PHILOSOPHICAL TRANSACTIONS B

Review

Behavioural and physiological limits to vision in mammals

Greg D. Field[1] and Alapakkam P. Sampath[2]

[1]Department of Neurobiology, Duke University School of Medicine, Durham, NC 27710, USA
[2]Stein Eye Institute, Department of Ophthalmology, UCLA, Los Angeles, CA 90095, USA

APS, 0000-0002-0785-9577

Human vision is exquisitely sensitive—a dark-adapted observer is capable of reliably detecting the absorption of a few quanta of light. Such sensitivity requires that the sensory receptors of the retina, rod photoreceptors, generate a reliable signal when single photons are absorbed. In addition, the retina must be able to extract this information and relay it to higher visual centres under conditions where very few rods signal single-photon responses while the majority generate only noise. Critical to signal transmission are mechanistic optimizations within rods and their dedicated retinal circuits that enhance the discriminability of single-photon responses by mitigating photoreceptor and synaptic noise. We describe behavioural experiments over the past century that have led to the appreciation of high sensitivity near absolute visual threshold. We further consider mechanisms within rod photoreceptors and dedicated rod circuits that act to extract single-photon responses from cellular noise. We highlight how these studies have shaped our understanding of brain function and point out several unresolved questions in the processing of light near the visual threshold.
This article is part of the themed issue 'Vision in dim light'.

1. Introduction

In 1905, Albert Einstein suggested that light may be composed of discrete particles [1]. This idea contradicted the firmly held nineteenth century view that light was a wave. It took more than a decade and several important advances in experimental physics for Einstein's view to gain broad acceptance [2]. Between 1907 and 1940, several experimental psychologists tried to estimate the number of photons required for a human observer to see under ideal dark-adapted conditions [3,4]. The answer was more difficult to obtain than one might guess. The number of photons impinging upon the front of the eye was straightforward to calculate given the intensity of a just-detectable flash; it was approximately 100 photons. However, it was much more challenging to determine the fraction of those photons that were absorbed by retinal photoreceptors. At stake in answering this question were two very different paradigms for conceptualizing the limits to vision. At one extreme, the nervous system could be noisy, thus requiring dozens of photons to be detected to overcome this noise. At the other extreme, one or a few photons may be sufficient, suggesting that cells and perceptual mechanisms in the brain have been optimized to a remarkable extent; they are capable of detecting an elementary particle in the universe [3]. This review tells how an answer to this question was reached and how vision scientists continue to grapple with a seemingly simple question, 'how many photons are required to see?' The answer changed the modern conception of vision and broadly influenced thinking about brain function.

2. Behavioural experiments indicate sensitivity to few photons

Initial attempts to estimate the number of photons required for seeing a just-detectable flash required a combination of carefully controlled behavioural measurements and precise calibration of the light stimulus [3,4]. Key factors

THE ROYAL SOCIETY PUBLISHING

were the size of the spot of light, the spectral composition of the stimulus, the duration of the flash and the energy of the flash. Without knowing all of these factors to high precision, uncertainties would propagate to the estimated number of photons required to see. Human subjects in these experiments needed to be fully dark-adapted and encouraged to maintain a minimal rate of reporting the presence of a flash when no flash was presented, so-called 'false positives' [3,5]. One additional challenge in these experiments was to define what constituted 'reliably seeing' the flash, as it was quickly realized that near threshold subjects were variable in their reports of seeing. Despite these challenges, it was generally accepted that approximately 100 photons at the cornea with a spot size covering 100–1000 rod photoreceptors was sufficient for a human observer to see the flash of light.

While this provided an initial estimate of the number of photons required, it begged the question of how many of these photons reached the retinal photoreceptors; the optical media of the eye will reflect, absorb and refract some fraction of these photons. Direct physical measurements of the optical properties of the cornea, lens and vitreous humour suggested that between 10 and 30% of photons at the cornea would reach the retina [6]. This fraction is frequently referred to as the quantum efficiency, or Q_e. Q_e, therefore, suggested that 10–30 photons were required to be absorbed at the retina for the subject to see a flash of light.

An alternative and clever method for estimating the minimum number of photons required for seeing was developed by Hecht, Schlaer and Pirenne in 1942 and improved upon by Barlow roughly a decade later [3–5]. The approach postulated that variability in seeing a flash of a given strength arose from Poisson variability in the number of photons absorbed from trial to trial. It was well appreciated by this point that the emission of photons from standard light sources obeyed Poisson statistics because each photon acted independently. Therefore, a flash that delivered on average 10 photons to the retina would vary from trial to trial according to a Poisson distribution with a mean of 10. Hecht and colleagues assumed a fixed threshold number of absorbed photons must be exceeded for an observer to see the flash. These two factors, a fixed threshold and trial-to-trial Poisson variability in photon number, provided a mathematical framework for explaining the psychophysical observation of a transition between flash strengths that are rarely detected to those routinely detected by human observers. This framework, therefore, allowed the threshold number of photons required for seeing to be estimated from the full psychometric frequency-of-seeing curves of human observers. The estimated threshold was a small number of photons: approximately five photons given a Q_e of approximately 5%.

A surprising realization emerged from these experiments and analysis. Given the size of the spot of light falling on the retina, these experiments indicated that approximately five to seven photons were being absorbed across a population of approximately 500 rod photoreceptors. The probability that any individual photoreceptor was routinely absorbing more than one photon was vanishingly small. Therefore, these behavioural experiments indicated that individual rod photoreceptors could respond reliably to individual photons—a prediction that was verified directly more than 30 years later by recording the electrical responses of individual rod photoreceptors to single photons [7,8].

The analysis and interpretation of the work by Hecht and co-workers, however, failed to explain one fundamental aspect of their (and other's) data [5]. This aspect was the infrequent but insuppressible phenomenon of subjects reporting that they saw a flash when no flash was delivered. Horace Barlow, inspired by developments in signal detection theory, proposed an important insight into the source of these 'false' reports. He assumed that the biological system itself had some insuppressible source (or sources) of noise that could occasionally masquerade as light. Barlow elaborated Hecht's description of the psychophysical data to include this source of noise, which he called 'dark light'. This was inspired from the nineteenth century description by German psychologists of eigengrau, or the 'grey' seen by human subjects in complete darkness. To estimate the amount of this noise present in the visual system, Barlow allowed subjects to report 'maybe seeing' in addition to 'seeing' and 'not seeing' the flash. This new choice allowed subjects to report seeing dimmer flashes of light more frequently, at the cost of also reporting more frequently the presence of a flash when none was delivered (also referred to as 'false positives'). This elaboration upon the experiments and analyses of Hecht and co-workers allowed Barlow to describe more fully the frequency of seeing data collected from human observers, while also estimating noise intrinsic to the visual system [5,9]. Unfortunately, this framework permitted a broad range of noise values that all explained the psychophysical data equally well (for a more complete description, see [10]). Nevertheless, Barlow's insights directed future physiology experiments to determine the mechanisms that optimize the nervous system to approach single-photon detection (see below).

It is difficult to overstate the impact of these quantitative behavioural studies on our modern perspective of the brain, and on biological systems more generally. The realization that neurons could be optimized to reach a physical limit imposed by nature was a revelation into the possibilities of natural selection. It spurred on a host of studies on the physical limits of audition, olfaction and sensory processing more generally [11]. Barlow's insights about the factors limiting absolute visual threshold probably influenced his later thinking about efficient coding and redundancy reduction by the nervous system, which itself was enormously influential [12]. The now commonplace assumption that the nervous system is optimized for any particular task has roots that intertwine with the discovery that dim-light vision is limited more by the discrete nature of light than by the complex and intrinsically noisy machinery of the brain.

3. Properties of rod photoreceptors that support reliable single-photon responses

The century of behavioural experiments summarized above indicates that rod photoreceptors respond reliably to single photons. As we will see below, photons initiate a biochemical cascade called phototransduction that leads to a change in the current flowing across the membrane of the rod outer segment (figure 1a; reviewed in [14]). Thus, the response generated by a single-activated rhodopsin molecule (R*) in an outer segment must be distinct from noise (or variability) intrinsic to phototransduction and other sources of cellular noise. Three important constraints can be derived from this statement. First, transduction noise in darkness needs to be mitigated such
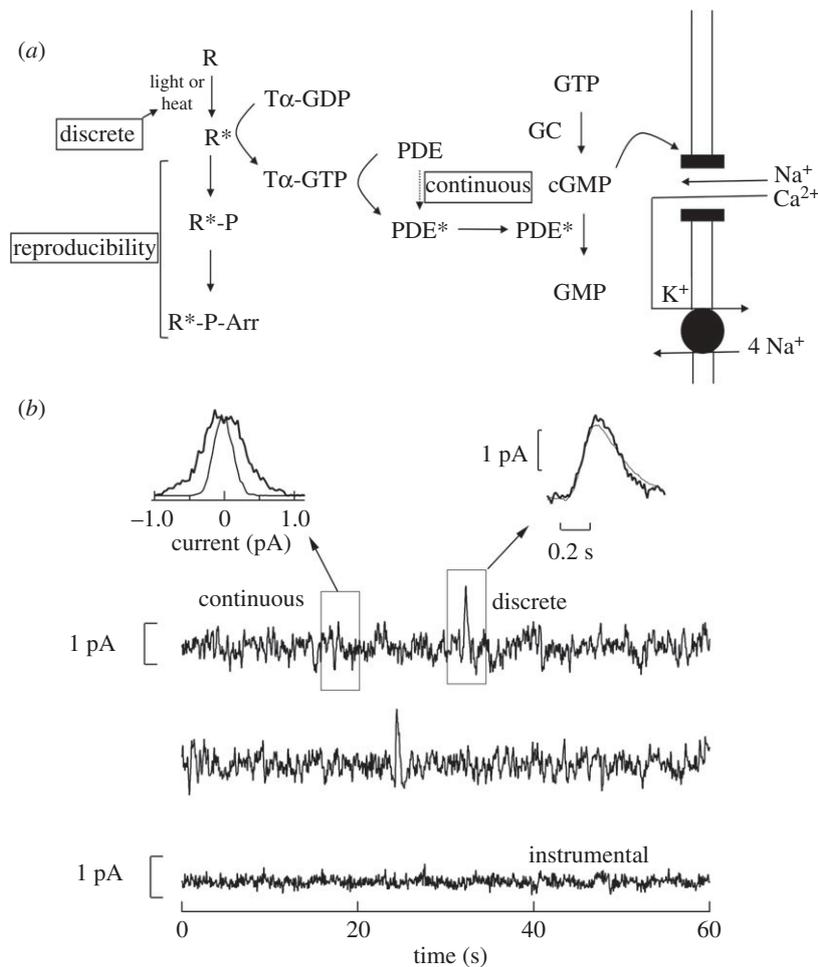
Figure 1. (*a*) The single-photon response is initiated in the rod outer segment when rhodopsin (R) absorbs a photon and is converted to activated rhodopsin (R*). R* in turn catalyses the exchange of GTP for GDP on the α subunit of the heterotrimeric G-protein transducin (T), which in turn stimulates the cGMP phosphodiesterase (PDE) to hydrolyse cGMP to GMP. This leads to the closure of cGMP-gated channels, which decreases the influx of Na^+ and Ca^{2+} in darkness and causes the photoreceptor to hyperpolarize. The deactivation of phototransduction occurs as the catalytic activity of R* is reduced by the phosphorylation of its C-terminus (R*-P) and the subsequent capping by visual arrestin (Arr). In addition, Tα-GTP must deactivate by hydrolysing GTP to GDP, and cGMP concentration must be restored through its Ca^{2+}-dependent production by guanylyl cyclase (GC). Also noted are that discrete noise is produced by the thermal activation of R, continuous noise is produced by the transducin-independent spontaneous activation of PDE, and reproducibility in the single-photon response must occur through the stereotyped shut-off of R*. (*b*) Dark noise in photoreceptors [13] as shown by suction electrode recordings from a primate rod. Present in darkness are two forms of noise: (i) infrequent discrete noise (right inset: thick trace) which is indistinguishable from the average single-photon response (right inset: thin trace), and (ii) omnipresent continuous noise (left inset: thick histogram) which is only suppressed when all the cGMP-gated channels are closed and instrumental noise can be isolated (left inset: thin histogram). Data are reproduced from Field *et al.* [10].

that single-photon responses are as large as possible relative to this noise. Second, the response-to-response variance in the signal produced by a single-photon absorption needs to be low to avoid mistaking a response for noise. Third, neurotransmitter release must be reliably controlled at the rod synaptic terminal to signal a photon absorption to downstream neurons. These constraints are met by several specializations that push against the physical and energetic limits of the system.

(a) Constraint #1: signal-independent phototransduction noise

Two major forms of noise compete with the single-photon response in rod phototransduction: *discrete* and *continuous* noise. Each form arises from distinct elements of phototransduction (figure 1*b*). Discrete noise is produced by the thermal activation of rhodopsin. This noise produces a change in the membrane current that is indistinguishable from that produced by a photo-activated rhodopsin molecule

(R*, figure 1*b*). This is the *dark noise* postulated by Barlow to limit vision and it can be easily expressed in terms of an equivalent light source [5]. While these noise events are relatively large, they are also rare. In mammalian rods they occur approximately once every 60–120 s [8,15]. Given a rod contains more than 10^8 rhodopsin molecules, the Poisson rate of spontaneous activation of an individual molecule is 1 every approximately 500–1000 years. This degree of molecular stability is outstanding for a biological enzyme and probably approaches a thermodynamic limit given the high-energy barrier required to activate the molecule [16].

Continuous noise is produced by the spontaneous activation of cGMP phosphodiesterase along with the rapid Ca^{2+}-dependent synthesis of cGMP by the guanylyl cyclase. While smaller in amplitude than discrete noise, continuous noise is omnipresent; thus, it dominates the rod photocurrent (figure 1*b*) [13]. The smaller the amplitude of these fluctuations, the higher the signal-to-noise ratio (SNR) of the single-photon response. Why then, has natural selection

not reduced the spontaneous activation of phosphodiesterase as it has reduced the spontaneous activation of rhodopsin? Such a reduction would presumably increase the detection sensitivity of vision and improve the fitness of the organism.

A possible answer to this question lies in the realization that the kinetics and amplitude of the single-photon response, and its recovery to baseline, are largely determined by the basal level of cGMP turnover [17]. Basal turnover of cGMP is largely set by its spontaneous hydrolysis to 5′-GMP by phosphodiesterase, which stimulates indirectly cGMP synthesis by guanylyl cyclase [15,18]. Thus, decreasing the spontaneous activity of phosphodiesterase would decrease the basal turnover of cGMP in darkness and cause a substantial slowing of the single-photon response. The amount of continuous noise is, therefore, intrinsically linked to the amplitude *and kinetics* of the single-photon response. The single-photon response is already relatively slow at approximately 0.25–0.5 s in duration in mammalian rods, and may limit the temporal resolution of night vision, which is only 2–4 Hz [19,20]. Thus, rod phototransduction has apparently settled on a balance between the temporal resolution and detection sensitivity of rod-mediated vision.

(b) Constraint #2: signal-dependent phototransduction noise

Discrete and continuous noise are observed in the absence of light. When stimulated by a photon, variability in the photoreceptor response cannot be fully accounted for by these two noise sources [7]. This indicates the presence of a signal-dependent noise source, resulting in added variability to the amplitude and kinetics of the single-photon response [21–24]. The presence of a signal-dependent source of noise is unsurprising; phototransduction is driven by the absorption of a single photon and the resulting activation and deactivation of a single molecule of rhodopsin. In the absence of any controlling mechanisms, the shut-off of rhodopsin is expected to be a first-order stochastic point (Poisson) process, such as the open-time for an ion channel. Such a process implies that the expected active lifetime of rhodopsin would obey an exponential distribution, and the single-photon response would have a coefficient of variation equal to unity.

Remarkably, while the single-photon response is variable, its coefficient of variation is three to four times less than expected from a stochastic point process. Therefore, some mechanism(s) must control the shut-off of rhodopsin to make it more stereotyped from trial to trial. The dominant mechanisms at work appear to be a series of phosphorylation events that occur on rhodopsin's C-terminus by rhodopsin kinase, which are followed by the binding of rhodopsin by arrestin [25–27]. Ca^{2+}-dependent feedback and saturation within transduction also probably play a limited role to reduce response variability [22–24]. A multistep shut-off mechanism has an important design feature, which is to delay most of the response variability to the falling phase of the single-photon response [23]. This delayed variability results in a highly reproducible rising phase and peak, which is largely what drives the response of downstream neurons [28–30]. Thus, even the mechanisms that terminate the catalytic activity of rhodopsin appear to be optimized to transmit single-photon responses to the retinal circuits that process these signals.

(c) Constraint #3: reliable neurotransmitter release

In addition to generating reliable single-photon responses, rod photoreceptors must reliably transmit these responses to their postsynaptic partner, rod bipolar cells [31]. Not surprisingly, several specializations have been identified at the rod spherule that support efficient, low-noise signal transmission. A key feature is that in darkness, rods are relatively depolarized with a membrane potential near −40 mV [32,33]. This is atypical for a neuron, and has two important implications for the release of neurotransmitter: (i) it causes a high rate of vesicular release of glutamate; (ii) at −40 mV Ca^{2+} channel opening depends steeply on membrane potential. Therefore, small changes in membrane potential produced by a single-photon response are tightly linked with changes in Ca^{2+} influx at the synapse and the resulting changes in vesicle release. Further supporting a tight coupling between phototransduction and transmitter release is the dominance of L-type Ca^{2+} channels ($Ca_V1.4$). These channels display little desensitization [34,35] and thus facilitate Ca^{2+}-dependent transmitter release to be high in darkness and may allow more accurate tracking of the single-photon response time course.

An important factor in determining the extent to which these properties of glutamate release at the rod synapse improve signal transmission is the variability in transmitter release for a given average rate. It is frequently assumed that vesicle release from the rod is Poisson, based on measurements from other synapses [36]. If correct, this implies that given a fixed proportional change in mean vesicle release between darkness and a single-photon response, higher basal release rates will produce a response with a higher SNR. For example, if basal release is four vesicles per unit time in darkness and drops to two (50% decrease) during a single-photon response, the SNR will equal approximately 0.8 given Poisson release (SNR = $(\mu_{darkness} - \mu_{response})/\sqrt{(\sigma_{darkness}^2 + \sigma_{response}^2)}$, *note* $\mu = \sigma^2$ for a Poisson process). However, if basal release is 20 vesicles per unit time and drops to 10 (again, 50% decrease), the SNR will equal approximately 1.8. Importantly, a possible function of the ribbon synapse and/or a consequence of high release rates is that vesicle release may be more regular than a Poisson process [37]. This could also support higher fidelity signalling between the rod and the rod bipolar cell, and thus remains an important factor to determine in future experiments.

4. How do retinal circuits relay single-photon responses to retinal ganglion cells?

(a) A circuit for efficient signal convergence

The mechanisms described above allow rod photoreceptors to generate single-photon responses that are reproducible from trial to trial. The task faced by downstream neurons is to process efficiently these signals while introducing minimal noise. A circuit of neurons that is specialized for this task is present in the mammalian retina, and called the rod bipolar (or primary) pathway (figure 2b). This circuit begins with the convergence of many rod signals onto rod bipolar cells, which receive exclusive input from rods. Signals from rod bipolar cells converge via excitatory (glutamatergic) synapses onto AII amacrine cells [31]. Signals can also converge across many AII amacrine cells because they are coupled electrically *via* gap junctions. Interestingly, these gap junctions are
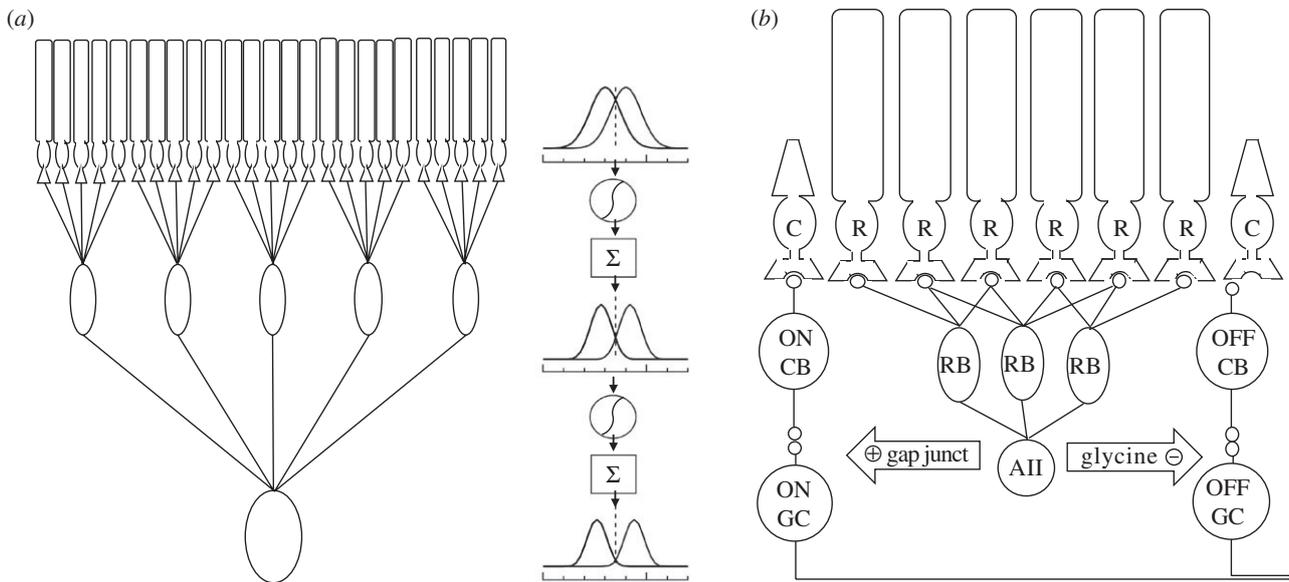
Figure 2. (a) The generalized motif for separating distributions of signals and noise that overlap substantially is to perform a nonlinear operation (thresholding) before summing them. Shown conceptually is pooling of rod photoreceptors in two stages. If distributions of signals and noise overlap substantially (right) then separating signals and noise by an appropriately placed threshold (dashed) before pooling at the subsequent stage can provide greater separation between the two distributions. A second round of nonlinear thresholding following by summation will allow further separation of the distributions. (b) The rod bipolar pathway in the mammalian retina is highly convergent as many rods (R) make selective connections with one class of rod bipolar cells (RB). In turn, many rod bipolar cells contact an AII amacrine cell that feeds signals through gap junctions to the ON cone (C) bipolar (CB) pathway and through glycinergic synapses to the OFF cone bipolar pathway.

modulated by both light and dopamine providing a point of control over the degree of signal convergence across the AII network [39].

Over these first few synapses, signal convergence is the dominant motif with 20–50 rods converging on a single rod bipolar cell and as many as 20 rod bipolar cells converging onto a single AII amacrine cell [40,41]. Divergence (e.g. one rod to many rod bipolar cells) is minimal. However, after AII amacrine cells, signals diverge to many cell types in the retina. Most of the approximately 10 types of cone bipolar cells receive input from AII amacrine cells, and at least one type of OFF ganglion cell receives direct glycinergic inhibition from AII amacrine cells [42,43]. What is the primary function of this dedicated rod circuit?

The answer to this question becomes clearer when considering the nature of processing a signal that is sparse across a large array of noisy detectors. At visual threshold, a tiny fraction of rods absorb a photon every quarter of a second (perhaps 1 in 10 000; reviewed in [44]). If the signal generated by a rod that absorbed a photon is pooled linearly with the noise generated by thousands of rods that did not absorb a photon, the signal will be swamped by the noise [8,36]. The major task of the rod circuit is to prevent these sparse signals from being overwhelmed by noise.

The circuit accomplishes this task by pooling rod signals nonlinearly [45]. This nonlinear pooling can be conceptualized as a threshold that is applied to the output of each rod. Signals below the threshold are suppressed to zero, while signals above the threshold are allowed to pass the synapse. Remarkably, the synapse between rods and rod bipolar cells appears to perform this operation nearly optimally given the SNR of rod photoreceptors and the dimmest light levels encountered in the natural (terrestrial) environments that support vision. One counterintuitive result of this thresholding nonlinearity is that many (approximately half) of single-photon responses in rods do not pass the synapse: only those responses with the largest

amplitudes pass. This is seemingly wasteful given that these signals are the result of numerous optimizations within rod phototransduction (see above). However, the cost of discarded signals is more than compensated for by the benefits of eliminating noise. The net effect of the thresholding nonlinearity at this synapse is a several-hundred-fold increase in SNR [45].

The mechanisms that achieve this nonlinear processing are located at the dendritic tips of the rod bipolar cells. This allows rod responses to be processed largely independently prior to their signals being combined at the rod bipolar soma. The mechanism leverages the high vesicular release rate of glutamate in darkness. Glutamate release is detected by the rod bipolar cell dendrites via type 6 metabotropic glutamate receptors (mGluR6; see [46]). Glutamate binding to mGluR6 drives activity of the heterotrimeric G protein, $G\alpha_o$ [47], which ultimately leads to the closure of cationic TRPM1 channels [48–50]. Thus, TRPM1 channels are opened and the rod bipolar cell is depolarized when a rod decreases its glutamate release. However, in darkness, the transduction machinery between mGluR6 and TRPM1 is held in saturation [51]. Thus, the rod bipolar cell is insensitive to small fluctuations in glutamate release at the synapse with each rod; only large decrements in glutamate release are sufficient to cause depolarization in the rod bipolar cell. Such an effect is seen in figure 3, which shows the increase in discriminability of the single-photon response compared with dark noise as signals flow from rods to rod bipolar cells. This confluence of mechanisms (rod glutamate release, mGluR6 signalling and second messenger saturation) collectively allows the synapse to perform a nonlinear thresholding operation, rejecting rod noise and small single-photon responses, while passing large rod responses.

This motif of nonlinear processing followed by signal pooling (figure 2a) is probably recapitulated through the retina. Rod bipolar cells are not noiseless, and thus it is probable that nonlinear processing also optimizes signal
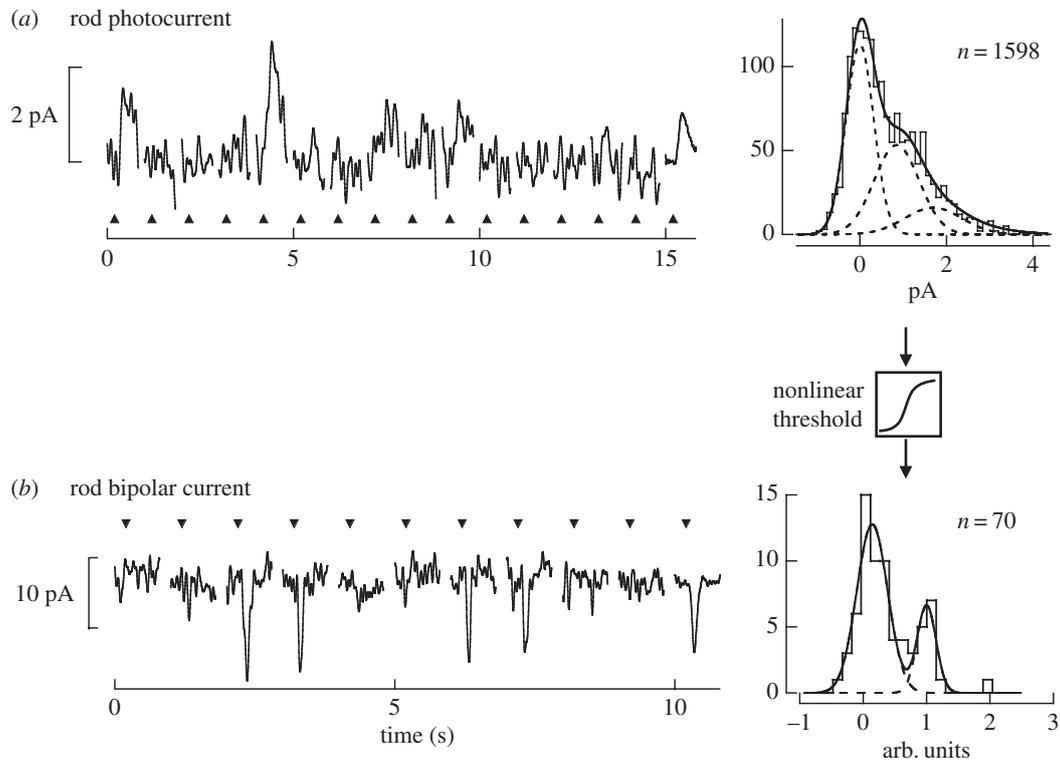
Figure 3. (a) The response of a rod photoreceptor held in a suction electrode to repeated presentations of a dim flash that yield less than one R* per trial. The average response per photon is at the far right. The amplitude of each trial was estimated by their correlation with the average response and plotted in a histogram. The histogram was fit as the sum of Gaussian distributions centred around the average amplitude of the dark noise, the single-photon response and the double-photon response. The areas of each of these distributions were scaled based on the Poisson probability of observing 0, 1 or 2 photons absorbed. Note the region of overlap between the distribution of noise and single-photon responses. (b) Similar analysis was done with rod bipolar cell responses, except those were collected during perforated-patch clamp recordings while holding the membrane potential at −60 mV. The greater signal-to-noise ratio is apparent from the reduced region of overlap in the distributions of noise and single-photon responses. Data are reproduced from Okawa et al. [38].

transmission to AII amacrine cells [52]. The motif may return again at signal summation in retinal ganglion cells, which receive convergent input from many AII amacrine cells and/or cone bipolar cells [53].

Further insights can be gained by comparing the processing of rod signals to that of cone signals. Cone signals immediately diverge to approximately 10 bipolar cell types [54–58]. This divergence establishes the parallel processing of visual information, a strategy that is carried forth throughout the brain. Cones typically operate in a stimulus regime with high SNR, which facilitates this approach. Rods typically operate in a stimulus regime with low SNR such that photon absorptions are sparse in space and time. Thus, prior to the divergence of rod signals across parallel pathways, they must be optimally (nonlinearly) pooled through multiple stages of convergence. Once rod signals reach AII amacrine cells, continuous noise in the rods and noise in the rod bipolar cells has been largely discarded. Only then are the signals sufficiently conditioned to support reliable parallel processing across cone bipolar cells and retinal ganglion cells.

5. What is the limiting noise source? Fundamental trade-offs in setting absolute visual threshold

It is commonly claimed that the spontaneous activation of rhodopsin limits the detection of photons at visual threshold [5].

However, a somewhat more nuanced perspective has arisen from recent work. As described above, continuous noise is omnipresent in the rod photocurrent and threatens to swamp rare photon detection events when signals from many rods are pooled in the retina and brain. Unless rod signals are processed with an appropriately set nonlinearity, visual threshold will be degraded by continuous noise such that limits imposed by the spontaneous activation of rhodopsin cannot be reached [38,45, 51]. This raises the question, does continuous noise or discrete noise limit vision? The answer depends on several factors.

One important factor is the extent to which the thresholding nonlinearity is set optimally given the distributions of signal and noise of the rod photoreceptor. For any given relationship between signal and noise in rod photoreceptors, there is a single best nonlinearity for optimally pooling rod responses at a given flash strength [46]. A mismatch between this nonlinearity and rod SNR has been shown experimentally to degrade the absolute sensitivity of vision [38]. These experiments used a genetically modified mouse line to alter the SNR distribution of the rods. In particular, the mean single-photon response became larger in amplitude and the variance of the continuous noise was increased; the net change was an overall improvement in the SNR of the single-photon response because the signal increased more than the noise. The simple prediction from this manipulation is that visual sensitivity would increase because the SNR of the photoreceptor was increased. However, quantitative behavioural experiments revealed a substantial decrease in behavioural sensitivity. This was largely because nonlinear thresholding at the
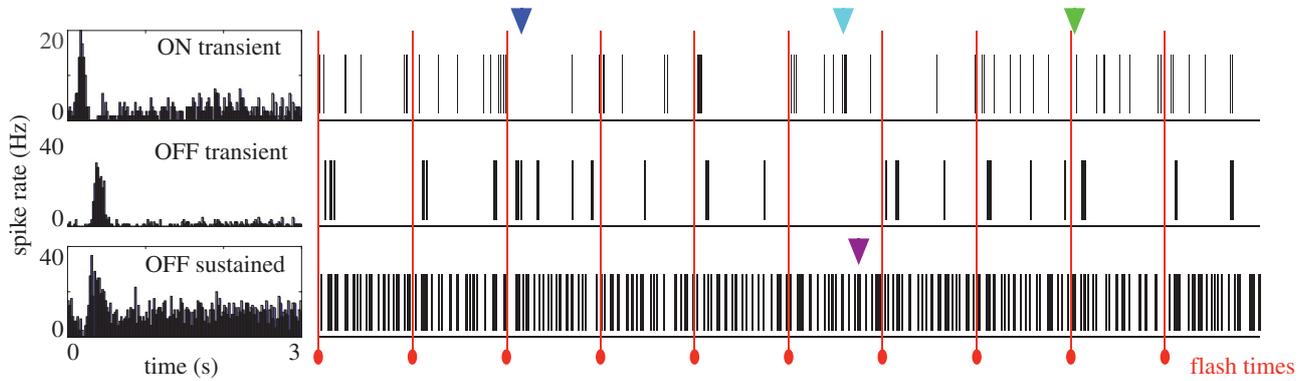
Figure 4. Retinal ganglion cell responses to dim flashes in darkness. Spike responses were measured on a high-density multielectrode array [62], which allows simultaneous measurement across hundreds of cells. The histograms on the left show the spike rate of three retinal ganglion cells following a 2 ms dim flash (at time zero) presented to the retina. The flash produced about 1 R* per 1000 rods. Each ganglion cell was classified as a distinct type according to its response properties measured at higher light levels: an ON transient cell (top), an OFF transient cell (middle) and an OFF sustained cell (bottom). Note the peak firing rates for the OFF cells are slightly delayed relative to that for the ON cell. To the right are raster plots indicating the timing of action potentials relative to individual flashes; flash times are indicated by the red lines. The green and blue arrowheads indicate flashes that produced clear responses in only one or two of the cells respectively. The cyan arrowhead indicates a false positive response in the ON cells when no flash was delivered, which is also seen in the OFF sustained cell (purple arrowhead). How the brain reads out signals of this nature across many distinct retinal ganglion cell types remains an open question.

rod-to-rod bipolar synapse no longer optimally segregated single-photon responses from continuous noise. In this view, continuous noise is potentially limiting because it necessitates the nonlinear threshold; and a suboptimal nonlinear threshold will limit sensitivity. Furthermore, even when the nonlinear threshold is optimal, it decreases the fraction of photons at the cornea that participate in vision after phototransduction, thus reducing quantum efficiency, Q_e.

A second important factor is the number of rods that are being pooled to drive the response of a downstream neuron or behaviour. If the number of rods is small (less than 100), then the spontaneous activation of rhodopsin is rare within the pool. Thus, changes to the rate of discrete noise events will minimally impact the sensitivity of a downstream neuron. On the other hand, for a population as small as 20 rods, the benefit of nonlinear pooling to mitigate continuous noise can be greater than 100-fold [45]. Under these conditions, small changes in the amount of continuous noise can have large effects on downstream sensitivity. Thus, continuous noise is limiting in this regime.

If the number of pooled rod signals is very large (greater than 100), then the spontaneous activation of rhodopsin becomes more frequent across the pool of rods. Therefore, small changes in this rate will have significant consequences on signal fidelity. Under these conditions, minimizing the discrete noise rate will have a large impact on behavioural sensitivity and can be considered limiting, in that context.

The consequence of these insights is that they imply the limiting rod noise source may depend on species and may even depend on retinal location within a species. For example, a retinal ganglion cell near the fovea of a primate pools over as few as 10–20 rod photoreceptors, while the pooling increases to approximately 5000 rods in the periphery [59]. Assuming the SNR properties of rods are constant over eccentricity, then the spontaneous activation of rhodopsin may limit retinal ganglion cell sensitivity in the periphery, but not near the fovea. By comparison, a retinal ganglion cell in the peripheral cat retina can pool signals over 100 000 rods, 10-fold more than in the primate retina [59]. Thus, the spontaneous activation of rhodopsin may be a more dominant noise source in cats where the convergence of rod signals is enormous [60].

The above arguments largely assume that noise downstream of the rods can be ignored and that nonlinear processing is optimized for rod signal and noise. However, recent work indicates that the noise downstream of rods is present in the retina [53,61]. The consequences of interactions between these downstream noise sources, rod noise and the number of rod signals that are being pooled is a potentially fruitful avenue for future research.

6. Vision at threshold: open questions

A century of behavioural studies and more than 50 years of neurophysiology have revealed much about the limits to vision. However, many open questions remain. For example, most neurophysiology experiments in this field have focused on neurons in the rod bipolar pathway (figure 2b) and a couple of retinal ganglion cell types. Thus, our understanding of how single-photon responses are represented across the 20–30 distinct ganglion cell types is minimal. Do a small number of ganglion cell types carry these signals to the brain, or are they distributed across many parallel pathways? Are ganglion cells acting as independent encoders of the signals at visual threshold, or does correlated activity across the population improve detection performance? An example of this ambiguity is shown in figure 4, which shows the spike response of three mouse retinal ganglion cells of different classes measured simultaneously. When a dim flash is delivered at regular intervals, it is clear that each cell responds to the flash (see histograms in figure 4). However, on individual trials there is significant variability in the response across the three cells. For example, there are trials in which the ON transient cell responds to the flash but the OFF transient cell does not, and vice-versa. Considering that mechanisms described above may cause the loss of single-photon responses *in lieu* of their discrimination from noise, distributing different thresholds across several classes of retinal ganglion cells in parallel may allow the visual system to maximize sensitivity. Furthermore, it is unclear if the rate of false positive signals can be mitigated by considering several retinal ganglion cell

8

rstb.royalsocietypublishing.org   Phil. Trans. R. Soc. B 372: 20160072

classes in parallel. Ultimately, answering these questions will provide a more complete view of how signals initiated in the rods ultimately drive the output of the retina.

Similarly, the focus on neurophysiology of retinal neurons near visual threshold has resulted in a dearth of information about the representation and processing of single-photon responses in downstream brain areas, such as the dorsal lateral geniculate nucleus, the superior colliculus and primary visual cortex. Are there mechanisms in visual cortical circuits that help to suppress noise when photons are scarce? To what extent does the sparsification of spikes as signals flow from RGCs to the lateral geniculate nucleus to visual cortex [63] control the SNR of single-photon responses? How does observer confidence that a dim flash of light has been perceived [5,9] relate to the neural activity in primary visual cortex or other cortical areas? Given that single-photon responses are an irreducible and highly relevant visual signal, determining how many neurons in visual cortex are activated by a single photon could reveal much about cortical architecture and signal processing.

Finally, previous research has focused largely on the problem of detecting photons at visual threshold. However, visual processing performs operations other than simple light detection. For example, identifying self-motion and the motion of objects are central computations performed in the retina, superior colliculus and visual cortex. To estimate accurately the direction and speed of motion, signals originating in rods must carry some information about the relative timing of photon absorptions and this information must be preserved by retinal circuitry. While it is appreciated that the rising phase of the single-photon response is preferentially transmitted through the retina near visual threshold, it remains to be seen how rod noise (e.g. continuous versus discrete) limits the temporal fidelity of signals at the retinal output. Does the temporal sensitivity of retinal processing reach limits imposed by either form of rod noise? If yes, what mechanisms achieve the efficient extraction of temporal information from single-photon responses? If no, what retinal noise sources degrade this information? Answering these questions will probably provide new insights into the mechanisms that both optimize and limit neural function.

References

1. Einstein A. 1905 On a hueristic viewpoint concerning the production and transformation of light. *Ann. Phys.* **17**, 132–148. (doi:10.1002/andp. 19053220607)

2. Millikan RA. 1924 The electron and the light-quant from the experimental point of view. *Nobel Lecture*, 23 May.

3. Hecht S, Shlaer S, Pirenne MH. 1942 Energy, quanta, and vision. *J. Gen. Physiol.* **25**, 819–840. (doi:10.1085/jgp.25.6.819)

4. Bouman MA. 1961 History and present status of quantum theory of vision. In *Sensory communication* (ed. WA Rosenblith), pp. 377–401. Cambridge, MA: MIT Press.

5. Barlow HB. 1956 Retinal noise and absolute threshold. *J. Opt. Soc. Am.* **46**, 634–639. (doi:10. 1364/JOSA.46.000634)

6. Donner K. 1992 Noise and the absolute thresholds of cone and rod vision. *Vision Res.* **32**, 853–866. (doi:10.1016/0042-6989(92)90028-H)

7. Baylor DA, Lamb TD, Yau KW. 1979 Responses of retinal rods to single photons. *J. Physiol.* **288**, 613–634.

8. Baylor DA, Nunn BJ, Schnapf JL. 1984 The photocurrent, noise and spectral sensitivity of rods of the monkey *Macaca fascicularis*. *J. Physiol.* **357**, 575–607. (doi:10.1113/jphysiol.1984.sp015518)

9. Sakitt B. 1972 Counting every quantum. *J. Physiol.* **223**, 131–150. (doi:10.1113/jphysiol.1972. sp009838)

10. Field GD, Sampath AP, Rieke F. 2005 Retinal processing near absolute threshold: from behavior to mechanism. *Annu. Rev. Physiol.* **67**, 491–514. (doi:10.1146/annurev.physiol.67.031103.151256)

11. Bialek W. 1987 Physical limits to sensation and perception. *Annu. Rev. Biophys. Biophys. Chem.* **16**, 455–478. (doi:10.1146/annurev.bb.16.060187. 002323)

12. Barlow HB. 1961 Possible principles underlying the transformation of sensory messages. In *Sensory communication* (ed. WA Rosenblith), pp. 217–234. Cambridge, MA: MIT Press.

13. Baylor DA, Matthews G, Yau KW. 1980 Two components of electrical dark noise in toad retinal rod outer segments. *J. Physiol.* **309**, 591–621. (doi:10.1113/jphysiol.1980.sp013529)

14. Arshavsky VY, Lamb TD, Pugh EN. 2002 G proteins and phototransduction. *Annu. Rev. Physiol.* **64**, 153–187. (doi:10.1146/annurev.physiol.64.082701.102229)

15. Burns ME, Mendez A, Chen J, Baylor DA. 2002 Dynamics of cyclic GMP synthesis in retinal rods. *Neuron* **36**, 81–91. (doi:10.1016/S0896-6273(02)00911-X)

16. Luo DG, Yue WW, Ala-Laurila P, Yau KW. 2011 Activation of visual pigments by light and heat. *Science* **332**, 1307–1312. (doi:10.1126/science. 1200172)

17. Rieke F, Baylor DA. 1996 Molecular origin of continuous dark noise in rod photoreceptors. *Biophys. J.* **71**, 2553–2572. (doi:10.1016/S0006-3495(96)79448-1)

18. Koch KW, Stryer L. 1988 Highly cooperative feedback control of retinal rod guanylate cyclase by calcium ions. *Nature* **334**, 64–66. (doi:10.1038/334064a0)

19. Conner JD, MacLeod DI. 1977 Rod photoreceptors detect rapid flicker. *Science* **195**, 698–699. (doi:10. 1126/science.841308)

20. Conner JD. 1982 The temporal properties of rod vision. *J. Physiol.* **332**, 139–155. (doi:10.1113/ jphysiol.1982.sp014406)

21. Rieke F, Baylor DA. 1998 Origin of reproducibility in the responses of retinal rods to single photons. *Biophys. J.* **75**, 1836–1857. (doi:10.1016/S0006-3495(98)77625-8)

22. Whitlock GG, Lamb TD. 1999 Variability in the time course of single photon responses from toad rods: termination of rhodopsin's activity. *Neuron* **23**, 337–351. (doi:10.1016/S0896-6273(00)80784-9)

23. Field GD, Rieke F. 2002 Mechanisms regulating variability of the single photon responses of mammalian rod photoreceptors. *Neuron* **35**, 733–747. (doi:10.1016/S0896-6273(02)00822-X)

24. Gross OP, Pugh ENJ, Burns ME. 2012 Calcium feedback to cGMP synthesis strongly attenuates single-photon responses driven by long rhodopsin lifetimes. *Neuron* **76**, 370–382. (doi:10.1016/j. neuron.2012.07.029)

25. Wilden U, Kuhn H. 1982 Light-dependent phosphorylation of rhodopsin: number of phosphorylation sites. *Biochemistry* **21**, 3014–3022. (doi:10.1021/bi00541a032)

26. Kühn H. 1984 Early steps in the light-triggered activation of the cyclic GMP enzymatic pathway in rod photoreceptors. *Prog. Clin. Biol. Res.* **164**, 303–311.

27. Doan T, Mendez A, Detwiler PB, Chen J, Rieke F. 2006 Multiple phosphorylation sites confer