The improvement of contrast sensitivity with practice is not compatible with a sensory threshold account

JOSHUA A. SOLOMON,[1*] CHRISTOPHER W. TYLER[1,]

[1]*Centre for Applied Vision Research, City University London, London EC1V 0HB, United Kingdom*
**Corresponding author: J.A.Solomon@city.ac.uk*

Received XX Month XXXX; revised XX Month, XXXX; accepted XX Month XXXX; posted XX Month XXXX (Doc. ID XXXXX); published XX Month XXXX

In forced-choice detection, incorrect responses are routinely ascribed to internal noise, because experienced psychophysical observers do not act as if they have a sensory threshold, below which all perceived intensities would be identical. To determine whether inexperienced observers have sensory thresholds, we examined psychometric functions (percent correct vs log contrast) for detection and detection in full-screen, dynamic visual noise. Over 5 days, neither type of psychometric function changed shape, but both shifted leftwards, indicating increased sensitivity. These results are not consistent with a lowered sensory threshold, which would decrease psychometric slope. Our results can be understood within the context of Dosher and Lu's (2000) 'stochastic' Perceptual Template Model, augmented to allow intrinsic uncertainty. Specifically, our results are consistent with a combination of reduced internal additive noise and improved filtering of external noise. © 2016 Optical Society of America

OCIS codes: *(330.1880) Detection; (330.4060) Vision modeling; (330.5510) Psychophysics.*

http://dx.doi.org/10.1364/AO.99.099999

1. INTRODUCTION

When measuring visual sensitivity, most psychophysicists adopt the *m*-alternative, forced-choice (*m*AFC) paradigm, because each response of the observer can be classified as either right or wrong. (Appendix A provides a list of the main symbols used in the text, and their meanings.) For invisibly faint targets, the probability of a correct choice is necessarily $1/m$. For very intense targets, 1 is theoretically the maximum probability correct, although most human observers occasionally lapse. That is, they choose incorrectly, even when the correct choice is unquestionably obvious [1].

Psychophysical literature offers two explanations for errors with faint targets. One possibility is that the observer didn't see anything that could have been the stimulus, was forced to guess and guessed wrong. This possibility can be formalized with the notion of a "sensory threshold," which is defined as the hypothetically least intense perceptual experience. Many studies have shown that a sensory threshold cannot explain all *m*AFC detection errors. For example, when $m > 2$, erroneous 'first-choice' responses can be followed by correct 'second-choice' responses with a probability greater than $1/(m-1)$ [2, 3]. Consequently, the alternative explanation is preferred for at least some *m*AFC detection errors. Signal-detection theorists [e.g., 4] call it "noise." Such noise is a stochastic process, which causes observers to see things that aren't there.

Although observer experience isn't always described in psychophysical literature, our own observers are typically encouraged to practice their visual tasks until they feel comfortable performing them. Only then do we begin to collect data. Validation of this *modus operandi* can be inferred from studies of "perceptual learning" [e.g., 5], which show that sensitivity (in this case, the reciprocal of the detection threshold) can indeed increase with practice [6].

We wondered why performance improves with practice. One hypothesis is that unpracticed observers act as if they have a sensory threshold, and consequently they do not see things that aren't there. Practice would affect sensitivity because it lowered or abolished the sensory threshold.

Alternative explanations for the effect of practice on detection threshold include a reduction in intrinsic uncertainty and an increase in signal-to-noise ratio. In Section 2 these ideas will be formalized within the context of a stochastic model, based on signal-detection theory. In Section 3 we will demonstrate that the effect of a lowered sensory threshold is very similar to the effect of reduced intrinsic uncertainty on psychometric functions for detection, making the psychometric slopes shallower. (The effect of elevated signal-to-noise ratios is different, improving sensitivity without affecting the psychometric slopes.)

To tease these alternatives apart, we measured psychometric functions for detection in the presence (as well as in the absence) of randomly generated texture. Such "external noise" can reduce intrinsic uncertainty when it selectively stimulates detection mechanisms. It is incapable of affecting intrinsic uncertainty when it has a broad spatial bandwidth, a broad temporal bandwidth, a wide spatial extent, and a long temporal extent, relative to the target [7–9]. Nonetheless, external noise can unquestionably elevate detection threshold above the range where a sensory threshold might operate. In other words, observers never see nothing; they see the external noise.

2. STOCHASTIC MODEL

2A. Signal-detection theory

Our hypotheses may be made concrete within the context of signal-detection theory [4]. In this model, the probability correct in any *m*AFC task can be described as

$$\Psi = (1-\delta)\psi + \delta(1-\psi)/(m-1), \quad \textbf{(1)}$$

where δ represents the lapse rate and

$$\psi = \int_{-\infty}^{\infty}\left[F_{\mathbf{N}}(x)\right]^{m-1} F_{\mathbf{S}}'(x)\,dx. \quad \textbf{(2)}$$

In Eq. 2, $F_{\mathbf{N}}(x)$ is the cumulative distribution (CDF) for the maximum signal $\mathbf{N}$, arising from each non-target and $F_{\mathbf{S}}'(x)$ is the derivative of the CDF (i.e., it is the density) for the maximum signal $\mathbf{S}$, arising from the target.

2B. Intrinsic uncertainty

In the absence of a sensory threshold, intrinsic uncertainty theory [10] would specify CDFs for the target (where $\mathbf{X}=\mathbf{S}$) and non-targets (where $\mathbf{X}=\mathbf{N}$) as the product of M CDFs, such that

$$F_{\mathbf{X}}(x) = \left[F_I(x)\right]^{M-K}\left[F_X(x)\right]^{K}. \quad \textbf{(3)}$$

In Eq. 3, K represents the number of "relevant micro-analyzers" [11] with equal sensitivity to the stimuli, $M-K$ represents the number of "irrelevant" micro-analyzers with no sensitivity to the stimuli. Micro-analyzers are conveniently assumed to be individual neurons or pools of neurons with similar receptive fields. The output from each micro-analyzer contains an independent sample of noise that can be assumed to be Gaussian. Consequently, micro-analyzer output can be described in terms of the Gaussian CDF:

$$F_X(x) = G(x;\mu_X,\sigma_X) = \frac{1}{2}\left[1+\operatorname{erf}\left(\frac{x-\mu_X}{\sigma_X\sqrt{2}}\right)\right]. \quad \textbf{(4)}$$

2C. Sensory threshold

Eq. 3 represents the limit for a vanishingly small sensory threshold, i.e., as $c \to -\infty$. For a finite sensory threshold c, we must construct a more general expression:

$$F_{\mathbf{X}}(x) = \begin{cases} \left[F_I(x)\right]^{M-K}\left[F_X(x)\right]^{K} & x > c \\ \left[F_I(c)\right]^{M-K}\left[F_X(x)\right]^{K} & x = c \\ 0 & \text{otherwise} \end{cases}. \quad \textbf{(5)}$$

Random variables *N, S,* and *I* are described in Section 2D, below.

2D. Power-law transduction

We have adapted Dosher and Lu's [12] parameterization for the relationships between μ_X, σ_X, and the stimulus. In the absence of external noise, the expected output of each micro-analyzer is a power function of stimulus contrast. Specifically, the expected signal is $\mu_X = (bc_X)^{\gamma}$, for $X = S, N,$ and I, where c_S is the contrast of the target, c_N is the contrast of the non-targets, $c_I = 0$ can be considered the expected input to the irrelevant micro-analyzers, and b and γ are arbitrary constants, greater than 0.

When external, Gaussian noise is present with amplitude N_{ext}, micro-analyzer outputs become doubly stochastic. Let θ denote a sample (of size 1) from a Gaussian random variable with zero mean and variance N_{ext}^2. Expected micro-analyzer output can be written as a function of this random sample:

$$\mu_X = \operatorname{sgn}\left(bc_X + A_{\text{f}}\theta\right)\left|bc_X + A_{\text{f}}\theta\right|^{\gamma}. \quad \textbf{(6)}$$

2E. Multiplicative noise and additive noise

Output variances can be described as the sum of two independent sources of internal noise. "Multiplicative noise" grows with the expected signal μ_X; "additive noise" does not. Since multiplicative noise and additive noise are assumed to be independent, their variances sum:

$$\sigma_X^2 = \left(A_{\text{m}}N_{\text{m}}\right)^2\left|\mu_X\right|^{2\gamma_2/\gamma} + \left(A_{\text{a}}N_{\text{a}}\right)^2, \quad \textbf{(7)}$$

where N_{a} is the standard deviation of the internal, additive noise (an arbitrary constant greater than 0) and N_{m} is also an arbitrary constant, greater than or equal to 0. The exponent γ_2 will be discussed below.

The three coefficients, A_{f}, A_{m}, and A_{a}, can be set equal to 1 with no loss of generality. In fact, we enforce this constraint when fitting the model to data from the last day of testing. On days prior to the last day of testing, best-fitting values for any (or all) of these coefficients thus might be greater than 1 (Our use of these three coefficients is opposite to that of Dosher and Lu [13], who constrained them to equal 1 when fitting the data from the first testing session in unpracticed observers.). If $A_{\text{f}} > 1$, it would imply an effect of practice on the amount of external noise passed by each micro-analyzer (e.g., via improved filtering [e.g., 13]); if $A_{\text{m}} > 1$, it would imply an effect of practice on the amount of multiplicative internal noise; and if $A_{\text{a}} > 1$, it would imply an effect of practice on the amount of additive internal noise.

2F. Contrast gain control

Psychophysical models of detection in the presence of "pattern masks" (e.g., sinusoidal luminance gratings) typically feature sigmoidal transduction (e.g., acceleration for low contrasts and compression for high contrasts) and eschew multiplicative noise. In the absence of external noise, these "contrast-gain-control" models are formally equivalent to models like ours, with power-law transduction and multiplicative noise [14]. For example, Legge and Foley's [15] popular non-linear-transducer model is equivalent to our model when γ = 2.4, γ_2 = 2, M = 1, and $c \ll 0$.

The exponent γ_2 must be slightly smaller than the exponent γ if our model (like Legge and Foley's) is to successfully fit supra-threshold contrast-discrimination data (i.e., "dipper" handles, see [16]). However, such data are not considered either in our study or that by Dosher and

Lu [13]. Consequently, in all our subsequent modelling, we constrain the two exponents to be equal: $\gamma_2 = \gamma$.

2G. Implementation

In previous research [e.g., 9, 12, 17] Monte Carlo simulations were required to estimate model performances when $\gamma \neq 1$ and $N_{\text{ext}} \neq 0$. All simulations reported in this paper, however, were computed analytically, in the sense that Monte Carlo simulations were not necessary. Where necessary, we substituted compound CDFs for the Gaussian CDFs in Eq. 4, i.e.:

$$F_X(x) = \int_{-\infty}^{x} du \int_{-\infty}^{\infty} G'(u;\mu_X,\sigma_X)\, G'(\theta;0,N_{\text{ext}})\, d\theta, \qquad \textbf{(8)}$$

where $G'(x;\mu,\sigma)$ is the density defined by the derivative of the Gaussian CDF defined in Eq. 4. Where necessary, these integrals were computed numerically.

3. ILLUSTRATIVE SIMULATIONS

Fig. 1 illustrates how a sensory threshold is expected to affect psychometric functions for 4AFC detection. Densities for target S and non-target N are shown for low (panel a) and high (panel b) values of the sensory threshold. By definition, non-targets in a detection task have zero contrast, i.e., $c_N = 0$. Panels (c) and (d) show how the probability correct Ψ varies with target contrast, when the latter is expressed in decibels, i.e., Strength (dB) $= 20\log_{10}(c_S)$.

The psychometric function in Fig. 1(c) was produced by computing the model's expected performance in 100 trials with each of 23 different target contrasts. The parameter values used for this simulation are the "baseline" values used for all of the other simulations described in this section, except where indicated otherwise. They were: $\delta = 0.01$, $K = 1$, $M = 1$, $c = -0.09$, $b = 3$, $\gamma = 1$, $N_{\text{m}} = 0$, $N_{\text{a}} = 0.03$, $A_{\text{f}} = 1$, and $A_{\text{a}} = 1$. In this case, where its value is $c = \mu_N - 3\sigma_N$, the sensory threshold has virtually no effect; i.e., the model performs identically when $c = -\infty$.

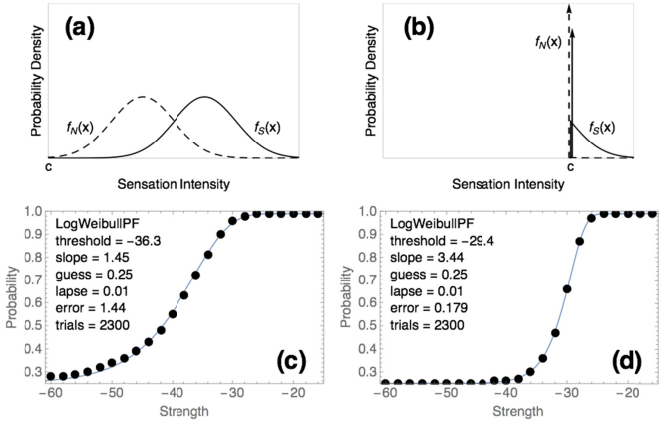

Fig. 1. Representative densities (with $\mu_S = u_N + 2\sigma_N$) for 4AFC detection target S and non-target N (a, b) and corresponding psychometric functions (c, d) with a high sensory threshold ($c = u_N + 3\sigma_N$; b, d) and a low sensory threshold ($c = u_N - 3\sigma_N$; a, c). Arrows in (b) represent delta functions, indicating high probability density at the sensory threshold c. Insets in panels (c) and (d) describe best-fitting Weibull distributions (smooth curves) to the model-derived accuracies (filled circles).

A comparison between Figs. 1(c) and 1(d) shows what happens when the sensory threshold c is increased. The psychometric function for detection shifts rightwards and gets steeper. To quantify these changes, we (maximum-likelihood) fit a cumulative Weibull distribution to the data in each panel. Its general form can be described as

$$\Psi = \frac{1}{m} + \left(1 - \frac{1}{m} - \delta\right)\left[1 - \exp\left(\left(\frac{c_N - c_S}{\alpha}\right)^{\beta}\right)\right]. \qquad \textbf{(9)}$$

In Eq. 9, the Weibull parameter α is typically considered to be the contrast threshold. In this paper threshold is given in decibels, i.e., $20\log_{10}(\alpha)$. The Weibull parameter β is typically used to specify the psychometric slope [18]. Best-fitting values for threshold and slope are shown in Figs. 1(c) and (d). Fixed values were assigned for the guess rate $1/m = 0.25$ and lapse rate $\delta = 0.01$. Finally, note that, although the maximum-likelihood fit is good, it is not perfect, i.e., the model produces non-Weibull psychometric functions. The error may be specified by a goodness-of-fit metric, proportional to the logarithm of maximum likelihood [19].

Fig. 2 illustrates the outcomes of four alternative models for the effect of practice on psychometric functions for detection (i.e., when $N_{\text{ext}} = 0$) and detection in noise ($N_{\text{ext}} = 0.2$). Each curve in the figure illustrates the effect of a single parameter, whose value decreased by a fixed amount, on each of 5 days' simulated practice. As noted in Section 1, the effect of a lowered sensory threshold is very similar to the effect of reduced intrinsic uncertainty on psychometric functions for detection: threshold and slope both drop. This is true even when intrinsic uncertainty remains high, as can be seen in Appendix C.

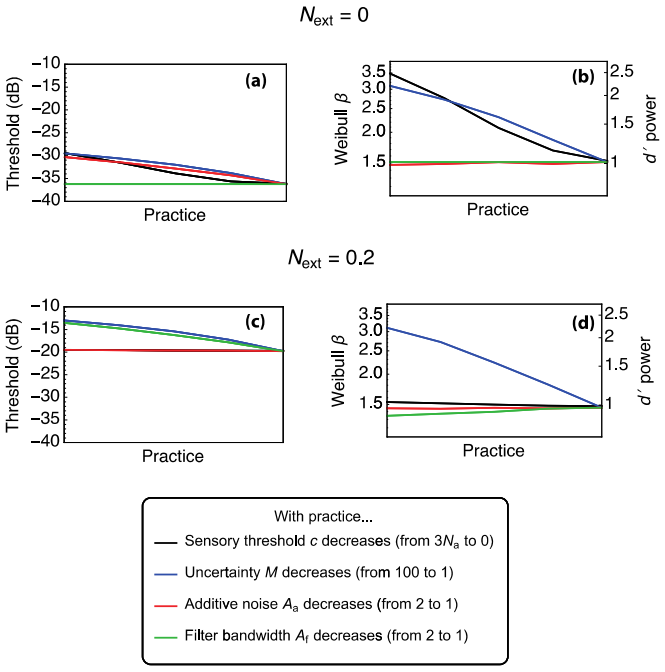

Fig. 2. (Online version in color). Values for threshold (a, c) and slope (b, d) from Weibull fits to simulated data from detection (a, b) and detection-in-noise (c, d) conditions. Note that high levels of uncertainty are required for detection-in-noise to have steep psychometric functions (i.e., high values of Weibull β or the approximately equivalent d' power).

External noise can elevate detection thresholds above the range where a sensory threshold might operate. Consequently, our simulations show no effect of a lowered sensory threshold on psychometric functions for detection in noise. On the other hand, a reduction of intrinsic uncertainty affects the psychometric function for detection in noise in the same way that it affects the psychometric function for detection: threshold and slope both drop.

In the absence of any sensory threshold or multiplicative noise, an increase in gain (parameter b) has the same effect on the model's psychometric function for detection as a decrease in additive noise (the product of parameters A_a and N_a). Fig. 2 illustrates this effect: threshold falls but slope does not. This differentiates an increase in the signal-to-noise ratio from a lowering of sensory threshold and a reduction of intrinsic uncertainty. A more thorough study of model behavior appears below.

4. METHODS

This experiment was approved by The National Research Ethics Service. Ten observers were recruited. Those who routinely wore corrective lenses for computer use wore them during this experiment. All ten observers claimed to have no previous experience with visual psychophysics, outside of routine visits to an optometrist or optician. Each observer provided written consent to participate in a non-invasive psychophysical experiment.

The methods were very similar to those used by Solomon [3], whose experienced observers did not produce evidence of a sensory threshold. A MacBook Pro running the PsychToolbox [20, 21] was used for stimulus generation and response recording. (Software will be made available upon request.) Stimuli were displayed on a 19-inch Sony Trinitron with a refresh rate of 120 Hz. True 14-bit grey-scale resolution was achieved using Cambridge Research Systems' Bits++ box. Luminance was linearized within 0.7%. Maximum and minimum luminances were 157.4 and 2.2 cd/m^2, respectively. The monitor's background luminance was set to the midpoint of these values, and the rest of the room was dark. Viewing distance was approximately 0.70 m. At this distance, there were 32 pixels per degree of visual angle. Observers were required to maintain fixation on the fixation point throughout each run.

Targets and non-targets were horizontal, cosine-phase Gabor patterns whose wavelength and spatial spread were $\lambda = 0.25^\circ$ and $\sigma = 0.18^\circ$, respectively. One Gabor appeared in each of four positions, marked by a black square (see Fig. 3). These black squares were visible throughout the experiment. Their centers formed a large ($5.6^\circ \times 5.6^\circ$), notional square centered on a black fixation spot. External dynamic noise, when present, was created by adding an independently determined, random intensity to each pixel on the 16° × 21° screen (except those depicting the black squares and fixation spot) on every refresh. Random intensities had zero mean and an r.m.s. contrast of 0.2, giving the noise a spectral density of 3×10^{-7} deg^2 s.

Four types of trial were randomly interleaved: detection, discrimination, detection in noise, and discrimination in noise. Each trial was initiated with a key press. Observers were allowed to respond (with another key press) after 1.0 s. On noise trials, dynamic noise was presented continuously throughout the trial. In the middle of each trial (i.e., between 0.41 and 0.59 s after the key press) the four Gabors (three non-targets and one target) were presented and the fixation spot removed for 0.18 s. On detection trials, the non-targets had zero contrast ($c_N = 0$). On discrimination trials, their Michelson contrast was $c_N = 0.04$. Target contrast was always higher (see below). Feedback in the form of a brief tone followed each incorrect response.

The discrimination trials did not promise any additional power for differentiating between candidate mechanisms of perceptual learning. Their inclusion was designed to encourage observers to compare all four alternatives (in "sensory trace mode" [22]), as our modeling assumes. Without a "roving" pedestal like this, some observers have a tendency to operate in "context-coding mode," in which each (or possibly only some) of the four alternatives is compared to an internal template. Thus, if one alternative provided a sufficiently close match to the internal template, the observer wouldn't need to examine any of the other alternatives when making a decision.

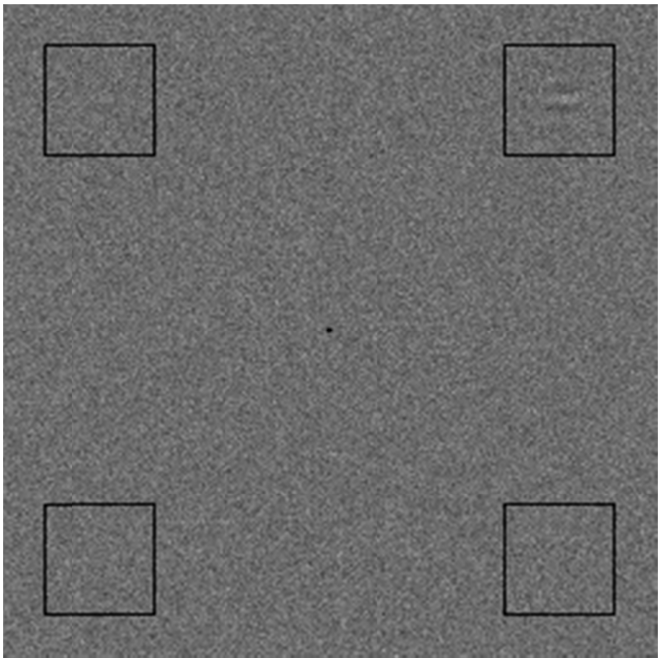

Fig. 3. Stimulus geometry used in the experiment. Gabor and noise contrasts have been adjusted for illustrative purposes.

To quickly and efficiently obtain estimates of threshold and slope while observers practiced detection, we used the adaptive psychometric staircase QUEST [23]. Prior to each trial, QUEST produced a (mean *a posteriori*) estimate for the observer's threshold [*A priori* distributions were Gaussian, with a mean of –27 dB and a standard deviation of 60 dB.]. This estimate was based on the conventional assumption of a Weibull psychometric function (Eq. 9), with the following parameters fixed: Weibull $\beta = 2$, $m = 4$, and $\delta = 0.01$. Of course, we did not expect these parameter values to represent all of our empirical psychometric functions. In particular, the slope (Weibull $\beta = 2$) was selected as a compromise, given the various simulations illustrated in Fig. 2.

As a means of accurately estimating the frequency of finger errors δ, the target was presented with a Michelson contrast of $c_S = 0.14$ (i.e., –17 dB) on one-ninth of the trials without noise. The precisions of our estimates for psychometric threshold and slope depend upon the extent to which our target contrasts elicit response accuracies that straddle the mid-point between psychometric floor (i.e., $1/m$) and psychometric ceiling ($1 - \delta$). Consequently, on the half the remaining trials, the target was presented 3.8 dB below QUEST's estimated threshold; on the other half, the target was presented 0.5 dB above QUEST's estimated threshold. These values were selected because they

correspond to response accuracies of $\Psi = 0.5$ and $\Psi = 0.75$, respectively, given the fixed parameter values described above.

Each observer performed four blocks of 88 trials each, on each of 5 consecutive days. The QUEST staircases were initialized at the beginning of each day.

5. RESULTS AND DISCUSSION

One observer was excused from the study after 2 days, because he was unable to attain a response accuracy greater than 50% correct with –17-dB targets in either of the "no-noise" conditions. From the psychometric data generated by each of the nine remaining observers, we obtained 20 (five days × four types of trial: detection, discrimination, detection in noise, and discrimination in noise) maximum-likelihood estimates of threshold and slope. Examples from a representative observer appear in Appendix D.

5A. Detection (in the absence of noise)

For each observer in each condition, we regressed maximum-likelihood estimates of threshold (in dB) and slope (Weibull β) against the day of testing. For each observer and each day, lapse rates (δ) were fixed at the values determined from the high-contrast/no-noise trials (or 0.01; whichever was larger). Without exception, the line best fit to every observer's detection thresholds in the absence of noise had a negative gradient (see Fig. 4a). Across the group, the mean and standard error of the fitted slopes were –0.72 and 0.15 dB/day, respectively. Thus, there can be no question that our methods were sufficient to elicit a significant facilitatory effect of practice on contrast sensitivity, $t(8) = -4.93$, $p < 0.001$.

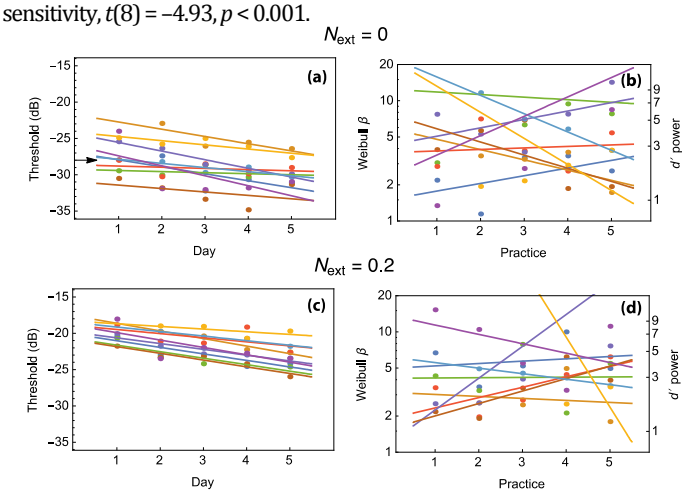

Fig. 4 (Online version in color). Values for threshold (a, c) and slope (b, d) from Weibull fits to empirical data from detection (a, b) and detection-in-noise (c, d) conditions. Each symbol represents the performance of a single observer on a single day. A regression line (color-coded online) has been fit to the data from each observer within each panel. The arrow in panel (a) indicates the (peri-threshold) pedestal amplitude that was used in all discrimination trials.

Psychometric slopes for detection, on the other hand, showed no systematic variation with day of testing, $t(8) = -0.77$, $p = 0.46$ (see Fig. 4b). The mean gradient was –0.07, with a standard error of 0.09. Recall that our main hypothesis was that practice would lower or abolish a sensory threshold. Section 3 demonstrated that a lowered sensory threshold [24] would manifest not only as a decrease in the detection threshold, but also as a decrease in psychometric slope. If this hypothesis were correct, then we would expect a (positive) correlation (across days of testing) between detection threshold and psychometric slope. For only three observers was there any hint (i.e., with Pearson's $r > 0.1$) of this correlation. It is perhaps noteworthy that these were also the only three observers for whom, on day 1, estimated lapse rates were too high (more than 25%) for adequate constraint of parameter values. In only one other instance (MP, day 2) did lapse rate exceed this value (i.e., $\delta > 0.25$). On the basis of these results, the data contain little evidence in support of our main hypothesis. Instead, our regression analysis of the psychometric parameters for detection seems more consistent with an effect of practice on internal, additive noise (see Fig. 2): thresholds dropped consistently over days, but slopes remained high.

5B. Detection in noise

As with the detection data, without exception, the line best fit to every observer's detection-in-noise thresholds had a negative gradient (see Fig. 4c). The mean and standard error were –0.79 and 0.08 dB/day, respectively. Thus, there can be no question that detection-in-noise improved significantly with practice, $t(8) = -10.48$, $p < 10^{-5}$, just as it did in the absence of noise.

Also similar to the detection data, the detection-in-noise data contained no evidence for a systematic effect of practice on psychometric slope, $t(8) = -0.35$, $p = 0.37$ (see Fig. 4d). Regression analysis indicated a mean gradient of –0.06, with a standard error of 0.18. Although these results are qualitatively similar to our detection results (i.e., thresholds dropped, but slopes did not) they cannot similarly be explained by a reduction in additive noise (see Fig. 2), which must be negligible in comparison to our massively suprathreshold external noise. Instead, they are consistent only with a reduction in the effect of this external noise on detection. In other words, practice seems to reduce the amount of external noise that is passed by the theoretical micro-analyzers.

Day-by-day averages appear, along with a fit of the model, in Fig. 5. These averages illustrate the aforementioned effects of practice on threshold, as well as the lack of a systematic effect of practice on psychometric slope. Visual comparison reveals two notable differences between Figs. 2 and 5. For one thing, the effect of external noise on our observers' contrast thresholds (an elevation of ~7 dB, on average) is much smaller than that predicted by our model (an elevation of ~16 dB), given the "baseline" parameter values used for Fig. 2. The other thing is that the psychometric slopes are much steeper than those produced by the model with these baseline parameter values.

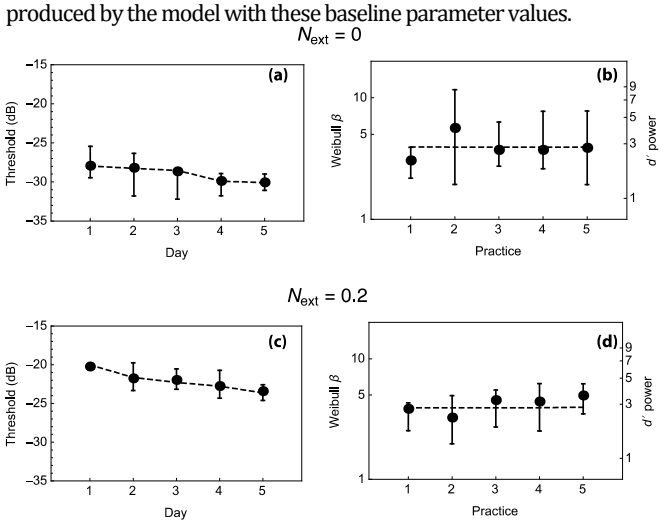

Fig. 5. Values for threshold (a, c) and slope (b, d) from Weibull fits to empirical data from detection (a, b) and detection-in-noise (c, d) conditions. Each symbol illustrates the median value, across 9 observers., ±1 quartile. Dashed curves illustrate a model fit. Parameter values were: $\delta = 0.01$, $K = 1$, $M = 870$

, $c = -\infty$, $b = 14$, $\gamma = 1$, $N_m = 0$, $N_a = 0.110$. A_f decreased from 1.58 to 1 (i.e., between day 1 and day 5) and A_a decreased from 1.28 to 1.

The one parameter in our model that governs the relationship between external noise contrast and threshold elevation is N_a. When this (additive) component of internal noise is large, then a correspondingly large amount of external noise is necessary for threshold elevation. Clearly, the baseline value of 0.03 was not sufficiently large. The model's threshold elevations were more consistent with those we obtained empirically when this parameter value was increased to $N_a = 0.110$.

The other discrepancy between our initial simulations and empirical results is perhaps more noteworthy. On the basis of previous research [7] with a slightly different paradigm (two temporally offset stimulus intervals in foveal view, rather than four spatially offset stimulus intervals in parafoveal view) we expected to record shallower psychometric functions for detection in our dynamic noise, which was both spatially and temporally extended, relative to the target and non-target Gabor patterns. Indeed, that finding has been recognized [e.g. 9, 16] as one of the chief indictments against attributions of steep psychometric slopes (in the absence of noise) to intrinsic uncertainty. Consequently, we attempted to replicate those historical methods as closely as possible in a supplementary experiment. Nonetheless, we were unable to reproduce the finding. Psychometric slopes for author J.A.S. remained high, even in high levels of full-screen, dynamic noise.

5C. High psychometric slopes for detection in noise

The model described in Section 2 contains three parameters that affect psychometric slopes for detection in noise. Slope increases with uncertainty *M* (as is illustrated in Fig. 2). It also increases with the exponent of the power-law transducer γ and decreases with the product $A_m N_m$. According to Birdsall's Theorem [25], when $\gamma > 1$, psychometric slope must decrease as the contrast of external noise increases. Our data contain no hint of this decrease, and consequently do not support a model containing non-linear transduction.

Without non-linear transduction, the only property that can account for the steep psychometric functions in noise is intrinsic uncertainty, which must be in the range of $M \approx 1{,}000$ to best fit (i.e., with minimum squared errors) the data. That number may seem to be implausibly high. However, Fig. 6 shows how poorly the model fits our psychometric slopes for detection in noise, when allowing non-linear transduction but constraining *M* to smaller values [26]. Without uncertainty (i.e., when $M = 1$) the model's best fit is very poor indeed. For a reasonable fit, with Weibull β close to the median value obtained in our detection-in-noise experiment, approximately micro-analyzers are required at each stimulus position, i.e. $M \approx 1000$.

5D. Contrast discrimination

As previously noted, the discrimination trials did not promise any additional power for diagnosing between candidate mechanisms of perceptual learning. Consequently, we did not analyze the data from our discrimination trials, other than to confirm what many previous authors (e.g., [27, 28]) have reported: when non-target contrasts are near the detection threshold (as in our experiment), discrimination thresholds are lower than detection thresholds and psychometric slopes for discrimination are lower than those for detection.

Signal-detection theory explains these results in one (or possibly a combination) of three ways. A sigmoidal transducer function (e.g., 27, 28) may suffice. If the transducer is merely a power-function (as in Eq. 6), then multiplicative noise must be non-negligible (i.e., $N_m > 0$; otherwise psychometric functions will be too steep [3]). The final alternative is to have an appreciable amount of intrinsic uncertainty (i.e., $M >> 1$ [10]). Since an appreciable amount of intrinsic uncertainty was required to fit our detection-in-noise results, we had little reason to explore model performance with non-zero multiplicative noise.

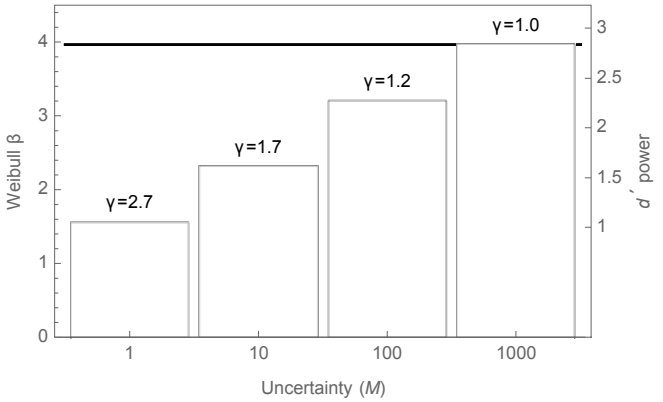

Fig. 6. Model performances with best-fitting values of *γ* (indicated on the figure), *b*, and N_a, for various fixed values of uncertainty *M*. All other model parameters were fixed at the baseline values, described in Section 3. The solid horizontal line shows the median values for Weibull β obtained in detection experiments. (In noise, the median was 4.2; in the absence of noise, it was 3.9. Geometric means were greater: 5.3 and 5.0, respectively. Arithmetic means were even greater.)

On the other hand, when non-target contrasts are *above* threshold, discrimination thresholds tend to be much *greater* than detection thresholds. There are only two ways of achieving this result within the context of signal-detection theory: either compressive transduction (or a sigmoidal combination of expansion and compression) or non-negligible multiplicative noise. Such options are beyond the scope of this project, but they are discussed at length elsewhere (e.g., [16]).

5E. Comparison with previous work

Yu, Klein, and Levi [5] found no effect of practice on contrast discrimination when non-target contrast was randomized. Low-contrast (and zero-contrast) non-targets were excluded from that experiment. However, they did report an effect of practice on contrast detection (i.e., when non-targets had zero contrast) in a situation amenable to context-coding mode, as discussed in Section 4. In this case, thresholds dropped at a rate of 1.4 dB/day, on average. Since this effect is almost twice as large as that reported in Section 5A, we cannot be confident that our results stem from the same sources of perceptual learning.

Dosher and Lu [12, 29] examined the effect of practice on contrast threshold for discriminating between two orientations of Gabor pattern, ±12° with respect to vertical. This task is fundamentally different from detection, however. For one thing, decisions are not governed by the activity in mechanisms best tuned to the candidate targets [30]. Consequently, it seems likely that contrast thresholds for discriminating between these orientations exceed contrast thresholds for mere detection, and thus could not reflect any possible role of a sensory threshold.

Nonetheless, our results are qualitatively similar to those reported by Dosher and Lu [12, 29]: contrast thresholds dropped with practice, in both the presence and absence of external noise, and psychometric slopes (inferred from Dosher and Lu's "threshold ratios") were unaffected by practice. The changes in signal processing that underlie the effects that we report are consequently comparable to those that

Dosher and Lu reported: a reduction in internal, additive noise and more effective filtering of external noise.

Our results are also consistent with Lu and Dosher's [31] study of contrast detection in experienced observers. Specifically, both our study and theirs indicate that psychometric slopes ("threshold ratios") are invariant with external noise contrast and are inconsistent with any model (e.g., a simple amplifier without a sensory threshold, intrinsic uncertainty, or non-linear transduction) in which detectability (specifically, the *d'* metric of signal-detection theory [4]) is proportional to signal contrast. Lu and Dosher fit their results with an approximation to the model described in Section 2, which they called "analytic PTM." Subsequently, Klein and Levi [9] noted that analytic PTM cannot produce the decrease in psychometric slope that Birdsall proved must accompany the introduction of external noise when $\gamma > 1$. Since neither Lu and Dosher's data nor ours contain any hint of this decrease, both sets of data support a model with linear transduction (i.e., $\gamma = 1$), and intrinsic uncertainty (i.e., $M > 1$). Within the context of our stochastic model, this uncertainty is the only feature capable of producing similarly steep psychometric functions for detection both in the presence and absence of external noise.

6. CONCLUSIONS AND CAVEATS

6A. Summary of findings

We sought evidence in favor of a lowering of sensory threshold with practice in a detection task, but found no such evidence. Instead, within the framework of signal-detection theory, accounting for the effects of practice required decreases in both additive internal noise and the external noise that passes through each micro-analyzer. For this model to best fit our data, intrinsic uncertainty had to be high and undiminished with practice.

6B. Physiological instantiation

Our results do not directly address the question of how practice reduces additive noise or how it improves micro-analyzer tuning, but these changes could be instantiated physiologically. If each micro-analyzer can be considered a pool of noisy neurons with similar (but not identical) receptive fields, then practice may effectively prune away some neurons from each micro-analyzer. Consequently, the micro-analyzer would have less intrinsic noise (i.e., the product of A_a and N_a would decrease) and it would process less external noise (i.e., the product of A_f and N_{ext} would decrease).

Of course, if ours were a physiological model, we would specify how A_a and A_f depend on the number of neurons in each micro-analyzer, and we might be able to explain all the effects of practice with one parameter. However, Lu and Dosher [32] reported that the effect of practice on orientation identification (in the absence of external noise) for foveal stimuli was negligible compared to its effect on orientation identification in noise. Thus, a different mechanism seems to be responsible for practice effects in the absence of noise, when they are found (e.g., outside the fovea).

6C. Conceptual issues with signal-detection theory

1. Attention

It seems unlikely that spatially focused attention is necessary for practice to facilitate detection. We know this because observers were required to compare the contents in multiple regions of the display, all at the same time. Perhaps micro-analyzer tuning and output variance ordinarily (i.e., in unpracticed observers) fluctuate with something more diffuse than spatial attention. Maybe "arousal" would be a better term. With practice, observers could converge on the appropriate level of this arousal for optimum task performance.

Another piece of evidence against an attentional explanation for these practice effects is their incompatibility with uncertainty reduction. Although it isn't strictly necessary to invoke attention in models of intrinsic uncertainty, it does seem implausible that observer performances would be affected by Gabor patterns that appeared at wildly inappropriate times (e.g., at the end, rather than in the middle of a trial) or positions (e.g., anywhere in the field other than in one of the four boxes shown in Fig. 3). Consequently, we believe it is safe to assume that attention can be involved in limiting the number of micro-analyzers used for a specific task. Nonetheless, it seems clear from our results that this number was unaffected by practice.

2. Multiplicative noise

In Section 5D we argued that it was reasonable to exclude multiplicative noise from our modeling (by setting $N_m = 0$) because we did not examine supra-threshold contrast discrimination (cf. Appendix D). However, multiplicative noise has been implicated in some detection experiments too. Swets, Tanner, and Birdsall [2] used it to explain the relationship between first and second responses in a 4AFC detection experiment. Solomon [Solomon], on the other hand, noted that the same results could be explained with intrinsic uncertainty (i.e., $M > 1$).

Burgess and Colbourne [33] invoked multiplicative noise to explain "observer inconsistency," over several trials in which 2AFC detections were limited by identical samples of external noise. Lu and Dosher [17] subsequently rejected an uncertainty-based model of detection in noise on the basis of "opposite demands" on *M*: large amounts of uncertainty were necessary to explain steep psychometric functions, but small amounts were necessary to explain the largely invariant ratio between percent correct and the probability of agreement on trials with identical samples of external noise.

As we did not employ Burgess and Colbourne's "double-pass" methodology, we have no comparable data to test this notion, but we suspect that the requisite relationship between probability correct and probability of agreement can indeed be obtained with large *M*, when the number of relevant micro-analyzers *K* is allowed to exceed 1, a possibility that was not explored by Lu and Dosher. Indeed, the practice of fixing the number of micro-analyzers relevant for each of the *m* forced-choice alternatives at $K = 1$ has become so standard that some authors seem to have forgotten it can have other values [e.g., 3, 17; cf. 34]. Nonetheless, inconsistency seems guaranteed to rise with *K* [35], and we can be confident that this parameter will have an essentially negligible effect on Weibull β, as long as $K < M/100$ [10].

3. Contrast gain control

In section 2F we noted that our stochastic model is formally equivalent to a model of contrast-gain-control, when external noise is absent. It is unclear how such a model of contrast-gain-control would behave when external noise is present. Dao, Lu, and Dosher [14] examined the behavior of an approximation called "cgcPTM." It replaces random variables with their expected values or standard deviations, just like analytic PTM (discussed in section 5E), to which cgcPTM is formally equivalent. However, the formal equivalence between cgcPTM and analytic PTM does not imply a formal equivalence between the models they approximate. It remains possible that there are some models of contrast-gain-control that are not formally equivalent to our stochastic model.

Given any accelerating transducer function (i.e. $\gamma > 1$) the psychometric function for detection necessarily gets shallower as the contrast of external noise increases [25]. However, that reduction in slope may be negligible when performance is limited by additive, internal noise. Baker and Meese [36] implicitly assumed this negligibility when they claimed that, "a pure gain control account of masking predicts no change in the psychometric slope because divisive suppression does not affect the form of contrast transduction." However, no one has yet described how a gain-control circuit would behave without substituting statistics (such as N_{ext}) for individual samples of external noise. This is a critical point, because Klein and Levi [9] demonstrated that the average output of a model with stochastically defined input might deviate qualitatively from the average output of an approximation based on the statistics of that input. Consequently, we are reluctant to form any conclusions on the basis of models that substitute statistics for individual samples of external noise [37].

Watson, Borthwick, & Taylor [38] considered a pure gain-control account of noise masking, but they rejected this idea on the basis of a comparison between "random" conditions, in which a unique sample of band-passed noise was used in each interval, and "fixed" conditions, in which the same sample was used in each interval and each trial. After practice, thresholds in the fixed conditions were considerably lower, suggesting that observers could re-tune their detection mechanisms to exploit the idiosyncrasies of some noise samples.

6D. Other model frameworks

Our data are clearly inconsistent with a model in which practice lowers (or eliminates) a sensory threshold for contrast, thereby imposing a new sensitivity limit based on one or more sources of Gaussian noise. However, it remains conceivable, though unlikely, that practice could lower a sensory threshold for contrast, if it also changed the nature of visual noise (i.e. the shape of its density). A comprehensive examination of models incorporating non-Gaussian visual noise is beyond the scope of this paper.

With sensory thresholds, intrinsic uncertainty, non-linear transduction, and multiplicative noise, our signal-detection model is fairly general. However, its homogenous population of micro-analyzers is not biologically plausible. Real neurons have a variety of response characteristics, particularly with respect to contrast. Indeed, it seems plausible that some neurons will reach their maximum firing rate before others have even begun to fire, even when both sets of neurons share the same receptive field. These two hypothetical sets of neurons can be said to form parallel "contrast channels," and psychophysical data have been fit with models containing an heterogeneity of mechanisms like this [39, 40] The behavior of these models in *m*AFC tasks of discrimination between supra-threshold contrast is quite similar to that of single-channel signal-detection models with compressive (or sigmoidal) transduction (see Fig. 16 of [41]). It is much harder to predict how these models would behave in tasks that are limited by external noise. For now, we must consider this to be an open question.

References and Notes

1. In this paper, we use the term "lapse rate" to describe our variable δ, such that $1-\delta$ is the maximum probability correct. Others have used the same term to describe an alternative variable, which we denote δ', such that $1-\delta'+\delta'/m$ is the maximum probability correct.
2. J. Swets, W. P. Tanner, Jr., and T. G. Birdsall, "Decision processes in perception," Psychol. Rev. **68,** 301 (1961).
3. J. A. Solomon, "Contrast discrimination: Second responses reveal the relationship between the mean and variance of visual signals," Vis. Res. **47,** 3247 (2007).
4. D. M. Green, and J. A. Swets, *Signal Detection Theory and Psychophysics* (Wiley, 1966).
5. C. Yu, S. A. Klein, and D. M. Levi, "Perceptual learning in contrast discrimination and the (minimal) role of context," J. Vis. **4,** 169 (2004).
6. With the exception of the sensory threshold, which is the hypothetically least intense perceptual experience, all psychophysical "thresholds" are measures of performance. Within the context of *m*AFC paradigms, they refer to the value of independent variable that allows observers to attain a criterion level of accuracy. Thus, the contrast threshold for contrast detection, which we shorten to the "detection threshold," refers to the contrast required for $(\delta-1)(e-1)+(e/m)$ accuracy. (This is the value of Eq. 9 when $c_S - c_N = \alpha$.
7. G. E. Legge, D. Kersten, and A. E. Burgess, "Contrast discrimination in noise," J. Opt. Soc. Am. A **4,** 391 (1987).
8. K. T. Blackwell, "The effect of white and filtered noise on contrast detection thresholds," Vis. Res. **38,** 267 (1998).
9. S. A. Klein, and D. M. Levi, "Stochastic model for detection of signals in noise," J. Opt. Soc. Am. A **26,** B110 (2009).
10. D. G. Pelli, "Uncertainty explains many aspects of visual contrast detection and discrimination," J. Opt. Soc. Am. A **2,** 1508 (1985).
11. N. V. S. Graham, *Visual Pattern Analyzers* (Oxford University Press, 1989).
12. B. A. Dosher, and Z.-L. Lu, "Mechanisms of perceptual attention in precuing of location," Vis. Res. **40,** 1269 (2000).
13. B. A. Dosher, and Z.-L. Lu, "Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting," Proc. Natl. Acad. Sci. **95,** 13988 (1998).
14. D. Y. Dao, Z.-L. Lu, and B. A. Dosher, "Adaptation to sine-wave gratings selectively reduces the contrast gain of the adapted stimuli," J. Vis. **6**, 739 (2006).
15. G. E. Legge, and J. M. Foley, "Contrast masking in human vision," J. Opt. Soc. Am. **70**, 1458 (1980).
16. J. A. Solomon, "The history of dipper functions," Attn. Percept. Psychophys. **3,** 435 (2009).
17. Z.-L. Lu, and B. A. Dosher, "Characterizing observers using external noise and observer models: assessing internal representations with external noise," Psych. Rev. **115,** 44 (2008).
18. Some researchers prefer to use the logarithm of signal-detection theory's sensitivity metric d' as the co-domain for psychometric functions, where the proportion correct is given by the equation $\Psi_{\text{SDT}}(d';m,\delta) = \delta + (1-2\delta)\int_{-\infty}^{\infty} G'(x;d',1)\left[G(x;0,1)\right]^{m-1} dx$ and psychometric data are summarised with the linear function $\ln d' = q\ln(c_S - c_N) + k$. This formulation produces a perfect fit to model behavior in the absence of intrinsic uncertainty and a sensory threshold (i.e. when $M = 1$, and $c \ll 0$). In other cases the fit is no longer perfect, but still very close. The relationship between Weibull β and d' power (q) is illustrated in Appendix B. For values of Weibull β between 1.3 and 5, the relationship is well approximated by the formula $q = 0.74\beta - 0.1$.
19. A. B. Watson, "Probability summation over time," Vis. Res. **19,** 515 (1979).
20. D. H. Brainard, "The psychophysics toolbox," Spat. Vis. **10,** 433 (1997).
21. D. G. Pelli, "The VideoToolbox software for psychophysics: transforming numbers into movies," Spat. Vis. **10,** 437 (1997).
22. N. I. Durlach, and L. D. Braida, "Intensity Perception. I. Preliminary Theory of Intensity Resolution," J. Acoust. Soc. Am. **46,** 372 (1969).
23. A. B. Watson, and D. G. Pelli, "Quest: A Bayesian adaptive psychometric method," Percept. Psychophys. **33,** 113 (1983).

24. Astute readers will note that changes in the sensory threshold *c* would go unnoticed if additive noise $A_a N_a$ also changed proportionately. In this section, changes in sensory threshold should be understood to mean "with respect to additive noise**.** "
25. D. J. Lasley, and T. E. Cohn, "Why luminance discrimination may be better than detection," Vis. Res. **21,** 273 (1981).
26. Each of these fits was obtained in three steps. First, we computed the exponent γ necessary for reproducing the median (across days) psychometric slope in the absence of noise (Weibull $\beta = 3.9$). (Collapsing our data across days allowed fits to be calculated with relative rapidity.) Using this value of γ, we computed the gain b necessary for reproducing the median contrast threshold for detection in noise (–22 dB). Finally, using these values of γ and b, we computed the additive noise N_a necessary for reproducing the median contrast threshold (–29 dB) for detection in the absence of noise. (This last step does not affect the performance of the model in noise, which is illustrated in Fig. 6)
27. J. Nachmias, and R. V. Sansbury, "Grating contrast: Discrimination may be better than detection," Vis. Res. **14,** 1039 (1974).
28. J. M. Foley, and G. E. Legge, "Contrast detection and near-threshold discrimination in human vision," Vis. Res. **21,** 1041 (1981).
29. Dosher & Lu, B. A. Dosher, and Z.-L. Lu, "Mechanisms of perceptual learning," Vis. Res. **39,** 3197 (1999).
30. J. A. Solomon, "Noise reveals mechanisms of detection and discrimination," J. Vis. **2,** 105 (2002).
31. Z.-L. Lu, and B. A. Dosher, "Characterizing human perceptual inefficiencies with equivalent internal noise," J. Opt. Soc. Am. A **16,** 764 (1999).
32. Z.-L. Lu and B. A. Dosher, "Perceptual learning retunes the perceptual template in foveal orientation identification," J. Vis. **4,** 44 (2004).
33. A. E. Burgess, and B. Colbourne, "Visual signal detection. IV. Observer inconsistency," J. Opt. Soc. Am. A **5,** 617 (1988).
34. A. S. Baldwin, D. H. Baker, and R. F. Hess, "What Do Contrast Threshold Equivalent Noise Studies Actually Measure? Noise vs. Nonlinearity in Different Masking Paradigms," PLoS. One **11,** e0150942 (2016).
35. With the product A_a N_a held constant, the expected maximum signal **S** must increase with *K*. This, in turn, would improve performance ψ. Therefore, for performance to remain constant, the product A_a N_a must fall as *K* increases. The probability of agreement varies with the ratio $N_{ext} / \left(A_a N_a \right)$, which consequently seems guaranteed to rise as *K* increases.
36. D. H. Baker, and T. S. Meese, "Zero-dimensional noise: The best mask you never saw," J. Vis. $12(10):20$ (2012).
37. Neither the model of Dao et al. [14] nor the models considered by Baker and Meese [36] allow output to fluctuate with each sample of independent, identically distributed external noise. In their Equation 9, Klein and Levi [9] described a gain-control model whose output would fluctuate with external noise. However, the gain-control itself does not fluctuate with external noise. It doesn't even fluctuate with the external signal. (Their σ_{tot} was computed from both the root-mean-squared contrast of the noise and the average contrast of the signal.) It isn't clear how the visual system could compute these statistics. In any case, the simulations described by Klein and Levi nonetheless indicate a "linearization" of psychometric slope in the presence of external noise.
38. A. B. Watson, R. Borthwick, and M. Taylor, "Image quality and entropy masking," SPIE Proceedings, 3016, paper 1 (1997).
39. P. C. Teo, and D. J. Heeger, "Perceptual image distortion," in Human Vision, Visual Processing, and Digital Display V, B. E. Rogowitz and J. P. Allebach, eds., Proc. SPIE 2179, 127–139 (1994).
40. K. A. May, and J. A. Solomon, "Connecting psychophysical performance to neuronal response properties I: Discrimination of suprathreshold stimuli," J. Vis. **15**(6):8 (2015).
41. A. B. Watson, and J. A. Solomon, "Model of visual contrast gain control and pattern masking," J. Opt. Soc. Am. A **14**, 2379 (1997).

Appendix A. TABLE OF SYMBOLS

Symbols appear in the first column and definitions in the second. The final column indicates the section in which the symbol is introduced.

c_N	Contrast of each non-target	**§2D**
c_S	Target contrast	**§2D**
N_{ext}	Amplitude of external noise	**§2D**
m	Number of alternatives in an *m*AFC paradigm	**§1 ¶1**
δ	Lapse rate	**§1 ¶1**
α	Scale parameter in the Weibull distribution	**§3 ¶3**
Weibull β	Shape parameter in the Weibull distribution	**§3 ¶3**
d'	Signal-detection theory's sensitivity metric	**§3 ¶3**
q	Gradient of ln *d'* vs ln ($c_S - c_N$)	**§3 ¶3**
Ψ	Probability correct in an *m*AFC paradigm	**§2A**
$F_X(x)$	Cumulative distribution function for random variable *X*	**§2A**
N	Random variable representing the maximum signal arising from the *M* micro-analyzers associated with each of the (*m*-1) non-targets	**§2A**
S	Random variable representing the maximum signal arising from the *M* micro-analyzers associated with the target	**§2A**

N	Random variable representing the signal arising from each relevant micro-analyzer associated with a non-target	**§2D**
S	Random variable representing the signal arising from each relevant micro-analyzer associated with the target	**§2D**
I	Random variable representing the signal arising from each irrelevant micro-analyzer	**§2D**
M	Number of micro-analyzers associated with each of the *m* alternatives (a model parameter)	**§2B**
K	Number of relevant micro-analyzers sensitive to the target (a model parameter)	**§2B**
c	Sensory threshold (a model parameter)	**§2C**
b	Gain (a model parameter)	**§2D**
γ	Exponent in the power-law transducer (a model parameter)	**§2D**
N_m	Model parameter governing the relationship between mean and variance in micro-analyzer output, after practice	**§2E**
N_a	Model parameter specifying the minimum standard deviation of micro-analyzer output, after practice	**§2E**
A_f	Model parameter governing the effect of practice on micro-analyzer tuning	**§2D**
A_m	Model parameter governing the effect of practice on multiplicative noise	**§2E**
A_a	Model parameter governing the effect of practice on additive noise	**§2E**

Appendix B. Comparing Weibull *β* with *d'* power

Fig. 7 illustrates the similarity between Weibull fits to the relationship between log contrast and percent correct and linear fits to the relationship between log contrast and log *d'*. The relationship described in panel f was used to calculate the right-hand axes in Figs. 2, 4, 5, and 6.

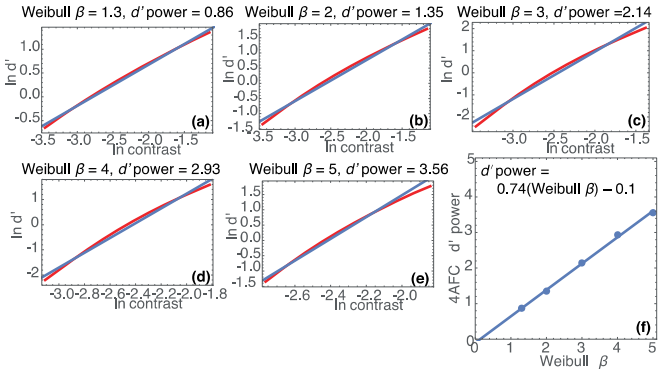

Fig. 7. Relationship between Weibull β and d' power. Each curve in panels a–e is a Weibull function of log contrast (Eqn. 9). It was generated with parameter values $m = 4, \delta = 0,$ and $\alpha = 0.1$. The value for β is indicated at the top of the panel. Over the range of values illustrated in each panel, the curve is well fit by a straight line (i.e. a constant d' power). Panel f summarizes the relationship between d' power and β.

Appendix C. Further illustrative simulations

Fig. 8 illustrates the outcome of various alternative models for the effect of practice. The layout is identical to that of Fig. 2. Whereas Fig. 2 illustrated reductions in sensory threshold, additive noise, and filter bandwidth in the absence of uncertainty (i.e. M = 1) and nonlinear transduction (i.e. γ = 1), Fig. 8 illustrates their reduction with uncertainty (M = 100; panels a–d), nonlinear transduction (γ = 2.4; panels e–h), and both (M = 100, γ = 2.4; panels i–l). In all cases, reducing the sensory threshold causes a reduction in the slope of psychometric functions for detection in the absence of noise.

Appendix D. Psychometric functions from a representative observer, with a maximum-likelihood model fit

RHDS was the first of nine observers to complete five days of testing. Fig. 9 contains all the raw data he generated on Days 1 (top row) and 5 (bottom row). Fig. 10 contains maximum-likelihood estimates of RHDS's threshold and slope from all days, along with a fit of the model.

Although RHDS was a typical observer in most respects (indeed, his detection thresholds were very close to the median values depicted in Fig. 5), he did produce the lowest lapse rates (δ) in our data set. His data were also noteworthy because maximum-likelihood estimates of Weibull β were consistently higher in the noise conditions than they were in the no-noise conditions. Although no combination of our model parameters is capable of reproducing this feature, it was not apparent in the data from any of the other observers.

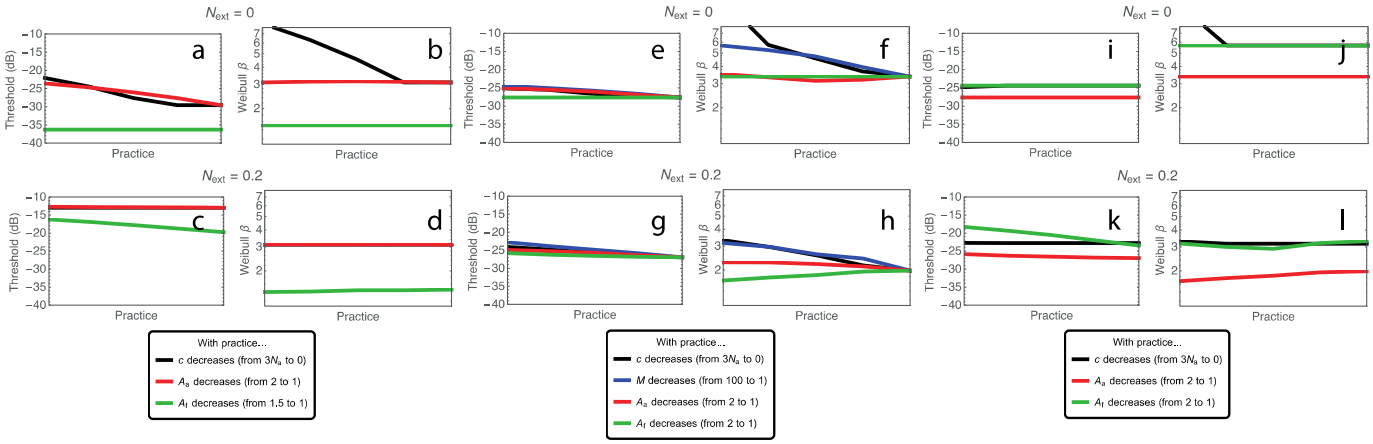

Fig. 8 (Online version in color). More values for threshold and slope from Weibull fits to simulated data from detection and detection-in-noise conditions. Each group of four panels (a-d, e-h, and i-l) has the same arrangement as Fig. 2. In panels a-d, M was fixed at 100 (and γ was fixed at 1). In panels e-h, γ was fixed at 2.4 (and M was fixed at 1), in panels i-l, M was fixed at 100 and γ was fixed at 2.4. Unless specified in the legends, the other parameter values were: $\delta = 0.01$, $K = 1$, $N_m = 0$, $c = -\infty$, $A_a = 1$, $A_f = 1$, $b = 11$, and $N_a = \beta^{\gamma}/10^{3.5}$.

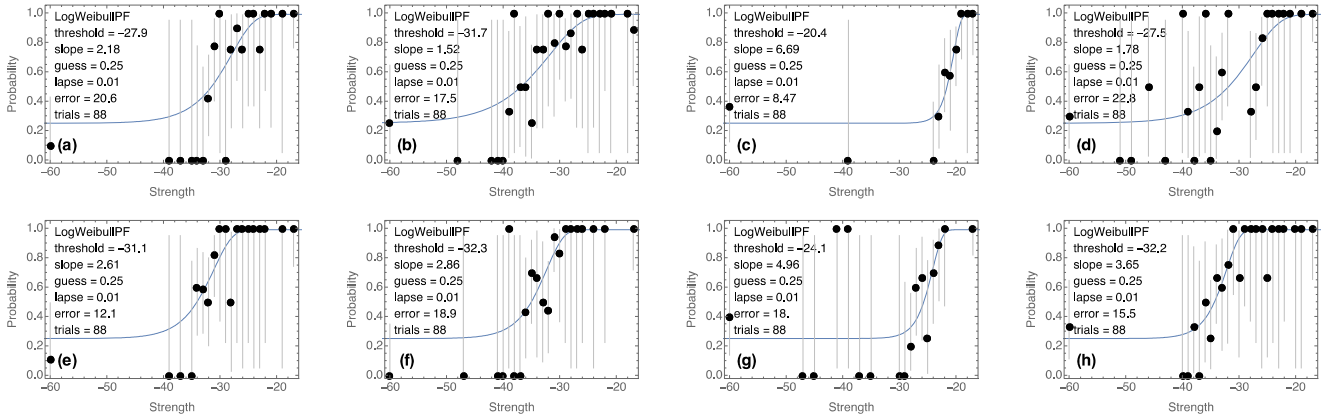

Fig. 9. Psychometric functions for 4AFC detection (a, e), contrast discrimination (b, f), detection-in-noise (c, g), and discrimination-in-noise (d, h) from a representative observer (RHDS) on Day 1 (a–d) and Day 5 (e–h). Insets specify the parameters of the best-fitting Weibull functions (smooth curves). Error bars contain 95% confidence intervals, derived from the binomial distribution.

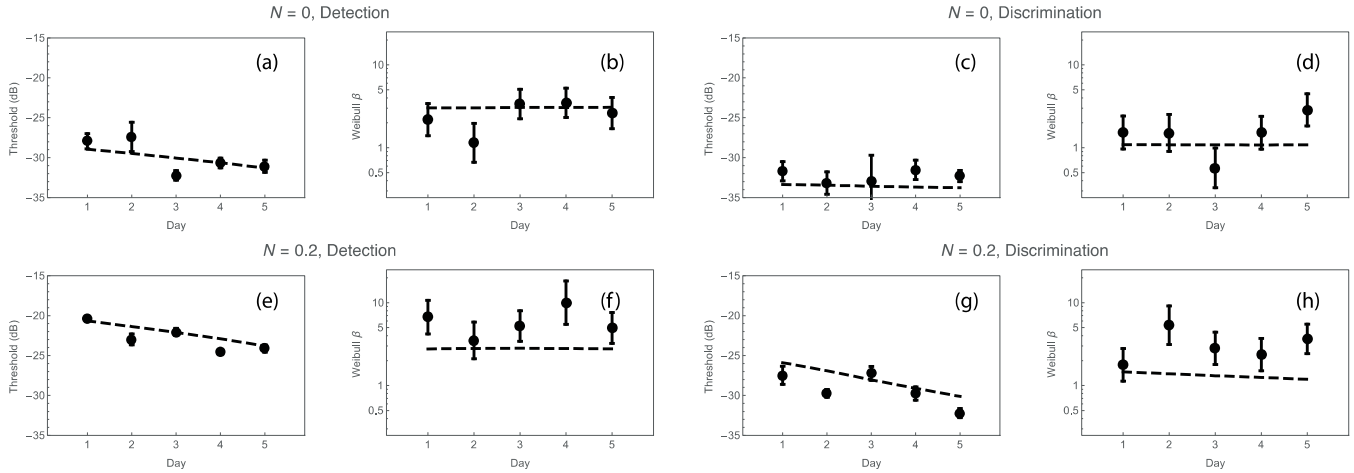

Fig. 10. Maximum-likelihood estimates of representative observer RHDS's threshold (a, c, e, and g) and slope (b, d, f, and h) from Weibull fits to empirical data from detection (a, b), discrimination (c, d), detection-in-noise (e, f), and discrimination-in-noise (g, h) conditions. Error bars contain two standard deviations of a parametric bootstrap distribution. Dashed curves illustrate a model fit. Parameter values were: $\delta = 0.01$, $K = 1$, $M = 1.7 \times 10^6$, $c = -\infty$, $b = 22$, $\gamma = 1$, $N_m = 0.26$, $N_a = 0.098$. A_m was held constant at 1, while A_f decreased from 1.46 to 1 (i.e., between day 1 and day 5) and A_a decreased from 1.30 to 1.