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

The quantum efficiency of vision

D. G. Pelli
Institute for Sensory Research
Syracuse University
Syracuse, NY13244-5290
U.S.A.

In *Vision: Coding and Efficiency*, Edited by Colin Blakemore, Cambridge University Press, Cambridge, pp.3-24, 1990.

1

The quantum efficiency of vision

How did you come to be interested in quantum efficiency?

H. B. Barlow: I was an undergraduate and doing Part II Physiology [at Cambridge University in the early 1940s]. I was directed to R. A. Fisher's book [1925, *Statistical Methods for Research Workers*]. The idea of statistical inference seemed relevant to understanding what the brain was doing. . . . It gave an objective handle to something the brain did rather well.

[August 27, 1987]

D. G. Pelli

Attempts to understand the quantum efficiency of vision have resulted in three distinct measures of efficiency. This chapter shows how they fit together, and presents some new measurements. We will show that the idea of equivalent input noise and a simplifying assumption called 'contrast invariance' allow the observer's overall quantum efficiency (as defined by Barlow, 1962a) to be factored into two components: transduction efficiency (called 'quantum efficiency of the eye' by Rose, 1948) and calculation efficiency (called 'central efficiency' by Barlow, 1977).

When light is absorbed by matter, it is absorbed discontinuously, in discrete quanta. Furthermore, it is absorbed randomly; the light intensity determines only the probability of absorption of a quantum of light, a photon (Einstein, 1905). This poses a fundamental limit to vision; the photon statistics of the retinal image impose an upper limit to the reliability of any decision based on that retinal image. An observer's overall *quantum efficiency* F is the smallest fraction of the corneal quanta (i.e. quanta sent into the eye) consistent with the level of the observer's performance (Barlow, 1958b, 1962a). (This is closely analogous to Fisher's (1925) definition of the efficiency of a statistic.) Surprisingly, the overall quantum efficiency of vision is very variable, and much smaller than best estimates of the fraction of photons absorbed by the photoreceptors in the retina.

At all reasonable luminances the fraction of corneal photons that excite photoreceptors is almost certainly quite constant. Barlow (1977) concluded that for rods it must be in the range 11% to 33% (for 507 nm light). This is independent of the size and duration of the signal, and independent of the background luminance, up to extremely high luminances. Yet, Barlow (1962b) had previously shown (see Fig. 1.1) that the highest overall quantum efficiency of vision is only 5%, and *that* only for a small brief spot on a very low luminance background. Figure 1.1 shows that increasing the background luminance reduces the overall quantum efficiency by many orders of magnitude.

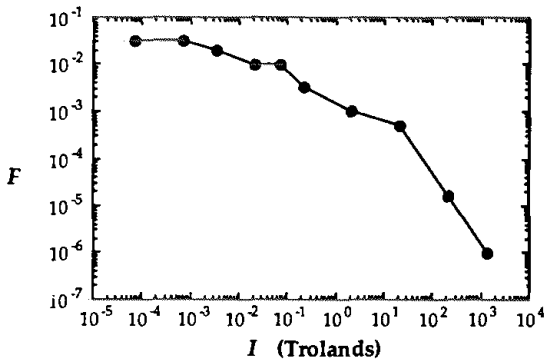

Fig. 1.1. Overall quantum efficiency versus retinal illuminance (in scotopic Trolands) for detecting a 46' diameter spot presented for 86 ms on a uniform background at 15° in the nasal field. Replotted from Barlow (1962b).

Figure 1.2 shows that the overall quantum efficiency is strongly dependent on both the area and duration of the spot. But, if the fraction of photons that excite photoreceptors is constant, why should the experimental conditions so dramatically affect the quantum efficiency?

The question still lacks a full answer, but this chapter will present a framework that will allow simple threshold measurements to distinguish between two classes of explanation for the variation in efficiency. The key idea is the notion of an observer's equivalent input noise, which is analogous to Barlow's (1956, 1957) dark light.

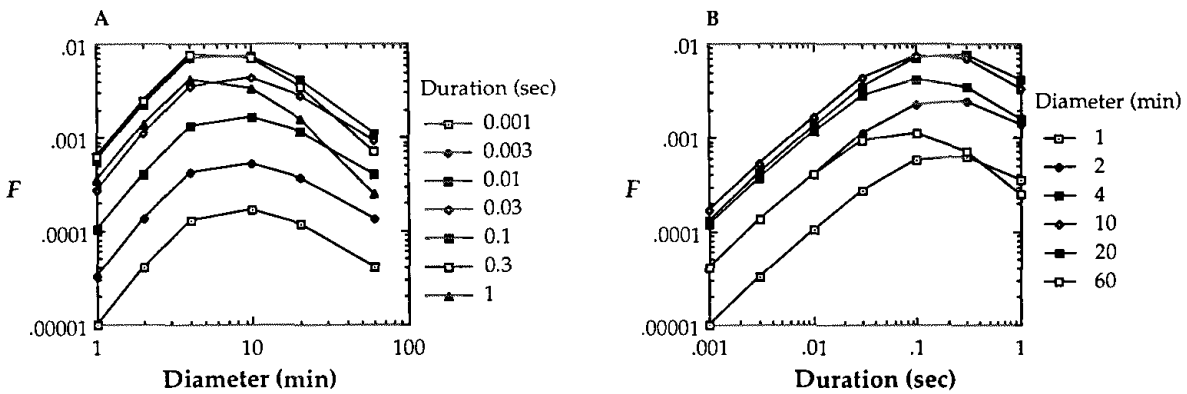

Fig. 1.2A and B. Overall quantum efficiency versus spot diameter and duration on a uniform background of 3.4 cd/m^2. Replotted from Jones (1959).

Barlow (1956) pointed out that Hecht, Shlaer & Pirenne's (1942) estimate of the fraction of photons absorbed was really only a lower bound, because there might be noise in the visual system which added to the photon noise. He went on to show that it was useful to attribute that noise to a 'dark light', which acts in the same way as a physical light (Barlow, 1957). Thus it is easier to understand the effect of scotopic backgrounds if one considers the *effective background* to be the sum of the actual background and the dark light. Graphs of increment threshold versus background luminance, e.g. Fig. 1.3, are consistent with this idea. When the background luminance I is below the level of the dark light I_{dark} then the dark light dominates, and threshold ΔI is constant. When the background luminance I is greater than the dark light I_{dark} then the background I dominates, and threshold ΔI rises as the background luminance rises.

Fortunately for our night vision, the dark light is very small (0.0028 td in Fig. 1.3) and is negligible at any luminance well above absolute threshold. However, it turns out that contrast thresholds in noise behave in an analogous way. At high luminances we usually don't worry about photon noise, but it is increasingly common to artificially add luminance noise to the display, typically by adding a random number to each pixel of an image. One can then measure threshold as a function of the amplitude of the noise. This approach is reminiscent of Barlow's dark light measurements. In fact, this sort of measurement is routine in electrical engineering, where it is called equivalent input noise measurement (e.g. Mumford & Schelbe, 1968). Equivalent input noise measurement dates back at least to the 1940s, and probably much before (North, 1942; Friis, 1944).

Consider traditional equivalent noise measurement. Suppose we are setting up an audio system. We have a record player, or some other audio signal source which inevitably has some noise. Suppose we know how much noise the source has. Now we want to choose an audio amplifier to amplify our signal. We want to choose an amplifier that has very little noise, so that it will reproduce the signal faithfully. Even the best amplifier has some noise. However, we don't want to buy an unnecessarily expensive low-noise amplifier, since it's enough for the amplifier's noise to be small relative to the noise already in our audio source. How shall we evaluate the amplifier? We could just measure the noise at the amplifier's output, but that would not be directly comparable with the noise in our original audio source. Or we could open the

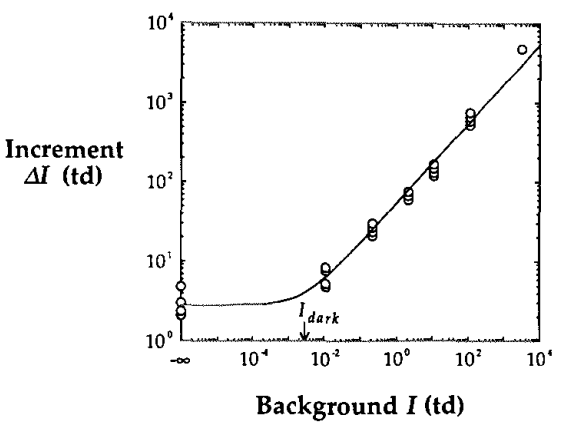

Fig. 1.3. Threshold increment ΔI for a small brief spot as a function of background luminance I, both in scotopic Trolands. The smooth curve represents the relation $\Delta I \propto (I + I_{dark})^{0.5}$, where $I_{dark} = 0.0028$ td. Replotted from Barlow (1957).

amplifier up and measure the noise at some point inside. But where? Nearly all electrical components generate noise.

Figure 1.4 illustrates the standard procedure, which is to measure the output noise level when a series of calibrated noise levels are applied to the input of the amplifier. Figure 1.5 shows a typical set of results, showing output noise level versus input noise level. Both scales are logarithmic, and are calibrated in deciBels. Noise should be described by its spectral density N, which is the power per unit bandwidth. In this example it has units of Watts per Hertz. The meters in Fig. 1.4 measure power, in Watts. The power spectral density N is calculated by dividing the measured power by the bandwidth, in Hertz, of the filter preceding the meter. *White noise* has the same spectral density at all frequencies of interest.

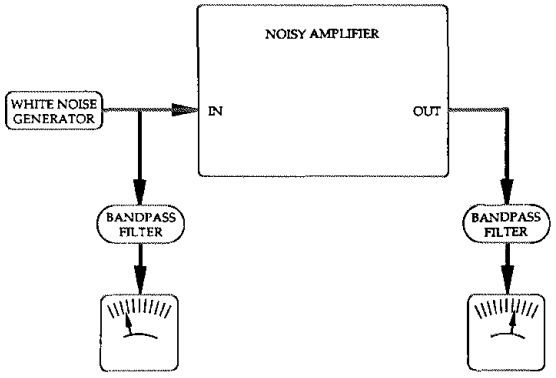

Fig. 1.4. The standard procedure for measuring equivalent input noise. The noise level (power spectral density) is measured at the input and output of the amplifier, at many levels of white noise input (Mumford & Schelbe, 1968). Power spectral density is measured by the combination of a bandpass filter and power meter.

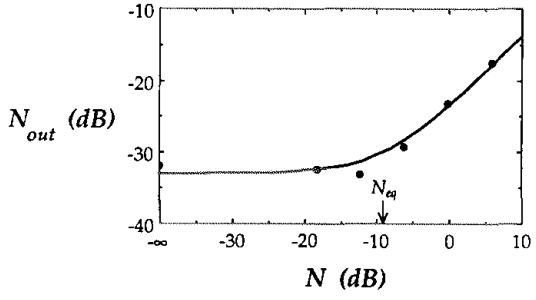

Fig. 1.5. Results from the experiment in Fig. 1.4: output noise level versus input noise level. Both decibel scales are relative to 1 μW/Hz. The smooth curve represents the relation $N_{out} \propto N + N_{eq}$, where N_{eq} is −9 dB re 1 μW/Hz.

In general, both the amplifier gain and the output noise power spectral density are frequency dependent, so measurements like those in Fig. 1.5 are made at each of many frequencies, by changing the center frequency of the bandpass filters. However, we will ignore the complication of frequency dependence for now, and come back to it later. The noise measured at the output contains contributions from both the external noise and the intrinsic noise of the amplifier. Low external noise levels are insignificant compared to the intrinsic noise, so the curve is flat at low external noise levels. At high external noise levels the *external* noise N dominates so the curve rises with a log–log slope of 1, indicating proportionality. The knee of the curve occurs when the effect of the external noise is equal to that of the intrinsic noise. That input noise level is called the *equivalent input noise* N_{eq}.

The smooth curve in Fig. 1.5 represents a simple relation. Output noise level is proportional to the sum of external and equivalent noise levels.

$$N_{out} \propto N + N_{eq} \tag{1}$$

Noises add for the same reason that variances add; the variance of the sum of two independent random variables is the sum of the variances of each variable. The proportionality constant and N_{eq} were determined so as to provide a maximum likelihood fit by Eq. 1 to the data. The value of N_{eq} is indicated by an arrow on the horizontal axis at the knee of the curve. The curve is asymptotically straight on either side of the knee and N_{eq} is the horizontal coordinate of the intersection of the two limiting lines. Note that N_{eq} is independent of the amplifier gain. With a higher gain the curve would be shifted higher on these logarithmic axes, but the horizontal position would be unchanged.

We have just measured the amplifier's intrinsic noise by finding its equivalent input noise. Figure 1.6, which illustrates this idea schematically, is a conceptual model of the situation we saw in Fig. 1.4. This is a *black-box* model, intended to model the input and output behavior of the system under study, without regard for its internal construction. The noisy amplifier is conceived to be made up of a hypothetical noiseless amplifier with a built-in noise generator at its input. The required internal noise level to match the intrinsic noise of the actual amplifier is called the *equivalent input noise*.

Thus we can measure the intrinsic noise by finding an equivalent input noise. This is sometimes called 'referring the amplifier's noise to its input'. Essentially the same approach can be applied to

vision. Indeed, this is analogous to Barlow's (1957) dark light measurements. By a similar analogy we can apply this idea to the contrast domain.

In order to measure the observer's equivalent input noise we need an output power meter. For this we will use the observer's squared contrast threshold, since it is well known that the squared contrast threshold is proportional to the noise level at the display (e.g. Stromeyer & Julesz, 1972; Pelli, 1981). The idea here is that, instead of building our own power meter, we use one that is already built into the visual system under study.

Figure 1.7 shows the model for the human observer. Like Fig. 1.6, this is a black-box model, intended to correctly model the behavior, but not the internal construction, of the observer. The stimulus to the observer is the sum of signal and noise. The signal is present to provide a dependent measure analogous to output noise level, as we'll see below. Note the fat arrows. They indicate transmission of a spatiotemporal function, a time-varying image, which could be represented by a three-dimensional array of numbers. For example the signal might be a flickering grating or a television program. The black-box model has an equivalent input noise ('contrast-invariant noise') which adds to the stimulus. We will call the sum of the stimulus and the equivalent input noise the *effective* stimulus. Finally, there is a transformation, 'calculation', which somehow reduces the spatiotemporal *effective* stimulus down to a simple decision, such as 'yes' or 'no', which can be represented by a single number, as indicated by the thin arrow. The observer's task is to detect the signal. The observer's level of performance is given by the average accuracy of the decisions, e.g. percent of decisions which are correct in a two-interval forced choice paradigm (i.e. two presentations in random order, one is signal plus noise, the other is noise alone; the observer is asked which one has the signal). An experiment consists of determining threshold, the squared contrast c^2 of the signal which yields a criterion level of performance, e.g. 75% correct.

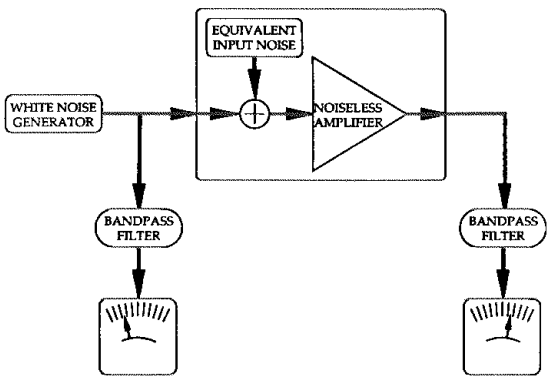

Fig. 1.6. A black box model for the noisy amplifier in Fig. 1.4.

In order to make the model as general as possible, yet still be able to measure its parameters, we need *three assumptions*, or constraints. First we assume that the observer's level of performance increases monotonically with the contrast of the signal (when everything else is fixed). This guarantees that there will be a unique threshold. Secondly, as indicated on the diagram by the prefix *contrast-invariant*, we assume that the calculation performed is independent of the

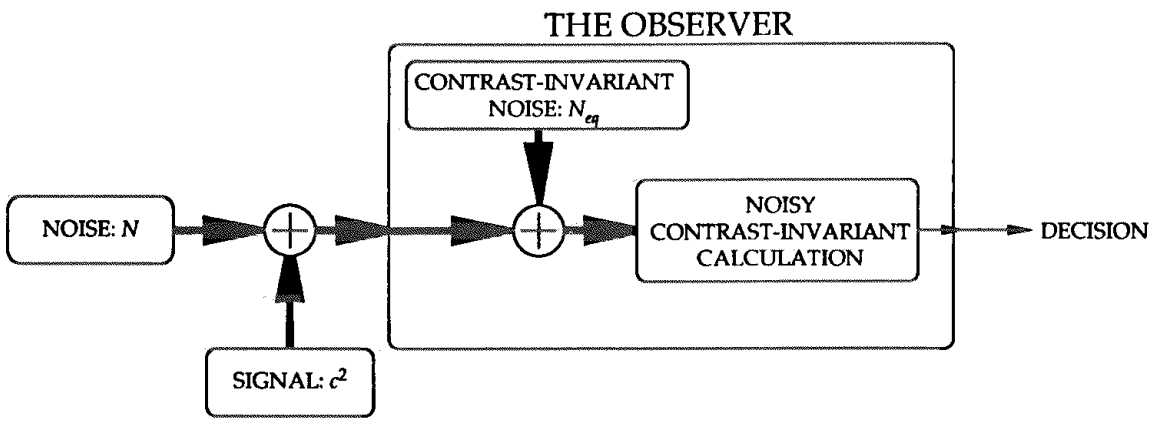

Fig. 1.7. A black box model for the human observer. The stimulus consists of the sum of noise with nose level N and a signal with a squared contrast c^2. The fat arrows indicate transmission of a time-varying image. The thin arrow indicates transmission of a scalar, the observer's decision.

contrast of the effective stimulus, which is its immediate input. Together, assumptions 1 and 2 are a linking hypothesis. They imply that the observer's squared contrast threshold (which we can measure) is proportional to the effective noise level (which is inaccessible). Thirdly, we assume that the equivalent input noise is independent of the amplitude of the input noise and signal, or at least that it too is contrast-invariant, independent of the contrast of the effective image. These assumptions allow us to use two threshold measurements at different external noise levels to estimate the equivalent noise level. In effect, the assumptions state that the proportionality constant and the equivalent noise N_{eq} are indeed constant, independent of the contrast of the effective image. These three assumptions are just enough to allow us to make psychophysical measurements that uniquely determine the parameters of our black-box model. Our model makes several testable predictions, as will be discussed below.

I should point out too, that while this model is required in order to measure the equivalent input noise (and 'transduction efficiency'), we will see below that the model is not required for measurement of either the 'calculation' or overall quantum efficiency of the observer, as these latter measurements are assumption-free. Those measures are assumption-free partly because each can be determined from a single threshold measurement. Measurement of an observer's equivalent noise level requires two threshold measurements and a linking hypothesis (our black-box model) to relate the two measurements.

Note that we have redefined *equivalent input noise* from its usual meaning in which it reflects all the noise in the system under study, to now only reflect noise that can be described by a contrast-invariant input noise. This subtlety is irrelevant in many simple systems, such as the linear amplifier in Figs. 1.4 and 1.6, where all the internal noise is contrast-invariant, but in general nonlinear systems can have an internal noise that is contrast-dependent (Legge, Kersten & Burgess, 1987; Ahumada, 1987). However, our model does allow the calculation to be noisy, provided the calculation is still contrast-invariant. This means that any noise introduced within the calculation would have to be proportional to the effective noise level at the input of the calculation. This would happen, if, for example, the observer used a randomly chosen receptive field on each trial to detect the same signal (Burgess & Colborne, 1988). Using the terminology of Lillywhite (1981), 'additive' noise (independent of the effective noise) is assigned by our model to the contrast-invariant noise, and 'multiplicative' noise (proportional to the effective noise) is assigned to the noisy calculation. From here on, *equivalent input noise* refers solely to the contrast-invariant input noise of our black box model in Fig. 1.7.

A wide range of models are consistent with our black box because they are contrast-invariant, e.g. all the ideal receivers derived by Peterson, Birdsall & Fox (1954) for various kinds of signal (known exactly or only statistically) in white noise, the uncertainty model shown by Pelli (1985) to provide a good model for many aspects of visual contrast detection and near-threshold discrimination, and the ideal discriminators of Geisler & Davila (1985).

Some popular models for detection, e.g. ones incorporating a high fixed threshold, are *not* scale invariant and are incompatible with our black-box model. However, these models are also incompatible with the empirical effects of noise (Pelli, 1981, 1985) and criterion effects (Nachmias, 1981). The desirable aspects of high-threshold-like behavior can arise from scale invariant models (Pelli, 1985) so incompatibility with such models is not a cause for worry.

Figure 1.8A shows a grating in white noise. The luminance of each pixel has an added random variation. In my experiments, these variations are dynamic and uncorrelated over space and time. That is what is meant by *white noise*. I measured contrast thresholds for the grating at many different noise levels, keeping the mean luminance of the display constant. Figure 1.8B illustrates another way to do this experiment, using a random dot display, such as used by Rose (1957) and van Meeteren & Boogaard (1973). A random-dot display models the way that photons are absorbed in the retina. The random-dot display in Fig. 1.8B appears very similar to the additive-noise display in Fig. 1.8A, but now the display consists solely of white dots on a black background, and the grating simply modulates the probability of occurrence of the dots. The most common random-dot display is the image intensifier, which is used to see at night (van Meeteren & Boogaard 1973).

> Once detected by the photocathode . . . photons can be made visible as bright specks on an image screen by electronic amplification. When the detected photon flux is sufficiently high, these specks combine in space and time to form a normal, smooth image. At lower light levels in object space, however, the detected photon flux can be so small that the corresponding specks are visible as such, especially when the amplification is high.

Many people have used such displays to 'bypass early levels of processing'. Barlow (1978) discussed,

> the demonstration by French, Julesz, Uttal and

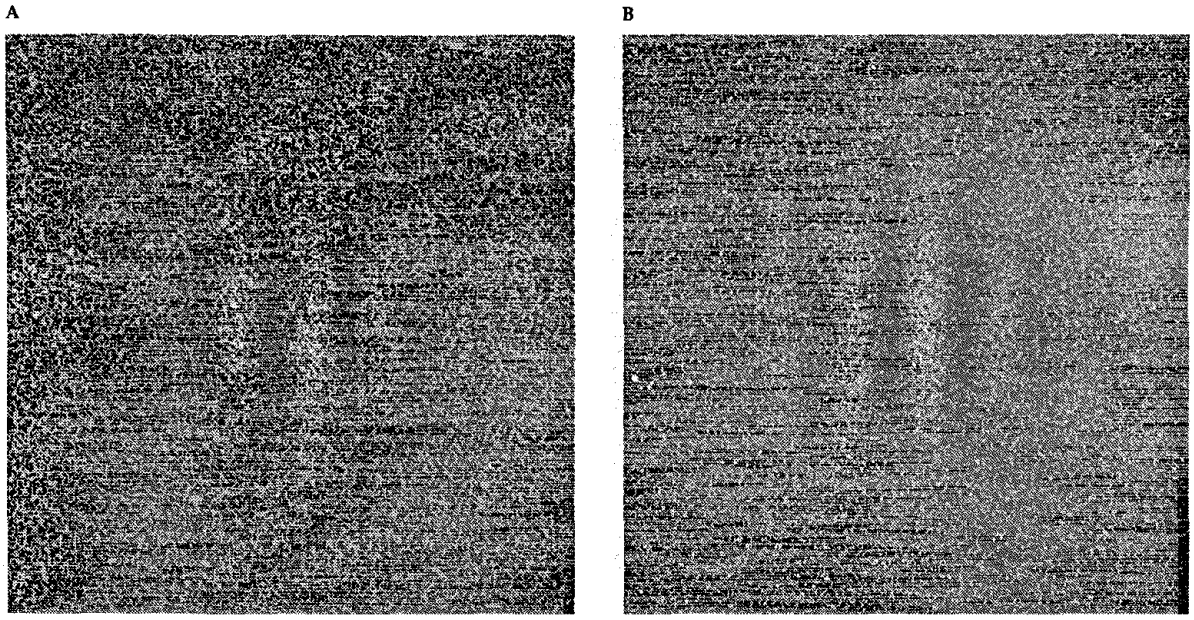

Fig. 1.8. Two ways to display a grating in noise. These illustrations are static, but in a real experiment the noise would be dynamic. The signal is a sinusoidal grating with 0.4 contrast, vignetted by a circularly symmetric Gaussian with a space constant of 1.5 grating periods, i.e. $s(x, y) = 0.4 \exp(-[(x+y)/1.5]^2) \sin(2\pi x)$. The main difference between these displays is that in A the variance is independent of the signal, and in B it is dependent, proportional to the expected luminance at each point. A. An additive display: each cell is the sum of signal and independent noise. The reflectance of each halftone cell is $(1 + s(x, y) + X_{x,y})L_{av}$, where $X_{x,y} = \pm 0.3$ is a random sample from a binomial distribution, determined independently for each cell, and the mean luminance corresponds to a 50% reflectance, $L_{av} = 0.5$. B. A random-dot display: each cell is a sample from a Poisson distribution whose mean is modulated by the signal. The reflectance of each halftone cell in the image is $X_{x,y}/11$, where $X_{x,y}$ is a random sample from a Poisson distribution with mean $11(1 + s(x, y))L_{av}$. The Poisson distribution was approximated by the sum of 11 samples from a binomial distribution yielding 1 with probability $(1 + s(x, y))L_{av}$ and 0 otherwise.

> others that patterns composed largely of random dots provide an opportunity to probe intermediate levels of processing in the visual system. The dots which compose the pattern are, it is suggested, reliably transduced and transmitted by the lower levels, and the limitations of performance result from the ways in which the nervous system can combine and compare groups of these dots at the next higher level.

It turns out that we can analyze both kinds of experiment (Figs. 1.8A and B) in the same way: the contrast threshold of the grating as a function of the spectral density of the noise (Pelli, 1981). Figures 1.8A and B show that although the methods of generation of the two kinds of display are quite different, the two displays look very much alike when they have the same noise level and luminance.

Since each experiment is done at a constant average luminance, it is convenient to define our measures of signal and noise in terms of the *contrast function*, which is a normalized version of the luminance function,

$$c(x, y, t) = \frac{L(x, y, t)}{L_{av}} - 1 \tag{2}$$

where (x, y, t) is a time and place on the display and L_{av} is the mean luminance (Linfoot, 1964). A complete set of definitions and a formal derivation of the theoretical results of this chapter appear in Appendix A.

Power spectral density N is the contrast power per unit bandwidth. Imagine filtering the image with a narrowband filter with unit gain in the passband. The *contrast power* of the filtered image is c^2_{rms}, the mean of the square of the contrast function. The ratio of the power to the bandwidth is an estimate of the power spectral density of the original image at the center frequency of the filter.

Figure 1.9 presents a typical set of results obtained with displays like those in Fig. 1.8. It shows the squared contrast threshold for a 4 c/deg grating in the presence of a noise background. The vertical scale is the contrast threshold. The horizontal scale is the noise level. All of the points are within two standard errors of the smooth curve. The curve represents proportionality of squared contrast threshold to the sum of the external noise and the equivalent noise.

$$c^2 \propto N + N_{eq} = N_{ef} \qquad (3)$$

At *low* noise levels the squared contrast threshold is nearly constant. At *high* noise levels the threshold rises in proportion to the noise level. This entire experiment can be summarized by Eq. 3 and its two fitted parameters: the observer's equivalent input noise N_{eq} (which we want to know) and the proportionality constant (which is not of interest here).

In the same way that earlier we talked of the *effective* luminance being the sum of the stimulus light and the dark light, we will now speak of the *effective* noise level N_{ef} being the sum of the displayed noise and the observer's equivalent noise.

It may not be obvious that Eq. 3 is dimensionally correct, nor why the contrast is squared. We are interested in squared contrast (i.e. contrast power), rather than contrast, because we are dealing with independent random quantities, so their variances add, and variance is proportional to the square of the amplitude. Recall that N represents power spectral density, which is proportional to the contrast power of the noise. c^2 is proportional to the contrast power of the signal. Thus, Eq. 3 basically says that the contrast power of the signal at threshold is proportional to the contrast power of the effective noise.

In our notation, upper case symbols like N are generally proportional to squared contrast and lower-case symbols like c are generally proportional to contrast.

Equation (3) is the principal prediction of our model. The model explicitly assumes that the performance is independent of the overall contrast of the effective image, so at a constant level of performance the squared contrast of the signal is proportional to the effective noise level, $c^2 \propto N_{ef}$. Equation (3) is not as direct as the analogous Eq. (1) used in electrical engineering because we cannot directly measure the observer's 'output' noise. Instead we measure a squared contrast threshold and assume that it is proportional to the observer's effective noise level (the 'output' noise). Equation (3) is obeyed by *every* published record or contrast threshold versus noise level at a constant luminance (Pelli, 1981).

Consider Figure 1.10, which shows thresholds for detection of a 4 c/deg grating at two luminances, 3.3 and 330 cd/m^2 (pupil size was not measured). Both conditions are well fit by Eq. 3. The effect of luminance on the two parameters is interesting. In the absence of noise the threshold is higher at the lower luminance. However, as the noise level increases, the curves eventually converge, so that at high noise level, the

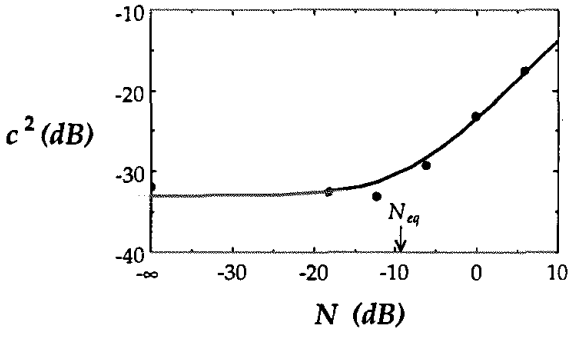

Fig. 1.9. Squared contrast threshold (in dB) versus noise level (in dB re 1 μs deg^2). The smooth curve represents the relation, $c^2 \propto N + N_{eq}$, where N_{eq} is −9 dB re 1 μs deg^2. Note that a deciBel is one tenth of a log unit change in *power*. Thus there are 10 dB per decade of squared-contrast and noise level, since they are both proportional to power.

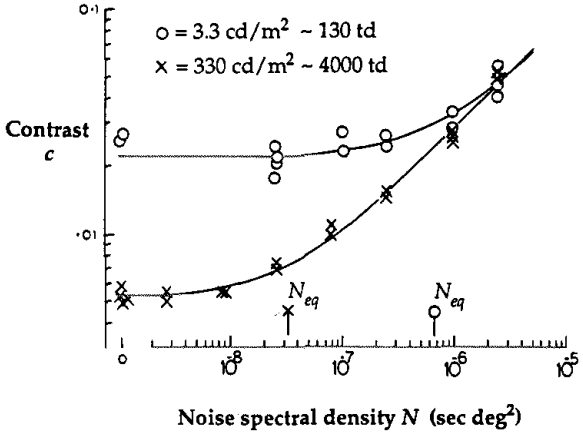

Fig. 1.10. Contrast threshold for a grating in noise at two luminances: 3.3 cd/m^2 and 330 cd/m^2 (about 130 and 4000 td, respectively). The smooth curve represents Eq. 3. The signal was a vertical 4 c/deg grating with a Gaussian envelope with a horizontal space constant of 2°, a vertical space constant of 5°, and a time constant of 70 ms. The noise was white up to 66 c/deg horizontally, 6.6 c/deg vertically, and 50 Hz temporally. From Pelli (1981).

hundred-fold change in luminance does not affect the contrast threshold. The two curves converge because they have the same proportionality constant. The only effect of the hundred-fold luminance increase is to reduce the equivalent noise level.

Figure 1.11 shows data from Nagaraja (1964), as re-analyzed by Pelli (1981), for detecting a disk on a uniform background of 0.34 cd/m^2 (pupil size was not measured). Each curve is for a different disk size: 5.7′, 14.7′, 32′. As one would expect, the smaller disks have higher thresholds. However, remarkably, all the disk sizes yield virtually the same estimate of the equivalent noise level.

Figure 1.12 shows data from van Meeteren & Boogaard (1973), as re-analyzed by Pelli (1981), for detecting a 4.5 c/deg sinusoidal grating on a random dot display. Each curve is for a different luminance: 0.035, 0.35, and 3.5 td. The curves tend to converge at higher noise level. Note that the equivalent noise level is higher at lower luminance.

Figure 1.13 shows data from van Meeteren (1973), as re-analyzed by Pelli (1981), for detecting a 2.2 c/deg sinusoidal grating on a random dot display. Each curve is for a different luminance: 0.0013, 0.013, 0.13, 1.3, and 13 td. Again, the curves tend to converge at higher noise level, and the equivalent noise level is higher at lower luminance.

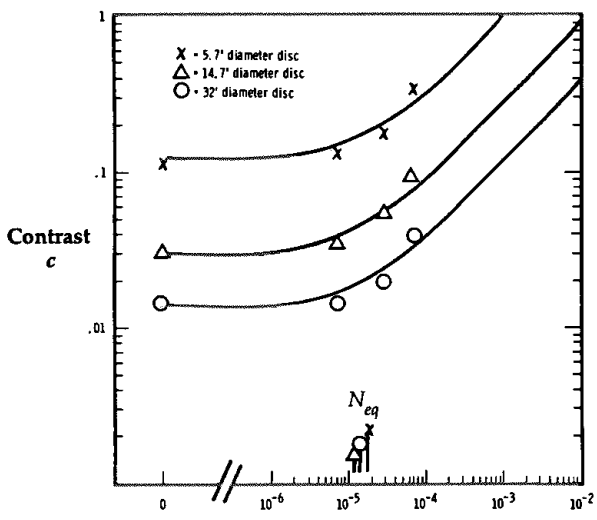

Fig. 1.11. Contrast threshold for a disk in noise. The disk diameter was 5.7′, 14.7′, or 32′ on a 0·34 cd/m^2 background (about 7 td). The smooth curve represents Eq. 3. From Pelli (1981), who re-analyzed data from Nagaraja (1964).

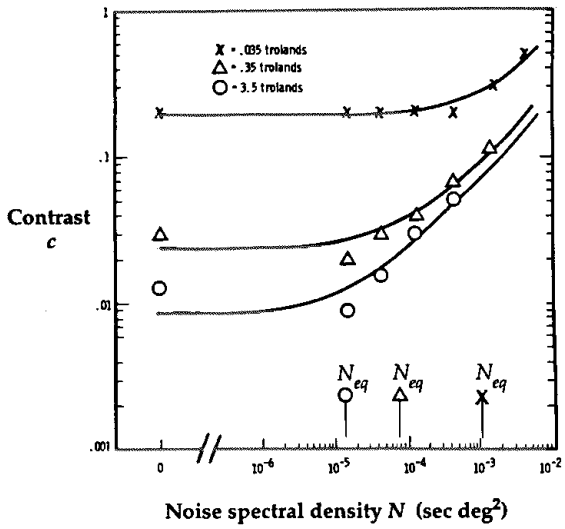

Fig. 1.12. Contrast threshold for a 4.5 c/deg grating on a random dot display at three luminances: 0.035, 0.35, and 3.5 td. The smooth curve represents Eq. 3. From Pelli (1981), who re-analyzed data from van Meeteren & Boogaard (1973). (The relatively poor fit at zero noise is probably due to the fact that these data were collected using a different apparatus, long after the other measurements.)

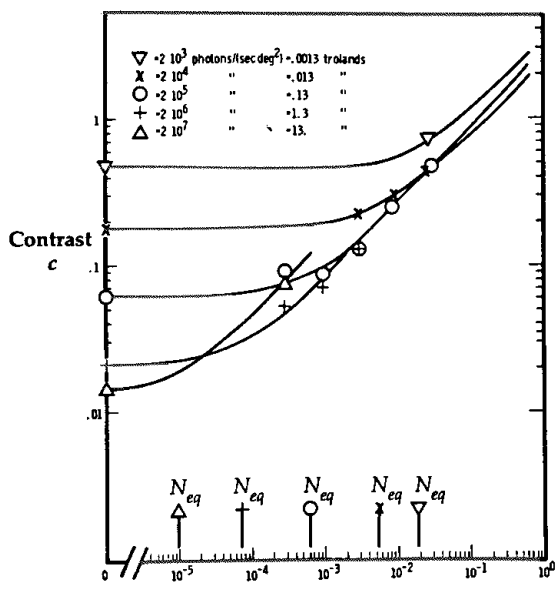

Fig. 1.13. Contrast threshold for a 2.2 c/deg grating at 7° nasal field on a random dot display at luminances of 0.0013, 0.013, 0.13, 1.3, and 13 td. The smooth curve represents Eq. 3. From Pelli (1981), who re-analyzed data from van Meeteren (1973).

Since all we care about in these graphs is the equivalent noise level, it is easy to summarize them all in a single graph. The various plotted symbols in Fig. 1.14 show the equivalent noise level for each experiment, versus the retinal illuminance, in Trolands. (Where necessary, pupil size was estimated from standard tables, Wyszecki & Stiles, 1967. Data are also included from Pelli, 1983, which will be described below.) Note that the data are quite orderly. At each luminance the equivalent noise levels are all nearly the same. As luminance increases the equivalent noise level decreases, in inverse proportion, as indicated by the slope of minus one.

It is interesting to compare the equivalent noise with photon noise, since photon noise may be an important component of the observer's equivalent noise. To make this comparison we need to calculate what noise level would result from full utilization of a given photon flux. The formula is

$$N_{\text{photon}} = \frac{1}{J_{\text{photon}}} \tag{4}$$

where J_{photon} is the photon flux (in photons per deg^2 sec) and N_{photon} is the *corneal-photon noise*, i.e. the noise level corresponding to use of all the photons, in contrast power per unit band width, in (c/deg)^2 Hz (see Appendix A). N_{photon} is the lowest possible equivalent noise level at retinal illuminance J_{photon}.

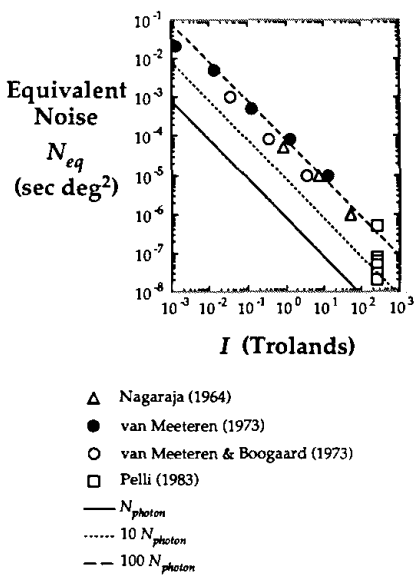

Fig. 1.14. Equivalent noise levels from Fig. 1.11 to 1.13 and from Pelli (1983). (The figure also includes more data of Nagaraja, 1964, as re-analyzed by Pelli, 1981, for a 14.7′ disk at luminances of approximately 0.9 and 50 td.) The lines show the equivalent noise level corresponding to use of 100% of the corneal quanta (solid line), 10% (finely dashed line), and 1% (coarsely dashed line).

Figure 1.14 shows the corneal-photon noise N_{photon} as a solid line. This corresponds to absorbing 100% of the corneal photons. Obviously the corneal-photon noise is much too low to account for the equivalent noise of the observers. Real observers fail to use all the photons and have additional sources of noise (Barlow, 1977), so their equivalent noise is higher than N_{photon}. If only 10% of the photons are absorbed, then, by the reciprocal relation between photon flux and noise level, the noise level will be ten times higher. That is shown as the middle line. Finally, if only 1% of the photons are absorbed, then the noise level will be 100 times higher, which is shown as the upper line. Note that most of the data points lie between the 1% and 10% lines. That means that these equivalent noise estimates can be accounted for by the absorbed-photon noise, if we are prepared to assume that somewhere between 1% and 10% are absorbed. (Or, alternatively, we could talk about the fraction that excite photoreceptors, and the exciting-photon noise.) Although the data points come from a variety of experiments by different laboratories, there is surprisingly little scatter.

The largest deviation occurs at 252 td, for the Pelli (1983) data, for which most of the points are clustered around the $10N_{\text{photon}}$ line, corresponding to 10% absorbed. Part of the reason for the discrepancy is that only these data have been corrected for the modulation transfer of the optics of the eye, so the plotted equivalent noise level is that at the retina. Whether the discrepancy among data points in Fig. 1.14 is real would be best determined by new experiments spanning the entire luminance range.

The data of van Meeteren (1973) are from the periphery and mostly at scotopic luminances so they mostly reflect rod activity; whereas the rest of the data are for foveal viewing, mostly at photopic luminances and thus mostly reflect cone activity. Even though Fig. 1.14 includes two receptor systems the data points are all consistent with a fairly constant fraction of photon absorption (1% to 10%), with little or no effect of background luminance, or parameters of the signal. Incidentally, this conclusion is consistent with that of the original authors, though the analysis presented here differs in important ways from that made by van Meeteren *et al.* (see Discussion).

At the beginning of this chapter we noted that we expect the fraction of photons absorbed to be constant, unlike the observer's overall quantum

efficiency, which is known to depend strongly on experimental conditions (e.g. Barlow's data in Fig. 1.1). Unfortunately, although the data in Fig. 1.14 span a wide range of conditions, they do not include the conditions used by Barlow (1962b) to measure quantum efficiency (Fig. 1.1).

The observer's equivalent input noise can only come from *two* sources, absorbed-photon noise, and neural noise. Neural noise arises in virtually all neural elements of the visual system from photoreceptor to cortex. Noise is ubiquitous in most physical systems. However, at least in man-made systems, it is a hallmark of good engineering that the dominant noise occurs at the first stage (e.g. North, 1942). This is achieved by having a high enough gain in each stage so that the noise of the first stage dwarfs the noise added in at each subsequent stage. North (1942) suggested that it would be useful to compare the equivalent input noise of sensitive radio receivers with the unavoidable thermal noise in the radio's antenna (the radio's input), and called this the radio's 'noise factor'. We will take a similar approach, comparing the observer's equivalent input noise with the unavoidable photon noise, and call it the observer's 'transduction efficiency'. (As we will see later, Rose (1948a) invented transduction efficiency, and though he did not point out the parallel to North's noise factor, he did acknowledge many discussions with North, who was also at RCA. Rose was developing sensitive television cameras; North was developing sensitive television receivers. Both were evaluating the sensitivity of their creations on an absolute scale.)

All of this will hopefully have convinced the reader that the observer's equivalent noise level is an interesting empirical quantity. Now we will need some definitions. The following narrative will emphasize an intuitive explanation for the equivalent noise idea that allows us to factor quantum efficiency into two components. Appendix A presents a parallel but more formal derivation.

For convenience, instead of referring to the contrast of the signal, it will be useful to refer to its contrast energy E, which is proportional to the squared contrast (see Appendix A; Pelli, 1981; Watson, Barlow & Robson, 1983).

Now let us consider several stages in the visual process. First consider the object, a luminance function at the visual field, or its image, a retinal illuminance function at the retina. In either case, it contains a target pattern with a certain contrast and thus a certain contrast energy. Similarly, the experimenter may introduce noise, e.g. by randomly varying the luminance of each pixel. Or, of course, there may be no noise at all. It is most convenient to describe an image by its contrast function (see Appendix A). The signal-to-noise ratio is the ratio of the signal contrast energy E to the noise level N, as shown in the upper left entry in Table 1.1. If there is no display noise, then the signal-to-noise ratio is infinite, as shown in the lower left entry.

The next stage that we can consider in the visual process is to imagine the image, instead of being a continuous luminance or illuminance function, as being a discontinuous collection of discretely absorbed photons. Call this the *photon* image. Knowing the retinal illuminance, we can calculate how much corneal-photon noise there will be (i.e. corresponding to absorption of 100% of the corneal photons). The photon noise adds to any noise in the contrast image, so the signal-to-noise ratio is $E/(N+N_{\text{photon}})$. If there is no display noise then the signal-to-noise ratio is just E/N_{photon}.

So far we have only considered the stimulus, without taking the observer into account at all. Now consider the *effective* image. We define the effective image as the sum of the original image (expressed as a contrast function) and the observer's equivalent input noise. The signal-to-noise ratio of the effective image is $E/(N+N_{\text{eq}})$. If there is no display noise then the signal-to-noise ratio is just E/N_{eq}.

Note that the absorbed-photon noise does not appear in the signal-to-noise ratio of the effective image in Table 1.1. The observer's equivalent noise already includes all the photon noise and any contrast-invariant neural noise. Besides, we do not know how much absorbed-photon noise there is in the observer's eye, because we do not know precisely what fraction of the photons is absorbed.

Finally, we can measure the observer's performance, which is the final result of the visual processing. We can calculate a signal-to-noise ratio d'^2 from the observer's percent of correct responses, or hit and false alarm rates (see Appendix A for definition of d').

A very important point is that these signal-to-noise ratios define the best possible performance with the information available at that stage (approximately; see 'signal-to-noise ratio' in Appendix A for a minor qualification.) Since the visual system cannot add information as the image goes through these various stages in the visual system, the signal-to-noise ratio can only go down. The ratio of the signal-to-noise ratios of any two stages is the *efficiency* of the transformation relating the two stages (Tanner & Birdsall, 1958).

Figure 1.15 shows the last three stages: the photon image, the effective image, and the perform-

Table 1.1. *Signal-to-noise ratios at four stages in the visual process.* See Appendix A for definitions.

	Image		Photon image		Effective image		Performance
In general:	$\frac{E}{N}$	$\geqslant$	$\frac{E}{N+N_{photon}}$	$\geqslant$	$\frac{E}{N+N_{eq}}$	$\geqslant$	d'^2
When $N=0$:	∞	$\geqslant$	$\frac{E}{N_{photon}}$	$\geqslant$	$\frac{E}{N_{eq}}$	$\geqslant$	d'^2

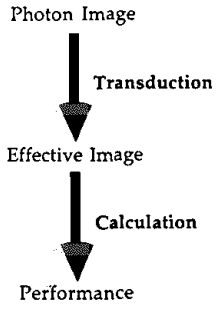

Fig. 1.15. Three stages of visual processing, and the efficiencies (in bold) of the transformations that relate them. For simplicity the diagram assumes no display noise, $N=0$.

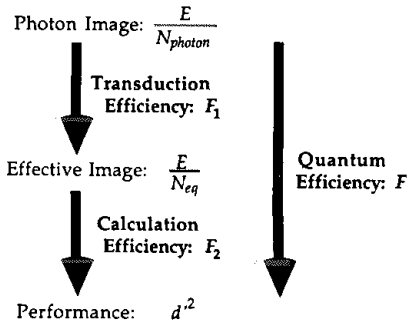

Fig. 1.16. The signal-to-noise ratios and efficiencies corresponding to Fig. 1.15.

ance, and names for the transformations that relate them. *Transduction* transforms the photon image (a physical stimulus) into an effective image (an internal effect), i.e. it substitutes the observer's equivalent input noise for the corneal-photon noise. This just makes the image noisier, because the equivalent noise reflects the failure to use all the corneal photons and the addition of neural noise. It includes all of the observer's contrast-invariant noise. *Calculation* transforms the noisy effective image into a decision, such as 'yes' or 'no'. The appropriate calculation depends on the signal and task. Our visual system, under our conscious guidance, reduces the spatiotemporal effective image (which could be represented by a large three-dimensional array of numbers) down to a single number, the decision, which is manifest as the observer's *performance*. The visual system's algorithm for making its decision may or may not be optimal, and may be constrained physiologically (e.g. by range of available receptive field sizes) and intellectually (e.g. by task complexity), and may vary from trial to trial.

Figure 1.16 shows the signal-to-noise ratios of the three stages, and the efficiencies of the transformations that relate them:

Transduction efficiency F_1, in the absence of display noise ($N=0$), measures how efficiently the visual system converts the photon image, which is just a stimulus, into the effective image, which incorporates the equivalent noise of the visual system (Pelli, 1981). This efficiency will be less than 1 if there is any increase in noise, either because of incomplete capture of the photons, or because of introduction of neural noise (or both). It might seem that 'transduction efficiency' ought to be defined physiologically, as a ratio of quantum bumps to incident photons. This is not a conflict. The black-box approach can be applied equally well to a cell as to an observer. We merely have to specify the input and output. We would naturally take the photoreceptor's input to be the incident light flux, and its output to be the bump count, which will yield the desired definition. Future research, comparing transduction efficiency of photoreceptor, retinal ganglion cell, and observer could reveal how our low equivalent noise level is achieved.

Calculation efficiency F_2 measures how efficiently the visual system converts the effective image into performance (Barlow, 1978; Pelli, 1981; Burgess, Wagner, Jennings & Barlow, 1981). This tells us how good a statistician our visual system is, how efficiently do we make decisions on the basis of noisy data? Calculation efficiency, like transduction efficiency, may be applied physiologically, but is not interesting for cells with

linear responses because it merely measures how well the stimulus and receptive field match.

Quantum efficiency F, in the absence of display noise ($N=0$), measures the whole thing, how efficiently the visual system converts the photon image into performance (Barlow, 1958b, 1962a). (This is closely analogous to Fisher's (1925) definition of efficiency of a statistic, presumably reflecting Barlow's early interest in Fisher's approach.)

Figure 1.16 shows the signal-to-noise ratios (in the absence of display noise), and the symbols which represent the efficiencies. The efficiency of the first transformation is F_1, the efficiency of the second transformation is F_2, and the overall efficiency is F. Finally, Table 1.2 lists the formulas for the efficiencies, which are just ratios of signal-to-noise ratios. Each entry in Table 1.2 is the ratio of two entries in the bottom row of Table 1.1. For example, transduction efficiency is the ratio of the signal-to-noise ratios of the effective image, E/N_{eq}, and the photon image, E/N_{photon}.

The beauty of these definitions is that the quantum efficiency is just the product of the transduction and calculation efficiencies.

$$F = F_1 F_2 \tag{5}$$

At the beginning of this chapter we saw that quantum efficiency varies over a very wide range, even though intuitively one might have expected it to be constant. Now that we can factor quantum efficiency into two components, we would like to know how much each varies.

Let us look at transduction efficiency first. We

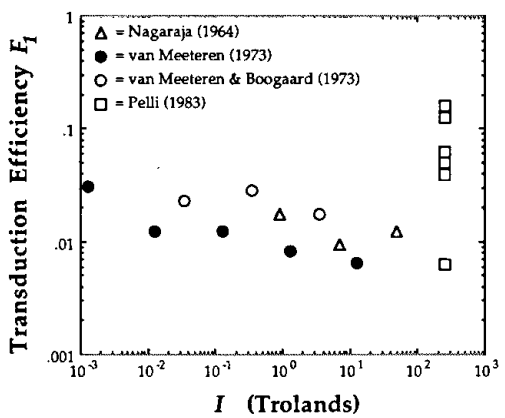

Fig. 1.17. The transduction efficiency, F_1, computed from the data in Fig. 1.14.

Table 1.2. *Transduction, calculation, and quantum efficiencies.* Note that $F = F_1 F_2$. In general (when N is not constrained to be zero), the calculation efficiency is $F_2 = d'^2/(E/(N+N_{eq}))$.

	Transduction	Calculation	Quantum
When $N=0$:	$F_1 = \frac{E/N_{eq}}{E/N_{photon}}$	$F_2 = \frac{d'^2}{E/N_{eq}}$	$F = \frac{d'^2}{E/N_{photon}}$

can see that, by its definition in Table 1.2, the signal energy appears in both the numerator and denominator, so it cancels out, and what is left is just the ratio of two noises, N_{photon}/N_{eq}. We already have the necessary data in Fig. 1.14.

Figure 1.17 shows the transduction efficiency, i.e. the ratio of the photon noise to the equivalent noise, at each luminance. Not surprisingly, given our comments on Fig. 1.14, we find that the transduction efficiency is relatively constant. Excluding the Pelli (1983) data, which will be discussed below, the entire variation is about a factor of four, yet the luminance ranges over nearly five orders of magnitude. Thus the variation is slight. Bear in mind too, that these experiments were conducted by different authors using various methods and that there are some uncertainties such as estimated pupil size, and accuracy of the noise and luminance calibrations.

We expect the fraction of photons absorbed to be constant, unlike the observer's overall quantum efficiency, which is known to depend strongly on experimental conditions, such as the signal area and duration and the background luminance (e.g. Figs. 1.1 and 1.2). However, Nagaraja's data show no effect of the disk size on the transduction efficiency. Unfortunately, although the data in Fig. 1.14 span a wide range of conditions, they do not include the conditions used by Barlow (1962b) to measure quantum efficiency (Fig. 1.1), so we cannot factor Barlow's measurements, to determine whether the enormous variation in quantum efficiency is due to variation of just transduction or calculation efficiency.

Over the years, various authors have measured calculation efficiencies, but, unfortunately, most of the experiments have been done with static noise, so they are not relevant here (Pelli, 1981; Burgess *et al.*, 1981). Another obstacle is that most of the experiments that one would like to use for such a calculation, such as those by van Meeteren, allowed unlimited viewing of the stimulus, so that we cannot sensibly calculate the contrast energy of the stimulus.

So far we have examined the dependence of the observer's equivalent noise on luminance. We have

not yet stopped to ask what is the spatiotemporal spectrum of this noise, $N(f_x, f_y, f_t)$. At first this might seem unanswerable, since, as pointed out before, we cannot measure the observer's equivalent noise directly; we can only measure its effect on the contrast threshold. However, we can take advantage of the phenomenon of the critical band (i.e. spatiotemporal frequency channels) whereby the threshold for a flickering grating is only affected by noise frequencies close to the signal frequency, in both spatial frequency (Greis & Rohler, 1970; Stromeyer & Julesz, 1972; Henning, Hertz & Hinton, 1981; Pelli, 1981) and temporal frequency (Pelli & Watson in Pelli 1981; Mandler & Makous, 1984). As a result, an equivalent noise measurement made with a flickering grating provides an estimate of the equivalent noise level in the vicinity of the spatiotemporal test frequency.

Referring back to Fig. 1.4, the output of the noise amplifier is analogous to the observer's noisy effective image, and the subsequent bandpass filter and power meter are analogous to the observer's spatiotemporal frequency channel and squared contrast threshold. A key point is that we do not need to know the bandwidth of the filter or channel, because that only affects the measured power by a constant proportionality factor (provided the spectrum is smooth over the bandwidth of the filter), and we only need to know the output power to within a proportionality factor (see Eqs. 1 and 3).

In this way, Pelli (1983) measured the spatiotemporal spectrum of the equivalent noise at 300 td. The results are shown in Fig. 1.18. The spectrum is flat (within experimental error) at about $10N_{\text{photon}}$ (i.e. corresponding to a transduction efficiency of 10%) except for a great rise at very low spatiotemporal frequency (0 c/deg, 0 Hz). Similar data (not shown) have been collected on a second observer and for the same observer six months later, after a complete re-calibration of the system.

The observer looked through a 1 mm artificial pupil, so the eye of the observer could be assumed to be diffraction limited (Campbell & Gubisch, 1966), allowing calculation of the retinal contrasts of the signal and noise as a function of spatial frequency. The plotted equivalent noise level is at the retina, which is most directly comparable to the photon noise, since the photon noise arises at the retina and thus is white (has a flat spatiotemporal spectrum) at the retina, but not at the visual field. Other methodological details are presented in Appendix B.

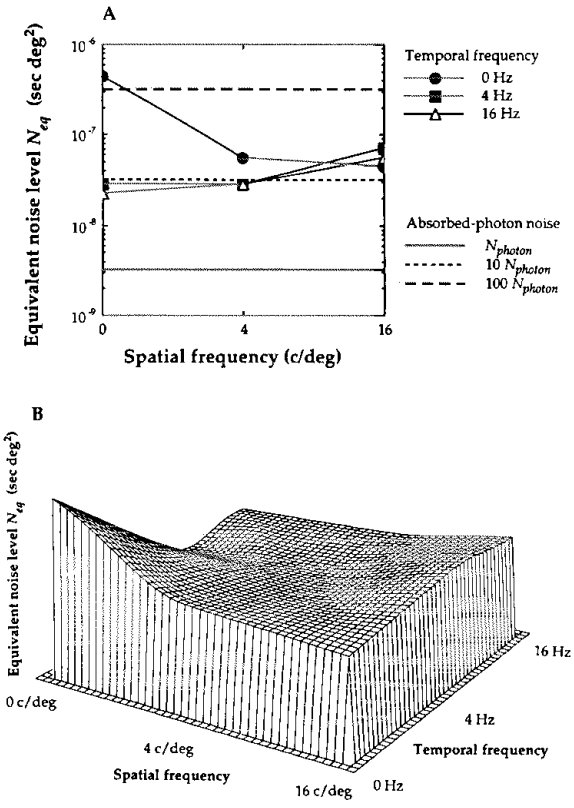

Fig. 1.18. A. The equivalent noise level at the retina as a function of spatiotemporal frequency at 252 td. Nine spatiotemporal frequencies were tested, all the combinations of 0, 4, and 16 c/deg and 0, 4, and 16 Hz. All signals were vignetted by a Gaussian with a width (and height) at 1/e of 0.75° and a duration of 0.75 s. The 95% confidence interval about each equivalent noise estimate is about ±0.3 log units. See Appendix B for methods. The data are from Pelli (1983). B. A smooth surface has been fit to the nine data points, and is shown in perspective. The floor seen at the edges represents the corneal-photon noise level N_{photon}.

Discussion

The source of the equivalent input noise

Absorbed-photon noise and neural noise are both sources for the observer's equivalent noise, but does one type dominate? Absorbed-photon noise N_{absorbed} has only one degree of freedom, the fraction of photons absorbed, otherwise it is fully determined. It is inversely proportional to luminance, and, because photon absorptions are uncorrelated over space and

time, it is independent of spatiotemporal frequency. Thus the hypothesis that the observer's equivalent noise is mostly absorbed-photon noise predicts both the luminance dependence seen in Fig. 1.14, and the spatiotemporal frequency independence seen in Fig. 1.18 (except for the deviation at the origin). Since noise levels cannot be negative – there is no such thing as negative variance – the equivalent noise level cannot be less than the absorbed-photon noise level. Since the absorbed-photon noise is frequency-independent, the absorbed-photon noise level can be no higher than the lowest equivalent noise level in Fig. 1.18 (except for measurement error). Thus the large rise in equivalent noise at low spatiotemporal frequency *must* be neural noise.

It is difficult to make a clear prediction for neural noise, as it depends on many physiological details that have yet to be worked out. However, we can make a few comments, making the simplifying assumption that at its site of origin the neural noise is independent of luminance. Neural noise that arises at a site distal to the site of light adaptation would act as a dark light and would become insignificant at luminances much above absolute threshold. Neural noise that arises at a site central to the site of light adaptation (i.e. after the contrast gain control, Rose, 1948a; Shapley, 1986) would produce an equivalent input noise (at the visual field) that is independent of luminance. This predicts that the low-frequency mountain seen in Fig. 1.18 is independent of luminance. Increasing the luminance would reveal more and more of the neural-noise mountain, as the absorbed-photon-noise sea was drained away. Reducing the luminance would eventually submerge the neural-noise under the rising sea of photon-noise.

The fact that at most spatiotemporal frequencies the observer's equivalent noise level can be accounted for (at least tentatively) by the absorbed-photon noise implies that the visual system exhibits that hallmark of good engineering that we described earlier in the context of man-made amplifiers and radio receivers. For the absorbed-photon noise to dominate the observer's equivalent noise implies that the gains at each stage of the visual system are high enough to amplify the absorbed-photon noise to exceed any contrast-invariant neural noise introduced at later stages. In the same way, the intrusion of neural noise at low spatiotemporal frequency implies that the gain is not high enough at those frequencies. It is well known that the visual system has very low sensitivity at low spatiotemporal frequency (Robson, 1966), probably as a side effect of light adaptation. All together this suggests that the familiar 'low-frequency cut' of the contrast sensitivity function, i.e. the relatively low sensitivity at low spatiotemporal frequencies is due to a higher equivalent noise level at those frequencies, due to the intrusion of neural noise above the absorbed-photon noise level.

The Rose–de Vries law

The finding of a fairly constant transduction efficiency in Fig. 1.17 is not a complete surprise, since when the overall quantum efficiency is constant, it is simplest to expect the component transduction and calculations efficiencies to be constant too. The Rose–de Vries law, $c \propto I^{-0.5}$ is equivalent to the statement that quantum efficiency is constant (provided that threshold is at a fixed d'; Rose, 1942, 1948a; de Vries, 1943; Barlow, 1977). The violations of the Rose–de Vries law (e.g. Fig. 1.1) seem to be more the exception than the rule. Thresholds for flickering gratings follow the Rose–de Vries law up to at least 1000 td (van Nes, Koenderink, Nas & Bouman, 1967; Kelly, 1972). Banks, Geisler, & Bennet (1987) show that, with respect to the retinal image (which is blurred by the eye's optics), the observer's overall quantum efficiency is independent of spatial frequency from 5 to 40 c/deg, from 3.4 to 340 cd/m^2, if the spatial-frequency gratings have a fixed number of cycles. However, quantum efficiency does depend on the number of cycles in a grating because of the inefficiency of 'probability summation' (Robson & Graham, 1981; Watson, Barlow & Robson, 1983; Banks, Geisler & Bennet, 1987).

The Rose–van Meeteren paradigm

We have chosen to present the ideas in this chapter in a logical progression, not in historical order. Now it is time to go back and point out where Rose's (1942, 1946, 1948a, b, 1957, 1977) ideas fit in the structure we have erected.

The early papers on quantum efficiency of vision (before 1956) all assumed that the definition of quantum efficiency as a psychophysical property was self evident, and defined it only implicitly, by their method of measuring it. There is no difficulty in understanding what Rose (1942, 1946, 1948a, b, 1957) meant when he made the first determinations of the quantum efficiency of a video camera and photographic film – signal-to-noise ratio out $(E/N)_{out}$ over signal-to-noise ratio in $(E/N)_{photon}$ – but it is not obvious how to apply this to the eye. As Barlow (1958b) pointed out, in the case of the eye one might wish to say *quantum efficiency* is

> the fraction of quanta sent through the pupil which are 'effectively absorbed' by the photosensitive materials subserving the mechanism under consid-

eration, but a difficulty arises in deciding what to understand by the words 'effectively absorbed'. In the case of rhodopsin a quantum can be absorbed without bleaching the molecule, and even if bleaching occurs it is by no means certain the rod is always activated. Further, if the rod is activated it is still not certain that this information is successfully transmitted to the place where the threshold decision is made.

Barlow (1956, 1962a, b) went on to give the definition of overall quantum efficiency used here, which amounts to computing the output signal-to-noise ratio from the observer's performance, $(E/N)_{out} = d'^2$. However, that was not what Rose (1948a, 1957) had in mind.

As we will see below, Rose's (1948a) 'quantum efficiency of the eye' is what we have here dubbed the *transduction efficiency*, so it is not surprising that his efficiencies (roughly 5%, independent of background luminance, in rough agreement with Fig. 1.17) were very different from Barlow's, which fall precipitously with background luminance, as seen in Fig. 1.1. In fact, this confusion of two fundamentally different measures bearing the same name has persisted for thirty years. While Jones (1959), Barlow (1962b), Cohn (1976), and Geisler & Davila (1985) measured quantum efficiency, Rose (1948a, 1957), Sturm & Morgan (1949), Nagaraja (1964), van Meeteren (1973), van Meeteren & Boogaard (1973), and Engstrom (1974) all measured what is here called transduction efficiency, though everyone claimed to be measuring 'quantum efficiency'. It is the purpose of this chapter to point out that both measures are important, and that it would be most informative to measure both over a wide range of conditions.

As discussed earlier, the paradigm advocated here for measuring the observer's equivalent noise level or transduction efficiency (and first used by Nagaraja, 1964) is to measure thresholds at more than one noise level, assuming contrast invariance. The contrast-invariance assumption implies that at threshold the two (or more) conditions will differ only in the overall contrast of the effective image.

The Rose–van Meeteren paradigm is to measure thresholds at more than one luminance (Rose, 1957; van Meeteren & Boogaard, 1973; Engstrom, 1974), implicitly assuming *luminance invariance*, i.e. assuming that the observer's performance depends solely on the contrast function of the effective image, independent of its luminance. In this approach, noise is added to the brighter display (i.e. at the higher luminance) to equate threshold with that of the dimmer display. The luminance-invariance assumption implies that this will equate the effective noise levels of the two displays, so that at threshold the two conditions will differ only in luminance, having identical effective (contrast) images. This paradigm lends itself particularly well to experiments in which the dimmer display has no noise (i.e. continuous) and the brighter display is a random dot display. The transduction efficiency is just the ratio of dot flux (on the bright display) to corneal photons (on the dim display). (This assumes that the bright display is sufficiently bright and noisy that the observer's equivalent input noise is negligible, which is easily achieved in practice.) However, as we have seen in Figs. 1.12 and 1.13, it is possible to analyze even these experiments by the Nagaraja paradigm, if several noise levels (dot fluxes) are used at one luminance. (This requires using neutral density filters to compensate for the increase in brightness as the dot flux is increased.)

The idea that vision is usually photon-noise limited has been widely discounted, as a result of dramatic violations of the Rose–de Vries law, $c \propto I^{-0.5}$ (Rose, 1942, 1948a; de Vries, 1943; e.g. Aguilar & Stiles, 1954; Barlow, 1958, 1962b). However, a constant transduction efficiency, by itself, does not require that the Rose–de Vries law be obeyed. Unfortunately, Rose's methods of measuring transduction efficiency forced him to make overly strong assumptions about the observer. Initially, Rose (1948a) boldly assumed that *all* thresholds have the same effective signal-to-noise ratio. The later, more sophisticated, Rose–van Meeteren paradigm requires only the luminance-invariance assumption, i.e. the assumption that thresholds for the same pattern at two different luminances have the same effective signal-to-noise ratio. (More precisely, they need the three assumptions of the Nagaraja paradigm, substituting luminance-invariance of performance for contrast-invariance of performance.) However, even this weaker assumption, when combined with the assumption of constant transduction efficiency, implies the Rose–de Vries law, $c \propto I^{-0.5}$, which is not generally true, so this approach cannot disentangle luminance-dependent variations of transduction and calculation efficiencies. In this chapter we use only the Nagaraja paradigm with its contrast-invariance assumption, which is consistent with all published measurements of threshold versus noise level at a constant luminance (Pelli, 1981). The Nagaraja paradigm can be used to determine, for example, whether the variation of overall quantum efficiency seen in Fig. 1.1 is due to variation of transduction or calculation efficiency (or both).

Future work

Extensive measurements of transduction and quantum efficiency are planned, both to re-analyze Barlow's (1962b) result shown in Fig. 1.1, and to examine the spatiotemporal spectrum at many luminances, to study the luminance-dependence of neural noise.

It will be important to test the model of Fig. 1.7. As noted before, the main prediction of the model is Eq. 3, $c^2 \propto N + N_{eq}$, which is widely confirmed (Pelli, 1981; Burgess *et al.*, 1981; Legge, Kersten & Burgess, 1987). But note that this equation must apply at any threshold criterion d'. This implies that the psychometric function, the growth of d' with contrast, must be contrast invariant too, i.e. when threshold is raised by external noise, the psychometric function must shift along the log contrast axis without changing shape. For gratings in dynamic white noise, this prediction has been confirmed by Pelli (1981), disconfirmed by Kersten (1984), and reconfirmed by Thomas (1985). More work is warranted.

Summary

The overall quantum efficiency of human vision is strongly dependent on many stimulus parameters. Quantum efficiency refers only to the stimulus and the observer's performance. However, in analogy to Barlow's notion of dark light, which is in the intensity domain, we can measure an observer's equivalent noise in the contrast domain. With this, and the contrast-invariance assumption, we can define and measure the signal-to-noise ratio at an intermediate stage in the visual process: the *effective* image, which includes the observer's equivalent noise. Introducing this intermediate stage allows quantum efficiency to be factored into two components: transduction and calculation efficiencies. This is derived formally in Appendix A. It appears that the transduction efficiency is quite stable, in the range 1% to 10%, over most conditions, confirming Rose (1948a).

While all these measures can be studied as a function of any experimental variation, it is particularly informative to study them as a function of spatiotemporal frequency. This is because the existence of spatiotemporal frequency channels (i.e. critical bands) allows us to interpret the variation in equivalent noise level N_{eq} with spatiotemporal frequency directly as the spatiotemporal spectrum $N_{eq}(f_x, f_y, f_t)$ of the equivalent noise. Preliminary results at 252 td indicate that the equivalent noise is dominated by absorbed-photon noise (corresponding to absorption of 10% of corneal photons) at all spatiotemporal frequencies in the range 0 to 16 c/deg and 0 to 16 Hz, except for very low spatiotemporal frequencies (near 0 c/deg, 0 Hz, i.e. signals that change slowly in space and time) where the equivalent noise is much higher, and therefore is of neural origin.

Appendix A: Quantum efficiency is the product of transduction and calculation efficiencies $F = F_1F_2$

This sequence of definitions is organized as a derivation, beginning with the physical stimulus, and ending with the factoring of quantum efficiency into transduction and calculation efficiencies.

L_{av} is the *mean luminance*, averaged over space and time.

$L_{SN}(x,y,t)$ is the *luminance function* of the image over space and time on a signal-plus-noise presentation. Note that because the noise is random, the luminance function is random, different on each presentation, drawn from an ensemble of possibilities.

$c_{SN}(x, y, t)$ is the *contrast function* of the image over space and time on a signal-plus-noise presentation (Linfoot, 1964),

$$c_{SN}(x, y, t) = \frac{L_{SN}(x,y,t)}{L_{av}} - 1$$

$s(x, y, t)$ is the *signal* expressed as a contrast function over space and time,

$$s(x, y, t) = \langle c_{SN}(x, y, t) \rangle$$

where the angle brackets $\langle\ \rangle$ indicate the ensemble average, across all possible instances (to average away the noise).

$n(x, y, t)$ is the *noise* expressed as a contrast function over space and time,

$$n(x, y, t) = c_{SN}(x, y, t) - s(x, y, t)$$

E is the contrast energy (i.e. integrated square contrast) of any contrast function, usually of the signal,

$$E = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} s^2(x, y, t)\,dx\,dy\,dt$$

(Pelli, 1981; Watson, Barlow & Robson, 1983).

$N(f_x, f_y, f_t)$ is the *power spectral density* of the noise, i.e. contrast power per unit two-sided bandwidth at spatiotemporal frequency f_x, f_y, f_t.

$$N(f_x, f_y, f_t) = \left\langle \lim_{X,Y,T\to\infty} \frac{1}{XYT} \times \left| \int_{-T/2}^{T/2}\int_{-Y/2}^{Y/2}\int_{-X/2}^{X/2} n(x, y, t)\, e^{-i\pi(xf_x + yf_y + tf_t)}\,dx\,dy\,dt \right|^2 \right\rangle$$

When $N(f_x, f_y, f_t)$ is constant over the band of interest, it will often be called simply the *noise level* N. Note that this definition uses *two-sided* bandwidth, including positive and negative frequency. A filter passing frequencies between 30 and 40 Hz also passes frequencies between −30 and −40 Hz and thus has a two-sided bandwidth of 10 + 10 = 20 Hz. In vision, unlike audition, we deal with noise varying in one, two, or three dimensions (though this chapter deals only with three), and we need this two-sided definition of N in order to make the expression for signal-to-noise ratio (see below) independent of the number of dimensions (Pelli, 1981).

c^2_{rms} is the *contrast power* (i.e. the mean square contrast) of the noise,

$$c^2_{rms} = \langle n^2(x, y, t)\rangle$$

or, more generally, if we don't want to assume the noise is stationary,

$$c^2_{rms} = \left\langle \lim_{X,Y,T\to\infty} \frac{1}{XYT} \int_{-T/2}^{T/2}\int_{-Y/2}^{Y/2}\int_{-X/2}^{X/2} n^2(x, y, t)\,dx\,dy\,dt \right\rangle$$

Note that, by Rayleigh's theorem (Bracewell, 1978), the integral of the power spectral density over all frequencies yields the contrast power,

$$c^2_{rms} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} N(f_x, f_y, f_t)\,df_x\,df_y\,df_t$$

E/N is the *signal-to-noise ratio*, a dimensionless quantity that determines the levels of performance for detection of any known signal with energy E in white noise with power spectral density N (Peterson, Birdsall & Fox, 1954). (Note that in audition, where the noise is always one-dimensional, the signal-to-noise ratio is usually written as $2E/N_0$, where $N_0 = 2N$, e.g. Tanner & Birdsall, 1958; Green & Swets, 1974, Eq. 6.37.)

d'^2 is the smallest signal-to-noise ratio consistent with the observed level of performance (Tanner & Birdsall, 1958),

$$d'^2 = \left(\frac{E}{N}\right)_{ideal}$$

In terms of the usual performance measures this works out to

$$d' = \begin{cases} \sqrt{2}\,z[P_{2afc}(c)] & \text{if task is two-alternative forced choice} \\ z[P_{yn}(c)] - z[P_{yn}(0)] & \text{if task is yes–no} \end{cases}$$

for a known signal in white noise (Elliott, 1964), where $P_{2afc}(c)$ is the proportion correct (at contrast c) in a two-alternative forced choice task, $P_{yn}(c)$ is the hit rate and $P_{yn}(0)$ is the false alarm rate in a yes–no task, and $z[P]$ represents the inverse cumulative normal, i.e.

$$P = \frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{z} e^{-u^2/2}\,du$$

N_{eq} is the *equivalent input noise level* of the observer, i.e. the noise that would have to be added at the display to model the observer as noise-free (Pelli, 1981; Ahumada & Watson, 1985; Ahumada, 1987). However, that is not measurable without some assumptions about the observer, so we re-define the equivalent noise as a parameter of the black-box model in Fig. 1.7. This model predicts that the squared contrast will be proportional to the sum of the noise and the equivalent noise,

$$c^2 \propto N + N_{eq}$$

Empirically we determine the observer's equivalent noise level by fitting this equation to our data (Nagaraja, 1964; Pelli, 1981).

N_{ef} is the *effective noise level* of the observer and display,

$$N_{ef} = N + N_{eq}.$$

The *effective image* is the sum of the stimulus $c_{SN}(x, y, t)$ and the observer's equivalent input noise. It has a signal-to-noise ratio of $E/(N + N_{eq})$.

N_{ideal} is the maximum noise level consistent with the observed level of performance,

$$N_{ideal} = \frac{E}{d'^2}$$

J represents the mean event flux of a Poisson process, e.g. photons per deg^{2}sec. The noise level of a Poisson process is the inverse of the flux (Papoulis, 1965; Pelli, 1981),

$$N = \frac{1}{J}$$

This statement is strictly true only for a uniform original image, or a very low contrast signal. However, it is an

extremely useful simplification, and a good approximation as long as the signal contrast is not too high. The simplification is required because, unfortunately, the signal-to-noise ratio E/N does not fully determine detectability of a known signal in Poisson noise (which is signal-dependent), as it does in white noise (which is independent of the signal). However, calculations show that if the signal contrast is less than 20%, $|s(x, y, t)| < 0.2$, then, for two-alternative forced-choice detection of a known signal in Poisson noise, the discrepancy between E/N and the best possible performance, $d'^2 = 2z^2[P_{2afc}]$, is at most ± 0.5 dB, which is negligible when one considers that contrast thresholds are generally measured with a standard deviation of 1 dB at best (also see Geisler & Davila, 1985).

J_{eq} is the minimum photon flux consistent with the measured equivalent noise level,

$$J_{eq} = \frac{1}{N_{eq}}$$

J_{ideal} is the minimum event flux consistent with the observed level of performance,

$$J_{ideal} = \frac{1}{N_{ideal}}$$

J_{photon} is the corneal photon flux.

N_{photon} is the noise level corresponding to the corneal photon flux,

$$N_{photon} = \frac{1}{J_{photon}}$$

J_{ef} is the minimum event flux consistent with the effective noise level,

$$J_{ef} = \frac{1}{N_{ef}}$$

The (corneal) *photon image* is the spatiotemporal pattern of absorption that would result if all corneal photons were absorbed. It is a sum of many delta functions (each representing one photon absorption) as a function of space and time. It has a signal-to-noise ratio of $E/(N + N_{photon})$.

Efficiency of any transformation is the ratio of the signal-to-noise ratios before and after (Tanner & Birdsall, 1958),

$$\frac{(E/N)_{out}}{(E/N)_{in}}$$

Transduction efficiency F_1 is the fraction of the corneal quanta required to account for the observer's equivalent input noise level (Nagaraja, 1964; Pelli, 1981),

$$F_1 = \frac{J_{eq}}{J_{photon}}$$

or, equivalently,

$$F_1 = \frac{E/N_{eq}}{E/N_{photon}}$$

which, in the absence of display noise ($N = 0$), is the observer's efficiency, from photon image to effective image (Rose, 1948a; Pelli, 1981). Rose called it the 'quantum efficiency of the eye'. Transduction efficiency, like the equivalent noise, is only measurable in the context of a model like that of Fig. 1.7, with its three assumptions (mainly contrast-invariance, in the Nagaraja paradigm, or luminance invariance, in the Rose–van Meeteren paradigm; see Discussion).

Calculation efficiency F_2 is the smallest fraction of the effective number of Poisson events on a random dot display consistent with the observed level of performance,

$$F_2 = \frac{J_{ideal}}{J_{ef}}$$

or, equivalently, it is the observer's efficiency, from effective image to response,

$$F_2 = \frac{d'^2}{E/N_{ef}}$$

(Barlow, 1977, 1978; Pelli, 1981; Burgess *et al.* 1981; Legge, Kersten & Burgess, 1987). Barlow (1977) defined calculation (which he called 'central') efficiency in terms of J and N, not J_{ef} and N_{ef}, but this is an insignificant difference when the noise level is high, $N \gg N_{eq}$, as is usually the case when calculation efficiency F_2 is measured. In Barlow's form, calculation efficiency can be measured without any assumptions.

Quantum efficiency F is the smallest fraction of the corneal quanta consistent with the observed level of performance (Barlow, 1958b, 1962a),

$$F = \frac{J_{ideal}}{J_{photon}}$$

or, equivalently, it is the observer's efficiency, from photon image to response,

$$F = \frac{d'^2}{E/N_{photon}}$$

Quantum efficiency can be measured without any assumptions. Note that this definition of F makes sense only when there is no display noise, $N = 0$, in which case the effective noise equals the equivalent noise, $N_{ef} = N_{eq}$, and

$$F = F_1 F_2$$

which is what we set out to prove.

Appendix B: Methods for Fig. 1.18

Procedure

A PDP-11/23 computer synthesized the signal and ran all the trials. The noise was always on and was at constant noise level throughout each trial.

Threshold was measured by the method of two-interval forced-choice, with feedback. Each trial consisted of two intervals, announced by a voice synthesizer (Intex talker), 'One . . . two'. The signal was presented during one or the other of the two intervals, pseudo-randomly determined. Then the observer indicated, by button push, which interval he thought contained the signal. The talker acknowledged the response ('one', 'two', or 'cancel') and gave feedback ('right' or 'wrong').

The QUEST staircase procedure determined the contrast of the signal on each trial, presenting the signal at the current maximum-probability estimate of threshold for that condition (Watson & Pelli 1983). After 40 trials, QUEST provided the maximum likelihood estimate of threshold (82% correct), based on the observer's responses. Each threshold was measured twice, and their average was used in subsequent analyses.

The observers

Two observers participated, one emmetrope and one well-corrected myope. Viewing was monocular. The observer sat in a light-proof booth, and viewed the display through a 1 mm pupil and a 550 nm narrowband filter (and the optical correction, if any). The position of the observer's head was maintained by a bite bar covered with dental impression compound, heat-molded to the observer's bite. The bite bar, in turn, was attached to a two-axis tool holder which allowed the observer's head to be moved horizontally and vertically relative to the fixed artificial pupil. The artificial pupil was carefully centered on the observer's pupil.

The display

Sinusoidal gratings were displayed using a television-like raster on a cathode-ray tube as described by Graham, Robson & Nachmias (1978). The screen was 30 cm wide and 20 cm high. Room lights were turned off to minimize intraocular scattered light. The screen displayed 100 frames per second, otherwise the space-average luminance in the absence of a signal, was constant throughout each experiment.

Noise

The noise was generated by a custom-designed Schottky TTL computer card which plugged into the PDP-11/23 bus. A 33 bit shift register was given exclusive–or feedback so as to produce a maximal-length sequence of states before repeating (Lipson, Foster & Walsh 1976, Peterson & Weldon, 1972). Each clock cycle shifted the EXCLUSIVE OR of three of the bits into bit 1. The clock rate could be programmed to any rate up to 40 MHz. There were $2^{33} - 1 \sim 10^{10}$ clock cycles before the feedback shift register sequence repeated: 3.6 minutes at a 40 MHz clock rate. One bit of the register drove an output producing one sample per clock cycle, either +1 or −1 volt, changing pseudorandomly at the beginning of each clock cycle. These samples were uncorrelated, and had a mean of 0 volts. This analog output was attenuated by a programmable attenuator and added to the signal at the display. The samples were uncorrelated, so the noise spectrum was nearly flat up to half the sample rate. The noise was free-running, not synchronized to the frames of the display or signal presentations, making it very unlikely that the observer could benefit by any memory of the previous noise cycle, more than 3 minutes ago. The noise was attenuated by a computer-controlled passive attenuator (FORESIGHT dB Attenuator), with a 75 Ω characteristic impedance. The attenuator directly drove a coaxial cable with a 75 Ω characteristic impedance terminated in its characteristic impedance at the display, where it was added to the signal. This arrangement resulted in a very flat transfer function; the gain was down by only 1 dB at 10 MHz.

Display calibration

The display (Joyce Electronics) exhibited no significant nonlinearity up to contrasts of 90%. The linearity was checked at the beginning of every experimental session by the simple expedient of counterphase flickering a sinusoidal grating at a rate above critical flicker fusion, and confirming that the display appeared uniform. Second harmonic distortion would produce a static grating of twice the original spatial frequency.

The display's luminance was 275 td. Because the display was viewed through a 550 nm narrowband filter, the luminance was computed from a radiance measurement by a radiometer which was NBS traceable (UDT 161), using the identity

$$1 \text{ lumen} = 4.1 \times 10^{15} \text{ quanta (555 nm)/s}$$

(The photopic efficiency at 550 nm is negligibly different from that at 555 nm, i.e. 0.995 versus 1.)

The Modulation Transfer Function (MTF) of the display $H(f_x)$ was measured by drifting a sinusoidal grating past a microphotometer (Schade, 1958). The microphotometer consisted of a photometer (UDT 161) which measured the light passed through a 1 mm aperture in the image plane of a 5:1 microscope objective (N.A. 0.11, Rolyn 80.3044), giving a 0.2 mm-wide collection area in the plane of the phosphor. All spatial frequencies were drifted at 1 Hz so as to

produce the same temporal frequency in the measuring system, to avoid confounding the spatial MTF of the display with the temporal MTF of the photometer. This calibration was checked at low spatial frequency (where the MTF is essentially flat) by very slowly drifting a low-spatial-frequency square wave at a nominal contrast of 20%, measuring the luminance in the center of the bright and dark bars, and calculating the contrast,

$$c = (L_{max} - L_{min})/(L_{max} + L_{min}),$$

with good agreement with the above procedure.

Noise calibration

It was not practical to directly measure the spatiotemporal spectrum of the noise. Therefore indirect methods were used. The method of generation of the noise produces a white spectrum (before attenuation and display). Since the attenuator Modulation Transfer Function (MTF) is flat over the relevant bandwidth the noise spectrum is proportional to the square of the display MTF:

$$N(f_x, f_y, f_t)/N(0, 0, 0) = H^2(f_x, f_y, f_t)$$

Since I have already measured the spatial MTF in the preceding section, all I need to do now is to measure $N(0, 0, 0)$, the noise level near zero spatiotemporal frequency.

The concept of *noise-equivalent bandwidth* is helpful in calibrating noise level. For a filter it is defined as follows. Imagine a bandpass filter with a passband gain equal to the peak gain of the actual filter, and zero gain outside the passband. Suppose we apply a white noise input to both the actual and the hypothetical filters. The equivalent bandwidth of the actual filter is defined as the bandwidth the hypothetical filter has to have in order to pass the same total noise power as the actual filter. Similarly we can consider any noise spectrum to be the result of passing white noise through some filter, and define the equivalent bandwidth of the noise as the noise-equivalent bandwidth of the imputed filter.

I calibrated the noise level at the display by measuring the output power of a narrowband spatiotemporal filter which accepted the display as its input. This procedure is a direct application of the definition of spectral density: the power passed by a narrowband filter divided by its equivalent bandwidth. The filter was synthesized by using a photometer to average light from a large aperture (4 cm by 8 cm). The photometer output was low-pass filtered (30 Hz cut off) and sampled by a 14-bit analog-to-digital converter at 100 Hz (once per frame). The computer smoothed these data by calculating the running average of the 10 most recent samples. Thus each smoothed sample represents the average luminance over an area of known width (4 cm) and height (8 cm) and duration (10 frames = 100 ms). These smoothed samples are the output of the synthesized spatiotemporal filter, whose input is the display. The filter has a (two-sided) noise-equivalent bandwidth of 1/(width × height × duration), centered on zero frequency. The contrast power c^2_{rms} of the smoothed photometer readings, divided by the noise-equivalent bandwidth of the filter yields an estimate of the noise level near zero frequency.

Acknowledgements

These ideas have developed over the last 12 years, and, though largely equivalent to the combination of what I presented in my 1981 Ph.D. dissertation and at the 1983 meeting of the Association for Vision and Ophthalmology, the notation and explanations have advanced significantly. Among the many people who provided helpful discussion, I would particularly like to thank Al Ahumada and Horace Barlow (on rigor of the analysis), Dan Kersten, Gordon Legge, Gary Rubin, and Beau Watson (on clarity), Dwight North and Al Rose (on their work at RCA), and Bob Wagner (on R. A. Fisher). I thank Dan Horne for locating the translation of Einstein (1905) and John Perrone for showing me how to produce Fig. 1.18B. This work was supported during 1978–9, by MoD contract 'Spatial noise spectra and target detection/recognition' to F. W. Campbell, and, during 1982–8, by National Institutes of Health Biomedical Research Grant to Syracuse University and National Eye Institute grant EY04432 to me. This chapter was written while on a sabbatical from Syracuse University to visit the Vision Group at the NASA Ames Research Center.

References

Aguilar, M. & Stiles, W. S. (1954) Saturation of the rod mechanism of the retina at high levels of stimulation. *Optica Acta*, **1**, 59–65

Ahumada, A. J. Jr & Watson, A. B. (1985) Equivalent-noise model for contrast detection and discrimination. *J. Opt. Soc. Am. A*, **2**, 1133–9.

Ahumada, A. J. Jr (1987) Putting the visual system noise back in the picture. *J. Opt. Soc. Am. A*, **4**, 2372–8.

Banks, M. S., Geisler, W. S. & Bennet, P. J. (1987) The physical limits of grating visibility. *Vision Res.*, **27**, 1915–24.

Barlow, H. B. (1956) Retinal noise and absolute threshold. *J. Opt. Soc. Am.*, **46**, 634–9.

Barlow, H. B. (1957) Increment thresholds at low intensities considered as signal/noise discrimination. *J. Physiol.*, **136**, 469–88.

Barlow, H. B. (1958a) Temporal and spatial summation in human vision at different background intensities. *J. Physiol.* **141**, 337–50.

Barlow, H. B. (1958b) Intrinsic noise of cones. *Visual Problems of Colour*. National Physical Laboratory Symposium, pp. 615–30. London: HMSO.

Barlow, H. B. (1962a) A method of determining the overall quantum efficiency of visual discriminations. *J. Physiol.*, **160**, 155–68.

Barlow, H. B. (1962b) Measurements of the quantum efficiency of discrimination in human scotopic vision. *J. Physiol.*, **160**, 169–88.

Barlow, H. B. (1977) Retinal and central factors in human vision limited by noise. In *Vertebrate Photoreception*, ed. H. B. Barlow & P. Fatt. New York: Academic Press.

Barlow, H. B. (1978) The efficiency of detecting changes of density in random dot patterns. *Vision Res.*, **18**, 637–50.

Bracewell, R. N. (1978) *The Fourier Transform and Its Applications*, New York: McGraw-Hill.

Burgess, A. E. & Colborne, B. (1988) Visual signal detection IV – observer inconsistency. *J. Opt. Soc. Am.*, (in press).

Burgess, A. E., Wagner, R. F., Jennings, R. J. & Barlow, H. B. (1981) Efficiency of human visual signal discrimination. *Science*, **214**, 93–4.

Campbell, F. W. & Gubisch, R. W. (1966) Optical quality of the human eye. *J. Physiol. (Lond.)*, **186**, 558–78.

Cohn, T. E. (1976) Quantum fluctuations limit foveal vision. *Vision Res.*, **16**, 573–9.

Einstein, A. (1905) Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Annal. Physik*, **17**, 132–48. Translation by Arons, A. B., & Peppard, M. B. (1965) Einstein's proposal of the photon concept – a translation of the *Annalen der Physik* paper of 1905. *American Journal of Physics*, **33**, 367–74.

Elliott, P. B. (1964) Tables of *d'*. In *Signal Detection and Recognition by Human Observers*, ed. J. A. Swets. New York: John Wiley.

Engstrom, R. W. (1974) Quantum efficiency of the eye determined by comparison with a TV camera. *J. Opt. Soc. Am.*, **64**, 1706–10.

Fisher, R. A. (1925) *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.

Friis, H. T. (1944) Noise figures of radio receivers. *Proceedings of the IRE*, **32**, 419–22.

Geisler, W. S. & Davila, K. D. (1985) Ideal discriminators in spatial vision: two-point stimuli. *J. Opt. Soc. Am. A*, **2**, 1483–97.

Graham, N., Robson, J. G. & Nachmias, J. (1978) Grating summation in fovea and periphery. *Vision Res.*, **18**, 815–26.

Green, D. M. & Swets, J. A., (1974) *Signal Detection Theory and Psychophysics*. Huntington, NY: Krieger.

Greis, U. & Rohler, R. (1970) Untersuchung der subjektiven Detailerkennbarkeit mit Hilfe der Ortsfrequenzfilterung. *Optica Acta*, **17**, 515–26. (A translation by Ilze Mueller with amendments by D. G. Pelli, 'A study of the subjective detectability of patterns by means of spatial-frequency filtering' is available from D. G. Pelli, Institute for Sensory Research, Syracuse University, Syracuse, NY 13244, USA.)

Hecht, S., Shlaer, S. & Pirenne, M. H. (1942) Energy, quanta, and vision. *J. Gen. Physiol.*, **25**, 819–40.

Henning, G. B., Hertz, B. G. & Hinton, J. L. (1981) Effects of different hypothetical detection mechanisms on the shape of spatial-frequency filters inferred from masking experiments: I. noise masks. *J. Opt. Soc. Am.*, **71**, 574–81.

Jones, R. C. (1959) Quantum efficiency of human vision. *J. Opt. Soc. Am.*, **49**, 645–53.

Kelly, D. (1972) Adaptation effects on spatio-temporal sine-wave thresholds. *Vision Res.*, **12**, 89–101.

Kersten, D. (1984) Spatial summation in visual noise. *Vision Res.*, **24**, 1977–90.

Legge, G. E., Kersten, D. & Burgess, A. E. (1987) Contrast discrimination in noise *J. Opt. Soc. Am. A*, **4**, 391–404.

Lillywhite, P. G. (1981) Multiplicative intrinsic noise and the limits to visual performance. *Vision Res.* **21**, 291–6.

Linfoot, E. H. (1964) *Fourier Methods in Optical Image Evaluation*. New York: Focal Press.

Lipson, E. D., Foster, K. W. & Walsh, M. P. (1976) A versatile pseudo-random noise generator. *IEEE Transactions on Instrumentation and Measurement*, **25**, 112–16.

Mandler, M. & Makous, W. (1984) A three channel model of temporal frequency perception. *Vision Res.*, **24**, 1881–7.

van Meeteren, A. (1973) *Visual Aspects of Image Intensification*. Report of the Institute for Perception, TNO. Soesterberg, The Netherlands.

van Meeteren, A. & Boogaard, J. (1973) Visual contrast sensitivity with ideal image intensifiers. *Optik*, **37**, 179–91.

Mumford, W. W. & Schelbe, E. H. (1968) *Noise Performance Factors in Communication Systems*. Dedham, MA: Horizon House–Microwave Inc. (See chapter 4.)

Nachmias, J. (1981) On the psychometric function for contrast detection. *Vision Res.*, **21**, 215–33.

Nagaraja, N. S. (1964) Effect of luminance noise on contrast thresholds. *J. Opt. Soc. Am.*, **54**, 950–5.

van Nes, F. L., Koenderink, J. J., Nas, H. & Bouman, M. A. (1967) Spatiotemporal modulation transfer in the human eye. *J. Opt. Soc. Am.*, **57**, 1082–8

North, D. O. (1942) The absolute sensitivity of radio receivers. *RCA Review*, **6**, 332–44.

Papoulis, A. (1965) *Probability, Random Variables, and Stochastic Processes*. New York: McGraw–Hill Book Co.

Pelli, D. G. (1981) Effects of visual noise. PhD thesis. Cambridge: University of Cambridge (unpublished).

Pelli, D. G. (1983) The spatiotemporal spectrum of the equivalent noise of human vision. *Invest. Ophthal. and Vis. Sci. (Suppl.)*, **4**, 46.

Pelli, D. G. (1985) Uncertainty explains many aspects of visual contrast detection and discrimination. *J. Opt. Soc. Am.*, **2**, 1508–32.

Peterson, W. W., Birdsall, T. G. & Fox, W. C. (1954) Theory of signal detectability. *Transactions of the IRE PGIT*, **4**, 171–212.

Peterson W. W. & Weldon E. J. Jr (1972) *Error-Correcting Codes*. Second Edition. Cambridge, Mass: MIT Press.

Robson, J. G. (1966) Spatial and temporal contrast-sensitivity functions of the visual system. *J. Opt. Soc. Am.*, **56**, 1141–2.

Robson, J. G. & Graham, N. (1981) Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Res.*, **21**, 409–18.

Rose, A. (1942) The relative sensitivities of television pickup

tubes, photographic film, and the human eye. *Proc. IRE*, **30**, 293–300.

Rose, A. (1946) A unified approach to the performance of photographic film, television pickup tubes, and the human eye. *Journal of the Society of Motion Picture and Television Engineers*, **47**, 273–94.

Rose, A. (1948a) The sensitivity performance of the human eye on an absolute scale. *J. Opt. Soc. Am.*, **38**, 196–208.

Rose, A. (1948b) Television pickup tubes and the problem of vision. *Advances in Electronics*, **1**, 131–66.

Rose, A. (1957) Quantum effects in human vision. *Advances in Biological and Medical Physics*, **5**, 211–42. (See the bottom paragraph on page 237.)

Rose, A. (1977) Vision: human versus electronic. In *Vertebrate Photoreception*, eds. H. B. Barlow & P. Fatt. New York: Academic Press.

Schade, O. H., Sr (1958) A method of measuring the optical sine-wave spatial spectrum of television image display devices. *Journal of the Society of Motion Picture and Television Engineers* **67**, 561–6.

Shapley, R. (1986) The importance of contrast for the activity of single neurons, the VEP and perception. *Vision Res.*, **26**, 45–61.

Stromeyer, C. F. & Julesz, B. (1972) Spatial frequency masking in vision: critical bands and spread of masking. *J. Opt. Soc. Am.*, **62**, 1221–32.

Sturm, R. E. & Morgan, R. H. (1949) Screen intensification systems and their limitations. *The American Journal of Roentgenology and Radium Therapy* **62**, 617–34.

Tanner, W. P., Jr. & Birdsall, T. G. (1958) Definitions of d' and η as psychophysical measures. *J. Acoust. Soc. Amer.* **30**, 922–8.

Thomas, J. P. (1985) Effect of static noise and grating masks on detection and identification of grating targets. *J. Opt. Soc. Am. A*, **2**, 1586–92.

de Vries, H. L. (1943) The quantum character of light and its bearing upon threshold of vision, the differential sensitivity and visual acuity of the eye. *Physica*, **10**, 553–64.

Watson, A. B., Barlow, H. B. & Robson, J. G. (1983) What does the eye see best? *Nature*, **302**, 419–22.

Watson, A. B. & Pelli, D. G. (1983) QUEST: A Bayesian adaptive psychometric method. *Perception and Psychophysics* **33**, 113–20.

Wyszecki, G. & Stiles, W. S. (1967) *Color Science*. New York: John Wiley and Sons Inc.