotted in these two panels represents a parameter combination that produces a probability of a correct response equal to .79 and illustrates how the two parameters for bias— $b$  for bias in detection and  $a$  for bias in identification—can be varied to produce equivalent outcomes across variations in detection and identification probabilities.

At this point, it is worth stating clearly that we are arguing that these changes in two logically distinct types of bias are only *potentially* involved in any given perceptual learning task. We are not arguing that either or both forms of bias are necessary components of perceptual learning. Instead, the central implication of our argument is that it is both critical and possible to assess the extent to which either or both forms of response bias may be present. In addition, we should make clear that in advancing

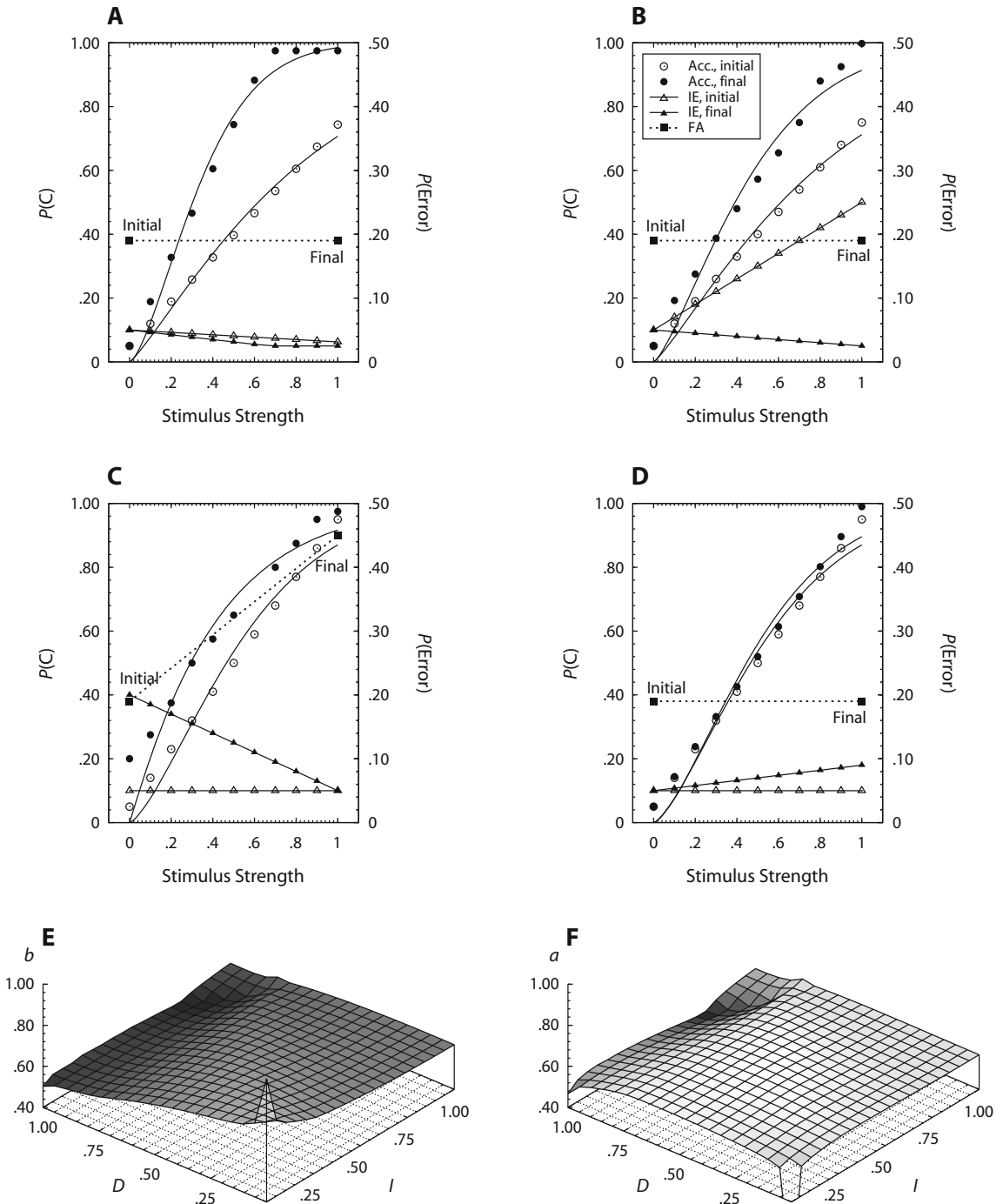


Figure 2. Predictions of the multinomial models. (A) Learning produces an increase in detectability. (B) Learning produces an increase in identifiability. (C) Learning produces a liberal bias in detection. (D) Learning produces an increase in the bias for identifying a detectable stimulus as being in state *A*. (E) Values of the detection bias parameter *b* that produce a constant accuracy level. (F) Values of the identification bias parameter *a* that produce a constant accuracy level. Acc. initial/final, accuracy level in the initial or final block; IE initial/final, interval error in the initial or final block; FA, false alarm.

this argument, we are not suggesting that the literatures on detection thresholds generally or perceptual learning specifically are critically compromised in any way. In fact, there is a long history of concern for the extent to which response biases may “contaminate” estimates of detection thresholds (e.g., Hatfield & Soderquist, 1969; Klein, 2001; Swets & Sewall, 1963). Instead, we suggest that empirical consideration of the potential for these two forms of response bias allows for external validation of core assumptions and for experimental tests of a broader range of alternative hypotheses regarding the functional circuitry and plasticity of the visual system than is currently the case.

In support of this argument, we will present a series of five experiments demonstrating that perceptual learning for contrast detection involves more than a change in perceptual sensitivity. Figure 3 provides an overview of the procedures used in each of the experiments. In Experiment 1, perceptual learning for contrast detection was compared across two methods for assessing detection thresholds: the commonly used adaptive staircase procedure and the classic MCS. The results of this experiment strongly suggest that reductions in threshold as a function of practice are accompanied by regular liberal shifts in observers' bias for detection. Experiment 2 combined the two methods, using an approach similar to that suggested by Kaernbach (1990). This experiment replicated the critical patterns in Experiment 1 and documented that this liberal shift occurs only for detection, with there being no change in response bias for identification. Experiments 3–5 documented that the liberal shifts in bias for detection are not an artifact of the need to make present/absent decisions, something that distinguishes both the MCS approach and the approach used in Experiment 2 from common uses of the 2AFC paradigm in studies of perceptual learning. The article will conclude with a consideration of a set of alternative hypotheses regarding the source of the shifts in bias for detection.

## EXPERIMENT 1

The purpose of Experiment 1 was to determine whether the qualitative changes observed in our earlier work (Copeland, 2003; Copeland & Wenger, 2003; Wenger & Rasche, 2006)—reductions in detection thresholds, accompanied by increases in false alarm rates—were in any way idiosyncratic to the method of threshold estimation (MCS) used in that work. Two sets of observers practiced the contrast detection task; one group used the MCS procedure, and the other practiced with the more frequently used 2AFC adaptive staircase procedure. We predicted that the observers in both groups would show evidence for exponential or power law decreases in detection thresholds as a function of more practice, at roughly equivalent magnitudes (excepting individual differences), suggesting no qualitative differences in perceptual learning due to experimental method. In addition, we predicted that (consistent with our earlier findings) reliable increases in false alarm rates would be obtained for those observers in the MCS condition.

## Method

### Participants

Six paid volunteers were recruited for participation. All the participants had normal or corrected-to-normal vision and unencumbered use of their hands. All the participants were naive as to the task. The participants were randomly assigned to one of two conditions, with an equal number of observers in each of the two conditions. In the first condition, a modified version of the traditional MCS was used in order to obtain both threshold estimates and false alarm rates. In the second condition, an adaptive staircase procedure was used in order to estimate detection thresholds.

### Materials

The stimuli consisted of achromatic Gabor patterns, specified by

$$I(x, y) = I_0 \left\{ 1.0 + c \sin \left[ 2\pi f (\cos \theta \pm y \sin \theta) \right] \exp \left( -\frac{x^2 + y^2}{2\sigma^2} \right) \right\}, \quad (1)$$

where mean luminance  $I_0$  was 71 cd/m<sup>2</sup> and tilt  $\theta$  from vertical was held constant at 0° for all the stimuli. Each Gabor subtended 1.54° horizontal  $\times$  1.54° vertical. Center frequency  $f$  was 2.3 cpd, and standard deviation  $\sigma$  was 0.385°. The stimuli were presented centered on a 2.0°  $\times$  2.0° square whose luminance was 75 cd/m<sup>2</sup>, with a background screen luminance of 46 cd/m<sup>2</sup>. Viewing distance was fixed at 76 cm. The stimuli were presented on a gamma-corrected computer monitor and were viewed through tachistoscopic goggles (Milgram, 1987). Responses were made on a custom-designed eight-button response box. All the displays were viewed in a completely darkened room.

### Procedure

The experiment required 10 days of testing per observer, with each testing session lasting approximately 60 min. Testing on Day 1 was designed to provide an initial (baseline) estimate of performance, while minimizing the amount of experience with the stimuli. Testing on each of the following days provided estimates of detection thresholds, by way of either the MCS or the adaptive staircase procedure (see Figure 3).

**Condition 1 (MCS).** The range of contrasts used in each session changed as each observer's performance changed. The method (see also Wenger & Rasche, 2006) was designed to maximize the number of levels of contrast in the range of contrasts over which accuracy was expected to change most rapidly. The stimuli used on the first day were 10 Gabor patches of unique contrasts, each of which was presented 10 times, along with 100 trials on which no Gabor patch was present, for a total of 200 trials. The order of the trials was randomized for each observer. The target stimuli ranged in (Michelson) contrast from 0.1984 to 0.3244, with a step size of 0.014. This range was determined (on the basis of pilot testing) to be wide enough so that part of the range should be above threshold and part should be below. In each subsequent block of testing, the range of contrast values was selected to lie symmetrically around the contrast value associated with an accuracy level of 70%, with seven contrast values above and below this contrast. The stimuli in each of the remaining nine blocks consisted of a set of 15 Gabor patches of unique contrasts, each of which was presented 20 times, along with a total of 300 trials on which no Gabor patch was present, for a total of 600 trials. The order of trials was randomized for each observer.

**Condition 2 (adaptive staircase).** Condition 2 used an adaptive staircase procedure (see, e.g., Klein, 2001; Leek, 2001; Levitt, 1971). This allows us to compare the pattern of results obtained in Condition 1 (MCS) with those obtained in the most frequently used paradigm in the literature. The first day of testing in Condition 2 began with the same starting contrast (0.2544) and changes in contrast step size (0.014) as those in Condition 1. Contrast decreased after a correct response was made three times in a row but increased when a single mistake was made (a 3–1 staircase, as in, e.g., Dosher & Lu, 1999). The staircase procedure was carried out until 200 trials were completed. In each of the remaining nine blocks, the initial

Task	Used in ...	Proportion of Total Trials	Events: Time						Correct Response
			1	2	3	4	5	6	
Method of constant stimuli	Experiment 1, Condition 1; Experiments 3-5 (Days 1 & 12)	.50	+						Present
		.50	+						Absent
Two temporal intervals, three responses, staircase	Experiment 2	.25	+						Interval 1
		.25	+						Interval 2
		.50	+						Absent
Two temporal intervals, two responses, staircase	Experiment 3 (Days 2-11)	.50	+						Interval 1
		.50	+						Interval 2
Two spatial locations, two responses, staircase	Experiment 1, Condition 2; Experiment 4 (Days 2-11)	.50	+		+				Left
		.50	+		+				Right
Two orientations, two responses, staircase	Experiment 5 (Days 2-11)	.50	+						"Tilt up"
		.50	+						"Tilt down"

Note: + = fixation, = noise, = contrast pattern (Gabor)

Figure 3. Overview of the experimental methods used in the five experiments.

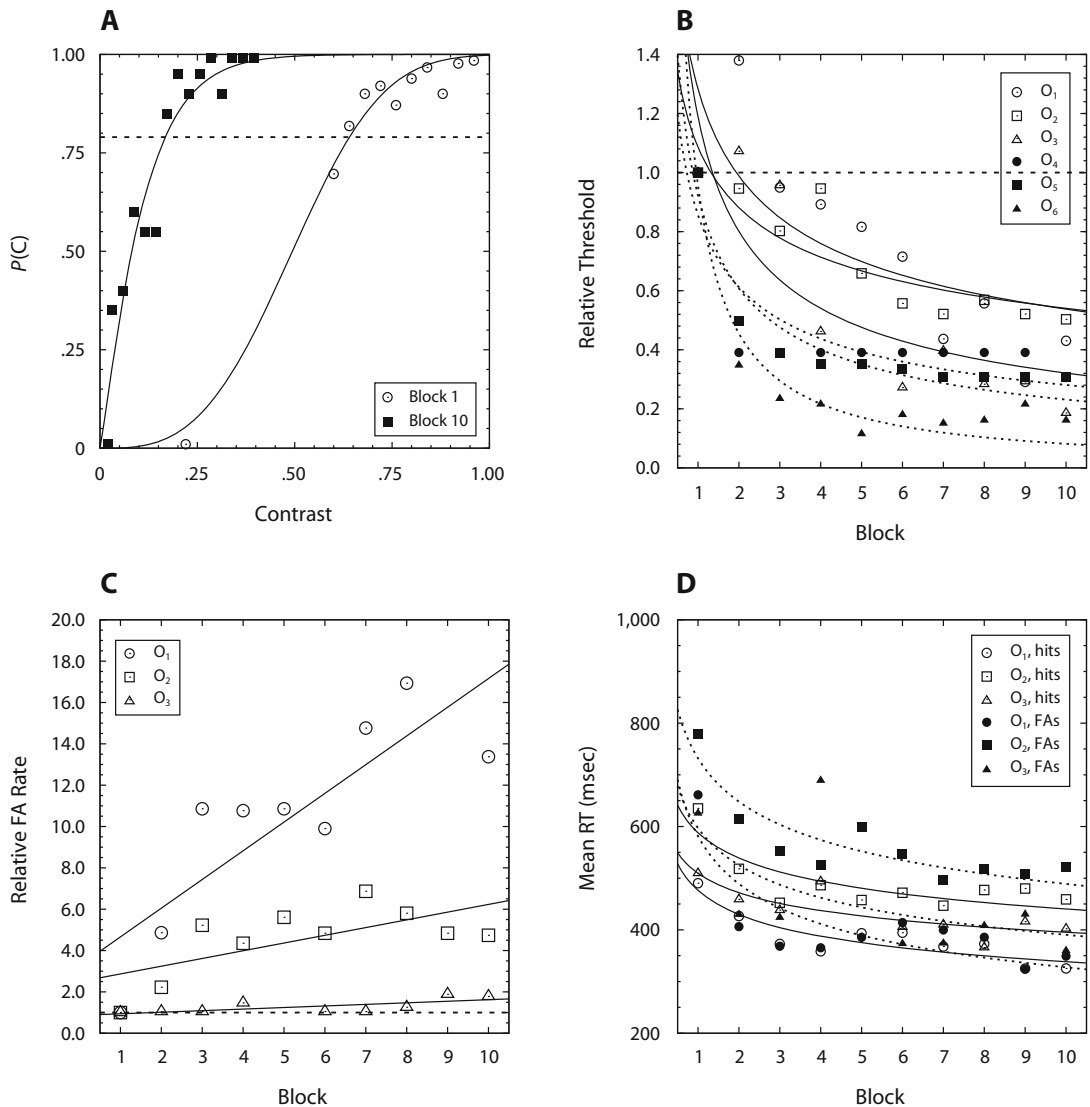
level of contrast for any block was equal to the threshold value estimated in the immediately preceding block of trials.

**Trial sequence.** Each trial began with the presentation of a fixation cross in the center of the screen for 500 msec. The fixation cross was then replaced by the test display. In Condition 1, this was either a Gabor patch (at one of the levels of contrast selected for the current block) or an identically sized frame of visual noise. In Condition 2, a Gabor patch possessing the appropriate level of contrast was presented immediately to one side of fixation, and an identically sized frame of visual noise was presented to the other side. The location of the Gabor patch was selected at random (with equal likelihood) on each trial. The observers responded using the index fingers of their dominant and nondominant hands. In Condition 1, the index finger

of the dominant hand was used for positive responses, and the index finger of the nondominant hand was used for negative responses. In Condition 2, the left index finger was used to indicate that the Gabor was perceived to the left of fixation, and the right index finger was used to indicate that the Gabor was perceived to the right of fixation. No feedback was given in either of the two conditions, and an intertrial interval of 1,000 msec separated each trial from the subsequent trial.

**Results**

We will begin with an examination of changes in detection threshold in each of the two conditions, in order to



**Figure 4.** Experiment 1. (A) Example psychometric functions for 1 observer obtained by fitting the linearized Weibull. (B) Changes in detection threshold as a function of trial block for 3 observers in Condition 1 (method of constant stimuli [MCS]) and Condition 2. (C) Changes in relative false alarm (FA) rates (rate on block  $n$  divided by the rate on Block 1) as a function of trial block for 3 observers in Condition 1 (MCS). (D) Changes in mean response times (RTs) as a function of block for hits and FAs, for the 3 observers in Condition 1 (MCS). For panels B–D, the symbols are the values of the mean RTs, and the lines are best-fitting power (panels B and D) or linear (panel C) functions.



demonstrate perceptual learning and examine the comparability of the qualitative pattern of change in the two conditions. We then will consider changes in false alarm rates and changes in response times (RTs).

**Changes in Detection Thresholds**

Threshold estimates for the observers in Condition 1 were obtained by using a two-parameter version of the Weibull (1951) distribution function as the psychometric function for variations in contrast (*c*):

$$F(c) = 1 - \exp\left[-\left(\frac{c}{\delta}\right)^\gamma\right]. \tag{2}$$

In order to estimate the parameters of the psychometric function, we began with the complement of the cumulative distribution function—the survivor function:

$$S(c) = 1 - F(c) = \exp\left[-\left(\frac{c}{\delta}\right)^\gamma\right]. \tag{3}$$

Taking the natural log of the negative of the natural log of the survivor function then gives

$$\ln\{-\ln[S(c)]\} = \gamma[\ln(c) - \ln(\delta)] = \gamma\ln(c) - \gamma\ln(\delta), \tag{4}$$

a linear function of log contrast. The threshold for each observer in each block in Condition 1 was estimated by first fitting Equation 4 to the log of the negative log of the empirical survivor function for accuracy, using linear regression; *R*<sup>2</sup> values for the obtained fits were greater than .91 in all cases. With the estimated slope for the regression line being  $\gamma$ , the other parameter for the Weibull is obtained by noting that the intercept (*a*) of the regression line is

$$a = -\gamma \ln(\delta).$$

Then, solving for  $\delta$ ,

$$\delta = \exp\left(-\frac{a}{\gamma}\right). \tag{5}$$

Threshold was then estimated by setting Equation 3 to 1 - .79 = .21 and solving for *c*. Threshold estimates for Condition 2 were obtained by taking the mean of the contrast values for the reversals for each observer in each block. Figure 4A plots example data from two blocks of trials for 1 observer, along with the estimated psychometric function.

Figure 4B plots the relative threshold values for each block (threshold in that block divided by the threshold estimate for Block 1), for each of the 3 observers in each of the two conditions. The form of change in both conditions is generally consistent with an exponential or power function reduction that is observed in a range of skills (see, e.g., Doshier & Lu, 2007; Heathcote, Brown, & Mewhort, 2000; Newell & Rosenbloom, 1981). Consequently, we used linear regression on the log of the threshold estimates as a function of the log of the block number to assess the reliability of the changes in threshold. The results of these analyses are presented in Table 2. The observers in both conditions showed reliable reductions in threshold as a function

**Table 2**  
**Experiment 1: Regression Analyses of Changes in Log Detection Threshold As a Function of Log Practice, for Each of the Observers in Each of the Two Conditions**

Condition	Observer	$\hat{\beta}$	SE	R <sup>2</sup>
1 (MCS)	1	-0.5275***	0.1335	.8612
	2	-0.3433***	0.0579	.8946
	3	-0.3765***	0.0937	.7982
2 (SC)	4	-0.3337***	0.0960	.8615
	5	-0.4667***	0.0653	.8645
	6	-0.7047**	0.1526	.7972

Note— $\hat{\beta}$ , estimated slope of the line relating log threshold to log block; SE, standard error of the estimate of  $\hat{\beta}$ ; R<sup>2</sup>, proportion of variance accounted for; MCS, method of constant stimuli; SC, staircase procedure. \*\**p* < .01. \*\*\**p* < .001.

of practice, with there being almost complete overlap in the range of estimated slopes across the two conditions.

**Changes in False Alarm Rates**

A central advantage of using the MCS to estimate thresholds, relative to the adaptive staircase procedure, is the ability to obtain false alarm rates as a function of practice.<sup>3</sup> Figure 4C presents the relative false alarm rates for each of the 3 observers in Condition 1 (MCS). Table 3 presents the results of fitting a linear regression model to the arcsin transform of the false alarm rates for each observer as a function of block of practice:

$$F' = \arcsin\left(\sqrt{F}\right), \tag{6}$$

where *F* is the observed false alarm rate. This transformation was used to adjust for the regular heterogeneity of variances that occurs for proportions below approximately .30 (Zar, 1999, p. 278). All 3 observers showed reliable increases in false alarm rates, with final rates being 2–14 times the initial rates. It is this increase in false alarm rates that strongly suggests that the observers experienced a liberal shift in detection criteria as they also decreased their detection thresholds.

**Changes in RTs**

Another common feature of skill acquisition is a reduction in RTs (e.g., Heathcote et al., 2000; Logan, 1988, 1992; Newell & Rosenbloom, 1981; Rosenbaum, Carlson, & Gilmore, 2001) that follows either an exponential or a power function. Although RTs are not typically considered in the literature on perceptual learning,

**Table 3**  
**Experiment 1: Regression Analyses of Changes in False Alarm Rates As a Function of Practice, for Each of the Observers in Condition 1 (Method of Constant Stimuli)**

Observer	$\hat{\beta}$	SE	R <sup>2</sup>
1	0.0613**	0.0130	.8355
2	0.0131*	0.0053	.7310
3	0.0023*	0.0010	.8402

Note— $\hat{\beta}$ , estimated slope of the line relating false alarm rate to block; SE, standard error of the estimate of  $\hat{\beta}$ ; R<sup>2</sup>, proportion of variance accounted for. \**p* < .05. \*\**p* < .01.

**Table 4**  
**Experiment 1: Regression Analyses of Changes**  
**in Response Times (RTs) As a Function of Practice, for Each**  
**of the Observers in Condition 1 (Method of Constant Stimuli),**  
**for Hits and False Alarms (FAs)**

Observer	Response Type	$\hat{\beta}$	SE	$R^2$
1	Hits	-0.1458***	0.0286	.8643
	FAs	-0.2055**	0.0580	.7620
2	Hits	-0.1121**	0.0302	.8329
	FAs	-0.1615***	0.0303	.8804
3	Hits	-0.1116**	-0.0296	.7709
	FAs	-0.1837*	0.0682	.6824

Note— $\hat{\beta}$ , estimated slope of the line relating log mean RT to log block; SE, standard error of the estimate of  $\hat{\beta}$ ;  $R^2$ , proportion of variance accounted for. \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

they are of interest here—at a minimum, to the extent to which they can speak to potential speed–accuracy trade-offs (Pachella, 1974; Pachella & Fisher, 1972; Pachella & Pew, 1968). We considered RTs from Condition 1 for correct responses on target-present trials (hits) on each block, averaged over all contrast levels used in that block, and RTs for incorrect responses on target-absent trials (false alarms) in each block, separately for each observer. Figure 4D presents the mean RTs for hits and false alarms for each of the 3 observers. Given that the changes in RTs generally followed the expected exponential or power function decreases across blocks, we assessed the reliability of those changes by using simple linear regression to relate changes in log RT to log block. The results of those analyses are presented in Table 4. For both hits and false alarms, the reductions in RT as a function of block were reliable for all observers.

### Discussion

Experiment 1 provides evidence that reliable reductions in contrast detection thresholds as a function of practice can involve more than just increases in perceptual sensitivity. Specifically, reliable reductions in threshold—the standard evidence for perceptual learning—are accompanied by reliable increases in false alarm rates. To obtain this evidence, we needed to adopt an experimental methodology (MCS) that has not always been used in the literature on visual perceptual learning. Consequently, in order to provide a point of comparison, we estimated thresholds using both MCS and the more common staircase procedure. Reliable reductions in threshold were obtained in both conditions, suggesting that both methods allowed us to observe the changes that have generally been used to document perceptual learning. The form of the reduction (exponential or power function) was the same in both conditions, and the magnitude of change was generally equivalent across methods for estimating threshold. In addition to the increases in false alarm rates, there was also evidence for reductions in RTs for both hits and false alarms, with the latter result suggesting a speed–accuracy trade-off. Although this evidence appears to point to a strategic shift in detection criterion, we will suggest in the General Discussion section that there may be other

mechanisms—some not requiring the inference of a volitional shift in response strategy—that are consistent with this pattern of results.

## EXPERIMENT 2

The results of Experiment 1 suggested that the evidence for liberal shifts in detection criterion—evidenced by increases in false alarm rates—can be obtained using an experimental approach (MCS) that also allows for reductions in detection thresholds that are similar in both form and magnitude to the reductions observed with the modal adaptive procedure. This evidence came by way of a design that varied threshold estimation procedures (MCS vs. staircase) across participants. Stronger evidence for the hypothesis that visual perceptual learning for contrast detection can involve changes in bias for detection and/or identification would come by way of a design that allowed both the adaptive procedure and the ability to obtain true false alarms. This type of design would also allow us to address the distinction, suggested in the introduction, between biases for detection and identification. The design of Experiment 2 was motivated by these two goals and followed from a design used by Kaernbach (1990). If the pattern we had observed to date was not an artifact of the specific experimental method, we would expect that we should be able to obtain reliable reductions in detection thresholds, along with reliable increases in false alarm rates, within observers.

### Method

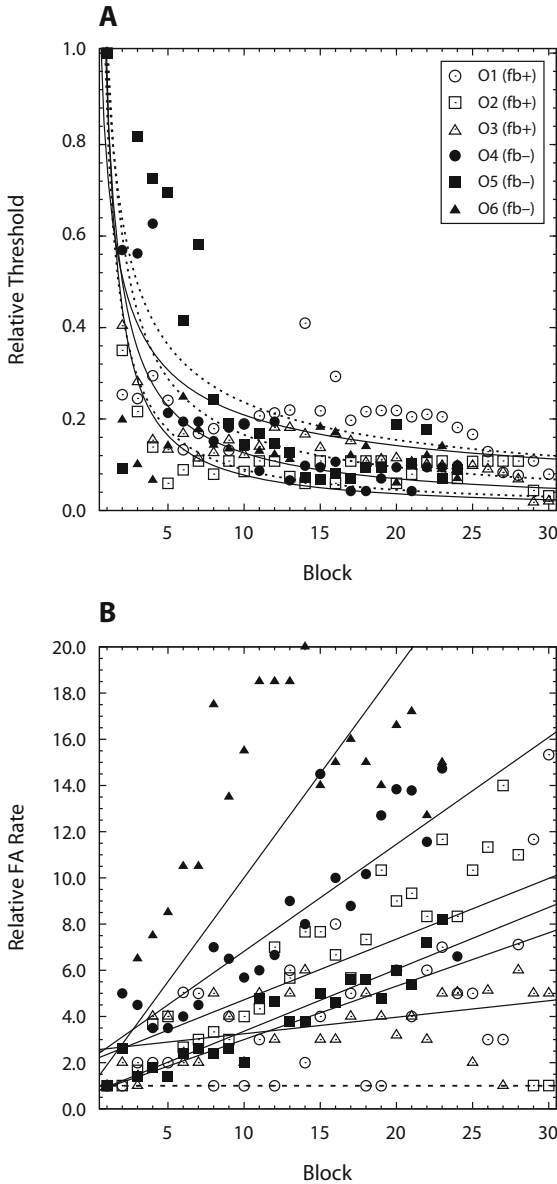
#### Participants

A total of 9 paid observers were recruited to participate in this experiment. The data from 3 observers were discarded prior to analysis because of failures to complete the protocol, due primarily to difficulties with the scheduling of testing sessions. All the observers had normal or corrected-to-normal vision and unencumbered use of their hands. All were right-handed, and none had participated in any previous studies of visual perceptual learning.

#### Materials, Design, and Procedure

The stimuli used in this experiment were the Gabor stimuli used in Experiment 1. Half of the total number of trials in each block contained a Gabor patch at contrast levels above 0. On half of these trials, the Gabor appeared in the first of two temporally distinct intervals, with the second interval being blank; on the other half of these trials, the Gabor appeared in the second interval, with the first interval being blank (see Figure 3). On the remaining half of the total number of trials, both intervals were blank. Three of the observers who had been retained practiced the task with trial-by-trial feedback, and the remaining 3 received no feedback. The feedback status was manipulated in order to address the possibility that changes in false alarm rates in Experiment 1 may have been due to lack of feedback.

Each trial was initiated by the observer. The initial display on each trial was a fixation cross, with a duration determined on each trial by obtaining a value from an exponential distribution over the range from 400 to 500 msec. The offset of the fixation cross was followed by the first of two display intervals, with a duration of 75 msec, followed by a noise mask for 75 msec and an interstimulus interval of 925 msec. This was followed by the second of the two display intervals and a second noise mask, both of which were present for 75 msec. The offset of the second noise mask was the cue for the observer to respond. The observers were instructed to use the button under their index finger if they perceived the Gabor to be present in the first interval, the button under their middle finger if they perceived the Gabor to be



**Figure 5. Experiment 2. (A) Changes in detection threshold as a function of trial block for 3 observers who received trial-by-trial feedback (fb+) and 3 observers who received no feedback (fb-). The symbols are the values of relative threshold; the lines are best-fitting power functions. (B) Changes in false alarm rates as a function of trial block for 3 observers who received trial-by-trial feedback (fb+) and 3 observers who received no feedback (fb-). The symbols are the values of relative threshold; the lines are best-fitting linear functions.**

present in the second interval, and the button under their ring finger if they perceived the Gabor to be absent from both intervals. For the observers in the condition in which feedback was provided, a single word (“Correct” or “Incorrect”) was presented in the center of the screen for 500 msec immediately following the response.

The range of contrast values for each block of trials was set on the basis of the observer’s threshold in the immediately preceding

block of trials. The increment separating contrast values was set by dividing the range between the previous block’s threshold and the lowest possible contrast value into 20 equal increments. The maximum value of contrast, and the starting value for that block of trials, was set by adding 10 contrast increments to the threshold value from the immediately preceding block. Contrast values in each block were changed using a three-down one-up rule, applied *only* to performance on the target-present trials. The observers completed a total of 240 trials in each block, and completed two blocks of trials in each session, with each session lasting between 40 and 45 min.

**Results**

**Changes in Detection Thresholds**

Detection thresholds were estimated for each observer in each block of trials as the geometric mean of the contrast values for the final 15 contrast reversals. Figure 5A plots threshold estimates for all 6 observers as a function of block of practice. Thresholds decreased in a manner consistent with either power or exponential functions. Reliability of the changes in threshold was assessed using simple linear regression with log threshold as the dependent variable and log block as the predictor; the results of this analysis are presented in Table 5. All 6 observers showed reliable decreases in contrast detection threshold, the standard evidence for perceptual learning.

**Changes in False Alarm Rates**

Data from the target-absent trials in each block were analyzed with respect to the relative frequency of false alarms across blocks, using transformed values of the observed false alarm rates (see Equation 6). The relative false alarm rate for each of the 6 observers is plotted in Figure 5B. Linear

**Table 5**  
**Experiment 2: Regression Analyses of Changes in Log Threshold As a Function of Log Block, for Each of the Observers in Each of the Two Feedback Conditions**

Feedback	Observer	$\hat{\beta}$	SE	R <sup>2</sup>
Present	1	-0.8227***	0.1003	.8215
	2	-0.5298***	0.0869	.8564
	3	-0.6119***	0.0721	.7203
Absent	4	-0.6233***	0.1058	.8052
	5	-0.5282***	0.0748	.7526
	6	-0.4806*	0.1806	.6280

Note— $\hat{\beta}$ , estimated slope of the line relating log threshold to log block; SE, standard error of the estimate of  $\hat{\beta}$ ; R<sup>2</sup>, proportion of variance accounted for. \**p* < .05. \*\*\**p* < .001.

**Table 6**  
**Experiment 2: Regression Analyses of Changes in Transformed False Alarm Rate As a Function of Block, for Each of the Observers in Each of the Two Feedback Conditions**

Feedback	Observers	$\hat{\beta}$	SE	R <sup>2</sup>
Present	1	0.0221***	0.0033	.8995
	2	0.0102***	0.0006	.9030
	3	0.0273***	0.0004	.8813
Absent	4	0.0883***	0.0083	.8436
	5	0.0107***	0.0009	.8723
	6	0.0240***	0.0017	.9419

Note— $\hat{\beta}$ , estimated slope of the line relating transformed false alarm rate to log block; SE, standard error of the estimate of  $\hat{\beta}$ ; R<sup>2</sup>, proportion of variance accounted for. \*\*\**p* < .001.

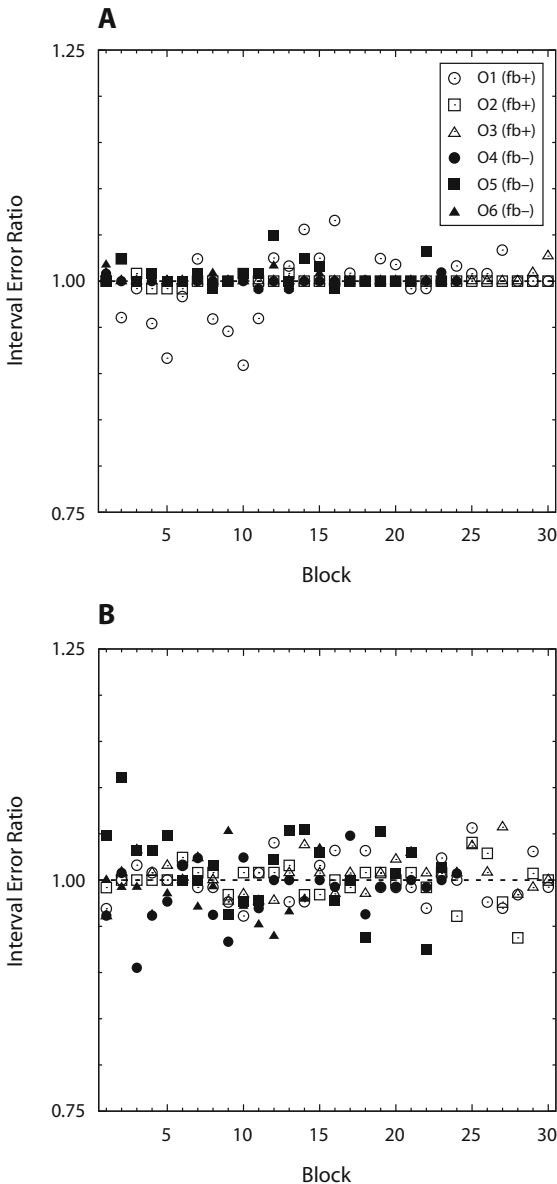


Figure 6. Experiment 2: Ratio of rate of erroneously choosing Display Interval 1 over Display Interval 2, for the target-present (A) and target-absent (B) trials, for 3 observers who received trial-by-trial feedback (fb+) and 3 observers who received no feedback (fb-).

regression was used to assess the reliability of the increases in transformed false alarm rates as a function of practice, with the results of this analysis presented in Table 6. All 6 observers showed reliable increases in false alarm rates, with final levels representing up to 15-fold increases.

**Changes in Interval Errors**

A benefit of the design used in Experiment 2 is that it is possible to distinguish a potential bias for reporting

the presence of the target stimulus from a potential bias for choosing one of the two intervals in which the target stimulus could occur. The first potential type of response bias is assessed using the false alarm rates, as described in the preceding paragraph. The second potential type of response bias is assessed using the relative rate of the observers' erroneously choosing Display Interval 1 versus Display Interval 2. We assessed this ratio separately for target-present and target-absent trials for the two feedback conditions, and the values of those ratios are presented in the four panels of Figure 6. Note that these are *not* relative ratios but, instead, are the true values of the relative rates at which Interval 1 was erroneously chosen over Interval 2. As can be seen in the figure, in all cases, the value of the ratio was very close to 1, suggesting no bias for response interval. The reliability of the difference in proportions (incorrect choice of Interval 1 vs. incorrect choice of Interval 2) was assessed using a normal approximation of  $\chi^2$  (Zar, 1999, pp. 555–556). In all cases, the differences were not reliable (all *ps* > .30). Thus, there is no evidence to suggest that the observers were biased with respect to response interval, either on target-present or target-absent trials, and no evidence suggesting any change in identification bias as a function of practice.

**Changes in RTs**

Given that levels of contrast on target-present trials were systematically varied across trials within a block (rather than randomized), we chose not to analyze the RTs for correct responses on target-present trials (hits). In addition, our primary concern with RTs was to document possible speed-accuracy trade-offs. Consequently,

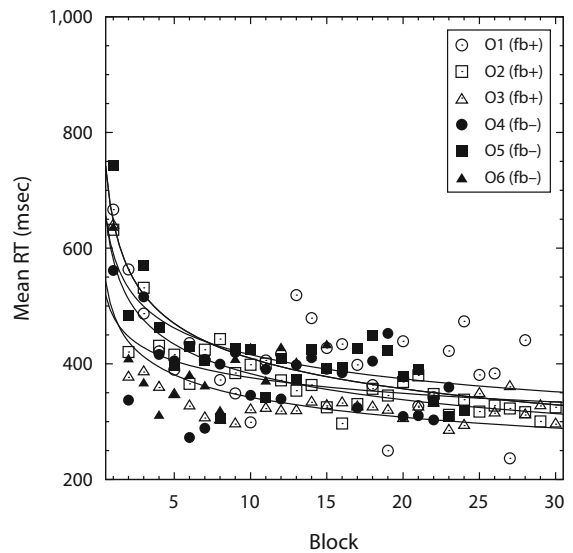


Figure 7. Experiment 2: Changes in mean response times (RTs) for false alarms as a function of trial block for 3 observers who received trial-by-trial feedback (fb+) and 3 observers who received no feedback (fb-). The symbols are the values of relative threshold; the lines are best-fitting power functions. The values in the legend are the initial level of false alarms for each observer.

**Table 7**  
**Experiment 2: Regression Analyses of Changes in Log Response Time (RT) for False Alarms As a Function of Log Block, for Each of the Observers in Each of the Two Feedback Conditions**

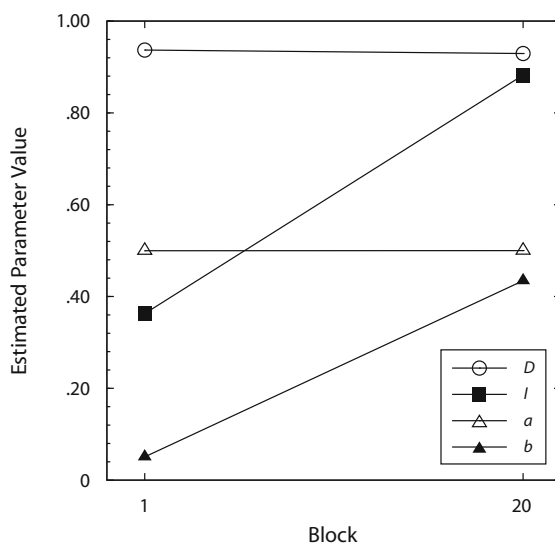
Feedback	Observer	$\hat{\beta}$	SE	$R^2$
Present	1	-0.1411**	0.0422	.8003
	2	-0.1733***	0.0173	.7822
	3	-0.1167***	0.0228	.8442
Absent	4	-0.1028	0.0423	.8161
	5	-0.0974*	0.2085	.7438
	6	-0.1756**	0.0707	.7221

Note— $\hat{\beta}$ , estimated slope of the line relating log RT to log block; SE, standard error of the estimate of  $\hat{\beta}$ ;  $R^2$ , proportion of variance accounted for. \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

we restricted our analysis to RTs for false alarms. Figure 7 presents the mean RTs as a function of practice, with those RTs showing very regular power function or exponential decreases, as one would expect for RTs for correct responses. The reliability of the changes in RT was assessed using simple linear regression for log RT as a function of log block, and the results of that analysis are presented in Table 7. All of the observers showed reliable decreases in mean RT for false alarms as a function of practice, consistent with the results of Experiment 1.

### Discussion

Experiment 2 was designed to investigate the extent to which the changes in both contrast detection thresholds and false alarm rates could be observed within observers in a task that combined the two methods used in Experiment 1. The design (see Figure 3) was a modification of one used by Kaernbach (1990). Adding target-absent trials to the standard 2I-2AFC design allowed us to estimate thresholds and false alarm rates for each observer in each block, using a trial sequence that is very similar



**Figure 8.** Estimated values of four parameters of the multinomial model as fit to the data from Blocks 1 and 20 for Observer 1, Experiment 2.

to those typically used in studies of perceptual learning. Consistent with the results of Experiment 1, we obtained reliable decreases in contrast detection thresholds, with these decreases accompanied by reliable increases in false alarm rates. Whereas the latter result suggests a liberal shift in response criterion for detection, the *absence* of any evidence for bias in the choice of presentation interval is consistent with our hypothesis that biases for detection and identification can exist independently of one another. Finally, the reliable reductions in RT for false alarms suggest, consistent with the results of Experiment 1, the presence of a speed-accuracy trade-off, a possibility we will consider further in the General Discussion section.

Before continuing, it is worth considering the extent to which the data from Experiment 2 can be used to demonstrate the basic viability of the simple multinomial model presented in the introduction (see Figure 1 and Table 1). To do this, we fit a restricted version of the model to the data from the first and last blocks of trials for Observer 1. The restrictions were motivated by the fact that the full model contains more parameters (eight) than there are degrees of freedom in the data (six). We constrained  $D_A = D_B$ ,  $I_A = I_B$ ,  $D_N = 0.01$ , and  $g = 0.50$ , resulting in four parameters that needed to be estimated. Parameters were estimated by minimizing the log likelihood measure  $G^2$  (e.g., Batchelder & Riefer, 1999). Figure 8 shows the changes in the values of the four parameters from the initial ( $G^2 = 0.066$ ) to the final ( $G^2 = 0.053$ ) block. Here, it can be seen that, for this observer, perceptual practice produced an increase in the parameter for identification ( $l$ ) and in the value of the parameter for detection bias ( $b$ ), with no changes in the parameter for detection ( $D$ ) or the bias for identification ( $a$ ).

### EXPERIMENTS 3-5

Although the results of Experiment 2 replicate the findings of Experiment 1 and reinforce the possibility that changes other than increases in perceptual sensitivity may be present in perceptual learning for contrast detection, it is still the case that the methods of Experiment 2 differed in a very critical way from standard practice in the literature. Specifically, as it is commonly implemented in studies of perceptual learning, the 2I-2AFC staircase procedure does not require observers to explicitly make present/absent judgments. This requirement was part of both the MCS procedure used in Experiment 1 and the 2I-3AFC procedure used in Experiment 2. Thus, it is possible that consistent practice in a task requiring present/absent judgments may be critical in producing the patterns that are suggestive of shifts in detection criterion.

Experiments 3-5 were designed to test this hypothesis, by requiring present/absent judgments only at the outset and conclusion of training. For the majority of each observer's practice in the contrast detection task, thresholds were estimated using one of three variations on the 2I-3AFC staircase procedure, thus allowing the majority of practice to be in a task context that did not require explicit present/absent decisions. If the requirement to make such decisions is critical in producing the patterns documented in Experiments 1 and 2, we should observe no increases

in false alarms and no reductions in RTs for false alarms, when comparing initial with final performance. Alternatively, however, if the patterns observed in Experiments 1 and 2 are not an artifact of requiring practice on the present/absent judgments, we should continue to observe increases in false alarms and reductions in associated RTs.

### Method

The three experiments used nearly identical methods, differing only in the implementation of the two display intervals. We selected three variations, consistent with commonly used methods in the literature (see Figure 3). Experiment 3 used two temporally distinct presentation intervals, Experiment 4 used two spatially distinct presentation intervals, and Experiment 5 defined the two response alternatives in terms of a dimension orthogonal to contrast (tilt).

### Participants

A total of 11 paid participants were recruited for Experiment 3, with data from 5 of those participants being discarded prior to analysis. A total of 7 paid participants were recruited for Experiment 4, with data from 1 of those participants being discarded prior to analysis. Finally, a total of 9 paid participants were recruited for Experiment 5, with data from 3 of those participants being discarded prior to analysis. The primary reason for discarding observers' data was difficulty in scheduling consistent testing sessions. All the participants had normal or corrected-to-normal vision and unencumbered use of their hands. All the participants were right-handed, and none had participated in any of the preceding experiments. Half of the participants in each experiment were assigned to a condition in which they received trial-by-trial feedback, and half were assigned to a condition in which they received no feedback.

### Materials, Design, and Procedure

The stimuli used in Experiments 3–5 were identical to those used in Experiments 1 and 2, with one exception for Experiment 5. In that experiment, the response alternatives were defined with respect to a dimension (tilt) orthogonal to contrast. In that experiment,  $\theta$  (tilt; see Equation 1) was  $53^\circ$  for one subset of the stimuli and  $37^\circ$  for the other subset. In all three experiments, the MCS procedure of Experiment 1 was used to estimate thresholds and false alarm rates in the first and final (12th) sessions. In Sessions 2–11, the 2I-2AFC procedure that is common in the literature was used. In all three experiments, contrasts in Sessions 2–11 were changed using the three-down one-up rule of Experiment 2.

Each trial in Sessions 2–11 of Experiment 3 proceeded as follows. The beginning of each trial was signaled with a dot centered in the display, and the trial was initiated with a button (pinky finger, left hand). The dot was then replaced with a fixation cross, centered in the display and present for between 400 and 500 msec, with the specific time on each trial set on the basis of an exponential distribution over that interval. This was followed by Display Interval 1, in which either a Gabor (of the appropriate contrast) or a blank was presented for 50 msec, followed by a randomly generated (on each trial) noise mask, also for 50 msec. Display Interval 2 followed 900 msec later, with the same timings for the Gabor or the blank and noise mask. The Gabor was present in only one of the two display intervals, with equal frequency for each interval in each block. The observers responded by using the index finger of their right hand to indicate that they perceived the target stimulus to have been presented in Display Interval 1 and the middle finger of their right hand to indicate that they perceived the target stimulus to have been presented in Display Interval 2. Following the response, feedback was provided to those observers assigned to the feedback condition, using the procedure in Experiment 2.

The trial events in Sessions 2–11 of Experiment 4 differed from those in Experiment 3 only in terms of the display intervals. In Experiment 4, the two display intervals were defined spatially, rather than temporally, with the first interval to the left of fixation and the second interval to the right of fixation. The test display thus consisted of the fixation cross, a Gabor patch at the appropriate level of contrast to

either the left or the right of the fixation cross, and a blank gray square on the opposite side of the fixation cross. The display remained visible for 50 msec, after which two randomly generated noise masks were presented in the two display locations. The fixation cross remained visible for the entire trial. Each of the two locations was used for the placement of the target with equal frequency in each block. The observers used the index finger of their right hand to indicate that they perceived the Gabor to be present to the left of the fixation cross and the middle finger of their right hand to indicate that they perceived the Gabor to be present to the right of the fixation cross.

In Sessions 2–11 of Experiment 5, the two display intervals were defined by the state of the stimulus on a dimension orthogonal to contrast: tilt. The test display consisted of a single Gabor, at the appropriate level of contrast, at one of two distinct levels of tilt. The two values of tilt were used with equal frequency in each block of trials. The observers responded with the index finger of their right hand to indicate that they perceived the Gabor at  $\theta = 53^\circ$  and with the middle finger of their right hand to indicate that they perceived the Gabor at  $\theta = 37^\circ$ .

## Results

### Changes in Detection Thresholds

Detection thresholds for the MCS blocks (first and final) were estimated using the linearized Weibull (Equation 4);  $R^2$  values for the obtained fits were greater than .92 in all cases. Detection thresholds for the 2I-2AFC staircase blocks (Days 2–11) were estimated by taking the geometric mean of the contrast values for the final 15 reversals. Relative contrast detection thresholds for all three experiments are presented in the three panels of Figure 9. As in Experiment 2, the overall pattern of reductions in thresholds followed power law or exponential decreases. The reliability of these reductions was assessed using simple linear regression relating log threshold to log block. The results of these analyses are presented in Table 8. The changes in threshold were reliable for all the observers in all three experiments, providing the standard evidence for perceptual learning.

### Changes in False Alarm Rates

The symbols at the far right of panels A–C in Figure 9 plot the relative value of the final false alarm rates (final divided by initial) for each of the 6 observers in each of the three experiments. The initial false alarm rates are indicated in the legend of Figure 9. Reliability of the differences between initial and final false alarm rates was assessed using a normal approximation of  $\chi^2$  (Zar, 1999), with all of the comparisons reaching standard levels of statistical reliability (all  $ps < .02$ ). Thus, even in the absence of making explicit present/absent judgments for the majority of their experimental experience, all of the observers showed large and reliable increases in false alarm rates. This strongly suggests that the patterns observed in Experiments 1 and 2 (and, by extension, those documented in Copeland, 2003; Wenger & Rasche, 2006)—particularly the increase in false alarm rates that suggests the liberal shift in response bias—are not artifacts of the method used to assess changes in threshold due to practice.

### Changes in Interval Errors

The design of Experiments 3–5, in comparison with that of Experiment 2, allowed us to assess changes in the

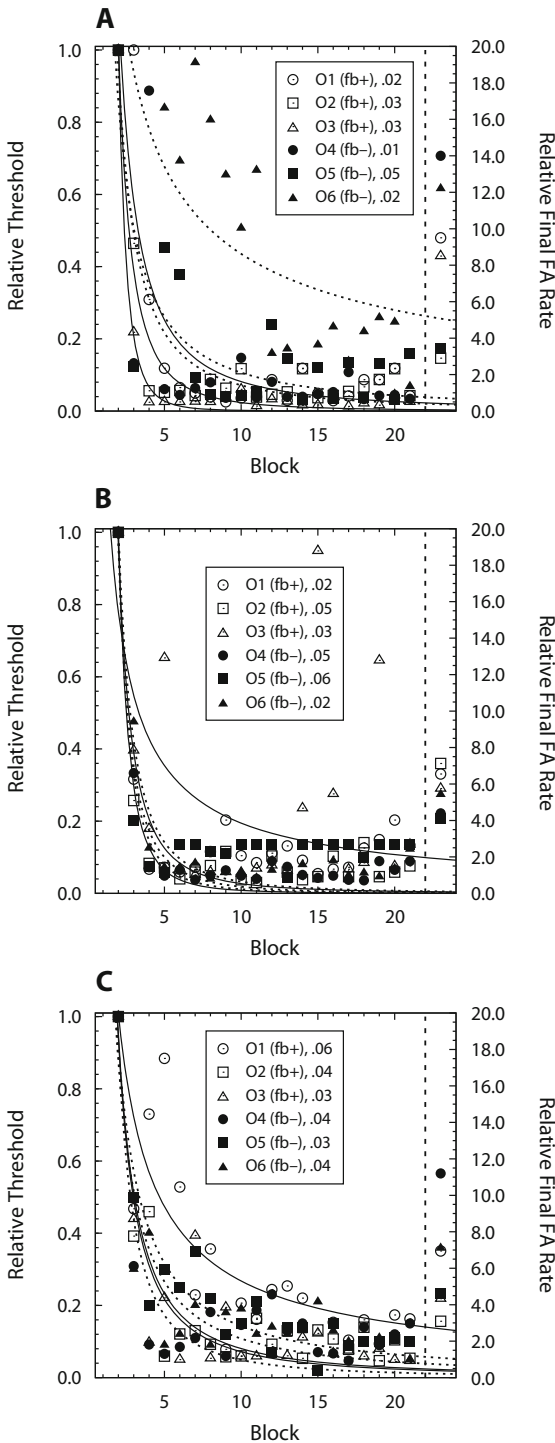


Figure 9. Experiments 3–5: Changes in relative threshold as a function of block of practice for the observers in Experiments 3 (A), 4 (B), and 5 (C), for the observers who received trial-by-trial feedback (fb+), and the observers who received no feedback (fb–). The symbols are the values of relative threshold; the lines are best-fitting power functions.

Table 8  
Experiments 3–5: Regression Analyses of Changes in Log Threshold As a Function of Log Block, for Each of the Observers in Each of the Two Feedback Conditions

	Feedback	Observer	$\hat{\beta}$	SE	R <sup>2</sup>
Experiment 3	Present	1	-0.8849*	0.3185	.7218
		2	-0.7535**	0.2361	.8614
		3	-1.0486***	0.2577	.7491
	Absent	4	-0.8714***	0.2098	.7036
		5	-0.5844*	0.1999	.7594
		6	-0.7073*	0.2831	.7525
Experiment 4	Present	1	-0.5225*	0.2191	.7112
		2	-0.7143**	0.2249	.8236
		3	-0.4637*	0.3513	.8820
	Absent	4	-0.7767**	0.2221	.8341
		5	-0.3376*	0.1784	.7160
		6	-0.8789***	0.1990	.8200
Experiment 5	Present	1	-0.6130*	0.2510	.7340
		2	-0.8205**	0.2459	.8429
		3	-0.7646***	0.1432	.7413
	Absent	4	-0.8578**	0.2337	.7422
		5	-0.8599*	0.5558	.6840
		6	-0.7553*	0.3187	.7794

Note— $\hat{\beta}$ , estimated slope of the line relating log threshold to log block; SE, standard error of the estimate of  $\hat{\beta}$ ; R<sup>2</sup>, proportion of variance accounted for. \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

relative bias for choosing one of the two response intervals only for the blocks of trials on Days 2–11. A check on the relative rates of choosing the two intervals allows us to distinguish a bias for detection (as revealed by the false alarm rates) from a bias for identification. The rates of erroneously choosing Interval 1, relative to Interval 2, are presented in the three panels of Figure 10. The reliability of the differences in these error rates was assessed using a normal approximation to  $\chi^2$  (Zar, 1999). None of the comparisons reached standard levels of statistical reliability (all  $ps > .21$ ). Thus, these results replicate those of Experiment 2 and (along with the increases in false alarm rates) provide evidence that an absence of bias for identification can exist alongside a liberal shift in bias for detection.

**Changes in RTs**

Finally, as in Experiments 1 and 2, we examined RTs for false alarms in the first and last blocks of the experiment (the two blocks in which the MCS was used). Mean RTs for each of the observers are presented in Figure 11. The reliability of the differences in mean RTs as a function of block for each of the observers was assessed using a one-tailed paired-sample  $t$  test, performed separately for each observer. The differences were reliable for all of the observers (all  $ts > 3.11$ , all  $ps < .005$ ), consistent with the results of Experiments 1 and 2.

**Discussion**

The results of Experiments 3–5 indicate that the basic patterns observed in Experiments 1 and 2—specifically, increases in false alarm rates accompanying reductions in detection thresholds—are regular features of learning in contrast detection and are not artifacts of a requirement to make present/absent judgments. In each of these three experiments, the majority of the observers’ practice with

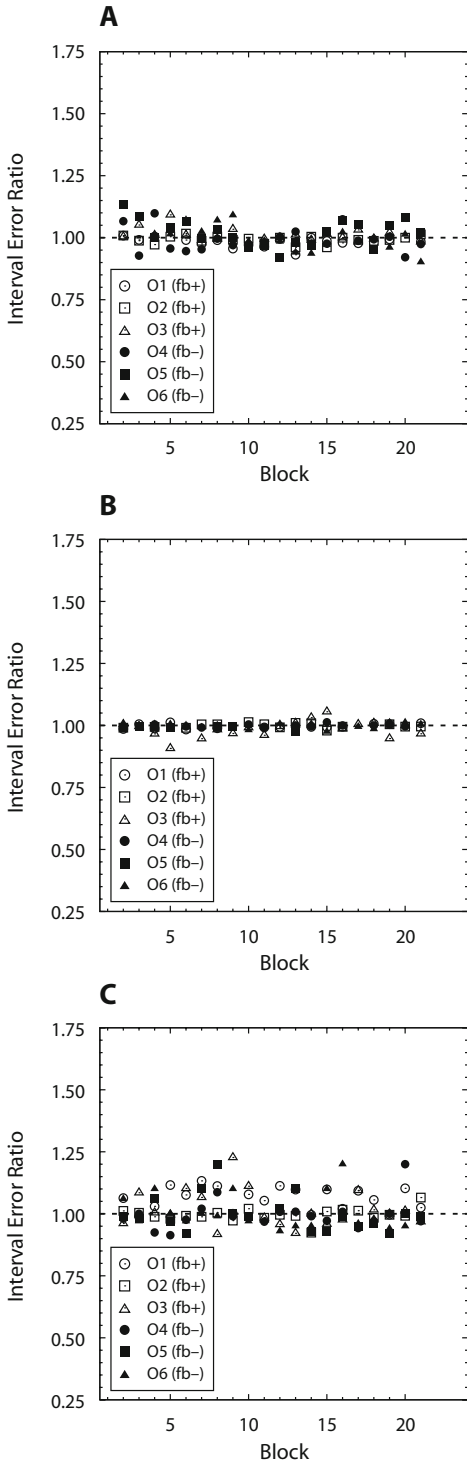


Figure 10. Experiments 3–5: Changes in the ratio of interval errors as a function of block of practice for the observers (O) in Experiments 3 (A), 4 (B), and 5 (C), for the observers who received trial-by-trial feedback (fb+) and the observers who received no feedback (fb-).

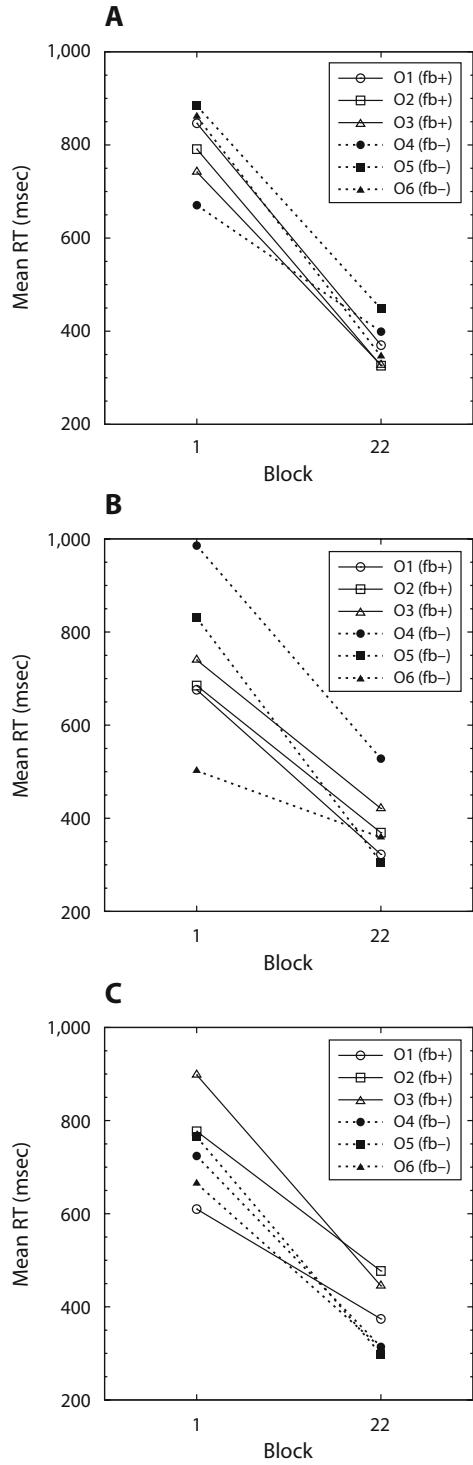


Figure 11. Experiments 3–5: Changes in mean reaction times (RTs) for false alarms a function of block of practice (first vs. last) for the observers (O) in Experiments 3 (A), 4 (B), and 5 (C), for the observers who received trial-by-trial feedback (fb+) and the observers who received no feedback (fb-).



the contrast detection task (10 of 12 sessions) did not require that the observers make present/absent judgments. Instead, the observers practiced the contrast detection task in three variants of the 2I-2AFC paradigm, in which the target stimulus was always present in one of the two intervals. In each case, there were three regularities: (1) contrast detection thresholds were reduced in a manner consistent with either power or exponential change as a function of practice; (2) false alarm rates increased linearly with practice, suggesting a liberal shift in bias for detection; and (3) preference for either of the two response alternatives was equal and stable as a function of practice, suggesting a lack of bias for identification. The pattern of results from these three experiments thus suggests that the patterns originally observed by Copeland (2003; Copeland & Wenger, 2003) and replicated in Experiments 1 and 2 (see also Seitz et al., 2005; Wenger & Rasche, 2006) cannot be explained by reference to any methodological artifact.

### GENERAL DISCUSSION

Studies of perceptual learning date to the origins of sensory science (see the discussions in, e.g., Ahissar & Hochstein, 1998; Sinha & Poggio, 2002), with interest over the past few decades being motivated, at least in part, by the possibility that the structure and function of early levels of the sensory systems may be highly plastic. Theoretical accounts have typically considered this plasticity in the context of feedforward systems, at levels well below those associated with attention, memory, and other cognitive processes. For example, perceptual learning has been distinguished from higher level (or cognitive) learning on the basis of the specificity of the effects of experience and the very low levels of transfer across stimulus variations (e.g., Doshier, Liu, Blair, & Lu, 2004; Doshier & Lu, 1999; Fahle & Poggio, 2002; Hawkey, Amitay, & Moore, 2004). However, this logic has been criticized for ignoring data suggesting that central and high-level processes can be at work in the absence of transfer (Mollon & Danilova, 1996; Vogels & Orban, 1994; Yu, Klein, & Levi, 2003, 2004). In addition, there are a number of parallels between perceptual and cognitive learning, including the power law or exponential improvements in performance (Doshier & Lu, 2007; Logan, 1988, 1992; Newell & Rosenbloom, 1981; Palmeri, 1999) documented in the present study. Finally, there is accumulating evidence that multiple levels of representation may be involved in perceptual learning and that tasks assumed to reflect low-level changes in perceptual sensitivity may involve a variety of high-level influences, as is the case for both cognitive and motor skills (e.g., Chase & Ericsson, 1981; Ericsson & Lehmann, 1996; Rosenbaum et al., 2001; Wenger & Payne, 1995).

The results of the present study also suggest the possibility of multiple levels of influence in perceptual learning. Most critically, the set of results presented here suggests that there may be liberal shifts in response criteria for detection across the course of learning, with these liberal shifts being evidenced by the regular increases in

false alarm rates for detection. Importantly, these liberal shifts in detection responses were not accompanied by any shifts in bias for identification. That is, although the rate of false positive detection responses increased as a function of practice, this increase was not accompanied by any detectable shift in the relative rates of choosing one of the two response alternatives used for the presence of the target stimulus.

### Stimulus and Task Factors

A modal interpretation of the changes in false alarm rates for detection would make reference to volitional, strategic, and/or attentional factors, many of which can be reliably influenced by a variety of stimulus and task manipulations. We will begin with a consideration of these factors and then will move to a set of additional possibilities.

Response feedback is one of the most obvious manipulations that can be considered with respect to shifts in response criteria (e.g., Atkinson, Carterette, & Kinchla, 1964; Green & Swets, 1966; Herzog & Fahle, 1997, 1999; Mackie, Wylie, & Smith, 1994; Tanner, Rauk, & Atkinson, 1970). Each of the experiments presented here involved a between-observers manipulation of the presence of feedback, with a consistent outcome: The presence of feedback made no reliable difference in the rate of false alarms as a function of practice. This is not altogether surprising, given the mixed evidence for the effects of feedback in perceptual learning (e.g., Fahle, 2002; Fahle & Edelman, 1993; McKee & Westheimer, 1978).

A second manipulation of potential interest involves the objective and subjective relative frequency of target-present and target-absent trials. Manipulating relative frequency has been shown, under various conditions (see, e.g., Craig, 1976; Green & Swets, 1966; Macmillan & Creelman, 2005; Parducci & Sandusky, 1965; Parks & Kellcutt, 1968; Schulman & Greenberg, 1970; Tanner, Haller, & Atkinson, 1967), to cause shifts in observers' response criteria. It is possible that, as learning proceeds, observers' subjective estimate of the relative frequency of target-present trials increases. Consequently, by reducing the relative frequency of target-present trials as a function of practice, it may be possible to attenuate or reverse the increase in false alarm rates.<sup>4</sup>

A third manipulation of potential interest involves the subjective value of correct and incorrect responses as a function of learning. It may be the case that the subjective benefit of being correct, given the presence of a weak signal, is far greater than the subjective cost of being incorrect, given the absence of a signal. Consequently, by manipulating explicit payoffs (e.g., Maddox & Bohil, 2003), possibly in conjunction with informational feedback and variations in relative frequencies, it may be possible to attenuate the increases in false alarm rates.

### Speed–Accuracy Trade-Offs

A standard signature of potentially confounding strategic influences in task performance is the speed–accuracy trade-off: a reduction in RT that comes at the cost of an increase in errors (e.g., Pachella, 1974; Pachella & Fisher,

1972; Pachella & Pew, 1968). If such a trade-off were at work in our data, we would expect to see consistent and reliable reductions in mean RTs on false alarm trials as the relative frequency of false alarms increased. This, in fact, was the pattern we observed in all five experiments (see Figures 4A, 7, and 11). Although the regularity of the trade-off suggests that simple strategic emphasis on detection responses is at the core of our findings, there is at least one reason to be skeptical of this as a comprehensive explanation. Specifically, if the changes we have documented are due to volitional, strategic choice, it would seem reasonable to expect that feedback would have some effect, with the trade-off present for observers who did not receive feedback and absent for those who did. However, the trade-off was present for all the observers, independently of the presence or absence of trial-by-trial feedback. Consequently, although the data are, at one level, consistent with the hypothesis that our effects reflect strategic trade-offs, we would assert that there are reasons to consider alternative explanations.

### Changes in Performance, Assuming Optimality

An alternative hypothesis for the source of the increases in false alarm rates and decreases in RTs for false alarms emerges from a recent analysis by R. D. Thomas (2006). This analysis makes the provocative suggestion that increases in false alarm rates may be obtained even when observers are responding optimally. The analysis takes into consideration the dynamics of the accrual of perceptual evidence for two-alternative response tasks. A key element of this analysis, for present purposes, is the notion that there is a subjective cost for waiting for the accrual of additional evidence (Edwards, 1965; Townsend & Wenger, 1996, 1999). Assume, for the sake of the present argument, that as observers increase their experience with the contrast detection task, they develop an expectation for the relative amount of evidence available as a function of time since the beginning of the trial. Furthermore, assume that as a consequence of this developing expectation, observers are increasingly less willing to wait for evidence to accumulate as their experience with the task increases. Such changes would be consistent with the nature of the task, since the exposure durations are short and relative contrasts decrease as practice proceeds. Finally, assume that at all points in practice, observers set their response criteria for the two alternatives in order to respond optimally (maximizing the expected value of the response on each trial).

Under these assumptions, R. D. Thomas (2006) has shown that it is possible to numerically determine the expected rate of false alarms and RT for false alarms, as a function of the discriminability of the two alternative stimuli and the subjective cost of waiting. Figure 12A presents the results of a set of such numeric simulations; the details of the simulation methods are presented in Appendix B. In Figure 12A,  $\Delta\mu$  is the difference (dissimilarity) of the two stimuli, and  $c$  is the subjective cost of waiting. Here, it can be seen that at low levels of dissimilarity between the two response alternatives (low levels of contrast in the

detection task), as the subjective cost of waiting increases, so does the probability of a false alarm. Figure 12B presents analogous results for the RTs for false alarms. Consequently, this analysis suggests that the patterns we have observed (here and in Copeland, 2003; Wenger & Rasche, 2006) can occur even when observers can be assumed to be responding optimally.

### Sustained Activation

The last possibility that we will consider is one that allows for the types of changes in behavior that we have documented but does not require any type of strategic or conscious control. The mechanism here is sustained activation in cortical regions responsible for coding the critical stimulus properties. Many of the cortical regions that can be reasonably assumed to be involved in coding for contrast detection—including the primary visual cortex, intraparietal sulcus, and dorsolateral prefrontal cortex—are laminar structures that have been shown to produce patterns of sustained activation that can act as short-duration memories (e.g., Bisley, Zaksas, Droll, & Pasternak, 2004; Brown & Bashir, 2002; Dehaene & Changeux, 2005; Del Giudice, Fusi, & Mattia, 2003; Harris, Miniussi, Harris, & Diamond, 2002; Major & Tank, 2004; McCormick et al., 2003; Ranganath, Cohen, Dam, & D'Esposito, 2004; Rigas & Castro-Alamancos, 2007; Roelfsema, 2005; Zhang & Alloway, 2004). The question of interest with respect to the present project is the extent to which such sustained activation might be operative in producing increases in false alarms with increasing experience.

To consider the potential role of sustained activation, we modeled the decay in sustained activation in a cortical population, using a simple exponential decay function for neuronal firing rate. Our goal was to determine whether a reasonable model for decay could be perturbed by the addition of noise in order to produce outputs consistent with increases in false alarm rates. The dependent measure of interest was the probability that the presentation of the target-absent stimulus would produce a level of activity consistent with that of the target-present stimulus with variations in onset asynchrony. Details of the simulation method are presented in Appendix C.

Figure 13 shows two characteristics of model performance following training. First, following offset of the target-absent stimulus, the model produces output that is consistent with the target's being present, with this activation declining slowly over time. Second, a 100-msec presentation of a target-absent (noise) stimulus is capable of driving this activation up, at onset asynchronies that are long enough to be consistent with the time between a target-present and a target-absent trial in a typical experiment. This suggests that such sustained activation could lead to increases in false alarms, with such increases being dependent on the time that has elapsed since the most recent target-present trial, and with this dependency being most pronounced following training. We thus examined our data for any evidence that such a pattern might be present. The nine panels of Figure 14 present the data from all five experiments. In each of the experiments, the probabil-

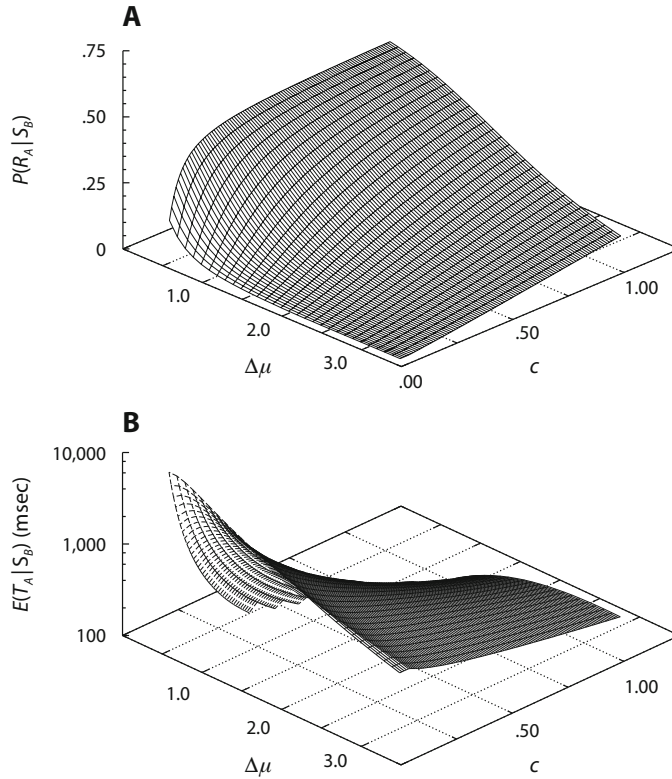


Figure 12. Results of numerical simulations of the probability of a false alarm [ $P(R_A | S_B)$ ] and the expected value of the RT for a false alarm [ $E(T_A | S_B)$ ], as a function of target–distractor dissimilarity ( $\Delta\mu$ ) and subjective cost of waiting ( $c$ ), based on the analysis in R. D. Thomas (2006). See Appendix B for details of the numeric simulations.

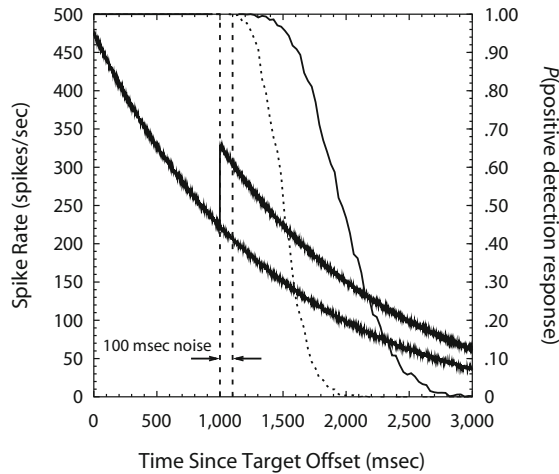


Figure 13. Effects of adding noise in a model of decay of sustained activation. The two jagged descending curves show the firing rate with (thicker line) and without (thinner line) the injection of 100 msec of noise at 1,000 msec post-stimulus-offset. The two smooth descending curves show the probability that the firing rate will exceed an arbitrary criterion for coding the presence of a target stimulus.

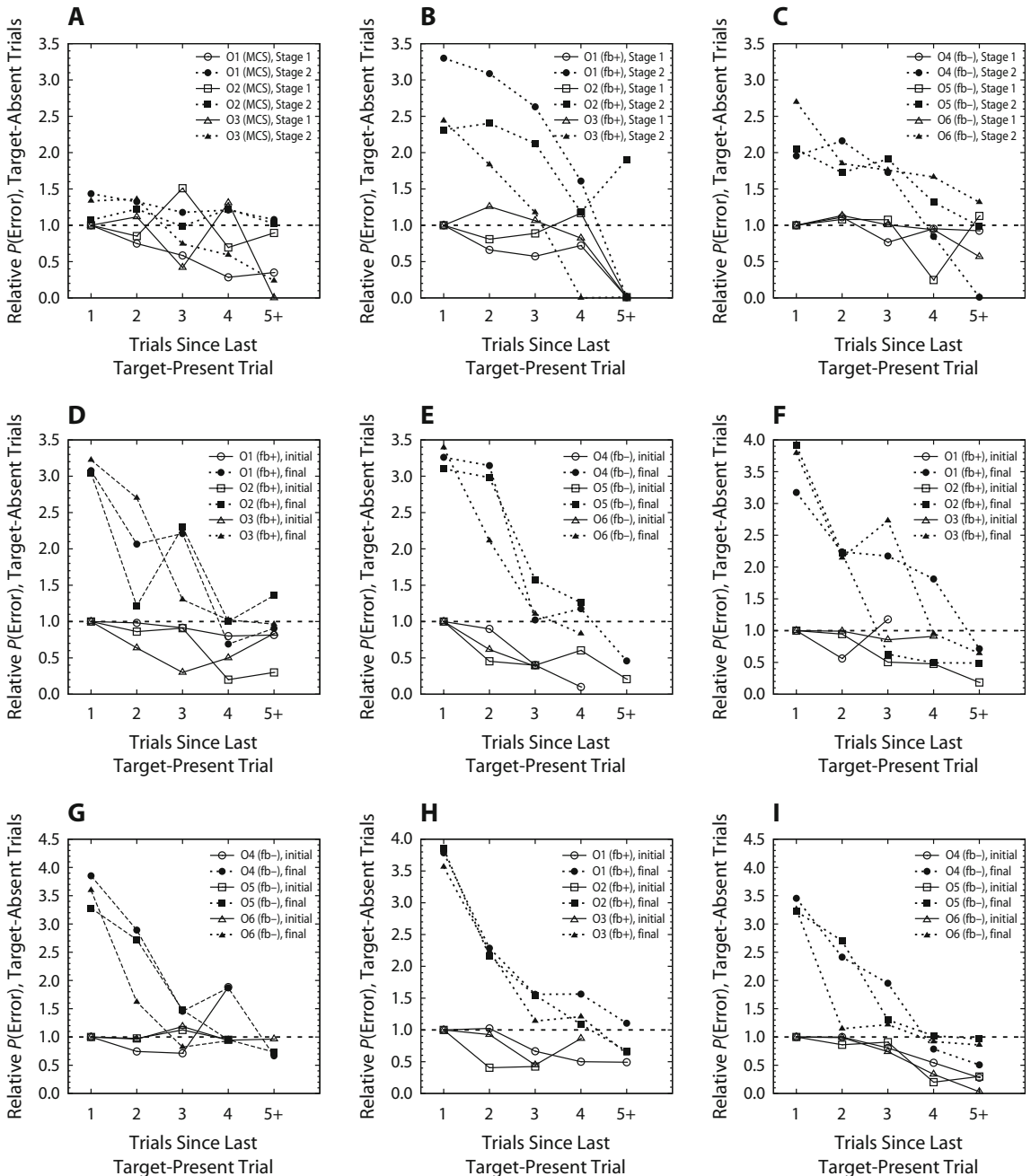


Figure 14. Probability of a false alarm as a function of the number of trials since the most recent target-present trial: (A) Experiment 1; (B) Experiment 2, with feedback; (C) Experiment 2, without feedback; (D) Experiment 4, with feedback; (E) Experiment 4, without feedback; (F) Experiment 5, with feedback; (G) Experiment 5, without feedback; (H) Experiment 6, with feedback; (I) Experiment 6, without feedback. MCS, method of consent stimuli; fb+, trial-by-trial feedback; fb-, no feedback.

ity of a false alarm decreases with increasing lag between the most recent target-present trial, with this effect being most pronounced following extensive practice on the contrast detection task. Thus, there is evidence from all of the experiments that is consistent with the pattern that would be produced by sustained activation (note that this pattern could also reflect factors related to sequential effects; see, e.g., Kornblum, 1969). The critical point here is that this pattern does not of necessity require any type of top-down, conscious, or strategic influence; instead, it may be an emergent functional property of the laminar structures.

## CONCLUSIONS

The theoretical and empirical work presented here strongly suggests that perceptual learning can involve more than improvements in perceptual sensitivity or identification abilities. Critically, and consistent with other recent reports in the literature (e.g., Seitz et al., 2005; Seitz & Watanabe, 2005; Wild & Busey, 2004), practice on contrast detection produces both reliable decreases in detection thresholds and reliable increases in false alarm rates. Critically, these increases in false alarm rates with practice occur in the absence of any changes in identification preferences and occur even when observers do not practice the task by making explicit present/absent decisions. Thus, the results of the five experiments reported here are consistent with the two central hypotheses of this study. The initial finding (Copeland, 2003; Wenger & Rasche, 2006) of a reliable liberal shift in response bias accompanying reliable decreases in detection thresholds was replicated in the task that originally revealed this pattern (MCS), in a modification of the more standard (for this literature) adaptive procedure (see also Kaernbach, 1990; Klein, 2001), and in an assessment following practice with three variations on the modal adaptive procedure. In addition, the data from Experiments 2–5 demonstrate that the increases in the false alarm rates for identification (the evidence for the liberal shift in detection thresholds) occur in the absence of any evidence for a bias, or a change in bias, for identification responses, consistent with the simple theoretical model presented in the introduction. Specifically, although false alarm rates for detection increased, there was no evidence suggesting that this increase in *target-present* responses was accompanied by any change—or any change in preference—for the response alternatives indicating the presence of contrast. Finally, these patterns occurred independently of the presence or absence of trial-by-trial feedback and, critically, the presence or absence of the requirement for explicit present/absent judgments (Experiments 3–5).

Although this evidence suggests the possibility of a shift in an explicit, volitional response strategy, and although there are patterns in the data that are consistent with this possibility, we believe there are viable alternatives to this inference, each of which finds some level of support in the data. In particular, there are alternative explanations, including changes that would be consistent with either optimal responding or sustained activation in regions of

the visual system, that would be capable of producing increases in false alarm rates without the need to posit any type of strategic change on the part of the observer. Further analysis and experimentation, of course, is needed in order to test among these possibilities. Whatever the outcome, we argue that the nature of the changes involved in perceptual learning is likely to be more complex than current theory might suggest, with the potential for involvement of a range of cortical regions beyond the visual and necessary recurrent connectivity from higher cortical regions (Ahissar & Hochstein, 1993, 2002, 2004; Hochstein & Ahissar, 2002; Li et al., 2004).

## AUTHOR NOTE

Sincere thanks are due John Palmer, Cathleen Moore, and Aaron Seitz for a set of very helpful discussions regarding this work. Correspondence concerning this article should be addressed to M. J. Wenger, Department of Psychology, Pennsylvania State University, 216 Moore Building, University Park, PA 16802 (e-mail: mjw19@psu.edu).

## REFERENCES

- ADINI, Y., & SAGI, D. (2001). Recurrent networks in human visual cortex: Psychophysical evidence. *Journal of the Optical Society of America A*, *18*, 2228-2236.
- ADINI, Y., SAGI, D., & TSODYKS, M. (2002). Context-enabled learning in the human visual system. *Nature*, *415*, 790-793.
- AHISSAR, M., & HOCHSTEIN, S. (1993). Attentional control of early perceptual learning. *Proceedings of the National Academy of Sciences*, *90*, 5718-5722.
- AHISSAR, M., & HOCHSTEIN, S. (1998). Perceptual learning. In V. Walsh & J. Kulikowski (Eds.), *Perceptual constancy: Why things look as they do* (pp. 455-498). Cambridge: Cambridge University Press.
- AHISSAR, M., & HOCHSTEIN, S. (2002). The role of attention in learning simple visual tasks. In M. Fahle & T. Poggio (Eds.), *Perceptual learning* (pp. 253-272). Cambridge, MA: MIT Press, Bradford Books.
- AHISSAR, M., & HOCHSTEIN, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, *8*, 457-464.
- AMIT, D. J., & BRUNEL, N. (1995). Learning internal representations in an attractor neural network with analogue neurons. *Network*, *6*, 359-388.
- AMIT, D. J., & BRUNEL, N. (1997). Dynamics of a recurrent network of spiking neurons before and following learning. *Network*, *8*, 373-404.
- ASHBY, F. G. (2000). A stochastic version of general recognition theory. *Journal of Mathematical Psychology*, *44*, 310-329.
- ATKINSON, R. C., CARTERETTE, E. C., & KINCHLA, R. A. (1964). The effect of information feedback upon psychophysical judgments. *Psychonomic Science*, *1*, 83-84.
- BATCHELDER, W. H., & RIEFER, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review*, *97*, 548-564.
- BATCHELDER, W. H., & RIEFER, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review*, *6*, 57-86.
- BISLEY, J. W., ZAKSAS, D., DROLL, J. A., & PASTERNAK, T. (2004). Activity of neurons in cortical area MT during a memory for motion task. *Journal of Neurophysiology*, *91*, 286-300.
- BROWN, M. W., & BASHIR, Z. I. (2002). Evidence concerning how neurons of the perirhinal cortex may effect familiarity discrimination. *Philosophical Transactions of the Royal Society B*, *357*, 1083-1095.
- CHASE, W. G., & ERICSSON, K. A. (1981). Skilled memory. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 141-190). Hillsdale, NJ: Erlbaum.
- COPELAND, A. M. (2003). *Multiple sources of evidence for the mechanisms of perceptual learning*. Unpublished master's thesis, University of Notre Dame.
- COPELAND, A. M., & WENGER, M. J. (2003). Investigating perceptual and decisional mechanisms for the dynamics of perceptual learning: Theory, models, and data [Abstract]. *Journal of Vision*, *3*(9), 676a.
- CRAIG, A. (1976). Signal recognition and the probability-matching decision rule. *Perception & Psychophysics*, *20*, 157-162.

- DEHAENE, S., & CHANGEUX, J.-P. (2005). Ongoing spontaneous activity controls access to consciousness: A neuronal model for inattentional blindness. *PLoS Biology*, *3*, 910-927.
- DEL GIUDICE, P., FUSI, S., & MATTIA, M. (2003). Modelling the formation of working memory with networks of integrate-and-fire neurons connected by plastic synapses. *Journal of Physiology*, *97*, 659-681.
- DI LOLLO, V., ENNS, J. T., & RENSINK, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General*, *129*, 481-507.
- DOSHER, B. A., LIU, S.-H., BLAIR, N., & LU, Z.-L. (2004). The spatial window of the perceptual template and endogenous attention. *Vision Research*, *44*, 1257-1271.
- DOSHER, B. A., & LU, Z.-L. (1999). Mechanisms of perceptual learning. *Vision Research*, *39*, 3197-3221.
- DOSHER, B. A., & LU, Z.-L. (2007). The functional form of performance improvements in perceptual learning: Learning rates and transfer. *Psychological Science*, *18*, 531-539.
- EDWARDS, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology*, *2*, 312-329.
- ERICSSON, K. A., & LEHMANN, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, *47*, 273-305.
- FAHLE, M. (2002). Learning to perceive features below the foveal photoreceptor spacing. In M. Fahle & T. Poggio (Eds.), *Perceptual learning* (pp. 197-218). Cambridge, MA: MIT Press, Bradford Books.
- FAHLE, M., & EDELMAN, S. (1993). Long-term learning in vernier acuity: Effects of stimulus orientation, range and of feedback. *Vision Research*, *33*, 397-412.
- FAHLE, M., & POGGIO, T. (Eds.) (2002). *Perceptual learning*. Cambridge, MA: MIT Press, Bradford Books.
- FECHNER, G. T. (1966). *Elements of psychophysics* (H. E. Adler, Trans.). New York: Holt, Rinehart & Winston. (Original work published 1860)
- GIBSON, E. J. (1953). Improvement in perceptual judgments as a function of controlled practice or training. *Psychological Bulletin*, *50*, 401-431.
- GIBSON, E. J. (1969). *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts.
- GIBSON, E. J., WALK, R. D., PICK, H. L., JR., & TIGHE, T. J. (1956). The effect of prolonged exposure to visual patterns on learning to discriminate similar and different patterns. *Journal of Comparative & Physiological Psychology*, *51*, 584-587.
- GOLD, J. [M.], BENNETT, P. J., & SEKULER, A. B. (1999). Signal but not noise changes with perceptual learning. *Nature*, *402*, 176-178.
- GOLD, J. M., SEKULER, A. B., & BENNETT, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science*, *28*, 167-207.
- GREEN, D. M., & SWETS, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- GRENT-T-JONG, T., & WOLDORFF, M. G. (2007). Timing and sequence of brain activity in top-down control of visual-spatial attention. *PLoS Biology*, *5*, 114-127.
- HARRIS, J. A., MINIUSSI, C., HARRIS, I. M., & DIAMOND, M. E. (2002). Transient storage of a tactile memory trace in primary somatosensory cortex. *Journal of Neuroscience*, *22*, 8720-8725.
- HATFIELD, J. L., & SODERQUIST, D. R. (1969). Practice effects and signal detection indices in an auditory vigilance task. *Journal of the Acoustical Society of America*, *46*, 1458-1463.
- HAWKEY, D. J. C., AMITAY, S., & MOORE, D. R. (2004). Early and rapid perceptual learning. *Nature Neuroscience*, *7*, 1055-1056.
- HEATHCOTE, A., BROWN, S., & MEWHORT, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review*, *7*, 185-207.
- HERZOG, M. H., & FAHLE, M. (1997). The role of feedback in learning a vernier discrimination task. *Vision Research*, *37*, 2133-2141.
- HERZOG, M. H., & FAHLE, M. (1998). Modeling perceptual learning: Difficulties and how they can be overcome. *Biological Cybernetics*, *78*, 107-117.
- HERZOG, M. H., & FAHLE, M. (1999). Effects of biased feedback on learning and deciding in a vernier discrimination task. *Vision Research*, *39*, 4232-4243.
- HOCHSTEIN, S., & AHISSAR, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, *36*, 791-804.
- HU, X., & BATCHELDER, W. H. (1994). The statistical analysis of engineering processing tree models with the EM algorithm. *Psychometrika*, *59*, 21-47.
- JACOBS, D. M., & MICHAELS, C. F. (2006). Lateral interception: I. Operative optical variables, attunement, and calibration. *Journal of Experimental Psychology: Human Perception & Performance*, *32*, 443-458.
- KAERNBACH, C. (1990). A single-interval adjustment matrix (SIAM) procedure for unbiased adaptive testing. *Journal of the Acoustical Society of America*, *88*, 2645-2655.
- KLEIN, S. A. (2001). Measuring, estimating, and understanding the psychometric function: A commentary. *Perception & Psychophysics*, *63*, 1421-1455.
- KORNBLUM, S. (1969). Sequential determinants of information processing in serial and discrete choice reaction time. *Psychological Review*, *76*, 113-131.
- LAMME, V. A. F., & ROELFSEMA, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*, 571-579.
- LEEK, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics*, *63*, 1279-1292.
- LEVITT, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, *49*, 467-477.
- LI, W., PIÉCH, V., & GILBERT, C. D. (2004). Perceptual learning and top-down influences in primary visual cortex. *Nature Neuroscience*, *7*, 651-657.
- LING, S., & CARRASCO, M. (2006). When sustained attention impairs perception. *Nature Neuroscience*, *9*, 1243-1245.
- LOGAN, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492-527.
- LOGAN, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *18*, 883-914.
- LU, Z.-L., & DOSHER, B. A. (2008). Characterizing observers using external noise and observer models: Assessing internal representations with external noise. *Psychological Review*, *115*, 44-82.
- MACKIE, R. R., WYLIE, C. D., & SMITH, M. J. (1994). Countering loss of vigilance in sonar watchstanding using signal injection and performance feedback. *Ergonomics*, *37*, 1157-1184.
- MACMILLAN, N. A., & CREELMAN, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.
- MADDOX, W. T., & BOHIL, C. J. (2003). A theoretical framework for understanding the effects of simultaneous base-rate and payoff manipulations on decision criterion learning in perceptual categorization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*, 307-320.
- MAJOR, G., & TANK, D. (2004). Persistent neural activity: Prevalence and mechanisms. *Current Opinion in Neurobiology*, *14*, 675-684.
- MCCORMICK, D. A., SHU, Y., HASENSTAUB, A., SANCHES-VIVES, M., BADOUAL, M., & BAL, T. (2003). Persistent cortical activity: Mechanisms of generation and effects on neuronal excitability. *Cerebral Cortex*, *13*, 1219-1231.
- MCKEE, S. P., & WESTHEIMER, G. (1978). Improvement in vernier acuity with practice. *Perception & Psychophysics*, *24*, 258-262.
- MILGRAM, P. (1987). A spectacle-mounted liquid-crystal tachistoscope. *Behavior Research Methods, Instruments, & Computers*, *19*, 449-456.
- MOLLON, J. D., & DANILOVA, M. V. (1996). Three remarks on perceptual learning. *Spatial Vision*, *10*, 51-58.
- MURESAN, R. C., & SAVIN, C. (2007). Resonance or integration? Self-sustained dynamics and excitability of neuronal microcircuits. *Journal of Neurophysiology*, *97*, 1911-1930.
- NEWELL, A., & ROSENBLUM, P. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.
- NISBETT, R. E., & WILSON, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.
- OLZAK, L. A., & THOMAS, J. P. (1981). Gratings: Why frequency dis-

- crimination is sometimes better than detection. *Journal of the Optical Society of America*, **71**, 64-70.
- PACHELLA, R. G. (1974). The interpretation of reaction time in information processing research. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 41-82). Hillsdale, NJ: Erlbaum.
- PACHELLA, R. G., & FISHER, D. (1972). Hick's law and the speed-accuracy trade-off in absolute judgment. *Journal of Experimental Psychology*, **92**, 378-384.
- PACHELLA, R. G., & PEW, R. W. (1968). Speed-accuracy tradeoff in reaction time: Effect of discrete criterion times. *Journal of Experimental Psychology*, **76**, 19-24.
- PALMERI, T. J. (1999). Theories of automaticity and the power law of practice. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 543-551.
- PARDUCCI, A., & SANDUSKY, A. (1965). Distribution and sequence effects in judgment. *Journal of Experimental Psychology*, **69**, 450-459.
- PARKS, T. E., & KELLICUTT, M. H. (1968). The probability-matching decision rule in the visual discrimination of order. *Perception & Psychophysics*, **3**, 356-360.
- PETROV, A. A., DOSHER, B. A., & LU, Z.-L. (2005). The dynamics of perceptual learning: An incremental reweighting model. *Psychological Review*, **112**, 715-743.
- RANGANATH, C., COHEN, M. X., DAM, C., & D'ESPOSITO, M. (2004). Inferior temporal, prefrontal, and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *Journal of Neuroscience*, **24**, 3917-3925.
- RIEFER, D. M., & BATCHELDER, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review*, **95**, 318-339.
- RIGAS, P., & CASTRO-ALAMANCOS, M. A. (2007). Thalamocortical up states: Differential effects of intrinsic and extrinsic cortical inputs on persistent activity. *Journal of Neuroscience*, **27**, 4261-4272.
- ROELFSEMA, P. R. (2005). Elemental operations in vision. *Trends in Cognitive Sciences*, **9**, 226-233.
- ROSENBAUM, D. A., CARLSON, R. A., & GILMORE, R. O. (2001). Acquisition of intellectual and perceptual-motor skills. *Annual Review of Psychology*, **52**, 453-470.
- SAGI, D., & TANNE, D. (1994). Perceptual learning: Learning to see. *Current Opinion in Neurobiology*, **4**, 195-199.
- SCHULMAN, A. I., & GREENBERG, G. Z. (1970). Operating characteristics and a priori probability of the signal. *Perception & Psychophysics*, **8**, 317-320.
- SEITZ, A. R., NANEZ, J. E., HOLLOWAY, S. R., KOYAMA, S., & WATANABE, T. (2005). Seeing what is not there shows the costs of perceptual learning. *Proceedings of the National Academy of Sciences*, **102**, 9080-9085.
- SEITZ, A. [R.], & WATANABE, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Sciences*, **9**, 329-334.
- SIGMAN, M., & GILBERT, C. D. (2000). Learning to find a shape. *Nature Neuroscience*, **3**, 264-269.
- SIGMAN, M., PAN, H., YAN, Y., STERN, E., SILBERSWEIG, D., & GILBERT, C. D. (2005). Top-down reorganization of activity in the visual pathway after learning a shape identification task. *Neuron*, **46**, 823-835.
- SINHA, P., & POGGIO, T. (2002). High-level learning of early visual tasks. In M. Fahle & T. Poggio (Eds.), *Perceptual learning* (pp. 273-298). Cambridge, MA: MIT Press, Bradford Books.
- STURM, W., SCHMENK, B., FIMM, B., SPECHT, K., WEIS, S., THRON, A., & WILLMES, K. (2006). Spatial attention: More than intrinsic alerting? *Experimental Brain Research*, **171**, 16-25.
- SWETS, J. A., & SEWALL, S. T. (1963). Invariance of signal detectability over stages of practice and levels of motivation. *Journal of Experimental Psychology*, **66**, 120-126.
- TANNER, T. A., JR., HALLER, R. W., & ATKINSON, R. C. (1967). Signal recognition as influenced by presentation schedules. *Perception & Psychophysics*, **2**, 349-358.
- TANNER, T. A., JR., RAUK, J. A., & ATKINSON, R. C. (1970). Signal recognition as influenced by information feedback. *Journal of Mathematical Psychology*, **7**, 259-274.
- THOMAS, J. P. (1985). Detection and identification: How are they related? *Journal of the Optical Society of America A*, **2**, 1457-1467.
- THOMAS, J. P., GILLE, J., & BARKER, R. A. (1982). Simultaneous visual detection and identification: Theory and data. *Journal of the Optical Society of America*, **72**, 1642-1651.
- THOMAS, R. D. (2006). Processing time predictions of current models of perception in the classic additive factors paradigm. *Journal of Mathematical Psychology*, **50**, 441-455.
- TOWNSEND, J. T., & WENGER, M. J. (1996). *Evidence monitoring theory: A dynamic extension of general recognition theory and cognitive stochastic processing theory*. Paper presented at the 29th Annual Meeting of the Society for Mathematical Psychology, Chapel Hill, NC.
- TOWNSEND, J. T., & WENGER, M. J. (1999, November). *Evidence monitoring theory: Foundation and experimental application*. Paper presented at the 40th Annual Meeting of the Psychonomic Society, Los Angeles.
- VOGELS, R., & ORBAN, G. A. (1994). Activity of inferior temporal neurons during orientation discrimination with successively presented gratings. *Journal of Neurophysiology*, **71**, 1428-1451.
- WEIBULL, W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics*, **18**, 293-297.
- WENGER, M. J., & PAYNE, D. G. (1995). On the acquisition of mnemonic skill: Application of skilled memory theory. *Journal of Experimental Psychology: Applied*, **1**, 194-215.
- WENGER, M. J., & RASCHE, C. (2006). Perceptual learning in contrast detection: Presence and cost of shifts in response criteria. *Psychonomic Bulletin & Review*, **13**, 656-661.
- WILD, H. A., & BUSEY, T. A. (2004). Seeing faces in the noise: Stochastic activity in perceptual regions of the brain may influence the perception of ambiguous stimuli. *Psychonomic Bulletin & Review*, **11**, 475-481.
- YU, C., KLEIN, S. A., & LEVI, D. M. (2003). Cross- and iso-oriented surrounds modulate the contrast response function: The effect of surround contrast. *Journal of Vision*, **3**, 527-540.
- YU, C., KLEIN, S. A., & LEVI, D. M. (2004). Perceptual learning in contrast discrimination and the (minimal) role of context. *Journal of Vision*, **4**, 169-182.
- ZAR, J. H. (1999). *Biostatistical analysis* (4th ed.). Upper Saddle River, NJ: Prentice Hall.
- ZENGER, B., & SAGI, D. (2002). Plasticity of low-level visual networks. In M. Fahle & T. Poggio (Eds.), *Perceptual learning* (pp. 177-196). Cambridge, MA: MIT Press, Bradford Books.
- ZHANG, M., & ALLOWAY, K. D. (2004). Stimulus-induced intercolumnar synchronization of neuronal activity in rat barrel cortex: A laminar analysis. *Journal of Neurophysiology*, **92**, 1464-1478.

## NOTES

1. Here, and again in the General Discussion section, we use a set of mathematical and computational models to either illustrate a theoretical point or demonstrate the viability of a theoretical alternative. We in no way wish to suggest that any of these models represent comprehensive or complete accounts of perceptual learning or alternatives to truly comprehensive models of perceptual learning (as in, e.g., Lu & Doshier, 2008; Petrov et al., 2005).

2. Note that we could also vary the value of the guessing parameter  $g$ , either alone or with  $a$ . We chose to restrict consideration here to modeling identification bias assuming perceptual evidence sufficient for identification.

3. Note that the lack of false alarm rates in the staircase procedure is not an inherent limitation. Rather, it reflects standard implementations of that task. Kaernbach (1990) implemented a modification of the staircase procedure that does allow for trials on which the target is absent from both presentation intervals, and we present a similar design in Experiment 2.

4. We thank an anonymous reviewer for this suggestion.

**APPENDIX A**  
**Methods: Use of the Multinomial Model**

The multinomial models for the example in the text (see also Kaernbach, 1990, and Experiment 2 in the present study) assume three possible states for the stimulus and three possible responses (see Figure 1). The equations for each possible response, given each state of the stimulus, are provided in Table 1. Specific model parameters (Table A1) were varied to show how practice-related changes in processing can affect the psychometric function (and thus, the threshold estimate), the rate of identification errors, and the rate of false alarms. Stimulus energy was modeled as a variable ranging from 0.0 to 1.0 in all cases. Improvement in detection (Case 1 in Table A1) was modeled by increasing the value of the parameters  $D_A$  and  $D_B$  at each level of stimulus energy. Improvement in identification (Case 2 in Table A1) was modeled by increasing the value of the parameters  $I_A$  and  $I_B$ , with the specific values being constant across levels of stimulus energy. The liberal shift in detection bias (Case 3 in Table A1) was modeled by increasing the value of  $b$ , with specific values being constant across levels of stimulus energy. The shift to a bias for identifying a stimulus as being in state  $A$  (Case 4 in Table A1) was modeled by increasing the value of  $a$ , with specific values being constant across levels of stimulus energy.

**Table A1**  
**Parameters for the Multinomial Models and Their Values, Ranges, and Constraints**

Parameter	Practice Level	Case			
		1	2	3	4
$D_A$	Initial	.000:.075:.750	.000:.100:1.000	.000:.100:1.000	.000:.100:1.000
	Final	.000:.150:1.000	.000:.100:1.000	.000:.100:1.000	.000:.100:1.000
$D_B$	Both	= $D_A$	= $D_A$	= $D_A$	= $D_A$
$D_N$	Both	.100	.100	.100	.100
$I_A$	Initial	.950	.500	.900	.900
	Final	.950	.950	.900	.900
$I_B$	Both	= $I_A$	= $I_A$	= $I_A$	= $I_A$
$a$	Initial	.500	.500	.500	.500
	Final	.500	.500	.500	.900
$b$	Initial	.100	.100	.100	.100
	Final	.100	.100	.400	.100
$g$	Both	.500	.500	.500	.500

Note—Ranges are noted as initial:step size:final. Case 1 models an improvement in detection, Case 2 models an improvement in identification, Case 3 models a liberal shift in detection bias, and Case 4 models a shift in identification bias in favor of response  $A$ .

We imposed the following constraints in producing the data. The detection parameters  $D_A$  and  $D_B$  were constrained to be equal, as were the identification parameters  $I_A$  and  $I_B$ . The detection parameter for the situation in which a stimulus was absent,  $D_N$ , was constant across all cases, as was the guessing parameter  $g$ .

**APPENDIX B**  
**Methods: Numeric Simulation From R. D. Thomas (2006)**

Here, we present the form of the numeric simulations used to obtain the predictions in Figure 12. These simulations are drawn from R. D. Thomas's (2006) analysis of the implications of factorial additivity for three general models of perception: signal detection theory with the latency distance hypothesis, stochastic general recognition theory, and a random-walk model of exemplar classification. The second of these three models provided the basis for considering how the increase in false alarms and reductions in RTs for false alarms could emerge even if it is assumed that observers are responding optimally. We recapitulate (with minor notational differences) the critical equations from R. D. Thomas; interested readers should consult that source for a complete description and extended contrasts.

Let  $S_0$  and  $S_1$  be the stimulus states corresponding to the absence and presence (respectively) of contrast, and let  $R_0$  and  $R_1$  be the responses indicating (respectively) presence and absence of contrast. Assume that the perceptual input to the eventual decision process is dynamic (as in, e.g., Ashby, 2000), reflecting an ongoing sampling of the proximal stimulus energies, and that it can be modeled as a stochastic process  $x_t(t)$  given stimulus  $S_t$ . Assume that for any fixed value of  $t$  following the onset of the stimulus, the distribution of  $x_t(t)$  is



APPENDIX B (Continued)

Gaussian with parameters  $\mu_{Di}$  and  $\sigma_{Di}^2$ . At each instant  $t$ , the decision process computes a discriminant function  $y_i(t) = h[x_i(t)]$  in support of a decision

$$\text{if } y_i(t) = h[x_i(t)] \begin{cases} > A, \text{ respond } R_1 \\ < -B, \text{ respond } R_0 \\ \text{otherwise, continue sampling,} \end{cases} \quad (B1)$$

where  $A$  and  $B$  are criterion levels of perceptual evidence. Assuming further that  $\mu_i$  and  $\sigma_i^2$  are the mean and variance for the discriminant function  $y_i(t)$ , the conditional response probabilities follow from the theory of standard diffusion processes:

$$P(R_1 | S_i) = \frac{1 - \exp\left(-\frac{2\mu_i B}{\sigma_i^2}\right)}{1 - \exp\left[-\frac{2\mu_i(A+B)}{\sigma_i^2}\right]} \quad (B2)$$

and

$$P(R_0 | S_i) = 1 - P(R_1 | S_i). \quad (B3)$$

Similarly, the expected conditional decision times  $T_d$  are given by

$$E[T_d | R_1 \cap S_i] = \frac{1}{\mu_i} \left\{ (A+B) \frac{1 + \exp\left[-\frac{2\mu_i(A+B)}{\sigma_i^2}\right]}{1 - \exp\left[-\frac{2\mu_i(A+B)}{\sigma_i^2}\right]} - B \frac{1 + \exp\left[-\frac{2\mu_i B}{\sigma_i^2}\right]}{1 - \exp\left[-\frac{2\mu_i B}{\sigma_i^2}\right]} \right\} \quad (B4)$$

and

$$E[T_d | R_0 \cap S_i] = \frac{1}{\mu_i} \left\{ (A+B) \frac{1 + \exp\left[-\frac{2\mu_i(A+B)}{\sigma_i^2}\right]}{1 - \exp\left[-\frac{2\mu_i(A+B)}{\sigma_i^2}\right]} - A \frac{1 + \exp\left[-\frac{2\mu_i A}{\sigma_i^2}\right]}{1 - \exp\left[-\frac{2\mu_i A}{\sigma_i^2}\right]} \right\}. \quad (B5)$$

Finally, assume that the observer computes an expected value for the trial, on the basis of the current payoff structure and the cost of waiting for additional information (continuing to sample). If we let the payoffs for the four stimulus–response pairs be  $V_{rs} > 0$ , with cost  $-v_{rs}$  when  $r \neq s$ , and let the cost per unit time of waiting be  $u < 0$ , the expected value for the trial is given by

$$EV = P(S_1) \left\{ P(R_1 | S_1) [v_{11} - uE(T_d | R_1 \cap S_1)] + P(R_2 | S_1) [-v_{21} - uE(T_d | R_2 \cap S_1)] \right\} \\ + P(S_2) \left\{ P(R_1 | S_2) [-v_{12} - uE(T_d | R_1 \cap S_2)] + P(R_2 | S_2) [v_{22} - uE(T_d | R_2 \cap S_2)] \right\}, \quad (B6)$$

where  $P(S_i)$  is the overall probability of stimulus  $S_i$ . The parameter values used in these numeric simulations are listed in Table B1. Note that the values obtained for Equations B2–B6 are exact once parameter values have been specified.

**Table B1**  
Parameters and Values Used in the Simulations of the Dynamic Stochastic Signal Detection Analysis of R. D. Thomas (2006)

Parameter	Description	Value(s)
$P(S_1)$	Overall probability of stimulus $S_1$	.50
$v_{11}, v_{12}, v_{21}, v_{22}$	Payoffs/costs	1.00
$\mu_0$	Mean, stimulus $S_0$	1.00
$\Delta\mu$	Difference in means, $S_1$ and $S_0$	0.50 to 3.00 by 0.10
$\sigma$	Variance, $S_1$ and $S_0$	1.00
$u$	Cost per unit time of waiting	0.03 to 3.00 by 0.01

**APPENDIX C**  
**Simulation of Decay of Sustained Activation**

The model for the decay of sustained activation is based on empirical and theoretical investigations of the role of sustained activation in perception and attention (e.g., Amit & Brunel, 1995, 1997; Grent-'t-Jong & Woldorff, 2007; Muresan & Savin, 2007; Sturm et al., 2006). A general result from this work is that sustained activation of structured patterns of cortical responses (e.g., following learning) decays in a roughly exponential manner, with the time course of that decay ranging from milliseconds to seconds. We modeled the overall activity (spike rate in a cortical population)  $R(t)$  as a function of time  $t$  since the offset of a trained stimulus, using a simple exponential decay function with additive noise:

$$R(t) = R_0 e^{-\lambda t} + N(\mu, \sigma; t), \quad (C1)$$

where  $R_0$  is the level of activity at the time of stimulus offset,  $\lambda$  is the constant for the rate of decay, and  $N(\mu, \sigma; t)$  is an independent sample drawn from a Gaussian distribution at time  $t$ . All parameter values were set on the basis of published empirical and modeling data (e.g., Amit & Brunel, 1995, 1997; Muresan & Savin, 2007).

The hypothesis we wished to test was whether sustained activation could be a possible contributor to the increase in false alarm rates observed in our experimental work. To test this hypothesis, we perturbed the decay by adding 100 msec of Gaussian noise at a 1,000-msec delay from the offset of the stimulus. Following the offset of the noise, the value of  $R_0$  was reset to the value of  $R(t)$ , and the decay was allowed to proceed from that value. We then asked whether that event (approximating a blank display) would cause the decay function to be perturbed in a way that would be consistent with a false alarm. We asked this question by comparing the activation function  $R(t)$  with an overall criterion  $\gamma$ :  $R(t) > \gamma$  was taken to indicate that the activation was consistent with the presence of the test stimulus. Table C1 lists the values of all of the parameters of the simulation that produced the data summarized in Figure 13.

**Table C1**  
**Parameters and Values Used in the Simulations**  
**of the Decay of Sustained Activation**

Parameter	Description	Value(s)
$R_0$	Initial population level of activity	500
$\lambda$	Decay rate constant	0.0007
$\mu$	Mean, noise distribution	0
$\sigma$	Standard deviation, noise distribution	20
$n$	Total number of simulation trials	500

(Manuscript received December 28, 2007;  
revision accepted for publication April 7, 2008.)