What limits performance in the amblyopic visual system: Seeing signals in noise with an amblyopic brain

Dennis M. Levi
School of Optometry and Helen Wills Neuroscience Institute, UC Berkeley, Berkeley, CA, USA

Stanley A. Klein
School of Optometry and Helen Wills Neuroscience Institute, UC Berkeley, Berkeley, CA, USA

Inning Chen
School of Optometry, UC Berkeley, Berkeley, CA, USA

Amblyopia results in a loss of visual acuity, contrast sensitivity, and position acuity. However, the nature of the neural losses is not yet fully understood. Here we report the results of experiments using noise to try to better understand the losses in amblyopia. Specifically, in one experiment we compared the performance of normal, amblyopic, and ideal observers for detecting a localized signal (a discrete frequency pattern or DFP) in fixed contrast white noise. In a second experiment, we used visibility-scaled noise and varied both the visibility of the noise (from 2 to 20 times the noise detection threshold) and the spatial frequency of the signal. Our results show a loss of efficiency for detection of known signals in noise that increases with the spatial frequency of the signal in observers with amblyopia. To determine whether the loss of efficiency was a consequence of a mismatched template, we derived classification images. We found that although the amblyopic observers’ template was shifted to lower spatial frequencies, the shift was insufficient to account for their threshold elevation. Reduced efficiency in the amblyopic visual system may reflect a high level of internal noise, a poorly matched position template, or both. To analyze the type of internal noise we used an “N-pass” technique, in which observers performed the identical experiment N times (where N = 3 or 4). The amount of disagreement between the repeated trials enables us to parse the internal noise into random noise and consistent noise beyond that due to the poorly matched template. Our results show that the amblyopes’ reduced efficiency for detecting signals in noise is explained in part by reduced template efficiency but to a greater extent by increased random internal noise. This loss is more or less independent of external noise contrast over a log unit range of external noise.

Keywords: reverse correlation, classification images, amblyopia, psychophysics, linear regression


Introduction

Amblyopia is a developmental disorder of spatial vision usually associated with the presence of strabismus, anisometropia, or form deprivation early in life. It affects visual acuity, contrast sensitivity, and position acuity (e.g., Ciuffreda, Levi, & Selenow, 1991; Kiorpes, 2006; Kiorpes & McKee, 1999; Levi, 1991, 2006). Amblyopia is clinically important because, aside from refractive error, it is the most frequent cause of vision loss in infants and young children; and amblyopia is of basic interest because it reflects the neural impairment that can occur when normal visual development is disrupted. However, the precise nature of the neural losses is not yet fully understood.

Measuring human visual performance in noise can provide important insights into the neural mechanisms and computations used to solve a visual task (Dosher & Lu, 1998, 1999; Gold, Murray, Bennett, & Sekuler, 2000; Pelli & Farell, 1999). In what is now considered the classical noise paradigm (Barlow, 1957; Pelli, 1990), external noise is added to the display screen while the observer performs a visual task (e.g., detecting or identifying a target). When the screen noise is low, human thresholds are more or less constant on a log–log plot. When the screen noise is high, exceeding the putative additive internal noise, thresholds increase in proportion to the screen noise (Figure 1). The “knee” in the TvN (threshold vs. noise) curve reflects the equivalent intrinsic noise—it is the noise on the screen that acts like the additive noise in the observer’s visual nervous system. This equivalent noise paradigm has been widely used in a number of domains to estimate and to quantify the limitations imposed by internal noise both in normal (Dosher & Lu, 1999; Eckstein, Ahumada, & Watson, 1997; Pelli, 1990) and in amblyopic (Kiorpes, Tang, & Movshon, 1999; Pelli, Levi, & Chung, 2004; Wang, Levi, Klein, & 1998; Watt & Hess, 1987) vision.

Humans with amblyopia show marked losses when tested with small letters in noise, and these losses have
been attributed both to a loss of efficiency (threshold elevation at all noise levels) and to additive internal noise (Pelli et al., 2004; Xu, Lu, Qiu, & Zhou, 2006).

As shown in Figure 1, similar results are obtained when detecting a “fuzzy bar” (discrete frequency pattern [DFP]) in noise (Levi & Klein, 2002, 2003). The inset in Figure 1 shows an example of the DFP in noise, and the top curves in Figure 3 (labeled ideal) show the spatial frequency (a) and spatial (b) profiles of the stimulus. Note that in Figure 1, the abscissa of the top panel is absolute noise contrast, and the position of the “kink” in the equivalent noise curves, representing the additive internal noise ($N_{eq}$), is shifted to the right in the amblyopic eye, suggesting increased additive noise. A rightward shift is associated with a larger threshold elevation at low than at high noise levels. In the lower panel, we have replotted the data with the noise specified in multiples of the noise detection threshold (i.e., noise threshold units or NTU). Plotted in this way, the curves are parallel, and $N_{eq}$ is almost identical in the two eyes at approximately 1; that is, noise has little impact on performance until it is visible, a result that we find to be quite common. We will return to this point and its implications for the equivalent noise model in the Discussion section.

The equivalent noise paradigm treats the visual system as a black box. Our interest is in opening up the black box and asking to what extent the efficiency loss can be attributed to an inappropriate template and to what extent can it be attributed to internal noise. For that purpose, one needs data beyond the TvN curves.

Earlier we outlined and tested new methods and a model for understanding the factors that limit the detection of noise added to noise (NvN curves) in normal (see Figure 2a of Levi, Klein, & Chen, 2005) and in amblyopic vision (Levi, Klein, & Chen, 2007). The model assumes multiple sources of noise that degrade human performance: an imperfect template (template noise), an additional consistent noise (due to higher order nonlinearities) that results in consistent errors (identical responses on repeated trials), and random noise. Amblyopes showed reduced sensitivity for detecting noise, particularly at high spatial frequencies, and much less loss for discriminating suprathreshold noise contrast. Internal noise was strongly dependent on the visibility of the external noise. For these previous NvN studies, compensating for the detection loss almost (but not quite) equated performance of the amblyopic and normal visual system. For the detection of white noise the raised thresholds of the amblyopic visual system could be attributed primarily to extra additive noise (Levi et al., 2007).

Understanding the amblyopic visual system’s response to white noise (NvN) is important because studies using white noise added to a signal (TvN) have reached different conclusions. For example, several studies have concluded that compared with normal observers, amblyopes have little or no elevation of additive internal noise (e.g., Kiorpes et al., 1999; Pelli et al., 2004). Other studies have suggested that amblyopes have increased additive internal noise (Xu et al., 2006), or that they have increased multiplicative noise (Levi & Klein, 2003). Typically, these studies use the same fixed physical noise contrast level(s) for both amblyopic and normal eyes. Our recent
study (Levi et al., 2007) showed that the amblyopic visual system does not respond in the same way to the noise as the normal visual system since the noise (or some components of the noise) may be less visible through the amblyopic eye. Thus, it is important to understand the response function of the amblyopic visual system over a range of noise contrast levels (e.g., Xu et al., 2006) and to scale the noise contrast to its detection threshold, as shown in Figure 1.

In this paper, we apply this approach to the case of known signals added to noise, and we ask how the sources of internal noise change as we vary the external noise in both the normal and the amblyopic visual system. We use classification images and an N-pass technique to quantify the observers’ decision template and internal noise.

Classification images provide an estimate of an observer’s “perceptive field” (for a recent review, see Neri & Levi, 2006). To the extent that performance in noise is similar to performance without noise, classification images can provide important insights into the mechanisms and strategies that an observer uses to detect or to discriminate a target. To date, only one study has measured classification images in the amblyopic visual system (Levi & Klein, 2003). We used a fixed relatively high level of noise to obtain measurable classification images; however, there is evidence that human performance may be different in high noise than it is in low noise (Lu & Dosher, 2004; Yu, Klein, & Levi, 2002). For one thing, if the limiting noise in the visual system is additive (e.g., Pelli, 1990), it will limit performance at low noise levels but not at high. On the other hand, multiplicative noise (i.e., noise that grows with the signal strength; Tolhurst, Movshon, & Dean, 1983) would limit performance in both high and low levels of noise. In the present paper, we used noise to derive the classification images used to detect a target over a wide range of noise levels.

Visual performance (both psychophysical and neuronal) is also limited by noise or variability (Parker & Newsome, 1998). To quantify this, we use an “N-pass” approach, in which the observer takes several passes through the identical stimuli and noise (Burgess & Colborne, 1988; Gold, Bennett, & Sekuler, 1999; Green, 1964; Levi & Klein, 2003; Levi et al., 2005, 2007). This enables us to measure response consistency, and since the stimuli and noise samples are identical in each pass, we are able to determine the ratio of random to consistent noise (sometimes referred to as the ratio of internal to external noise by other authors) at different levels of external noise. In addition to a mismatched template, consistent noise may arise through a variety of sources. For example, applying a point-wise nonlinearity to an image prior to using a linear template would result in a distortion, which would result in the observer making consistent errors. Surprisingly, our results show that consistent errors beyond a mismatched template play little role in limiting the detection of known signals in noise.

A brief report of some of our preliminary results with fixed high contrast noise has been previously published (Levi & Klein, 2003). However, as noted above, amblyopes may be substantially less sensitive to noise than are normal observers, and, as shown elsewhere (Levi et al., 2005, 2007), the observers’ internal noise depends on the visibility of the noise. Thus, in the present paper, we provide new data from a large group of amblyopes over a wide range of noise contrasts and spatial frequencies; we use a new N-pass method to estimate internal noise directly, and we apply new analyses to learn about the factors that limit the detection of signals in noise.

### Methods

Our methods have been described elsewhere (Levi & Klein, 2002, 2003; Levi et al., 2005, 2007) and will only be briefly discussed here.

### Observers

Five normal control observers (including two of the authors) and 17 amblyopic observers (6 anisometropic; 6 strabismic and 5 with both strabismus and anisometropia) participated in this study. Their clinical details are given in Table 1, and in most of the figures, the type of amblyopia is coded by color (strabismic—red; anisometropic—green; and mixed—blue). The acuities listed here were determined via Bailey–Lovio or Davidson–Eskridge charts. Fragmentary data were obtained on a sixth normal observer. Viewing was monocular, with appropriate optical correction. All experiments were performed in compliance with the relevant laws and institutional guidelines.

### Stimuli

Our signals and noise consisted of sums of sinusoids and are described in detail in Levi and Klein (2002, 2003) along with details of the ideal observer modeling. Briefly, the signal is a discrete frequency pattern (DFP)—a bar-like pattern given by $c \cos^{10}(\pi y) \cos(2\pi 6y)$ (see top curve of Figure 3b) and composed of 11 harmonics all added in cosine phase (see top curves of Figures 3a and 3b for the spatial frequency and spatial profiles). Our choice of $\cos^{10}(\pi y)$ as the envelope was to limit the signal to eleven harmonics. The noise is a one-dimensional grating consisting of the same 11 harmonics with phases and amplitudes randomized with each harmonic having equal variance. In Experiment 1, the standard deviation of each noise component had a fixed contrast of 0.04 (see Levi & Klein, 2002), and the DFP and the noise contained spatial...
frequencies from 1 to 11 c/deg. In the second experiment, we used visibility-scaled noise (described below), and the stimuli were viewed from one of three viewing distances so that they provided three ranges of signal and noise spatial frequencies (0.5 to 5.5 c/deg; 1 to 11 c/deg; 2 to 22 c/deg). The target and the noise (inset in Figure 1) were presented for 0.75 s, in a square field with a mean luminance of 42 cd/m² with a dark surround.

### Psychophysical methods

We used signal detection methodology to measure the observers’ performance ($d'$)—which is a measure of the observers’ signal-to-noise ratio—and linear regression to compute the classification coefficients (Levi & Klein, 2002, 2003). The signal to be detected was either a blank (0 contrast) or one of three near-threshold stimuli, and the observer responded with numbers from 1 (confident the signal was a blank) to 4 (highest perceived contrast). Observers were given auditory feedback following each trial. The contrast levels were chosen to yield $d'$s between approximately 0.5 and 2, based on preliminary trials. Thresholds, corresponding to the contrast at which $d' = 1$, were determined by fitting a transducer (power) function to the $d'$ vs. contrast data using Matlab’s lsqnonlin. One advantage of our rating-scale method is that it enables us to collect data over a reasonable range of $d'$ values.

### Table 1. Observer characteristics.

<table>
<thead>
<tr>
<th>Observer</th>
<th>Age (years)</th>
<th>Gender</th>
<th>Eye</th>
<th>Refractive error</th>
<th>Line letter acuity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strabismic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MR</td>
<td>22</td>
<td>F</td>
<td>L EsoT 14°</td>
<td>−3.25/−2.50 × 175</td>
<td>20/20°</td>
</tr>
<tr>
<td>JT</td>
<td>52</td>
<td>F</td>
<td>L EsoT 5°</td>
<td>−1.00/−0.50 × 10</td>
<td>20/16°</td>
</tr>
<tr>
<td>SF</td>
<td>20</td>
<td>F</td>
<td>L EsoT 6-8°</td>
<td>−0.75/−0.50 × 90</td>
<td>20/63°</td>
</tr>
<tr>
<td>RH</td>
<td>32</td>
<td>M</td>
<td>L EsoT 2°</td>
<td>−1.00/−0.50 × 170</td>
<td>20/15</td>
</tr>
<tr>
<td>DH</td>
<td>20</td>
<td>M</td>
<td>L EsoT 5°</td>
<td>+0.25/−0.50 × 100</td>
<td>20/15</td>
</tr>
<tr>
<td>AH</td>
<td>19</td>
<td>F</td>
<td>R EsoT 12°</td>
<td>+0.50/−1.00 × 93</td>
<td>20/50</td>
</tr>
<tr>
<td>Anisometric</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JW</td>
<td>22</td>
<td>F</td>
<td>None</td>
<td>+1.75</td>
<td>20/80°</td>
</tr>
<tr>
<td>SC</td>
<td>27</td>
<td>M</td>
<td>None</td>
<td>−2.00</td>
<td>20/20</td>
</tr>
<tr>
<td>VG</td>
<td>31</td>
<td>M</td>
<td>None</td>
<td>+0.50</td>
<td>20/16°</td>
</tr>
<tr>
<td>MLR</td>
<td>44</td>
<td>F</td>
<td>None</td>
<td>+4.00/−1.00 × 31</td>
<td>20/125°</td>
</tr>
<tr>
<td>RJ</td>
<td>53</td>
<td>M</td>
<td>None</td>
<td>+0.50/−0.25 × 95</td>
<td>20/14</td>
</tr>
<tr>
<td>AM</td>
<td>22</td>
<td>F</td>
<td>None</td>
<td>+2.50/−1.00 × 10</td>
<td>20/50</td>
</tr>
<tr>
<td>Strabismus and anisometropia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SM</td>
<td>55</td>
<td>F</td>
<td>Alt. ExoT 18°</td>
<td>+2.75/−1.25 × 135</td>
<td>20/40 (20/25°)</td>
</tr>
<tr>
<td>JD</td>
<td>19</td>
<td>M</td>
<td>L EsoT 3°</td>
<td>+2.50</td>
<td>20/16</td>
</tr>
<tr>
<td>AW</td>
<td>22</td>
<td>F</td>
<td>R EsoT 4–6° &amp; hypoT 4°</td>
<td>+2.75/−1.00 × 160</td>
<td>20/80°</td>
</tr>
<tr>
<td>DM</td>
<td>40</td>
<td>F</td>
<td>L ExoT 3°</td>
<td>+0.50/−0.25 × 92</td>
<td>20/20</td>
</tr>
<tr>
<td>DS</td>
<td>26</td>
<td>M</td>
<td>R EsoT 8°</td>
<td>+2.25 DS</td>
<td>20/40</td>
</tr>
</tbody>
</table>

Note: EsoT = esotropia; EXoT = exotropia; HypoT = hypotropia.
Data were collected in runs of 410 trials, preceded by 20 practice trials. Each classification image is based on the results averaged over either 3 or 4 separate runs (1230–1630 trials). In Experiment 1, classification images were obtained for a fixed noise contrast (0.04). In Experiment 2, they were obtained at a range of noise levels, from 2 to 20 times the noise detection threshold. Noise detection thresholds were measured prior to the DFP experiments (see Levi et al., 2005, 2007).

### Classification images

The coefficients of the classification images were estimated in the frequency. Specifically, the observers’ rating responses were regressed on the eleven cosine noise components. To check whether the sine components contributed significantly, we also calculated the full template using odd as well as even symmetric components. We found less than a 1% difference in template efficiency between the two template computations. The rating-scale methods have been shown to be able to improve the quality of classification images (Murray, Bennett, & Sekuler, 2002). Template efficiency was estimated directly from the coefficients.

### Response consistency

In Experiment 2, we used an N-pass method to determine our observers’ response consistency (Burgess & Colborne, 1988; Gold et al., 1999; Levi & Klein, 2003; Levi et al., 2005, 2007). In this method, identical stimuli are used in each pass. Specifically, we saved the random seed from the initial run and re-used it so that the noise was identical in either 2 (double pass), 3 (triple pass), or 4 runs (quadruple pass). In the double-pass case, we ran two separate double passes. For triple and quadruple passes, we analyzed and averaged all possible pairings. Although the same stimuli were presented in multiple runs, the order of presentation was randomized across runs to avoid response order contingencies. In our initial experiments, we used $N = 2$ (twice) or 4. However, we found that $N = 3$ provided approximately equal reliability with fewer trials, and most of the response consistency data reported here was with $N = 3$ passes.

We used two methods for estimating the correlation $q^2$ that is the ratio of consistent response variance to total response variance (discussed in detail in Levi et al., 2007). We use $q^2$ rather than $q$ as a reminder that we are discussing ratios of variances. The first method involves finding a best fit to the data in terms of the estimated $d'$ values for the different stimulus levels and the placement of criteria. To estimate the correlation, $q^2$, we replaced the assumptions of independent Gaussian noise on each pass with bivariate Gaussian noise that was correlated across pairs of passes. The correlation was a free parameter that was varied to get the best fits to the double pass data. All pairings of multiple passes were examined, and the $q^2$ from each pairing was averaged. The search for the optimal $q^2$ was done with the $d'$ and criteria constrained to their best values.

The second method for estimating $q^2$ was much simpler. We simply calculated the standard correlations of the responses for all pairs of passes, $p$. This analysis was done separately for each stimulus level. We found that the two methods gave similar estimates of $q^2$, so our estimates of the ratio of consistent to total noise are based on the average of the two methods.

### General response model

Our general model for both detection of noise added to noise (NvN) and known signals added to noise (TvN) is given elsewhere (see Appendix and Figure 2a of Levi et al., 2005, 2007). The model assumes various sources of noise that degrade human performance. These include an imperfect template (template noise), consistent noise (higher order nonlinearities), and random noise. Figure A1 in Appendix A illustrates the model predictions for the manner in which each of these noise sources and observer efficiencies depend on external noise contrast.

### Results

#### Experiment 1: Fixed high contrast noise

In the first experiment, we asked 7 amblyopic observers (2 anisometric, 3 strabismic and 3 with both strabismus and anisometropia) to detect a fuzzy bar (DFP SF 1–11 cpd) that was presented in fixed high contrast noise (noise contrast of 0.04). As is evident in Figure 2, most of the seven amblyopic observers had similar efficiency (human/ideal) in their amblyopic (solid symbols) and non-amblyopic eyes (open square shows the mean of the non-amblyopic eyes), which was a slightly lower root efficiency than the normal controls (open circle). Two of the amblyopes (DS and DH) show markedly reduced root efficiency.

#### Classification images in fixed high contrast noise

Figure 3a shows the classification coefficients for detection of the DFP (averaged across stimulus levels), obtained using linear regression to determine how each of the spatial frequency components of the noise affected the observer’s rating response. The classification coefficients are shown at each spatial frequency for the normal
controls and for each of the amblyopic eyes. Note that the more severe amblyopes (e.g., DM and DH) do show a shift in the peak of the spatial frequency tuning toward lower spatial frequencies (Levi, Waugh, & Beard, 1994). The classification images (Figure 3b) are the Fourier transforms of the classification coefficients and show the average detection spatial templates of the normal controls and each of the ambylopic eyes. The ambylopic detection templates are somewhat broader in space than those of the normal observers, corresponding to the shift in spatial frequency tuning toward lower spatial frequencies (Levi et al., 1994). Interestingly, the ambylopic with the most reduced detection efficiency (DH) shows a very poorly defined template with essentially no spatial frequency tuning.

To determine whether the reduced detection efficiency of the ambylopic was due to a poorly matched template, we compared the ambylopic observers’ root efficiency (Figure 2, solid symbols) to the root efficiency of their templates (Figure 2 horizontal bars, see Methods section). Note that the template efficiencies were estimated directly from the measured classification image coefficients in Figure 3a (not from the fits). For ambylopic observers RJ and DS, the template efficiencies seen in Figure 2 are actually higher than the template efficiencies of the normal eyes. The explanation of this high template efficiency is seen in Figure 3. The classification coefficients of the normal eyes place too much weighting on the high frequency coefficients relative to the ideal observer (the top curves in Figure 3 are the classification coefficients and the classification image of the ideal observer). The high frequency falloffs of RJ and DS are well matched to the profile of the ideal coefficients. The average template efficiency for the ambylopic eyes was about 75%, comparable to the normal observers (81%). Note that the difference in the efficiency of the human observers and the template observers (i.e., the distance between each symbol and the corresponding horizontal bar in Figure 2) reflects various sources of internal noise. It is clear from this figure that most of the detection efficiency loss cannot be simply attributed to a mismatched template—rather, it must reflect noise in the visual system.

**Experiment 2: Visibility-scaled noise**

In this section, we explore to what extent the relatively low loss of efficiency in most ambylopic is due to the fact that the fixed contrast (0.04) noise is less visible (and therefore less effective) in the ambylopic eye. As noted above, the amount of internal noise may depend upon the visibility of the external noise.

To accomplish this, we measured noise detection thresholds in one of three spatial frequency bands (0.5–5.5 cpd, 1–11 cpd, or 2–22 cpd). Each panel of Figure 4 represents one spatial frequency band and plots the ambylopic eye’s noise detection threshold (in rms contrast—abscissa) against the threshold for detecting a DFP (peak contrast) in noise that is twice the noise detection threshold (ordinate). We chose noise at 2 NTU because we were able to measure thresholds at this noise level even at the highest spatial frequencies in most observers. For comparison, the mean noise detection thresholds of the normal control observers and the preferred eyes of ambylopic are shown by the open circles and squares, respectively. The figure shows clearly that while the noise detection thresholds of the normal observers show only a small dependence on spatial frequency band, those of the ambylopic eyes show a marked dependence. They are within about a factor of three (5 of 7 within a factor of two) of normal at 0.5–5.5 cpd and are more elevated at 2–22 cpd. It is also interesting to note that with the exception of anisometric ambylopic JW, the ambylopic loss for the DFP is greater than the loss for noise. This result is not surprising since the noise can be detected...
based on the first two harmonics (2 and 4 cpd for a DFP with spatial frequencies of 1–11 c/deg), whereas the DFP only has substantial energies above the third harmonic (8 cpd and above for the same DFP) where the amblyopic sensitivity can be much reduced. In interpreting Figure 4, it is worth noting that the rms contrast of a DFP pattern is 0.515 times the peak contrast. If the ordinate had been plotted as rms contrast, the DFP threshold would be quite close to the noise threshold.

We use these noise detection thresholds to scale the noise in the experiments that follow.

Specifically, we measured thresholds for detecting a DFP with the noise set at a given multiple of the observer’s noise detection threshold (2, 5, 10, and 20 NTU [noise threshold units]). Figure 5 shows the DFP detection threshold versus noise contrast (in NTU) for each of the amblyopic eyes (solid symbols), and for comparison, the mean data of normal controls (open circles) and the non-amblyopic fellow eyes (open squares). Each panel represents a different DFP and noise spatial frequency band. The leftmost points show DFP thresholds with no noise. Note that the thresholds for each observer at 2 NTU are the thresholds shown in Figure 4 at different rms contrast levels. The data show three interesting features:

1. At each scale, the amblyopic eye thresholds are uniformly elevated, and this elevation is generally most marked at the higher spatial frequencies. For example, JD (blue circles) shows about a 2.5-fold loss (relative to normal controls) at 0.5–5.5 c/deg, a ≈6-fold loss at 1–11 c/deg, and an 8-fold loss at 2–22 c/deg. Note that many of the amblyopes could not be tested at the highest spatial frequency range because they needed more contrast to reach threshold than we could generate. At each spatial scale, the uniform threshold elevation (i.e., thresholds are elevated about the same amount at each noise level) implies that the loss is at least in part due to multiplicative noise. In the sections that follow, we explore to what extent this noise contrast independent loss is due to a mismatched template or to other sources of noise.

Figure 3. (a) Classification coefficients for detection. The regression coefficients (symbols) are plotted as a function of spatial frequency for each of the amblyopic eyes and averaged across the dominant eyes of three normal observers. The smooth curves are the four parameter exponentiated difference of Gaussians (E-DOG) given by $p_1(\exp(-p_2f^2) - \exp(-p_3f^2))^{p_4}$ that are the best fit to the eleven data points. An additional datum at 30 c/deg with a value of zero was added to account for the expected falloff at very high spatial frequencies. The exponentiation ($p_4$) was needed to sharpen the tuning and provide a good fit to the ideal observer data (top curve). (b) Classification images for detection. The dotted lines are the raw classification images for each of the amblyopic eyes, and the average across the dominant eyes of three normal observers. The solid curves are the Fourier transforms of the (E-DOG) curves fit to the regression coefficients in the left panel, assuming all frequency components are in cosine phase. The top curves in both panels are the ideal observer. The classification coefficients and images have been vertically offset from one another for clarity.
Figure 4. Noise detection thresholds. Each panel represents one spatial frequency band and plots the amblyopic eye’s noise detection threshold (rms) along the abscissa, plotted against the threshold (peak contrast) for detecting a DFP in noise (at twice the noise detection threshold—ordinate). For comparison, the mean noise detection thresholds of the normal control observers and the preferred eyes of amblyopes are shown by the open circles and squares respectively. The dotted line has unit slope.

Figure 5. DFP detection threshold versus noise contrast (in NTU) for each of the amblyopic eyes (solid symbols). For comparison, the mean data of normal controls (open circles) and the non-amblyopic fellow eyes (open squares). Each panel represents a different DFP and noise spatial frequency band. The left most points show DFP thresholds with no noise. The curves show the equivalent noise fit to the data. The dotted line has unit slope.
2. The kink in the curves representing the additive equivalent internal noise ($N_{eq}$) occurs at about the noise detection threshold ($\approx 1$ NTU) for all observers. Figure 1 (lower panel) shows an example of this. The mean value of $N_{eq}$ was 1.18 ± 0.04 NTU (SD = 0.23NTU), based on the 29 fits (12 amblyopic eye fits, 5 normal eye fits, and 12 non-amblyopic eye fits—note that Figure 5 shows only fits to the mean normal and non-amblyopic eye data). This finding, that the noise becomes effective as a masker when it becomes visible, is not terribly surprising, but it was not inevitable and surprisingly, it has not been pointed out before. We will return to this point in the Discussion.

3. At the two highest spatial frequencies, thresholds for the non-amblyopic eyes were slightly elevated relative to the normal controls.

Human efficiency at different noise contrast levels

As in Experiment 1, we are interested in what limits observers’ efficiency. Figure 6 shows root efficiency plotted as a function of noise contrast (in NTU). Note that the calculation of efficiency is done based on the absolute noise contrast (as in the top panel of Figure 1), but the values are plotted in terms of the noise contrast in threshold units (as in the lower panel of Figure 1). Each column shows a different spatial frequency range. The top row shows the mean human efficiency for the normal control observers (large circles) and for the non-amblyopic eyes of our amblyopic observers (large squares). For comparison, the ideal observer (efficiency = 1) is shown by the gray dotted line (we will discuss the small symbols and solid lines later). For normal observers (and non-amblyopic eyes), root efficiency is more or less independent of noise contrast. The thin gray dotted line is the best fitting line to the normal efficiency vs. noise contrast data, and for reference, we replicate these lines in every panel in the 7 lower rows, which show the efficiency for each of the amblyopic eyes (large solid symbols). The second, third, and fourth rows show amblyopic observers who were tested at each of the three frequency ranges (all panels in the red box were tested at the three ranges); the bottom 4 rows are amblyopic observers tested either at the two highest frequency ranges (panels in the green box), the two lowest frequency ranges (panels in the blue box), or only at a single frequency range (unboxed).

Inspection of Figure 6, which represents over 150 Kiloteas (i.e., over 100000 datasets), shows that in general the efficiency loss (compare large colored symbols with thin gray dotted line) shows little dependence on noise contrast but strong dependence on signal spatial frequency (the loss increases going from the left-most column to the rightmost). In the lowest spatial frequency range, the loss is negligible. Note that at the higher frequencies, most of the amblyopic observers could not be tested at high noise contrasts since we could not produce sufficient signal plus noise contrast to measure performance.

Classification images at different spatial frequencies

Below, we ask to what extent the efficiency loss can be attributed to an inappropriate template and to what extent it can be attributed to internal noise. We assessed the observer’s template by measuring their classification images using linear regression (regressing the subject’s ratings against the 11 spatial frequency amplitudes). Examples of the classification coefficients for the blanks (zero contrast DFP), normalized, and averaged across external noise levels (from 2 to 10 or 20 times the noise detection threshold) are shown in Figure 7. Each panel shows a different spatial frequency range. Each curve shows how each of the spatial frequency components of the external noise affected the observer’s responses. The classification coefficients for the amblyopic eyes are shown by the solid symbols. For comparison, we show the mean classification images of the normal control observers (open circles) and non-amblyopic eyes (open squares).

It is clear that the peaks of the curves shift systematically to the right as the center frequency of the DFP increases from 3 (range 0.5–5.5) to 6 (range 1–11) cpd. The peak does not shift proportionally higher when the DFP center frequency is 12 (range 2–22) cpd. The classification coefficient curve for the high spatial frequency DFP peaks at about 8 cpd in the normal controls. Interestingly, while the peaks of the amblyopic eyes are lower than those of the normal or preferred eyes at the two lower frequencies, several of the amblyopes actually peak at higher spatial frequencies than normal in the highest spatial frequency range. This can be seen in Figure 8, where the peak spatial frequency of the classification image for all of the observers is plotted against the noise mid-spatial frequency.
We are interested in whether and how the classification image depends on external noise contrast. Almost all previous studies have used high levels of noise to obtain measurable classification images. To the extent that performance in noise is similar to performance without noise, classification images can provide important insights into the mechanisms and strategies that an observer uses to detect or discriminate a target; however, there is evidence that human performance may be different in high noise than it is in low noise (Lu & Dosher, 2004; Yu et al., 2002). Figure 9 shows how the classification images

**Classification images at different noise contrast levels**

We are interested in whether and how the classification image depends on external noise contrast. Almost all previous studies have used high levels of noise to obtain measurable classification images. To the extent that performance in noise is similar to performance without noise, classification images can provide important insights into the mechanisms and strategies that an observer uses to detect or discriminate a target; however, there is evidence that human performance may be different in high noise than it is in low noise (Lu & Dosher, 2004; Yu et al., 2002). Figure 9 shows how the classification images
for the blanks (zero contrast DFP spatial frequency 0.5–5.5 c/deg) depend on noise level (from 2 to 20 NTU). Second-order nonlinearities (regression with the quadratic terms), shown by the tiny symbols bunched between 6 and 12 c/deg, are on average nonzero but are very small relative to the first-order coefficients. It is clear from inspection of Figure 9 that the classification image does not change systematically with noise contrast, except for a slight broadening of the tuning width at the higher contrasts as might be expected from saturation. This is confirmed in Figure 10, which shows the peak spatial frequency of the best fitting Gaussian functions, averaged across all observers, plotted as a function of noise contrast. Noise contrast has only a small effect on peak spatial frequency, while the spatial frequency range (coded by symbol size) has a striking effect. The peak spatial frequency is actually slightly higher than the peak stimulus frequency at the lowest spatial frequency and somewhat lower at the highest.

Similar contrast invariance is evident in the preferred eyes of amblyopic observers (Figure 11, top and bottom panels) and in the amblyopic eyes (the center panel of Figure 11 shows data of JD an example). However, like most (but not all) of the amblyopic observers, JD’s amblyopic eye shows a shift in template tuning toward lower spatial frequencies (as seen in Figure 8).

Our main interest is to what extent amblyopic performance is limited by the observer’s template. The lines in Figure 6 show the thresholds estimated for the template observer (i.e., an ideal observer with the amblyopic
Internal noise

By repeating the experiment several times with identical noise sequences shown in a randomized order, we estimated $q^2$, the ratio of consistent to total noise power. The amount of response disagreement between the $N$ tests allows the system’s total noise to be parsed into random noise that is independent across multiple presentations of the identical stimulus and noise that is consistent (100% correlated) across multiple presentations (see Appendix A for details).

Figure 12 shows that, in normal control observers, $q^2$ increases from about 40% to 60%, as external noise contrast increases from 2 to 20 times threshold. While some of the amblyopes show similar increases in the lowest spatial frequency range, at the higher frequencies both the non-amblyopic eyes (open squares) and the amblyopic eyes show much lower values of $q^2$, particularly as noise contrast increases. This indicates that the amblyopic visual system has increased random noise (relative to normal). Note that at 0 NTU, $q^2$ must fall to zero. The small symbols in Figure 13 show the contribution (if any) of consistent noise to human performance (beyond the effects of consistent noise resulting from the template). Consistent noise results in little or no loss beyond the mismatched template, so the small symbols are rarely seen. The negligible role for consistent noise other than a mismatched template rules out the possibility that there are unusual suppression effects whereby some inappropriate spatial frequencies or phases disrupt the visibility of a DFP in a consistent but nonlinear manner across multiple passes.

These effects are seen clearly in the summary figure (Figure 13), which shows the mean data of the normal controls and of the amblyopic eyes plotted as the inverse of efficiency vs. noise contrast. The inverse of root efficiency are the human and the template observer thresholds divided by the ideal observer threshold. They are like TvN plots rotated by 45 deg. Here it is evident that, on average, the amblyopic loss increases with spatial frequency (note the change in the ordinate scale as spatial frequency increases from top to bottom), and that both the increase in template-mismatch noise (light gray zones) and the loss due to random internal noise (dark gray zones) vary remarkably little across external noise levels (NTU). Consistent internal noise beyond the template mismatch (smallest symbols) plays little or no significant role in limiting performance in either normal or amblyopic vision. Note that unlike Figure 6, Figure 13 is plotted in linear axes, and this makes two specific points. One is that human performance changes dramatically between 0 and 2 NTU and remarkably little from 2 to 20 NTU. The other point is that the template-mismatch noise of the amblyopic eye increases substantially as spatial frequency increases. Plotting the data as inverse efficiency (rather than root efficiency) also highlights the large increase in random noise in the amblyopic visual system.

The non-amblyopic eye

At several junctures, we have pointed to the non-amblyopic eye, on average being somewhat poorer than normal particularly at the higher spatial frequencies (for example, see Figures 6 and 12), suggesting the possibility of a central deficit. Therefore, it is of some interest to
compare on a subject-by-subject basis the efficiency of the two eyes. Figure 14 provides such a comparison, plotting both human (top panel) and template (bottom panel) efficiencies for each observer’s amblyopic eye against that of their fellow eye. For both human and template efficiencies, the data tend to hover close to but slightly above the equality line. For all but the least efficient observers, the loss in both eyes is similar. For the least efficient observers, the amblyopic eye loss is substantially more severe than the loss in the non-amblyopic eye.

**Discussion**

Our results, showing a marked loss of efficiency for detection of signals in noise, confirm and extend the results of a number of previous studies using noise to try to understand the mechanisms of amblyopia, and we discuss these and their implications for understanding the neural mechanisms of amblyopia below.

**Relationship to previous studies**

A large number of studies have applied the noise paradigm to exploring internal noise in amblyopia. Here we focus on those that have used white luminance or pixel noise (Kersten, Hess, & Plant, 1988; Kiorpes et al., 1999; Levi & Klein, 2003; Nordmann, Freeman, & Cassanova, 1992; Pelli et al., 2004; Xu et al., 2006) as opposed to, for example, positional noise (Wang et al., 1998; Watt & Hess, 1987).

An influential model, the linear amplifier model [LAM], treats internal noise as if it were equivalent input noise...
added to the screen display (Barlow, 1957; Pelli & Farrell, 1999). A more recent version of this model is the linear uncertainty model \(\text{LUM}\) (Eckstein et al., 1997). Because the equivalent input noise is additive, it limits performance at detection threshold and at low noise levels, but not at high external noise levels, where reduced efficiency degrades threshold. Previous studies of grating contrast sensitivity (Kersten et al., 1988; Kiorpes et al., 1999; Nordmann et al., 1992) and letter recognition (Pelli et al., 2004) have reported both increases in equivalent input noise and, more often, large reductions in efficiency in amblyopes as evidenced by the substantial threshold elevations at high noise levels.

In a detailed vision model, noise need not be additive (Dosher & Lu, 1999; Eckstein et al., 1997; Levi et al., 2005; Lu & Dosher, 2004). Indeed, there is strong physiological evidence that in cortical neurons noise increases in proportion to the signal strength (Shadlen, Britten, Newsome, & Movshon, 1996). This signal-dependent intrinsic noise is not made explicit in the standard additive noise model (it is factored into the efficiency loss). One such model that has recently been applied to amblyopia is the perceptual template model (PTM; Lu & Dosher, 2004; Xu et al., 2006). Below we attempt to clarify the relationship between LUM and PTM by discussing how threshold versus noise functions are specified and modeled.

A key difference between LUM and PTM relates to whether the noise is applied early or late (see Chung, Levi, & Tjan, 2005). Our finding that \(N_{\text{eq}} \approx N_{\text{th}}\) can be used to eliminate the class of models in which the independent additive noise follows multiplicative noise. Suppose the amblyopic loss is purely multiplicative. In that case, thresholds would be elevated at large \(N_{\text{ext}}\). Assuming the multiplicative noise comes before the additive noise thresholds would be untouched at low \(N_{\text{ext}}\). The result is that the kink point, \(N_{\text{eq}}\) of the \(\text{TvN}\) curve would shift leftward (see Equation 3). However, given that \(N_{\text{eq}}\) is tied to \(N_{\text{th}}\) that shift is to the right for degraded amblyopic vision. This simple logic is a problem for models with independent late additive noise such as the PTM model. One can always rescue the model by saying that the late additive noise is not independent and must increase as the multiplicative noise increases, in which case it is indistinguishable from a model with early additive noise. In a model with independent early additive noise, an increase in multiplicative noise, such as a mismatched template, lifts the log–log \(\text{TvN}\) plot vertically. Additional additive noise would increase the \(\text{TvN}\) floor and thereby shift \(N_{\text{eq}}\) rightward as would be expected from a concomitant increase in \(N_{\text{th}}\).

From this brief discussion, it is clear that distinguishing between multiplicative and early vs. late additive noise is model dependent. However, for models in which the noise sources are independent, to avoid multiplicative noise lowering noise detection thresholds, the early noise model may be preferable. For now we remain agnostic on this issue.

### Threshold versus noise (TvN) functions

The class of functions that have been commonly used to fit \(\text{TvN}\) data, including the PTM and LUM models discussed previously, can be written as

\[
\text{Th}(N_{\text{ext}}, d') = \text{Th}_0 \ ICRF(d')(1 + (N_{\text{ext}}/N_{\text{eq}})^p)^{1/p},
\]

\(\text{Figure 14. A comparison of the efficiency of the amblyopic (abscissa) to that of the fellow eye for each observer. The top panel shows the human efficiency. The lower panel, the template efficiency. The spatial frequency range is coded by symbol size (larger symbols for higher frequencies). The solid line shows equality.}\)
where \( Th \) is the threshold contrast that is a function of two experimental parameters, \( N_{\text{ext}} \) the external noise, and \( d' \) is the signal detection measure of detection. The contrast response function \( \text{ICRF}(d') \) (inverse contrast response function) specifies how the threshold depends on the \( d' \) that is chosen for defining threshold. The special feature of the class of models specified by Equation 1 is that the dependence on \( N_{\text{ext}} \) and \( d' \) is separable. Although later in this discussion, we will show that our data provide strong evidence against a version of LUM with a linear CRF shape of the \( \text{TvN} \) function. Lu and Dosher provide difficult to pin down the pooling exponent, \( p \), three parameters, \( \text{N}_{\text{eq}}, \text{Th}_0, \) and \( p \), are evident: \( \text{N}_{\text{eq}} \), the equivalent input noise, shifts the \( \text{TvN} \) curve horizontally on log–log axes. This parameter is has been popularized by Barlow (1957) and Pelli (1990). The threshold parameter \( \text{Th}_0 \) and the associated ICRF shifts the curve vertically on log–log axes. We typically normalize the CRF\((d')\) function as CRF(1) = 1 so that Equation 1 gives

\[
\text{Th}(0, 1) = \text{Th}_0. \tag{2}
\]

That is, \( \text{Th}_0 \) is the threshold for zero external noise at \( d' = 1 \). For \( d' \) values other than unity, the ICRF simply shifts the \( \text{TvN} \) curve vertically on log–log axes without changing the shape of the function. The Minkowski exponent, \( p \), has a slight effect on the shape of the \( \text{TvN} \) function when \( N_{\text{ext}} \) is close to \( N_{\text{eq}} \) but the effect is so small that an unreasonable amount of data is needed to pin it down.

One version of Equation 1 that has been popularized by Lu and Dosher (1999) is the perceptual template model (PTM) in which the pooling exponent, \( p \), is coupled to the transducer exponent of the CRF. Another version of Equation 1 is the linear uncertainty model (LUM) of Eckstein et al. (1997) in which the pooling exponent is \( p = 2 \) and is decoupled from the CRF. As stated above, it is very difficult to pin down the pooling exponent, \( p \), based on the shape of the \( \text{TvN} \) function. Lu and Dosher provide evidence against a version of LUM with a linear CRF (the linear amplifier model, LAM). However, it is well known that the inverse contrast response function is nonlinear and is approximately \( \text{ICRF}(d') = d'^{p/2} \) for low values of \( d' \) and it accelerates above \( d' = 1 \) (Nachmias & Sansbury, 1974; Stromeyer & Klein, 1974). This behavior of the inverse CRF (contrast as a function of \( d' \)) is the inverse behavior of the usual contrast response function where \( d' \) is a function of contrast.

We have written Equation 1 using \( N_{\text{eq}} \) a parameter from the LUM. From the numerator of Equation 16 of Lu and Dosher (1999), we see that \( N_{\text{eq}} \) can be written in terms of PTM parameters as

\[
N_{\text{eq}}^2 = N_A^2/(1 + N_M^2), \tag{3}
\]

where \( N_A \) is the magnitude of the PTM additive noise, \( N_M \) is the magnitude of the PTM multiplicative noise, and the exponent \( p = 2\gamma \) where \( \gamma \) is the power of the nonlinear stage of the PTM. Note that Lu and Dosher use the same convention we do with \( N \) standing for noise contrast rather than noise power (for Pelli, on the other hand, \( N \) means power).

Our approach can be thought of as a black box modification of Equation 1

\[
\text{Th}(N_{\text{ext}}, d') = \text{Th}_0 \text{ICRF}(d')(1 + a_{\text{cons}}N_{\text{ext}} + (N_{\text{ext}}/N_{\text{eq}})^2)^{1/2}. \tag{4}
\]

We have replace the exponent \( p \) of Equation 1 with \( p = 2 \) since that Minkowski exponent is very hard to pin down from the shape of the \( \text{TvN} \) curve. We have also added a consistent noise term, arbitrarily chosen to add linearly, that may contribute, especially at low levels of external noise (Levi et al., 2005). Further clarification of Equation 4 with plots is presented in Appendix A.

The discussion above is relevant to the recent study of Xu et al. (2006). They measured orientation discrimination in white noise and applied the PTM model of (Dosher & Lu, 1998, 1999). In agreement with Levi and Klein (2003) and with the present study, they concluded that amblyopes have deficient perceptual templates for their task. However, they also suggest that they have raised levels of additive internal noise. This is consistent with the elevated thresholds for noise detection that we recently reported (Levi et al., 2007); however, as noted above, the question of whether the noise is additive or multiplicative is model dependent, and we prefer to remain agnostic at this stage. Our present results also show that the raised internal noise is not simply a consequence of using a fixed level of external noise contrast.

It could well be that the shape of the transducer function is similar at low and at high noise levels because of a similar amount of uncertainty. It could also be the case that the shape of the transducer function at low and high noise levels is different. To our knowledge, that hypothesis has never been adequately tested since previous tests (Eckstein et al., 1997; Lu & Dosher, 1999) were done over too small a contrast range to have sufficient statistical power.

We used our rating scale signal detection methodology to determine the shape of the transducer function (the exponent relating \( d' \) to stimulus contrast, also called the ICRF in Equation 1) for each observer and condition in the present study. We found no significant or systematic differences across the three spatial frequency ranges, so we averaged them. These exponents are plotted as a function of noise contrast (in NTU) in Figure 15. While the exponents are substantially higher (\( \approx 2.5 \)) when there is no noise, in nonzero external noise the exponent is \( \approx 1.3–1.6 \) and shows only a slight change as external noise increases from 2 to 20 NTU. The sharp decrease in the CRF exponent when noise is present violates the separability of the class of models (PTM, LUM) specified by Equation 1.
We note that the high exponent may be specific to our detection task and stimulus and may not be true for the orientation identification task used by Xu et al. (2006) and by Lu and Dosher (1999).

The template for signals in noise

Our classification images reveal the observer’s decision template, showing which aspects of noise influence our observers’ responses. Our results show the amblyopic template, like that of normal observers, is band-pass in shape, and in some (but not all) amblyopes it is tuned to lower spatial frequencies than that of either the fellow eye or of the normal control observers. In this respect, there is considerable individual variation. Some observers (e.g., SM) show a marked leftward shift in peak spatial frequency (toward lower spatial frequencies), while others (e.g., SC) if anything show either no shift or a shift toward slightly higher spatial frequencies (in the 2–22 c/deg range). Most importantly, our measurements and modeling show that the amblyopic template is only slightly less efficient than that of normals for low spatial frequency stimuli but significantly less efficient at the highest frequencies. Because the template efficiency is more or less independent of noise level, it evidently produces internal noise that is proportional to the external noise. Most importantly, our results show that the amblyopic template for detecting a signal in noise only accounts in part for the amblyopic loss.

Internal noise and response consistency

We used an N-pass method (Burgess & Colborne, 1988; Gold et al., 1999; Green, 1964; Levi et al., 2005, 2007) to estimate the ratio of random to consistent noise in the observers’ visual system. Our results, consistent with our previous work using signals known exactly in fixed contrast noise, show that increased random noise is a critical factor limiting amblyopic performance in detecting and discriminating the position of signals in noise (Levi & Klein, 2003). The present results provide further support that both the observer’s template and random internal noise limits detection of signals in noise.

We do not yet understand the origin of the high fraction of random noise in the amblyopic visual system. Earlier we (Levi & Klein, 2003) suggested the increased fraction of random noise in the amblyopic cortex might be a consequence of a variable or noisy template (McIlhagga & Paakonnen, 1999). Noisy templates can be achieved in a variety of ways, e.g., by including randomly selected, but irrelevant, neurons (Shadlen et al., 1996) or by uncertainty (Pelli, 1990) in which a multiplicity of mechanisms (e.g., shifted templates) are monitored. We pointed out that a multiplicity of shifted templates would lead to a broader template, particularly at high spatial frequencies and would degrade the bar position discrimination task more than the bar detection task and importantly would lead to an increased proportion of internal noise.

We note that uncertainty would change as we go from low to high external noise and that would account for the changes we observe in the transducer exponent; however, there may still be uncertainty at high external noise levels in that the observer may not know which of the multiple internal highly active templates to attend to.

Indeed, to test this point we also measured response consistency (using our N-pass method) for discriminating

---

Figure 15. Top: The exponent relating $d'$ to stimulus contrast versus external noise contrast (NTU). The lines are the best fit to the mean normal (grey dotted), NAE (black dashed), and amblyopic (black solid) data. Bottom: exponent in noise vs. exponent in no noise. The numbers after each symbol show the number of observer/conditions represented in the mean data.
the position of the DFP (using the same visibility-scaled noise and setting the target contrast at 2.5 times its detection threshold in noise). Our position discrimination task is described elsewhere (Levi & Klein, 2002, 2003). Figure 16 compares for each amblyopic observer $q^2$ (the ratio of consistent to total noise power) for detection (lower abscissa) versus $q^2$ for position (left ordinate). The top and right axes show the proportion of random noise power $(1 - q^2)$. The open symbols are normal controls (circles) and the non-amblyopic eyes. Solid symbols are amblyopic eyes. DFP spatial frequency range is coded by symbol size.

As noted in the Introduction section, a large number of previous studies have used external noise in an attempt to better understand the internal noise that limits performance in the amblyopic visual system (Levi & Klein, 2003; Nordmann et al., 1992; Pelli et al., 2004; Wang et al., 1998; Watt & Hess, 1987; Xu et al., 2006). In most of these studies, external noise was added to a stimulus, and the observer’s task was to detect or to identify the stimulus. In some studies (e.g., Xu et al., 2006), the amount of external noise was varied; however, ours is the first study to equate noise visibility in normal and amblyopic eyes.

**Noise visibility**

The amblyopic visual system shows reduced sensitivity to white noise, even when the components of the noise are within the band-pass of the amblyopic eye (Levi et al., 2007). Previous studies of the detection of signals in noise in amblyopia have used a single fixed rms noise contrast (e.g., Levi & Klein, 2003; Pelli et al., 2004). The visibility of the noise would be different for normal and amblyopic observers and would vary amongst observers depending on the degree of amblyopia. As shown in Figure 12, the ratio of consistent to total noise depends strongly on noise visibility. If the external noise was close to the observer’s threshold, one might mistakenly conclude that they had high degrees of random noise relative to normal observers for whom the noise is highly visible (Levi & Klein, 2003). Our results show that even after accounting for noise visibility, many of the amblyopic observers show lower values of $q^2$, indicating increased random noise, particularly for high spatial frequencies.

It is interesting to compare the present results for detection of signals in noise with our earlier results for detection and discrimination of white noise (Levi et al., 2007) in the same observers. Our noise detection/discrimination task relies only on noise energy, so position uncertainty (and shifted templates) would not be expected to have any impact on performance. As noted above, a multiplicity of shifted templates could lead to a broader template (McIlhagga & Paakonnen, 1999), which would degrade the detection of a bar as shown by Levi and Klein (2003) and by the present study using localized patterns, leading to an increased proportion of internal noise. Moreover, this explanation also provides a natural account for the strong effect of spatial frequency shown here.

**Mechanisms of amblyopia**

The three most frequent explanations for the losses seen in humans with amblyopia are as follows: (1) a spatial scale shift, i.e., a loss of contrast sensitivity at high spatial frequencies, consistent with the loss of contrast sensitivity of small (high spatial frequency) receptive fields in area V1 in monkeys with experimental amblyopia (Kiorpes,
Kiper, O’Keefe, Cavanaugh, & Movshon, 1998). This explanation, which has found broad agreement, is consistent with a coarse (low spatial frequency) template for detection. However, the loss of neural contrast sensitivity is too small to fully account for the behavioral losses of contrast sensitivity in monkeys with amblyopia, and the spatial scale shift hypothesis also cannot fully explain the loss of position acuity in humans with strabismic amblyopia (Levi & Klein, 1982; Levi et al., 1994). To account for this “additional” loss, two other explanations have been suggested: (2) cortical undersampling (i.e., a reduced complement of cortical neurons; Levi, 1991; Levi & Klein, 1986) and (3) uncalibrated topographical jitter (i.e., mis-wiring of cortical neurons; Field & Hess, 1996; Hess, 1982). There has been considerable debate surrounding undersampling versus jitter (Field & Hess, 1996; Levi & Klein, 1996; for a review, see Kiorpes & McKee, 1999). Our finding of increased random multiplicative noise offers a new and different account, which may lead to déténte. If the information that V1 neurons of the amblyopic cortex send to higher levels is subject to random noise, from trial to trial different samples of the target will be more or less effective. On one trial one set of neurons might provide reliable signals, while on the next trial, a different set of neurons might provide reliable signals from the same target. The net effect would be equivalent to combining undersampling with positional jitter. Indeed, in previous work (Levi, Klein, & Wang, 1994), we showed that strabismic amblyopes show a uniform loss of Vernier acuity over the entire range of target contrasts (a multiplicative loss). The only stimulus manipulation that produced this pattern in normal observers consisted of undersampling the target combined with random positioning of the samples from trial to trial. Undersampling the stimulus in a regular predictable way does not mimic the multiplicative pattern of loss, and we had suggested that strabismic amblyopia may involve elevated levels of central noise. As noted earlier, the effect of amblyopia may be a combination of early noise (which affects detection of both noise and signal) and late noise due to a template-mismatch and random noise. This late noise may help to explain why the behavioral losses of contrast sensitivity are greater than the neural losses in V1 (Kiorpes, 2006). Our present study adds several new pieces to the puzzle: First, it quantifies this noise and parcels it into three parts—random noise, template noise, and consistent noise not due to the poorly matched template. Secondly, it shows that the predominant noise in the amblyopic visual system is random internal noise. Consistent noise (beyond the poorly matched template) appears to play little or no limiting role in either normal or amblyopic detection of signals in noise.

We do not yet understand the origin of the high degree of random internal noise. However, our finding that both strabismic and anisometropic amblyopes show high random noise when viewing with both the amblyopic and the preferred eyes suggests that this elevated random noise is central and is likely related to the absence of correlated binocular visual experience early in life (Kind et al., 2002; McKee, Levi, & Movshon, 2003).

Summary and conclusions

Our results show that amblyopes have reduced root efficiency for detecting known signals in noise. The loss of efficiency increases with the spatial frequency of the signal. This loss is explained in part by reduced template efficiency, but to a greater extent by increased random internal noise. This efficiency loss is more or less independent of external noise contrast over a log unit range of external noise. Combining threshold measurements with trial-by-trial analysis (Green, 1964) allows us a unique way of dissecting the sources of noise in the visual nervous system that limit amblyopic vision.

Appendix A

It is instructive to present analytic examples of our various types of plots for clarity.

We consider four TvN curves.

Ideal observer: \( I(N_{\text{ext}}) = a_I N_{\text{ext}} \) \hspace{1cm} (A1a)

Template observer: \( T(N_{\text{ext}}) = a_T N_{\text{ext}} \) \hspace{1cm} (A1b)

Consistent observer: \( C(N_{\text{ext}}) = \sqrt{a_C N_{\text{ext}} + (a_T N_{\text{ext}})^2} \) \hspace{1cm} (A1c)

Human observer: \( H(N_{\text{ext}}) = \sqrt{C(N_{\text{ext}})^2 + a_R^2 N_{\text{ext}}^2} \) \hspace{1cm} (A1d)

where \( a_I \) is obtained by a mathematical ideal observer calculation. For detection of signals in white noise (the present situation given our stimuli), the matched filter template is ideal so the calculation is quite simple (Levi & Klein, 2002). The value \( a_T \) is given by \( a_T = a_I r \), where \( r \) is the correlation between the ideal template and the human template. The values of \( a_C, a_R, \) and \( a_C \) are obtained from measuring the human threshold and from measuring the ratio of consistent to total noise in a multi-pass experiment. The human threshold at zero external noise is a measure of \( a_0 \). The parameter \( a_C \) is determined by ratio \( C(N_{\text{ext}}) / H(N_{\text{ext}}) = q(N_{\text{ext}}) \), where \( q^2 \) is the correlation between responses of the two passes of a double pass experiment, as discussed in the text. The forms of the “black box” curves in Equations A1c and A1d were...
chosen for their simplicity and their compatibility with the data rather than because of any deeper principle.

We have taken the consistent noise term in Equation A1c to come in linearly rather than quadratically to show that the black box can be complicated with various multiplicative noise contributions contributing with different powers and interaction terms (not shown). The assumptions of independent Gaussian multiplicative noise made by Lu and Dosher (1999) were made to get simple results rather than out of a belief that the assumptions were correct, as they pointed out. We chose to have the consistent noise contribute linearly rather than quadratically because some of our earlier data, for a different task (Levi et al., 2005), indicated that the consistent noise contribution was stronger at low levels of external noise and was negligible at higher levels.

Note that Equation A1d includes a term from Equation A1c. The connection of Equations A1c and A1d to Equation 4 is

\[
a_0 = 1 \quad \text{(that is, we factor } a_0 \text{ into } Th_0 \text{ of Equation 4)};
\]

\[
a_{cons} = a_C^2;
\]

\[
1/N_{eq} = a_T^2 + a_R^2.
\]

For Figure A1, we chose the parameters to be

\[
a_i = 0.09 \quad \text{(a 10% Weber fraction)}
\]

\[
a_T = 0.1 \quad \text{(a 10% Weber fraction for the consistent part of the human threshold at high noise levels)}
\]

\[
a_C = 0.2 \quad \text{(this term contributes mainly at low noise levels)}
\]

\[
a_0 = 1 \quad \text{(an additive input noise = 1 means the human threshold = 1 for } N_{ext} = 0).
\]

\[
a_R = 0.1 \quad \text{(the internal random noise is about equal to the internal template noise)}.
\]

(A2)

Figure A1. Response to external noise at four stages of visual processing. The curves in panels a–c are the TvN curves for the ideal, template, consistent and human observer specified in Equation A1, with parameter values specified in Equation A2. The human threshold at zero external noise is, for example, given by \(a_0 = 1.0\). The ordinate in panels a–c is the linear (panel c) or quadratic (panels a and b) response activity. The abscissa in panel a is the external noise squared on in contrast squared units. In panels b–d, the units are normalized to noise threshold units (NTU) based on our finding that the detection threshold for seeing noise is at \(N_{eq}\) the kink point (close of \(N_{eq}\) of Equation 1) of the TvN curve on log axes. The ordinate for panel d is root efficiency, with the panel d curves simply the ideal curve of panel c divided by the template and total curves.
At asymptotically low \( N_{\text{ext}} \), the human threshold is \( a_0 = 1 \). At asymptotically large \( N_{\text{ext}} \), the human threshold is 
\[
(a_I^2 + a_R^2)^{1/2} = 0.141 N_{\text{ext}}.
\]
Thus, the asymptotic behavior is captured by the function

\[
\text{Thresh} = a_0 (1 + (a_I^2 + a_R^2) N_{\text{ext}}^2)^{1/2} = a_0 (1 + N_{\text{ext}}^2 / N_{\text{eq}}^2)^{1/2},
\]
where \( N_{\text{eq}} = 0.02^{-1/2} = 7.1 \) is the equivalent input noise. Thus, the kink point of the TvN curve of Figure A1a is defined to be at an external noise square of 
\[
1 / (a_I^2 + a_R^2) = 50.
\]

As discussed in the text, \( N_{\text{eq}} \) has been found to be very close to the threshold for seeing the noise. We have found it useful for many of our plots to have the abscissa expressed in noise threshold units. That effectively places \( N_{\text{eq}} = 1 \). Panel b of Figure A1 is identical to panel a, except that the \( x \) axis has been shifted by a factor of 50 so that the abscissa is expressed in \( N_{\text{eq}} \) units (noise threshold units squared). Panels c and d are also expressed in NTU (but not squared).

Panel c is a replot of panel b, on log–log axes. This is the type of plot shown in the text. It has the advantage of making the shape of the low \( N_{\text{ext}} \) regime more visible. For example, the slightly shallower slope of the consistent noise contribution becomes visible. The log–log axes have the disadvantage of masking the near linearity of TvN functions when plotted as threshold squared vs. noise squared as in panels a and b. Panel d is similar to panel c except that it is root efficiency that is on the ordinate. Root efficiency is the ratio of the thresholds to the ideal observer’s threshold. It can be seen that above \( 2 N_{\text{eq}} \), the root efficiencies are approximately constant.

A word about the units of panel a may be useful. The external noise and the four observer responses are all in the same rms contrast units. For \( N_{\text{ext}} \) and the ideal and template observers, the rms power is easily obtained because there is a direct connection between the template output (an ideal template for the ideal observer) and \( d' \).

One of the more confusing aspects of our approach is the specification of the consistent noise in terms of our double pass methodology. Figure A2 may be helpful in clarifying our procedure. The plot represents the internal activity of the visual system on which the decision is based. The two axes are the activity in the first and second pass. The inner blue circle is the activity produced by the random internal noise; the red ellipse is the blue activity plus the observer’s template. The black ellipse is the red activity plus any further consistent activity not specified by the template response.

The left and right hand panels are for the cases of minimal and moderate amounts of external noise. In the right panel, the inner circle and the ellipses are substantially larger than those of the left panel. The larger size is caused by the increased external noise together with the system’s multiplicative noise. The elongation of the

---

Figure A2. The circles and ellipses depict bivariate Gaussian density distributions for double pass response activity. The abscissa represents the response activity for the first pass (the quantity depicted in Figure A1 above), and the ordinate is the similar activity for the second pass. The left panel is for very weak external noise, where there is minimal activity difference in the two passes. The right panel has stronger activity because it is for the case of stronger external noise. The blue, red, and black ellipses represent the noise distributions for the ideal, template, and human curves. The magenta ellipse is for the signal plus noise distribution of the human response. The vertical and horizontal dashed lines represent the criteria in our double pass, rating scale signal detection method.
ellipses indicates that the consistent portion of the neural activity is greater than the random portion.

The vertical and horizontal dashed lines represent the observer’s criteria that are used to convert the floating point response activity to integer levels that correspond to which of the five response buttons get pressed. Our signal detection methodology is able to determine the location of the response criteria and relative separation of the signal (magenta ellipse) from the noise (black ellipse). The correlation of the button press responses gives a good approximation to the elongation of the ellipse that provides a measure of the ratio of consistent to random noise.

**Acknowledgments**

This research was supported by Research grants R01EY01728 and R01 EY04776 from the National Eye Institute, NIH. We are grateful to Hope Queener for developing the processing tools and to Roger Li for helpful comments on the paper.

Commercial relationships: none.
Corresponding author: Dennis M. Levi.
Email: dlevi@berkeley.edu.
Address: School of Optometry, University of California, Berkeley, CA 94720-2020.

**References**


