J. Physiol. (1969), **202**, *pp.* 421–436
Printed in Great Britain

QUANTUM EFFICIENCY AND FALSE POSITIVE RATE

BY P. E. HALLETT
From the Department of Physiology, University of Alberta, Edmonton, Alberta, Canada

(*Received* 27 *November* 1968)

SUMMARY

1. This paper presents an analysis of the efficiency of performance at the *absolute threshold* of human vision. The data are from the same series as the previous papers (Hallett, 1969*b*, *c*) and consist of frequency-of-seeing curves, thresholds, false positive rates and equivalent background measurements, accumulated as small samples over a number of days.

2. Quantum efficiency is defined here as the ratio of the thresholds of an ideal and a real detector performing the same task with *the same sampling error*. This avoids the problem as to whether the frequency-of-seeing curve of the real detector is exactly a Poisson sum or not.

3. The long-term quantum efficiency can be low (about 0·04) as a result of drifts in the mean threshold.

4. The average short-term quantum efficiency is in the region of 0·1, which is roughly the physiological limit set by Rushton's (1956*b*) measurements of rhodopsin density in the living rods. If this is correct, then the absorption of a quantum, and not the bleaching of a rhodopsin molecule, is sufficient for the generation of a neural event.

5. Application of a simple signal/noise theory to the data gives solutions close to those suggested by Barlow (1956) and shows that false positives almost invariably arise from errors subsequent to the signal/noise decision process.

INTRODUCTION

As is well known, largely as a result of work by Pirenne and colleagues (e.g. Hecht, Shlaer & Pirenne, 1942; Pirenne, 1956; Pirenne & Marriott, 1959), the performance of a human observer at the absolute threshold of vision has a strong resemblance to the behaviour of an ideal quantum counter which fires upon the receipt of c or more quanta. In both cases the probability of a response increases with increasing signal energy in a sigmoid fashion and a mean threshold and S.E. of mean can be defined on the basis of this frequency-of-seeing curve (e.g. Hallett, 1969*c*). In one respect man and an ideal coincidence counter differ: a noise-free detector will never respond to a flash of zero intensity (or 'blank') but a human observer occasionally will.

It is natural to ask how close a human observer can approach to perfection. With respect to vision the practical limit is set by an important investigation (Rushton, 1956*a*, *b*) of the optical density of rhodopsin in the human rods. Rushton concluded that it is most likely that 0·1 of blue-green (507 nm) light striking the cornea is absorbed by the rods. To what extent is the absorbed information about the external world utilized by the eye and brain? How much wastage is there?

A number of attempts to answer the questions raised by the work of Pirenne and of Rushton are reviewed by Barlow (1956, 1957, 1962*a*, *b*) who has introduced useful new types of measurement and analysis. In the previous paper it was shown that frequency-of-seeing curves, sampling errors, and, by implication, quantum efficiencies depended upon whether the measurements were made on a short-term or on a long-term basis. The present paper is of interest because five types of measurement (short-term and long-term frequency-of-seeing curves, $f(\log N)$ and $g(\log N)$, mean log thresholds m, false positive rates $g_{F.P.}$ and equivalent background estimates X of the hypothetical intrinsic noise) have been obtained for three observers. These quantities have not been previously measured in the same observer and it is therefore possible to re-examine the general problems of quantum efficiency and intrinsic noise with relatively few assumptions.

METHODS

The absolute threshold measurements formed a small part of an extensive series of threshold measurements. The general methods have already been described (Hallett, 1969*a*). The fully dark-adapted observer viewed with his left eye a green light (530 nm) of 12′ subtense and 1·5 msec duration which appeared 18° nasally to the fixation point. The test light entered at the nasal margin of the dilated pupil.

Method of constant stimuli. This method of determining thresholds and frequency-of-seeing curves has already been described (Hallett, 1969*c*). Suffice it to say that the stimuli were randomized and that the experimenter did not have access to previous measurements. The theoretical purpose of the experiments was not revealed to the observers.

Calibrations. These have been described (Hallett, 1969*a*). The standard errors of the daily photometric calibrations were 0·02 log and the S.E. in daily shutter setting is assumed to be 0·02 log. Redundancy in the nature and number of the measurements makes it likely that the limiting errors are systematic (e.g. brightness of standard lamps and accuracy of density scale on photometer).

Definitions

sample — a frequency-of-seeing curve for $n = 5$ presentations at each of a series of $i = 10$ intensities, varying in steps of $\Delta(\log N) = 0{\cdot}087$. The samples are obtained on k different days.

$\log N$ — the $\log_{10}$ energy axis.

$g(\log N)$ — a frequency-of-seeing curve obtained by simply adding of the number of 'seen' responses at a given intensity over the k days and dividing the number of presentations at each intensity by $5k$.

$f(\log N)$ a frequency-of-seeing curve obtained by displacing the samples along the $\log N$ axis to a common sample mean threshold, accumulating the responses 'seen' in bins of width $\Delta(\log N)$ and plotting the frequencies in intervals of $\Delta(\log N) \times \{(n/n-1)\}^{0.5}$ in order to avoid bias.

f, g used as suffixes to denote quantities derived from, or as abbreviations for, $f(\log N)$ or $g(\log N)$.

threshold usually the mean threshold on a log scale of measurement which can be defined in several ways on the basis of the f-seeing sample or curve, e.g.

$$\text{mean} = \log N_H + \Delta(\log N)\left(0.5 - \sum_i p_i\right), \tag{1}$$

where N_H is the largest flash energy used, $\Delta(\log N)$ is the step width and p_i is the observed frequency of seeing (probability) at the ith flash energy.

x_i the mean threshold of a single series, i.e. a f-seeing curve sample based on one presentation at each intensity. Note that the symbol x is also used elsewhere unambiguously as the mean number of intrinsic noise events.

$\bar{x}_i$ the mean threshold of a sample of $n = 5$ series.

σ_f the S.E. of mean threshold x_i of a f-seeing sample based on one presentation at each intensity and defined by

$$\sigma_f = \Delta(\log N) \times \left(\frac{n}{n-1}\right)^{0.5} \times \left(\sum_i f_i(1-f_i)\right)^{0.5} \tag{2}$$

on the hypothesis that the responses at each intensity i are independent and of constant probability f_i. This corresponds to a within-sample estimate of error.

σ_g the S.E. of mean threshold of a f-seeing curve $g(\log N)$ based on one presentation at each intensity and defined by

$$\sigma_g = \Delta(\log N) \times \left(\sum_i g_i(1-g_i)\right)^{0.5}. \tag{3}$$

$\sigma_{\bar{x}_i}$ the between sample estimate of the S.E. of the observed mean log threshold $\bar{x}_i$ of a sample.

σ_μ $\sigma_\mu = (\sigma^2_{\bar{x}_i} - \sigma_f^2\, n^{-1})^{0.5}$, the S.E. of the true mean log threshold μ which is distributed with mean m and variance σ^2_μ.

c the Poisson parameter of the $p(\log N)$ curve of an ideal coincidence counter.

c' the parameter of an ideal, stable, noise-free coincident counter which would yield the observed sampling error σ_f if the counter or its responses replaced the observer in all experiments and calculations.

F the variance ratio of two samples.

F quantum efficiency generally.

F_g the long-term over-all quantum efficiency, obtained from the f-seeing curve $g(\log N)$. This efficiency will be lower than $\mathscr{E}(F_f)$ if the observer's mean threshold drifts with time.

$\mathscr{E}(F_f)$ the average short-term overall quantum efficiency, obtained from the f-seeing curve $f(\log N)$, is a better estimate of the instantaneous overall efficiency than F_g if the observer's mean threshold drifts.

F' The over-all quantum efficiencies F_g and $\mathscr{E}(F_f)$ are practical measures which are less than F' the physiological, or primary, quantum efficiency (corresponding to n/N in Barlow, 1956, or F in Barlow, 1957). F' is the fraction of 507 nm $h\nu$ at the cornea which gives rise to nervous events. $\mathscr{E}(F_f)$ is less than F' because f is plotted against log (light), rather than log (light plus noise) and F_g is lower still as a result of drifts in the threshold.

$X\alpha\tau$ — X is the equivalent background estimate of the hypothetical intrinsic light of the visual processes, obtained by Barlow's (1957) procedure. The units of X are external background units, e.g. $h\nu$ (507 nm) cornea deg^{-2} sec^{-1}. α deg^2 is the spatial integration constant and τ sec the temporal integration constant so $X\alpha\tau$ represents the mean number of noise events in the same (space) × (time) sample as the test flash, expressed in external units: $h\nu$ (cornea). Thus $X\alpha\tau$ is related to the mean intrinsic noise x by the primary quantum efficiency F'.

$Q(\chi^2)$ — a measure of goodness of fit: probability that the χ^2 criterion for ν degrees of freedom exceeds the observed value. The null hypothesis is accepted if Q is neither too high nor too low.

Relations

Standard deviations. The relations between the frequency-of-seeing curves and various estimates of error are defined or derived in the previous paper (Hallett, 1969*c*).

Modified Barlow signal/noise analysis. Barlow's (1956) analysis gives rise to a set of criterion versus noise curves which should intersect to give a point solution if the theory is correct.

Consider a stable coincidence counter which responds whenever the Poisson distributed intrinsic noise (mean of x) exceeds the criterion, c events. Then the false positive rate is

$$g_{F.P.} = \sum_{y=c}^{\infty} e^{-x}x^y/y! \tag{4}$$

The c versus curve x corresponding to (4) is easily obtained from Molina's (1942) tables. The normal approximation to (4) is *not* satisfactory if the false positive rate is 0·02 or less so that the mathematical form of the noise distribution is a matter of some importance.

Barlow obtained his second solution curve from the slope of the f-seeing curve at its mean: this should be a Poisson sum on a scale of (say) log (light + noise) but is flattened if plotted against log (light). This relation is correct but in practice it is subject to sampling errors of the mean and of the slope, and the relative intensity calibration of the apparatus near the mean intensity must be very accurate. The quantity σ_f is more satisfactory as it avoids problems of interpolating probabilities around the mean, utilizes information from the whole frequency-of-seeing curve, and the step width $\Delta(\log N)$ can reasonably be obtained from the mean of the several step widths close to the mean threshold.

A Poisson sum on a scale of N has mean $c-x$ and variance c. On a scale of $\log N$ then

$$\sigma' = c^{0\cdot5}\log_{10} e/(c-x).$$

Thus for a step width identical to that for the f-seeing curve $f(\log N)$, viz. $\Delta(\log N).(n/n-1)^{0\cdot5}$,

$$\sigma_f = \{\Delta\log N) \times (n/n-1)^{0\cdot5} \times 0\cdot564 \times 0\cdot4343 \times c^{0\cdot5}/(c-x)\}^{0\cdot5}$$

(similar relations are developed in Hallett, 1969*c*, Methods). The c versus x solution curve may be obtained by substituting (2) for σ_f and simplifying.

A third solution curve is the straight line

$$c \approx F'(10^m + X\alpha\tau), \quad x = F'X\alpha\tau,$$

which represents the fact that the mean of a Poisson sum on a scale of light plus noise is c.

Equivalent background measurements. Barlow (1957, 1958) has described the history of the *Eigengrau* concept and has reported his own analysis and measurements. This type of approach has been powerfully applied by Rushton (1965) to the problems of dark-adaptation.

Briefly in increment threshold experiments if the total noise from all sources is Poisson distributed and tending to the normal distribution then

mean signal = K × (mean noise in space × time sample)
from background plus nervous system$^{0\cdot5}$,

where K, the 'signal/noise ratio', sets the rate at which noise events are inevitably mistaken for signal. Using the notation given in *Definitions*, and if $B\,h\nu$ (507 nm) cornea $\sec^{-1}\deg^{-2}$ is the background intensity, then this relation becomes

$$F'10^m = K(F'\alpha\tau X + F'\alpha\tau B)^{0\cdot5},$$

whence

$$10^m = K(\alpha\tau/F')^{0\cdot5}(X+B)^{0\cdot5}.$$

If $B \gg X$ then 10^m is proportional to $B^{0\cdot5}$ and if $B \ll X$ then 10^m = constant is proportional to X. Thus the intersection of the asymptotes to the log increment threshold versus log background curve is at $B = X$.

The argument is not correct because usually 10^m is proportional to B^{ω}, where ω is not constant, is usually greater than 0·5 and is a function of background intensity and test size and duration. However, Barlow (1957) and Blakemore & Rushton (1965) have noted that the intersection of the asymptotes is independent of test size and duration and background intensity if one restricts attention to the range of background intensities well below the beginning of rod saturation. This has not been my experience in preliminary experiments but even if the asymptote intersection is easily definable the relation (5) may be in error because the distribution of noise events may not tend to normality, in which case K is not that value of a standard normal deviate which cuts off an upper tail equal to the false positive rate.

None the less, relation (5) properly used is valuable. The increment threshold data of Hallett (1969*b*) have been analysed for the individual observers and X found by regression analysis. The values of $X\alpha\tau$ are given in Table 1.

RESULTS

Over-all quantum efficiencies, F_f and F_g

Following the proposal of Barlow (1962*a*), the performance of an observer relative to an ideal detector is expressed by the *over-all quantum efficiency F* defined here by the ratio

$$\frac{\text{threshold of ideal detector}}{\text{threshold of observer}}, \qquad (6)$$

it being understood that the ideal detector performs exactly the same task as the observer with *the same accuracy*. 'Threshold' and 'accuracy' must, of course, be defined in the same way in the two cases.

In principle this definition amounts to placing a filter in front of an ideal coincidence counter. The transmission F of the filter and the coincidence level of the counter c' are both adjusted until the frequency-of-seeing versus log energy curve of the counter plus filter $p(\log a - \log F)$ matches the human frequency-of-seeing curve $f(\log N)$ or $g(\log N)$ in shape and position, e.g.

$$p(\log a - \log F) = \sum_{y=c'}^{\infty} e^{-a} a^y/y! = f(\log N),$$

where the middle term is the cumulative Poisson probability that c' or more counts are registered when the mean count is a.

In practice (*a*) the two *f*-seeing curves may never exactly match in shape so that c' may depend on the criterion of matching, and (*b*) an experiment cannot be performed on an ideal detector. If this were possible the absolute energy scale of the apparatus would be irrelevant and F could be determined with good accuracy. As it is, the threshold of an ideal detector is given by probability theory but the threshold of an observer is susceptible to possibly large errors arising from the difficulties of photometric measurements at low intensities. In addition the shape of an *f*-seeing curve is subject to sampling errors which are by no means trivial.

Identification of c'. Table 1 summarizes the frequency-of-seeing curves $f(\log N)$ and $g(\log N)$ obtained for the three observers by procedures already described in detail (Hallett, 1969*c*).

A frequency-of-seeing sample, consisting of $n = 5$ presentations at each of ten intensities varying in steps of $\Delta(\log N) = 0{\cdot}087$, plus about seven blanks, was obtained either at the beginning or end of the day's 3 hr observation period. Such a sample can be collected in 5–6 min. If the frequencies at a given intensity are simply added for the twenty to twenty-one samples collected over a 3-month period one obtains the 'long-term' *f*-seeing curve $g(\log N)$ on an intensity scale of $\Delta(\log N)$. Suppose now that the samples are displaced along the $\log N$ axis to a common sample mean, and frequencies added in bins of width $(\log N)$ extending from that mean, and the final accumulated frequencies plotted on a scale of

$$(\log N) \times \{n/(n-1)\}^{0{\cdot}5}$$

so as to eliminate bias. This curve, $f(\log N)$, will be identical with $g(\log N)$ if there are no trends or drifts which distinguish short-term from long-term performance. These trends do exist. They have been analysed (Hallett, 1969*c*) and their likely nature has an important bearing on the definition of quantum efficiency (*v.i.*).

Table 1 shows estimates of c' according to two criteria. The calculation of σ_f has already been given (relation (2)). This quantity combines the measured frequencies of seeing into a single number so that by treating the observed and ideal curves in identical ways it is easy to identify a c' which yields the observed σ_f. Relations between σ_f and c' have been given (Hallett, 1969*c*) and, of course, it does not matter if σ_f calculated in this way is not an accurate estimate of the spread of the thresholds x_i (which may occur in the observer's responses at an intensity vary in probability) since σ_f is simply used to 'write the signature' of c' and reduce labour. The sampling error (σ_f or σ_g) criterion for c' is to be preferred because (*a*) the matched ideal detector then performs with the same sampling error as the observer, which is very much in keeping with the spirit and practicality of Barlow's definition of over-all quantum efficiency, and (*b*) the σ_f

TABLE 1. Summary of frequency-of-seeing curves by three observers

Observer...	D.B.		B.S.		M.G.	
Mean threshold 10^m	99 $h\nu$ (cornea)		110		83	
Mean noise $X\alpha\tau$(1)	31 $h\nu$ (cornea)		68		23	
False positive rate:						
Absolute threshold	2/148 = 0·014		3/163 = 0·018		4/166 = 0·024	
Over-all(2)	50/3433 = 0·015		67/3761 = 0·018		72/3843 = 0·019	
$\sigma_{\bar{x}_i}$(3)	0·219 log		0·205		0·119	
σ_f(3)	0·081 log		0·089		0·082	
σ_g(3)	0·108 log		0·105		0·085	
Apparent Poisson parameter c'(3)	Short term	Long term	Short term	Long term	Short term	Long term
From σ_f or σ_g(4)	14	3	9	3	13	7
From Barlow's probit method(5)	10	4	10	4	10	9
Over-all quantum efficiency						
From σ_f, or σ_g, and m	0·135	0·025	0·077	0·023	0·150	0·078
From Barlow's probit method	0·095	0·032	0·081	0·025	0·114	0·115
Exponential term(6)	1·13	—	1·11	—	1·03	—
Provisional signal/noise solution(7)						
From σ_f, m and $X\alpha\tau$: c	23	—	23	—	21	—
x	5·5	—	9·0	—	4·5	—
F'	0·18	—	0·13	—	0·20	—
Expected F.P.R.	10^{-6}	—	$2\times10^{-}$	—	10^{-6}	—

(1) Expressed $h\nu$ (507 nm) cornea. sec^{-1}. $(\frac{1}{4}\pi)$ deg^{-2}. X is derived as in Methods, α taken as $(\frac{1}{4}\pi)$ deg^2 and τ as 0·1 sec. This value of α must be regarded as an underestimate (Hallett, Marriott & Rodger, 1962). τ is as measured for these observers (Hallett, 1969*b*).

(2) Lumped for all sorts of viewing conditions.

(3) Defined in Methods.

(4) Calculated from the relations given in Methods of Hallett (1969*c*).

(5) Barlow's (1962*a*) two-point probit method.

(6) Short-term quantum efficiency could be increased by this factor (see text).

(7) Modification of Barlow's (1956) method (described in Methods).

estimate utilizes statistical information from the whole of f-seeing curve, not from local regions such as the tails or mid point. Also shown in Table 1 are estimates of c' from Barlow's (1962*a*) two-point probit method, based on frequencies near to 0·05 or 0·95, which are very close.

Calculation of the long-term over-all quantum efficiency F_g. There is no difficulty in calculating the long-term quantum efficiency F_g. In effect the f-seeing curve $g(\log N)$ was measured in a single experiment spread over a 3-month period and the samples mentioned above were no more than ordinary pages of the data. Using the definition (6) for a logarithmic measurement scale F_g is given by $(c'_g - \delta)/10^m$ where $c - \delta$ is the antilog mean log threshold of the ideal detector and δ is close to 0·5 (Pirenne, Marriott & O'Doherty, 1957, p. 77) but is in any case easily found by direct calculation. The value for c' based on σ_g gives F_g = *ca.* 0·04.

Clearly the long-term over-all quantum efficiency F_g falls short of the upper limit of about 0·1 set by Rushton (1956) and in very large measure this is due to the variation of the mean μ, because if this variation is eliminated f-seeing curves are steeper and the over-all quantum efficiency is much higher, as will now be shown.

Calculation of the average short term over-all quantum efficiency, $\mathscr{E}(F_f)$. A rigorous definition of the average short-term efficiency is dependent on a hypothesis about the nature of the long-term trends or drifts in performance since

$$\mathscr{E}(F_f) = \mathscr{E}\left(\frac{c'_f - \delta}{10^{\mu}}\right).$$

For many purposes it suffices to assume that the f-seeing curve $f\mathscr{E}(\log N)$ drifts along the $\log N$ axis without change in shape. If this is exactly true then

$$\mathscr{E}(F_f) = (c'_f - \delta)10^{-m}\exp\tfrac{1}{2}(\sigma_{\mu}\ln 10)^2, \tag{7}$$

where the exponential term arises from the need to consider a mean threshold which is not the same as the antilog of the mean log threshold. Table 1 shows that the contribution of the exponential term is small and is probably best ignored. The lowest observed threshold energy was 35 $h\nu$ (507 nm) cornea in a sample for D.B., and if the curve $f(\log N)$ were not considerably flattened F_f would be $(14 - 0{\cdot}5)/35 = 0{\cdot}39$, which is considerably greater than Rushton's limiting value of 0·1 and is only possible as a result of some spurious factor, e.g. a brightening of the apparatus so that the mean threshold was actually higher than 35 $h\nu$ but this possibility is excluded by the daily calibrations.

The average short-term quantum efficiency shown in Table 1 corresponds to definition (7), the exponential term being ignored, and uses estimates of c' from σ_f. It will be noted that for two of the three observers the values are higher than 0·1. Are these values sufficiently high to be fallacious? Almost certainly not. One important source of error is the limited accuracy of photometric calibrations at low intensities; these have been extensively reworked (see Hallett, 1969*a*) and the likely systematic error is 0·05 log which would be compatible with an $\mathscr{E}(F_f)$ which is 26 % lower (or higher), but the absolute thresholds of the present observers are unremarkable when compared with the values obtained by other workers. What is more remarkable is the steepness of the curves $f(\log N)$ and the correspondingly large values of c'_f which arise from the elimination of an important source of biological variation, the drift in the mean μ. It is in fact difficult to fit values of c' when $c' > 8$ because the shape of Poisson sums on semilog plots, and hence σ_f, do not then change very rapidly with c'. As a fair illustration consider that the ratio of the σ_f^2's corresponding to $c' = 9$ and $c' = 14$ is 1·25, which is about the upper 5 % point of Fisher's F ($\nu_1 = \nu_2 = 80$), i.e. if c' is truly 9 it may appear to be $\geqslant 14$ in one experiment in twenty. Finally, Rushton's (1956*b*) measurements and the likely variation between individuals do not exclude the possibility that slightly more than 0·1 of blue-green light striking the cornea is absorbed by the rods. In summary the average short-term quantum efficiency at the absolute threshold is sufficiently high for it to be very likely that there is very little loss of efficiency subsequent to the absorption of light.

Comparison with previous work. Only Hartline and McDonald (in Pirenne *et al.* 1957) have accumulated a $f(\log N)$ frequency-of-seeing curve from small samples displaced to a common mean. Their curves represent the responses of eleven observers to a 3° subtense test located 9° from the fixation point and is effectively based on 250 flashes at each intensity, collected from a number of samples with five flashes at each intensity. The parameter c' is given as 7, though 8 would be as good, but this reduces to $c' = 5$ after correction for bias. There is no information as to whether some of their observers gave f-seeing curves as steep as the present ones, but the over-all 'reliability' of their observers was the same, namely about 14 false positives for 1200 blanks.

Hartline & McDonald's curve cannot reasonably be considered to conflict with the present results since it is inappropriate to compare the average data of a few observers with results for individuals. One should perhaps also note that the fitting of an f-seeing curve by a Poisson curve is no proof of the homogeneity of the data, e.g. the $g(\log N)$ curves for the three observers can be pooled to give an excellent curve for $c' = 5$ (shown in Hallett, 1969*c*, Fig. 1). In addition the sample shifting procedures used in obtaining the curves $f(\log N)$ have a smoothing effect so that the smoothness of these curves gives no reliable estimate of the sampling errors. Hecht *et al.* (1942) and Baumgardt (1960) have obtained frequency-of-seeing curves from single long experiments with about fifty presentations at each intensity, using small short-duration tests appearing about 20 degrees from the fixation point. It is convenient to consider these two excellent and meticulous investigations together. Both sets of data are tabulated by Baumgardt. The parameter c' ranges from 5 to 8 (median 7) and the threshold $\dot{N}$ from 80 to 130 hv (510 nm) cornea (median 102). The thresholds $\dot{N}$ are not mean thresholds but are defined for 0·55 f-seeing so quantum efficiency is given by $c'/\dot{N}$ and ranges from 0·038 to 0·088 (median 0·073).

Barlow (1956) has demonstrated that the position and perhaps shape of frequency-of-seeing curves can be changed by voluntary alterations in criterion. Barlow (1962*a*) has linearized the curve of the ideal detector by transforming the f axis to probit and the energy axis to square root energy. If an actual f-seeing curve can be exactly mimicked by an ideal coincidence counter, it suffices to measure co-ordinates at only two points (P_1, N_1; P_2, N_2) on the f-seeing curve, which for maximum sampling efficiency should be near $P_1 = 0{\cdot}95$ and $P_2 = 0{\cdot}05$. This is exactly what has been done in Table 1 and it was noted that this criterion gave slightly low estimates of efficiency relative to the σ_f method. Now one suspects that if real f-seeing curves depart from the ideal then this is most likely true in the tails of the curves: even if the detector is perfect and c' stable a small error such that 2% of 'seen' are mistakenly called 'not seen' and vice versa will lead to 2% false positives and greater prominence in the tails of the f-seeing curve.

Barlow (1962*b*) has applied his two point probit method to a slightly different problem: he has not asked his observers whether a given flash (either N_1 or N_2 at random) is visible but whether it is the brighter (N_1) or the dimmer (N_2) of the pair. When N_1 and N_2 are near the threshold of detection variations in brightness are slight and the 'over-all quantum efficiency of discrimination' reduces to the over-all quantum efficiency of detection. For these conditions, using a small brief flash 15 degrees from the fixation point Barlow estimated the over-all quantum efficiency as 0·05 (S.E. 0·015) which is a little less than the combined estimate from Hecht *et al.* (1942) and Baumgardt (1960) given above.

Summary. Previous work on frequency-of-seeing curves has been based on long single experiments for which the over-all quantum efficiency of detection at the absolute threshold is in the region of 0·06. Now Rushton's (1956*a*, *b*) measurements of rhodopsin density show that about 0·1 of 507 nm light striking the cornea is absorbed by the rods and there is evidence (Wald & Brown, 1953; Hagins, 1954; Rushton, 1956*a*, *b*) that the efficiency of bleaching *in vitro* or *in vivo* is about 0·6. One is inclined to suspect, then, on the basis of this sort of evidence that the bleaching of rhodopsin is necessary for the generation of a nervous event and that at least 0·4 of the absorbed information about the external world is lost. The present data are of interest because the average short term quantum efficiency is so high (about 0·1) that it is likely that nearly every absorbed quantum is counted and, as Wald & Brown (1953) supposed, the bleaching of a rhodopsin molecule to retinene is not necessary for the generation of a nervous event. Very little, if any, of the information that the retina absorbs from the world is, on average, lost to the brain. Of course this high efficiency can only be appreciated because a factor which can flatten frequency-of-seeing curves and give rise to spuriously low efficiencies (drift in the mean μ) has been eliminated.

False positive rate and other considerations. The method of constant stimuli allows the observer to adopt his own criterion as to what constitutes the presence of the test flash. In this connexion the proportion of blanks which are mistakenly 'seen' is a quantity of considerable interest which has received little experimental study. False positives may arise in two ways: (i) as a result of intrinsic noise or noise of the real background (if present) exceeding the observer's criterion (this sort of mistake is *unavoidable* and will be made by even the most perfect statistician), (ii) as a result of the messages for 'not seen' being distorted into 'seen' and vice versa (this sort of error is, in principle, *avoidable*). Two questions arise: (*a*) is the false positive rate invariant; (*b*) to what extent is the false positive rate made up of avoidable and unavoidable errors?

The first answer is that the false positive rate is probably heterogeneous, *if* one analyses the false positives collected from all the various viewing conditions (zero, transient and steady backgrounds) of Hallett (1969*b*). The second answer is that most false positives are avoidable errors.

The number of blanks presented in a sample was usually $n = 7$, 8 or 9 (the actual number being decided by the experimenter) and the number seen was $s = 0$, 1 or 2. The estimated false positive rate varies with sample size in a striking fashion (see Table 2) and it is no surprise that the over-all false positive rates of Table 1 and the binomial distribution give a poor prediction of Table 2: $Q(\chi^2) = 10^{-7}$, although if the false positive rate is calculated for each sample size for each observer the prediction is excellent: $Q(\chi^2) = 0{\cdot}48$. Does the striking effect of sample size indicate that the experimenter was sometimes able to recognize unreliable performance and detect a temporarily increased false positive rate by making the sample size larger? If so the range of the thresholds x_i in the sample was of little use as a clue because average sample range does not correlate with the size of the blank sample. And if the experimenter did have some tactical plan what would have been its effect if the false positive rate was actually homogeneous? These questions cannot be easily answered and it is probably best to eliminate the possible effects of tactics by lumping the samples for each observer. The numbers of the samples with $s = 0$, 1, 2 blanks 'seen' can then be satisfactorily described by a *negative* binomial distribution (using $p = -0{\cdot}031$, $n = -3{\cdot}29$, obtained from the over-all mean and variance of s): $Q(\chi^2) = 0{\cdot}19$. The fit indicates that false positive rate is probably heterogeneous, and would, perhaps, be improved if p and n could have estimated from the individual observers' data—but the degrees of freedom would then have vanished.

The second question is more difficult to answer. On a signal/noise approach frequency-of-seeing curves, mean threshold, false positive rate and equivalent background estimates of noise are all related by formulations of the type given by Barlow (1956, 1957). In principle the present data contain more information than is necessary to solve the original problems posed by Barlow but the analysis of this and the previous paper have shown that the problems are actually far more complex: the positions and shapes of frequency-of-seeing curves, and probably false positive rates, must all be considered to vary with time. Nevertheless, it can be strongly argued that false positives are not usually unavoidable errors arising from the confusion of noise with signal but are avoidable errors of the sort which may arise when 'yes' is occasionally called when 'no' is appropriate.

Barlow (1957) has considered the case of a coincidence counter which operates with fixed criterion c and is subject to intrinsic noise which is Poisson distributed about a fixed mean x. The false positive rate and f-seeing curve of this counter are, of course, invariant in time. It is easily shown by simple manipulations of Poisson's exponential limit (see Methods) that corresponding to some particular false positive rate or sampling error σ_f is a solution curve, c versus x. It is important to realize that the solution curve for the observed false positive rate and the curve for the observed sampling error σ_f do *not* intersect and this means that the observed false positive rate is inflated by avoidable errors.

Table 1 shows the observed values of σ_f, c' and false positive rate. The locus of the solutions from σ_f corresponds to expected false positive rates of $\leqslant 10^{-6}$ for D.B., $\leqslant 2\times10^{-4}$ for B.S. and $\leqslant 2\times10^{-5}$ for M.G., whereas the observed rates are in the region of 10^{-2}. The locus of the solutions from the observed false positive rates corresponds to σ_f appropriate

TABLE 2. False positive rate: the number of samples of size n with s blanks seen

Observer	D.B.				B.S.				M.G.			
Sample size (n) Blanks seen (s)	7	8	9	Pooled	7	8	9	Pooled	7	8	9	Pooled
$s = 0$	129	143	89	361 (349)	161	130	69	360 (361)	127	171	61	359 (369)
$s = 1$	0	11	13	24 (34)	4	10	22	36 (36)	4	16	25	45 (36)
$s = 2$	0	0	1	1 (2)	0	0	3	3 (2)	0	1	3	4 (2)
Total number of samples	129	154	103	386	165	140	94	399	131	188	89	408
False positive rate	0	0·009	0·016		0·003	0·009	0·033		0·004	0·012	0·039	

The pooled values are the observed row totals for each observer and the bracketed values are predictions using a negative binomial distribution.

to $c' \leqslant 5$ for D.B., $\leqslant 5$ for B.S. and $\leqslant 4$ for M.G. whereas the observed σ_f corresponds to c' of 9–14. So large are these discrepancies that even if the signal/noise model is too simple it still seems likely that very few false positives arise inevitably from noise.

The question now arises as to whether point solutions can be obtained from the intersections of the c versus x curves derived from the sampling error σ_f and from mean threshold and equivalent background measurements (see Methods). Such a solution to be plausible must not yield a primary quantum efficiency F' much larger than the value of 0·1 to 0·15, established by Rushton's (1956*a*, *b*) ophthalmoscopic method. Provisional solutions are indicated in Table 1.

The present results are in remarkable agreement in most respects with those of Barlow. The signal/noise solutions in Table 1 may be compared with the values $c = 19$, $x = 8{\cdot}9$, $F' = 0{\cdot}14$ and expected f.p.r. = 0·002 given in Barlow (1956), yet the present results are calculated from the mean and variance of f-seeing curves and an equivalent background estimate of dark light, whereas Barlow's results are calculated from mean and slope of f-seeing curves and a plausible false positive rate. Using an equivalent background method Barlow's (1957) estimate of dark light at 6·5° eccentricity from the fovea amounts to $-2{\cdot}66$ log scotopic trolands and the present over-all estimate at 18° eccentricity to $-2{\cdot}98$ log scotopic trolands. The number of noise events is about 38 $h\nu$ (cornea) in each case, if one assumes that integration extends over 0·1 sec of time and 0·4 $\deg^2$ of field, which is appropriate to Barlow's observers and conditions (Barlow, 1958), and over 0·1 sec and $\frac{1}{4}\pi$ $\deg^2$ which is appropriate to the present conditions (Hallett, 1969*b*; Hallett *et al.* 1962).

It does seem plausible that something like Barlow's signal noise analysis is true, provided one allows that the false positive rate may be inflated by errors which arise later than the stage of signal/noise decision.

DISCUSSION

Methods of averaging. The frequency-of-seeing curves $f(\log N)$ and $g(\log N)$ are derived from the same data but are averaged in different ways. In which circumstances is the one curve and its corresponding over-all quantum efficiency appropriate and in which circumstances the other?

The curve $g(\log N)$ is the most simply derived curve. The 'long term' quantum efficiency F_g is of interest since it shows the way in which drifts in position and shape of the instantaneous f-seeing curve reduce quantum efficiency but F_g and $g(\log N)$ are of no direct relevance to (say) Barlow's (1957) signal/noise analysis. The reason is very plain if one considers the hypothesis, approximately true for many purposes, that the instantaneous f-seeing curve drifts without change in shape along the log energy axis: $g(\log N)$ is the average of the spread out family of instantaneous curves and is necessarily shallower than an instantaneous curve, but it is the spread and position of the latter curve which is of direct relevance to quantum efficiency. By the same token the curve $f(\log N)$ is a better estimate of the instantaneous curve for it is calculated by displacing the samples to a common mean and would be an unbiased estimate of the instantaneous curve if this did not change its shape during its shifts along the log energy axis. $f(\log N)$ and $\mathscr{E}(F_f)$ are the quantities of interest when over-all quantum efficiency is compared with limits set by the optical density of rhodopsin in the rods. $f(\log N)$ and its mid-slope or sampling error σ_f are the quantities relevant to Barlow's signal/noise analysis. These points are important because lower over-all quantum efficiencies, comparable to those of Barlow (1962*b*) are obtained from $g(\log N)$, and the discrepancies between the signal/noise solutions for observed false positive rate and sampling error σ_f are less striking if $g(\log N)$ is used.

It may be helpful to point out an analogy between the spread of the f-seeing curve of a human observer and the fluctuations in the base line of a sensitive amplifier. The spread on the short-term view ('noise') may differ considerably from the spread on a long-term view ('noise' plus drift), but much may depend on the individual observer or amplifier and on uses or abuses during operation.

Sampling errors. The sampling errors in frequency-of-seeing curves have not received much attention and only Barlow (1962*a*) has considered the sampling errors in over-all quantum efficiency (his estimates do not include the uncertainties in absolute energy calibrations). If the responses at each intensity are independent of each other and of constant probability then the binomial distribution applies and confidence limits are easily found, e.g. if twenty-five out of fifty presentations are seen then the approximate 95 % confidence limits to the estimated frequency of seeing are $0{\cdot}50 \pm 0{\cdot}14$. This sort of scatter is not obvious in results of Hecht *et al.* (1942) or Baumgardt (1960) which must be regarded as fortunate samples in some way or another, but presumably the confidence limits for the apparent Poisson parameter c' are not so small as ± 2 even in these meticulous experiments. In the present case c' has been derived from σ_f or σ_g, since this is most appropriate to the definition of over-all quantum efficiency, and limits to c' can easily be found from consideration of Fisher's **F**, as already shown. The limits are quite wide: the short term quantum efficiency for observer M.G. reduces from 0·15 (Table 1) to 0·08 if one allows that c' is truly 9 and the energy calibration is 0·1 log low.

The usefulness of the quantum efficiency concept. As Barlow (1962*a*) has remarked the initial interest in the quantum efficiency concept arose from the possibility that a large quantity of human performance data could be reduced to the statement that quantum efficiency was high and constant. This rapidly proved not to be the case; quantum efficiency falls with light-adaptation and is also dependent on the details of the task (Barlow, 1962*b*). Probably the main use of the concept is in the comparison of the performance data obtained by different authors by different methods, since it combines into one number information about the spread and the position of frequency-of-seeing curves. In this respect the average short-term quantum efficiencies of the present observers (average 0·11) are higher than the mean value of 0·05 obtained by Barlow (1962*b*), possibly slightly higher than the value of 0·07–0·08 which is appropriate to Hecht *et al.* (1942) and

Baumgardt (1960), but about the same as 0·1, the most likely value of the fraction of blue-green light incident on the cornea which is absorbed by the rods (Rushton, 1956*b*). The present high values are not surprising, because an important source of variation which can affect frequency-of-seeing curves has been eliminated. It seems that little of the information absorbed from the external world is lost. Exactly how much is lost will depend upon the results of a more extensive application of Barlow's (1956) signal/noise analysis than the provisional analysis given in this paper.

The work reported in this paper was supported by the Medical Research Council of Canada and the Defence Research Board of Canada, grants MRC MA 1981 and DRB 9310 122.

I am indebted to the Cairns Memorial Fund and to the Department of Physiology, University College London, for much of the basic optical equipment, and to undergraduate students of the University of Alberta and many others who have acted as observers or given assistance.

REFERENCES

BARLOW, H. B. (1956). Retinal noise and the absolute threshold. *J. opt. Soc. Am.* **46**, 634–639.

BARLOW, H. B. (1957). Increment thresholds at low intensities considered as signal/noise discriminations. *J. Physiol.* **136**, 469–488.

BARLOW, H. B. (1958). Temporal and spatial summation in human vision at different background intensities. *J. Physiol.* **141**, 337–350.

BARLOW, H. B. (1962*a*). A method for determining the over-all quantum efficiency of visual discriminations. *J. Physiol.* **160**, 155–168.

BARLOW, H. B. (1962*b*). Measurements of the quantum efficiency of discrimination in human scotopic vision. *J. Physiol.* **160**, 169–188.

BAUMGARDT, E. (1960). Mesure pyrométrique du seuil visuel absolu. *Optica Acta* **7**, 305–316.

BLAKEMORE, C. B. & RUSHTON, W. A. H. (1965). Dark-adaptation and increment threshold in a rod monochromat. *J. Physiol.* **181**, 612–628.

HAGINS, W. A. (1954). The photosensitivity of mammalian rhodopsin *in situ*. *J. Physiol.* **126**, 37*P*.

HALLETT, P. E. (1969*a*). Rod increment thresholds on steady and flashed backgrounds. *J. Physiol.* **202**, 355–377.

HALLETT, P. E. (1969*b*). Impulse functions for human rod vision. *J. Physiol.* **202**, 379–402.

HALLETT, P. E. (1969*c*). The variation in visual threshold measurement. *J. Physiol.* **202**, 403–419.

HALLETT, P. E., MARRIOTT, F. H. C. & RODGER, F. C. (1962). The relationship of visual threshold to retinal position and area. *J. Physiol.* **160**, 364–373.

HECHT, S., SHLAER, S. & PIRENNE, M. H. (1942). Energy, quanta and vision. *J. gen. Physiol.* **25**, 819–840.

MARRIOTT, F. H. C. (1958). Neural 'dark noise' and retinal quantum-events. *J. Physiol.* **142**, 49–50.

MOLINA, E. C. (1942). *Poisson's Exponential Binomial Limit.* New York: D. Van Nostrand Co. Inc.

PIRENNE, M. H. (1956). Physiological mechanisms of vision and the quantum nature of light. *Biol. Rev.* **31**, 194–241.

PIRENNE, M. H. & MARRIOTT, F. H. C. (1959). The quantum theory of light and the psychophysiology of vision. In *Psychology: A Study of a Science*, ed. KOCH, S., vol. **1**, pp. 288–361. McGraw-Hill.

PIRENNE, M. H., MARRIOTT, F. H. C. & O'DOHERTY, E. F. (1957). Individual differences in night-vision efficiency. *Med. Res. Coun. Spec. Rep. Ser.* no. 294. With a section on 'The frequency of seeing at low illumination' by HARTLINE, H. K. & MCDONALD, P. R.

Rushton, W. A. H. (1956*a*). The difference spectrum and the photosensitivity of rhodopsin in the human eye. *J. Physiol.* **134**, 11–29.

Rushton, W. A. H. (1956*b*). The rhodopsin density in the human rods. *J. Physiol.* **134**, 30–46.

Rushton, W. A. H. (1965). Visual adaptation. *Proc. R. Soc.* B **162**, 20–46.

Wald, G. & Brown, P. K. (1953). The molar extinction of rhodopsin. *J. gen. Physiol.* **37**, 189–200.