Computational Vision
U. Minn. Psy 5036
Daniel Kersten
Lecture 2: Limits to Vision

Introduction

■ Variability

We noted in Lecture 1 that a central challenge to visual perception is dealing with variability in the image data received by the eye. Your eye's retinal image of a signal (say your friend's face) is never the same from one time to the next. In this lecture, we'll start off with the simplest case of variability in vision--the intensity of a single point of light.

Imagine you have a flashlight with a high and low beam setting, and you want to send a signal (high or low) to a distant friend by flashing a light with either the high or the low beam on (e.g. "high if by land, low if by sea"). In a perfect world,, this would be a one bit communication system. But the world isn't perfect and suppose the flash light switch is a bit flaky. The result is that occasionally the amount light emitted with the high beam is lower than for the low beam. Your friend won't be able to reliably tell whether a flash corresponded to the high or the low setting. You'd be stuck with a communication system that on average transmits less than one bit of information.

Now suppose you work really hard to make the most reliable flashlight possible. It turns out that no matter what you do, you'll always be a bit short of the ideal one bit information capacity. Normally you wouldn't notice, because you could make the error rate very tiny by increasing the intensity difference between the high and low settings, but if the intensity differences for the two switch settings are small, you'd always find some residual variability due to photon fluctuations.

■ Discrimination

We are going to ask the psychophysical question: How well can one discriminate a slight change in brightness of a spot given that from trial to trial the number of photons landing in the eye (or transduced by the retina) is not consistently the same? This is the problem of intensity discrimination, and we'll develop a model to account for human discrimination thresholds given variability. We'll see that photon fluctuations fundamentally limit human discrimination thresholds.

■ Detection

But first we consider detection, a special case of discrimination. On the best of nights when the sky is clear and moonless, you can see about 2000 stars. Why can't you see more? There are millions more there, but they are just too dim (the fact that you can't see any in the daytime is a problem in discrimination). In other words, what are the limits to just detecting a faint point of light in the dark? This question was addressed in a classic study by Hecht, Schlaer and Pirenne in what is probably the first, study that combined psychophysics, neuroscience, and computational theory in vision. Hecht, Schlaer and Pirenne measured human absolute threshold for faint small flashes of light. We will take a look at their experiment with the following goals in mind:

- Learn about the physiology of the human eye
- Quantify limits to the eye's ability for light detection
- Devise a computational theory of detection and discrimination, called Signal Detection Theory (SDT)
- Obtain a preview of how to generalize SDT to the study of perception as inference.

What are the fundamental physical limitations to vision?

Paul Dirac, one of the principal architects of quantum mechanics, was known for being laconic. He made a famous statement about photons,that implies the fundamental limits to spatial resolution, and visual light intensity discrimination:

"*Each photon interferes only with itself. Interference between photons never occurs.*"

These two sentences imply that: 1) Interference or diffraction limits the spatial resolution of any imaging device, including the human eye; 2) Statistical independence of photon emission and absorption limits detection and discrimination of light intensity. The first will provide us with a springboard for developing linear systems analysis of vision. The second limitation provides us with a springboard for developing signal detection theory. In this lecture,we study the second limitation. Later, we'll return to the consequences of interference for human spatial resolution.

Within several of decades of the establishment of the quantum basis of light (the quanta being photons) in the first decade of the 20th century, and the subsequent development of quantum mechanics in the 1920s, by the 1940s it was natural to ask:"What is the least number of photons a human observer can detect?" Hecht, Schlaer,and Pirenne, then in a biophysics lab at Columbia University set out to answer this question which was published in their classic paper of 1942: "Energy, quanta, and vision.".

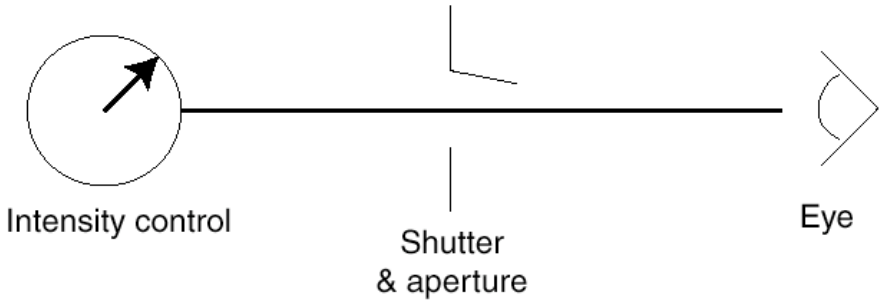
Basic question: What is threshold?

First we'll introduce workable definitions of two basic concepts in psychophysics: the psychometric function, and threshold. Our definitions won't be final, but will be sufficient to understand the 1940s experiment of Hecht, Schlaer and Pirenne. Later we'll make use of the development of Signal Detection Theory in the 1950s to fine-tune these definitions.

■ Experimental set-up

Imagine setting up the experiment as follows. We set up an apparatus to flash spots of light and we position an observer to view the flashes. We take steps to control the wavelength, size, duration and intensity of the light. All variables are fixed except for the light intensity control. This is our experimental **independent variable**, as it is called. Let's call it **highmean**. What we'd like to know is the smallest value of **highmean** that observers can see--in other words, we want to know the observer's **threshold.** Sounds simple, but there is a subtlety, due to variability.

Suppose we set **highmean** to a fixed value (say 4). Then we begin a **block** of **trials** in which we flash the light using the shutter. For concreteness, let's suppose the block has 100 trials. Our dependent variable is the proportion of times the observer says "yes, I saw the flash". Let's simulate the process of flashing photons (quanta) into the eye, but we will use dots on the screen to represent individual photons.



■ Photon flash simulation: Did you see the light flashed or not?

Whenever it is convenient, we will make use of predefined functions in *Mathematica*. **DiscreteDistributions.m** contains the definitions of several standard discrete probability models. These provide us the means to write some code to draw samples (i.e. drawing the number of photons flashed "from a hat"). Select the following cell and evaluate it (i.e. hit ENTER) to read in the file that defines **PoissonDistribution[]**:

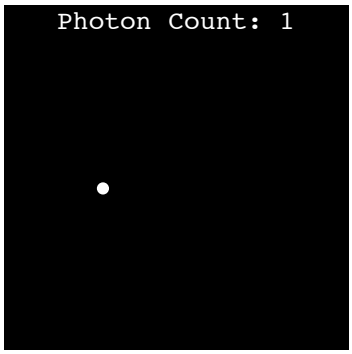
```
<< "Statistics`DiscreteDistributions`"
```

Let's define a Poisson distribution with a mean (**highmean**) of 4 photons. And let **numberofphotons[mean]** be a function that determines the actual number of photons measured at the receiver (eye). Execute this next cell:

```
numberofphotons[mean_] := Random[PoissonDistribution[mean]];
dotsize = 0.04;
```

Now let's simulate a trial presentation of a light flash. Select the following cell and evaluate it:

```
highmean = 4; numhighsample = numberofphotons[highmean];
highsample = Table[{Random[], Random[]}, {numhighsample}];

highg = Graphics[{PointSize[dotsize], Point /@ highsample},
   AspectRatio → 1, Frame → False, FrameTicks → None,
   Background → GrayLevel[0.0], PlotRange → {{-.2, 1.2}, {-.2, 1.2}},
   PlotLabel → "Photon Count: " <> ToString[numhighsample]];
Show[Graphics[highg], Frame → False];
```



Now, select it again, and repeat 30 times. Keep count of how many times you saw ANY dots at all.

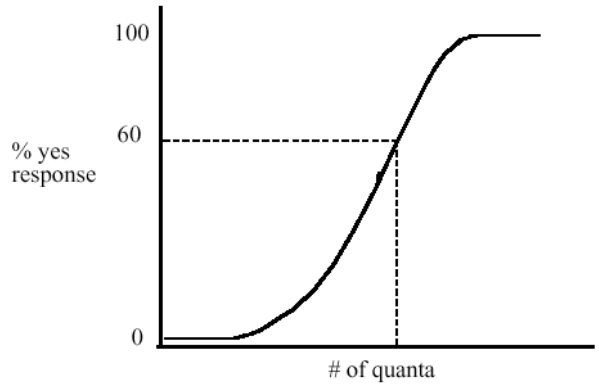
Did you always see dots?

If not, why not?--we will see answers to this question later.

What fraction out of 30 was it? This fraction would correspond to a psychophysically measured dependent variable. We will call it **% yes responses**. (Later on, when we introduce signal detection theory, we will call it the "hit rate").

■ Psychometric function

Now imagine, that we increase the intensity a bit, say to **highmean** = 7. We could run another block, and measure the proportion seen. If we do this again for another mean intensity, and so on, we can generate what is called a **psychometric function** which plots %yes responses vs. intensity (# of quanta). We would expect that if the intensity is low, the percent would be zero (later on we will see that this isn't always a good assumption), and if the intensity is high, the percent should be 100% (also not an infallible assumption). But because of variability the psychometric function doesn't look like a step-function (i.e. going abruptly from 0 to 100% as some threshold value), but typically looks like this:



--it is called a sigmoid, because of its s-like shape.

■ So what is threshold?

We see there is not a specific value of the independent variable, intensity (# of quanta), at which the light goes from being invisible to visible. So we'll do what Hecht et al. did, and arbitrarily decide to define threshold as: "the intensity at which the %yes responses is 60%".

Hecht et al. measured the detectability of flashes of light in a dark room. Before running the experiment,they looked for the optimal stimulus conditions (size, duration, position, wavelength) and state of adaptation of the subjects. The goal was to measure the lowest possible threshold.

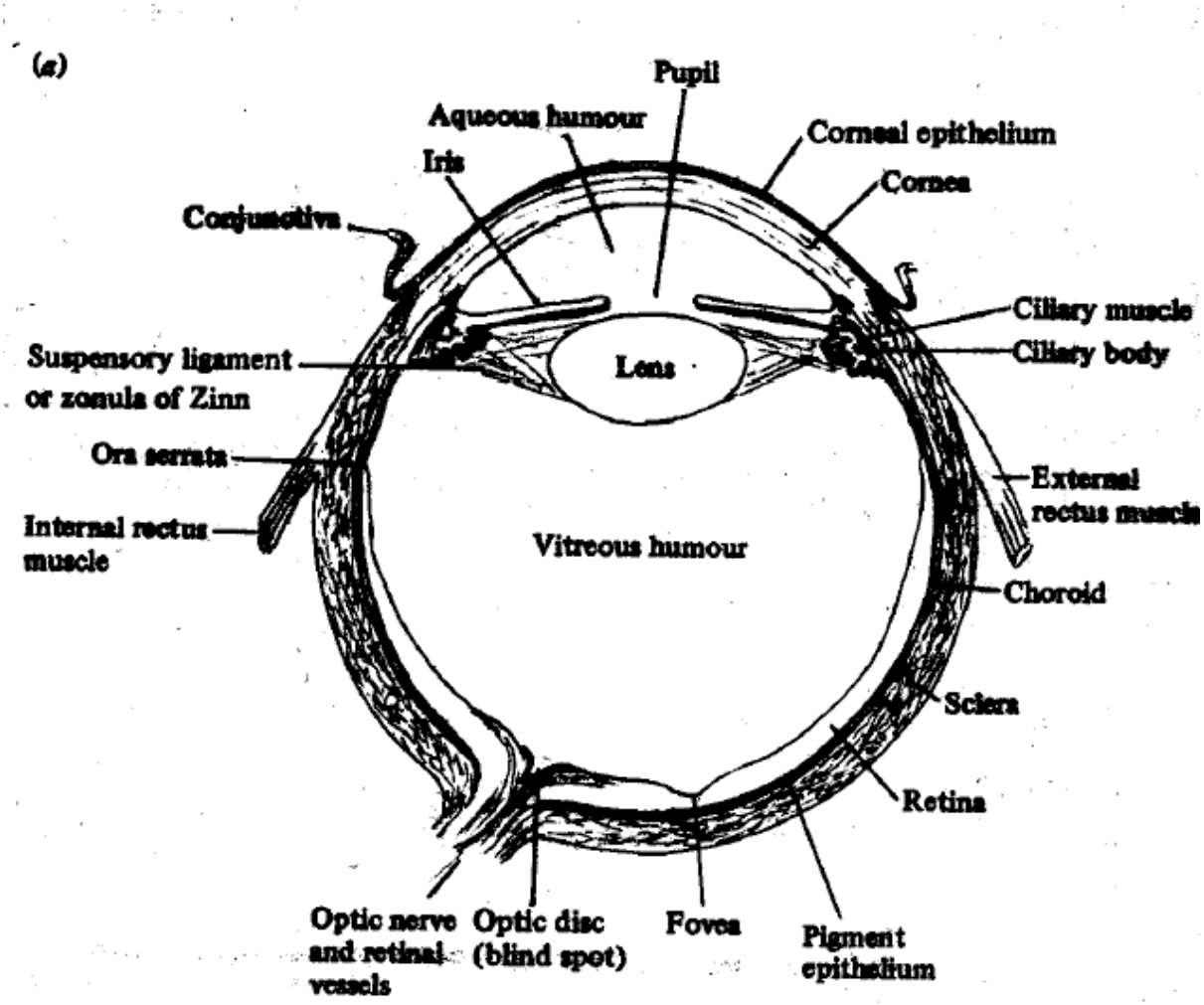
Before getting started, let's get an overview of the visual anatomy to see where the eye is in relationship to the visual pathways.

Human Visual Anatomy: Brief Overview of the "front-end"

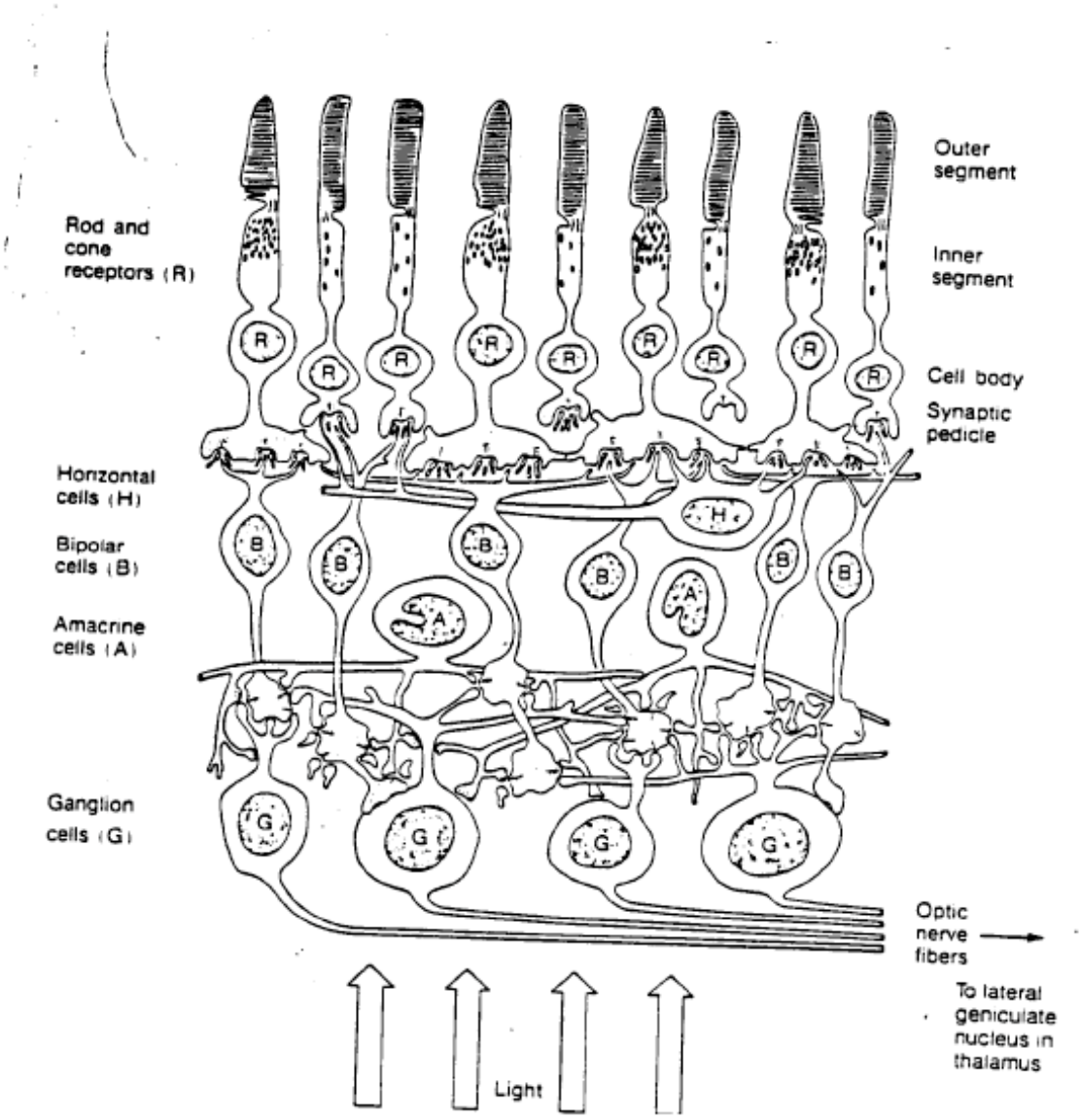
Light travels through the eye to the retina where it gets transduced to nerve impulses that leave the eye by the optic nerve to the lateral geniculate nucleus (branches also go off to the superior colliculus), and from there to the visual cortex.

For the next few lectures, we are going to concentrate on the eye itself. The light enters the eye at the cornea which does the bulk of the refraction, passes through the pupil formed by the iris, through the lens which controls accommodation (if you are young enough), and through the tissue of the retina to the photosensitive parts of the receptor cells called the outer segments.

■ **Eye schematic**



■ **Retina schematic**



■ Eye to V1 schematic

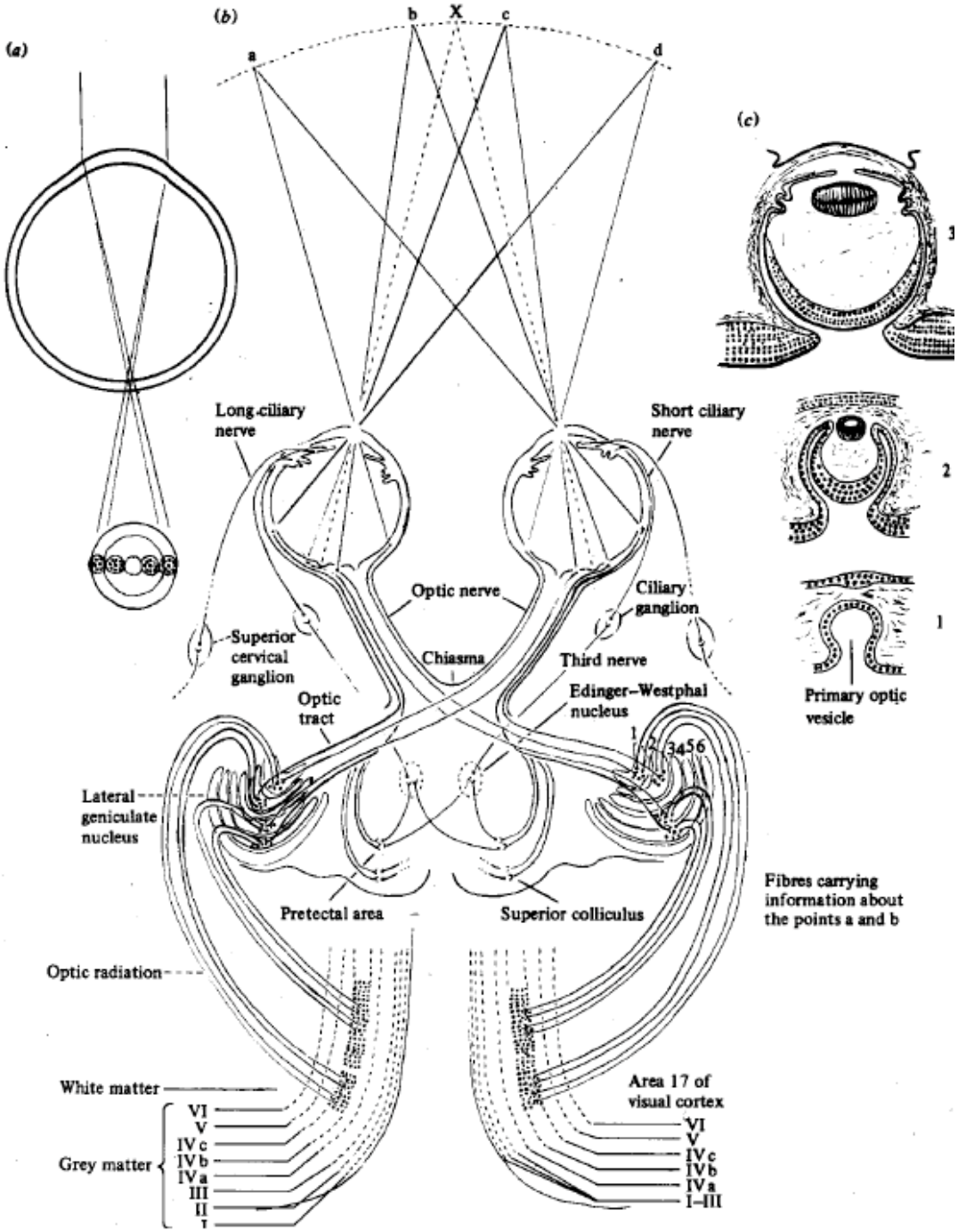


Figure from: Barlow, H. B., & Mollon, J. D. (1982). The Senses. Cambridge: Cambridge University Press.

The experiment of Hecht, Schlaer & Pirenne

Optimal conditions for detection

In order to find the least number of photons that could be seen, it was essential for Hecht, Schlaer, and Pirenne to find the optimal conditions for detection. The needed information about the optimal:

- state of the subject, i.e. adaptation state
- flash placement on the retina
- flash size
- flash duration
- wavelength of the light

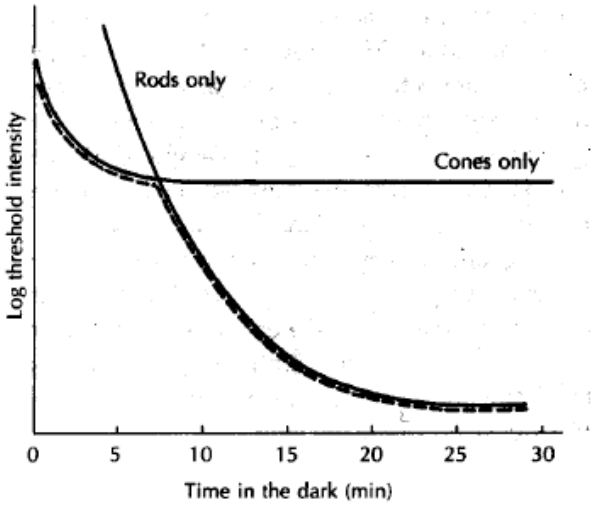
■ Adaptation of the subject

When coming into a dark room after being outside on a bright day, it takes awhile to be able to detect changes in brightness, and your sensitivity gradually improves as you get used to the dark. This is called dark adaptation.

Full dark adaptation can take up to 30 minutes, and involves a change from cone vision (photopic) to rod vision (scotopic). The theory of two receptor systems goes back to the 19th century and is called duplicity theory. Adaptation also involves the regeneration of rod photopigment which gets bleached in response to light.

Much psychophysical work has gone into characterizing human dark adaptation. The figures show psychophysical measurements of human threshold to a flash of light as a function of time the subject has been sitting in the dark. Psychophysicists either present their dependent variable as a threshold (e.g. number of photons), or as the reciprocal of threshold called sensitivity.

■ Threshold vs. time in dark



What is the mechanism of adaptation? At the time of this experiment (1940's), it was thought that the percentage of unbleached rhodopsin molecules directly reflected sensitivity. Rhodopsin has a half-life of 5 minutes, meaning that if all of the molecules are bleached by light, then half of them would be restored in 5 minutes (75% in 10 minutes, etc.) We now know that adaptation involves retinal processes other than just pigment bleaching (see Figure below). Note the hundred-fold drop in threshold (increase in sensitivity) between 15 minutes and 30 minutes, even though 90% of the rhodopsin has already been regenerated.

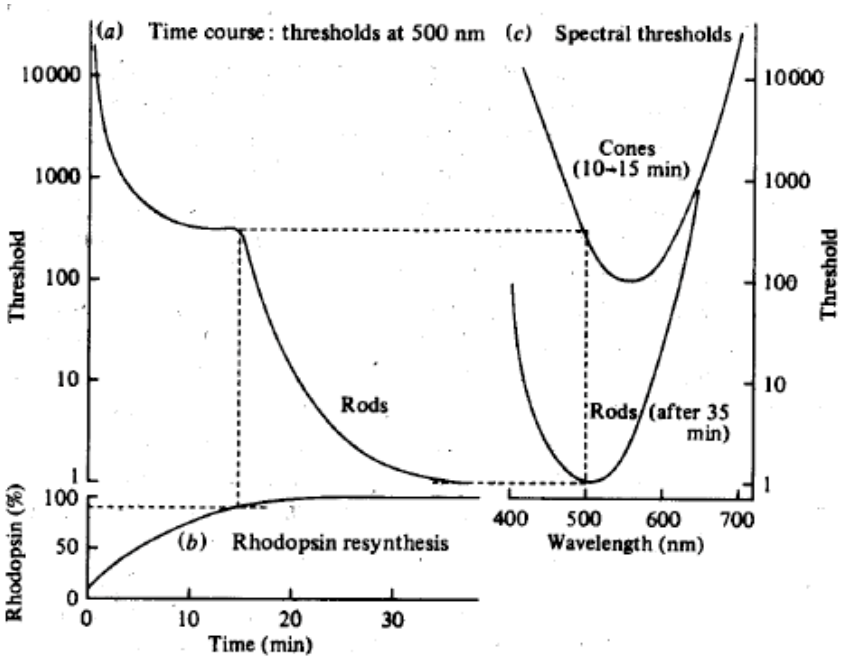


Fig. 7.3. Dark-adaptation curves, the resynthesis of rhodopsin, and spectral sensitivity curves showing the Purkinje shift. (*a*). The drop of log threshold with time following adaptation to a light that bleached a high proportion of rhodopsin in the rods. (*b*). The time course of resynthesis of rhodopsin on the same time scale; notice that less than 10% of the rhodopsin escaped the initial bleach, but 90% had been regenerated after 15 min, when the break occurred in the dark-adaptation curve. (*c*). Two sets of thresholds plotted against wavelength. The upper curve was obtained between 10 and 15 min after starting dark adaptation, and it represents the spectral sensitivity of the cone system. The lower curve was obtained after full dark adaptation and corresponds to the rods. Note that the peak of sensitivity moves (Purkinje shift). These curves represent similar data to V_λ and V'_λ plotted as relative sensitivity in Fig. 5.8, though the latter are obtained by different methods. The fact that the hundred fold drop in rod threshold occurs while the final 10% of the rhodopsin is being resynthesised shows that the change in the threshold is not a simple matter of increased light absorption in the photopigment. (After Rushton and others.)

Figure from: Barlow, H. B., & Mollon, J. D. (1982). The Senses. Cambridge: Cambridge University Press.

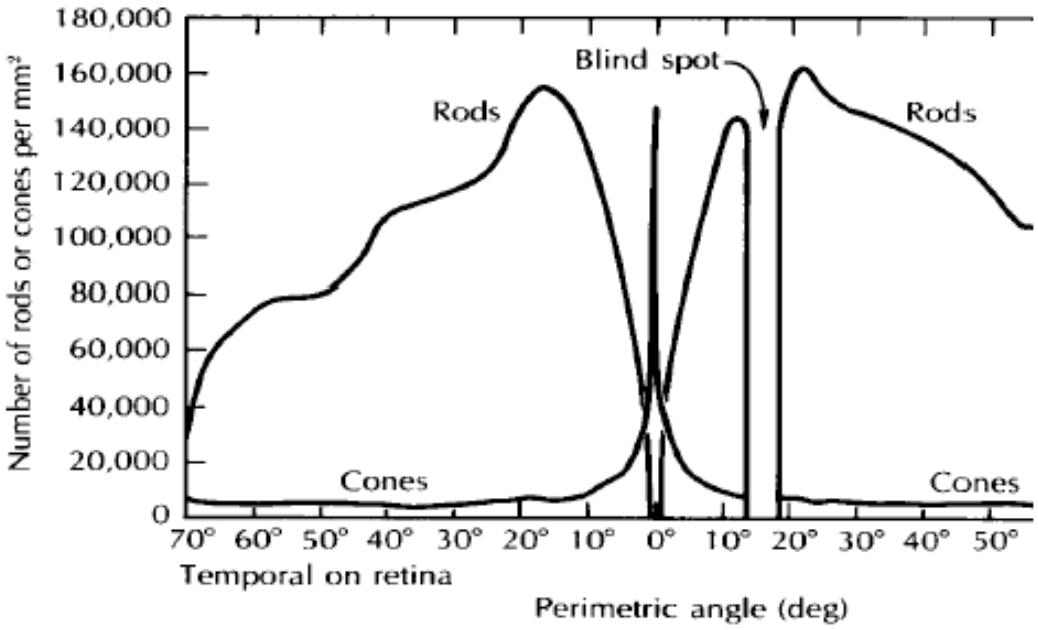
Hecht, Schlaer, and Pirenne decided that the observer should be completely dark adapted--in other words, the poor subjects had to sit in a dark room for over a half hour before the data collection could begin

■ Flash placement

The next consideration for the Hecht, Schlaer, and Pirenne experiment is where to put the image of flash on the retina, or in other words: where should the subject look?

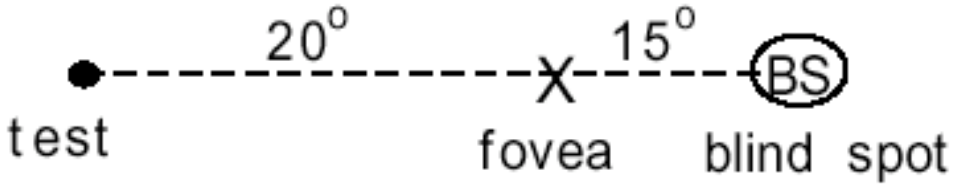
There are two spots to avoid. First, one must avoid the blind spot in the eye, the region where visual information leaves the eye enroute to the thalamus (and other destinations) of the brain. This is the optic disk, and source of the optic nerve. It consists of axons of the retinal ganglion cells (we will say more about these cells later). If a small light flash is restricted to in this region, it will not be detected. It is curious to note that the blind spot wasn't discovered until 1668 by Mariotte--despite the fact that it is right before your very eyes! Because the brain completes missing information through the blind spot, you can only detect it in your own eye by carefully looking for it. Mariotte had reason to look for it in his own eye, because anatomists had noticed that when they dissected eyes, there was this curious white spot where the optic nerve joined the eye at about the same location in all the eyes

■ Photoreceptor density vs. eccentricity



But, perhaps somewhat surprisingly, the observer shouldn't be looking directly at the flash either. A dim star is more likely to be seen if you look just to the side of it, rather than directly at it (in which case it would fall on the fovea). This is because the maximum concentration of rods is not at the fovea, where there are virtually no rods, but at about 20 deg of visual angle away from the fovea (see figure).

■ Figure of placement

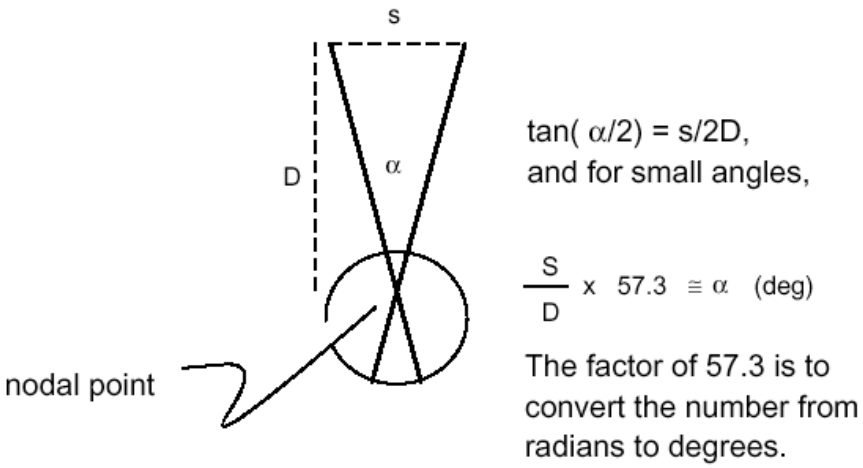


Hecht et al. decided to place the flash at 20 degrees of visual angle to the left of the fixation point (the blind spot being at about 15 degrees to the right of fixation) for right eye viewing. This would avoid the blind spot, and place the flash close to where the density of rod receptors is highest.

By the way, vision scientists measure size in terms of visual angle, rather than meters or feet. This is because visual angle is proportional to size on the retina, and is easily measured if one knows the size and distance of an object. Visual angle (a) is defined in the following way.

■ Visual angle definition

We'd like to measure extent on the retina, not the extent or size of the external stimulus on, say a screen. There isn't an really easy direct way of measuring how big an image is on the retina; however, we can easily measure **visual angle**--which is proportional to retinal size. Then if we know the distance from the nodal point of the eye to the retina, the visual angle will tell us the extent on the retina.



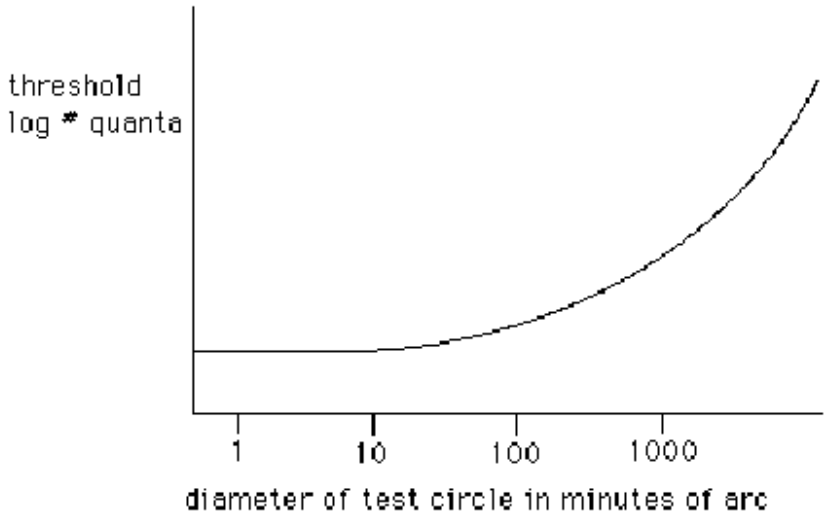
Both the moon and sun subtend about 1/2 degree of visual angle, which makes a solar eclipse a particularly singular event. See if one of your fingers subtends a degree at arm's length. This can come in handy, for example to estimate the distance to an object of known size (e.g. a 5.5 foot person off in the distance).

■ Flash size

How large should the spot of light be?

Below is a characteristic plot of the way human light thresholds vary with the diameter of a flash (in minutes of arc, where 60 minutes = 1 degree).

■ Threshold vs. spot size



There is little difference in the number of quanta required for detection of spots of light of various sizes as long as their diameters are less than 10 min (1/6 deg). 1000 quanta spread over 1 min is as easily detected as 1000 quanta spread over 10 min; but 1000 quanta spread over 1000 min requires more quanta to be detected. This is called spatial integration.

Hecht, Schlaer, and Pirenne decided that a spot with diameter of less than 10 min would guarantee that detection would not be limited by the eye's inability to spatially integrate light energy. Note that the degree of spatial integration does depend on where in the retina it is measured, state of adaptation, and duration. Hecht et al. used a size suitable for the conditions of their experiment.

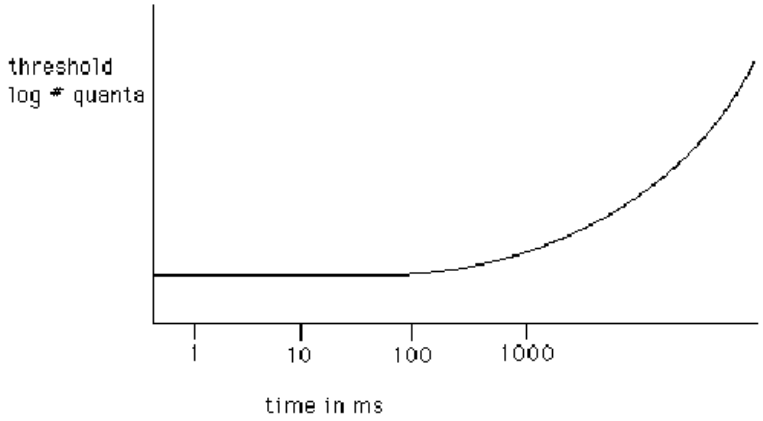
■ Flash duration

What should the duration of the stimulus be? Below is a characteristic plot of how human thresholds vary as a function of the duration of a flash. (4.5 min spot at 15 deg)

Analogous to spatial integration, the eye also shows temporal integration. In this case, there is little difference in the threshold in terms of the number of quanta for a durations under 100 milliseconds. 1000 quanta is detected as easily at 1 ms and at 100 ms. A useful number to remember is that under scotopic conditions, the temporal integration time of the eye is about 200 msec, but is significantly shorter under photopic conditions (~100 msec).

Hecht, Schlaer, and Pirenne played it safe and used a 1 millisecond duration for their flash.

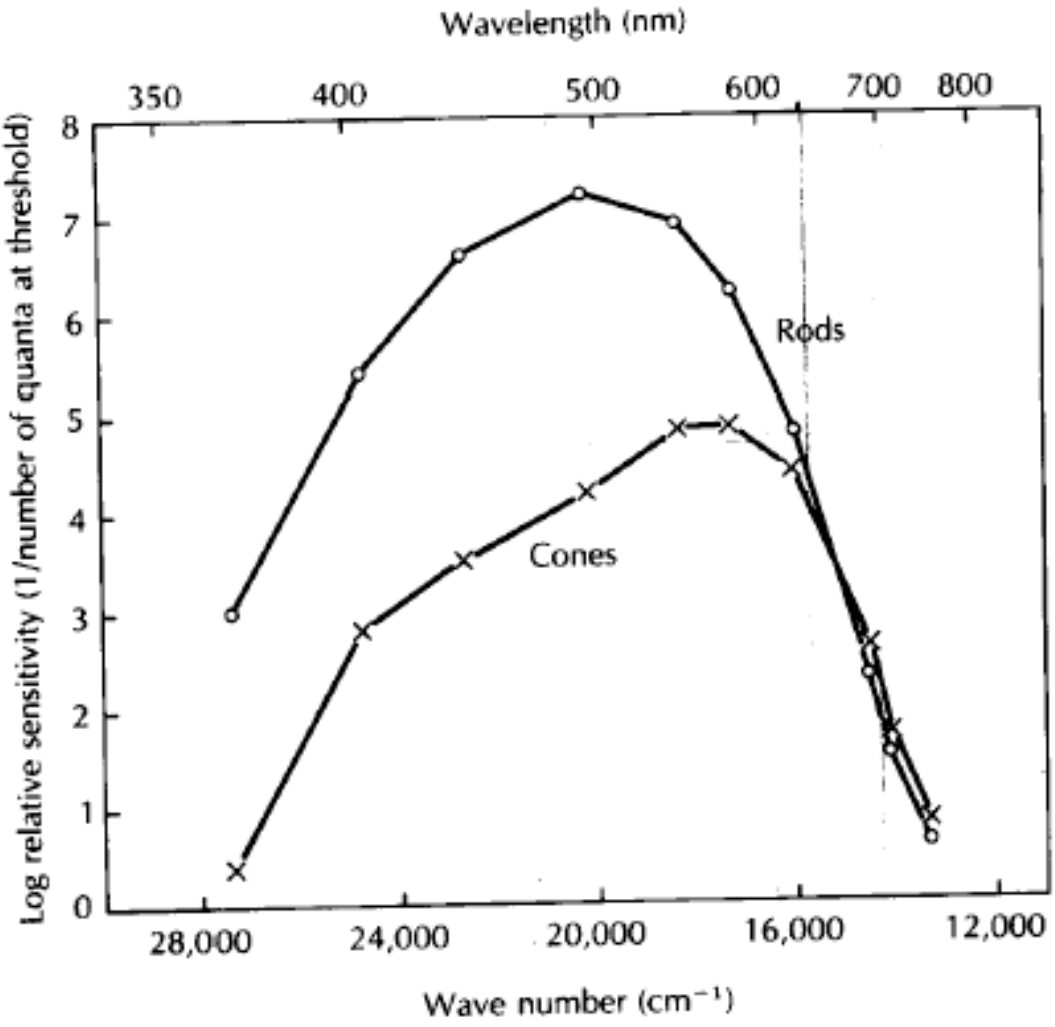
■ Threshold vs. duration



■ Flash wavelength

What should the wavelength of the light be? Psychophysical measurements again provide the answer.

■ **Sensitivity vs. wavelength**



As you can see from the figure, human thresholds vary with wavelength. Different wavelengths of light vary in the efficiency with which they are transduced, and thus detected. This is a direct consequence of the absorption properties of the photoreceptors. For scotopic (rod) vision, around 500 nm is optimal (Hecht used 510nm). (For photopic vision, 555nm is optimal).

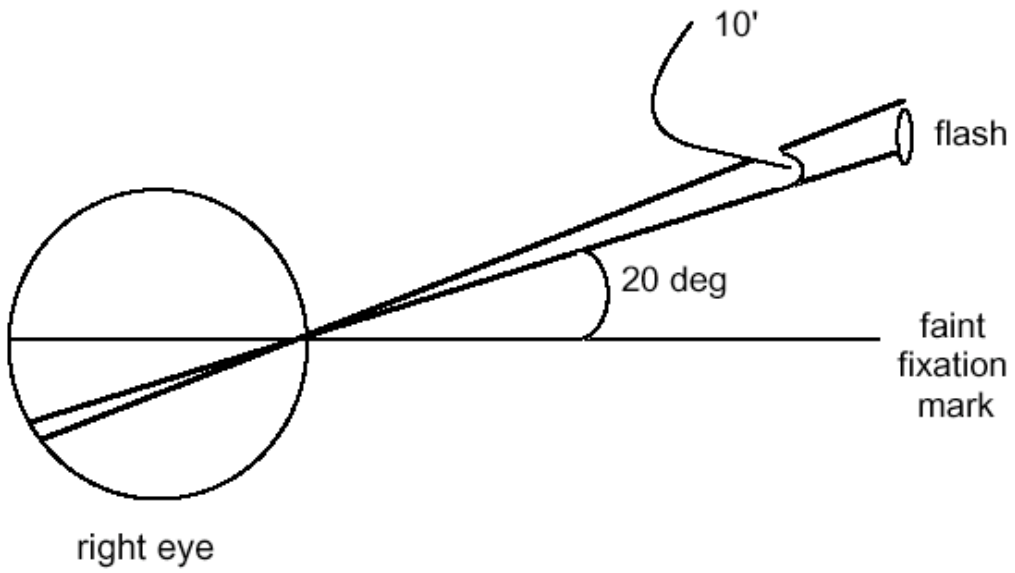
Summary of experimental conditions

To summarize the conditions chosen by Hecht et al.:

- The subjects were dark adapted and viewed a disk of light against a dark background
- The flash was placed 20 degrees to the left of the fixation mark (right eye)
- The flash size was 10 minutes of arc
- The duration of the flash was 1 ms.
- The wavelength was 510 nanometers.

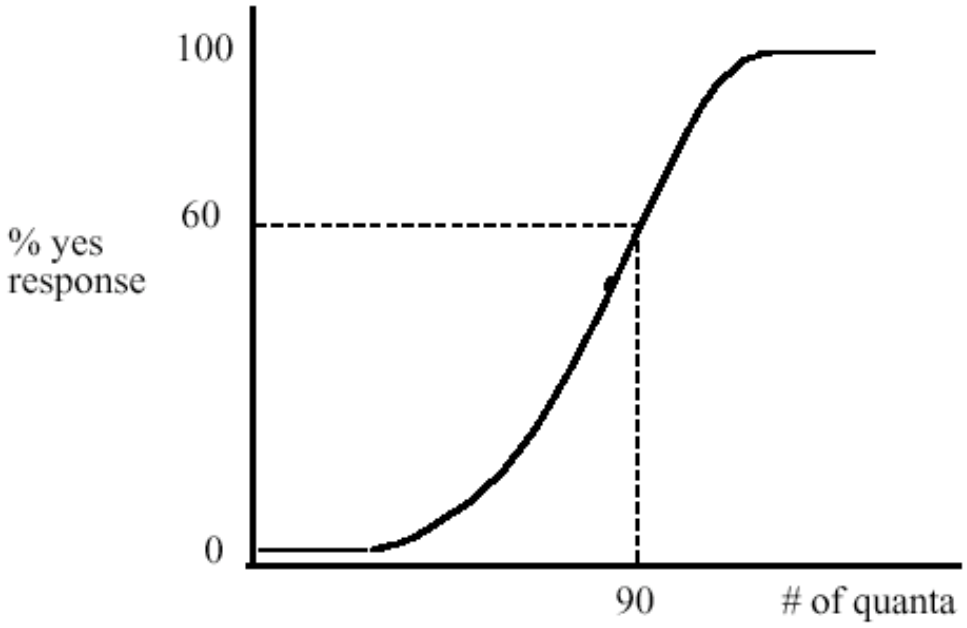
Hecht et al., did not have to make the measurements necessary to determine these conditions themselves. The data to make the decisions for their experiment were all in the existing literature of the time. If we repeated the experiment today, we could make some minor improvements in their choices, but the results would not be substantially different.

■ **Figure: Spot placement relative to eye**



Results of the experiment

■ **Threshold measured**

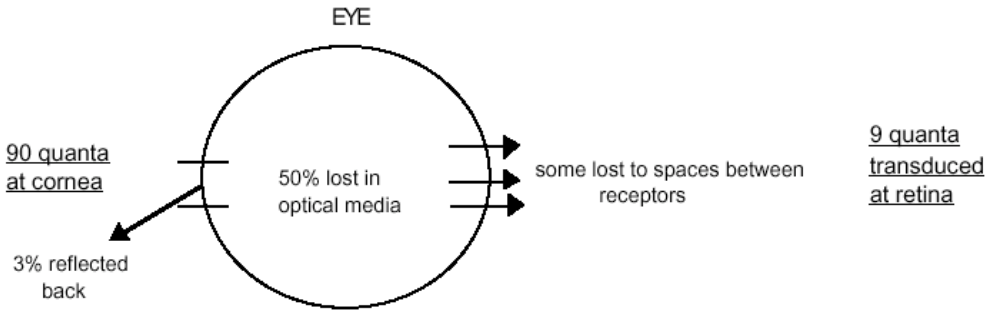


For the criterion of 60% yes responses, about 90 quanta (on average) were required at the cornea. This is an amazingly small figure. But we're not done.

What is the significance of a threshold of 90 quanta?

Of the 90 quanta, about 3% is reflected from the cornea. Then 50% of the remaining light is actually transmitted through the vitreous fluid of the eye. Finally, 20% of this is lost in the spaces between receptors (not transduced). The result is that they concluded that only about 9 photons (on average) are actually required at the receptors to "see" a flash.

■ **Photon losses in optic media**



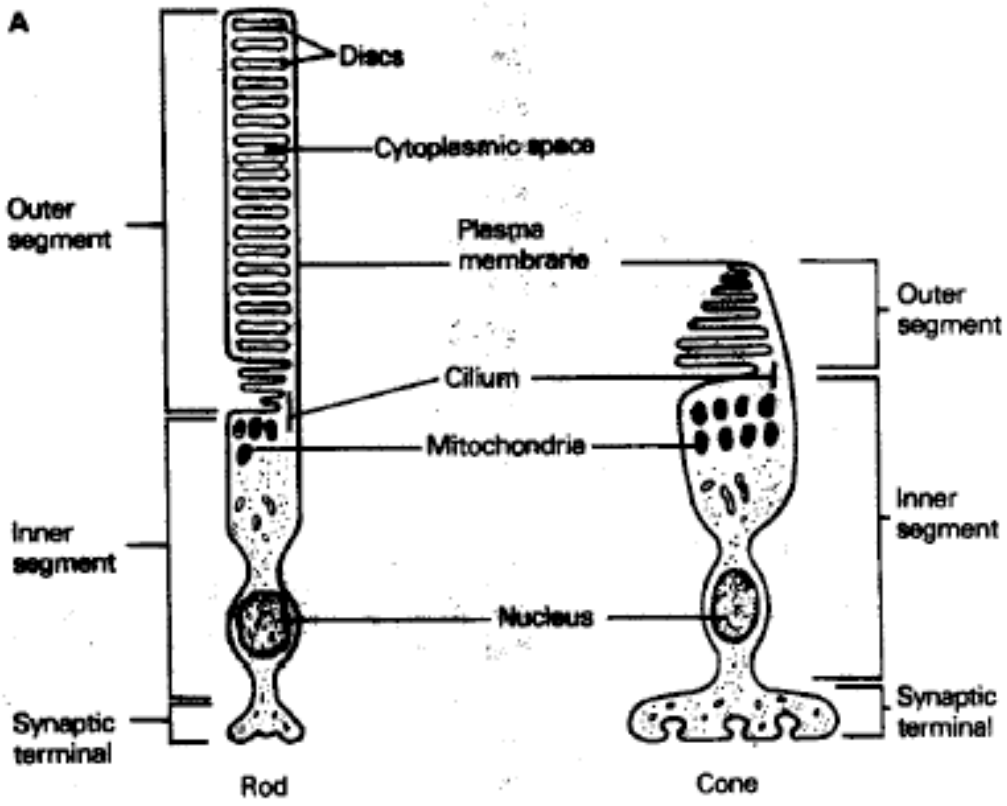
Hecht et al. estimated that a single rod can process a single photon. Because the 9 photons are spread over about 500 rods (their estimate), the chances that 2 photons are absorbed by one rod is very small. (Current estimate of the number of rods in this part of the retina in a 10' diameter spot is closer to 350). In any case:

One photon is sufficient to activate one rod.

Four decades later: "Going inside the box"

Almost 50 years later, the basic results and conclusions of this experiment still stand. In 1979, Baylor et al. were able to suck single toad rods into electrodes and measure current responses to single photon events (see figure)

■ **Rods and cones schematic**



■ Photon "bumps" in rod current

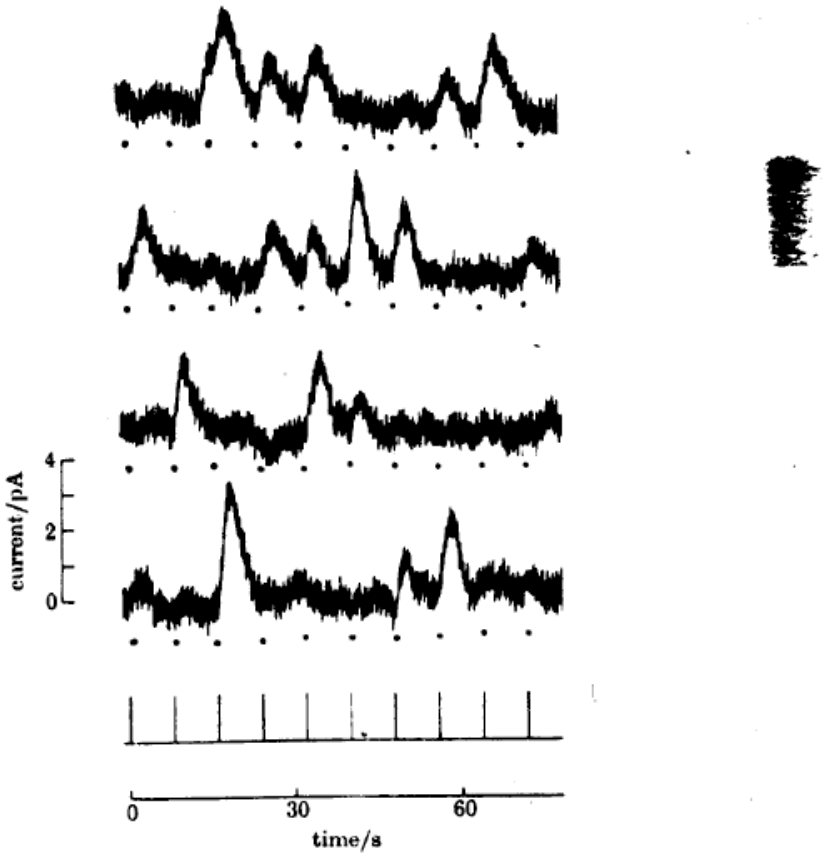


FIGURE 4. Responses of a toad rod to single quantal absorptions. Dots indicate the times of delivering a flash of light causing an average of 0.92 isomerizations of the photosensitive pigment. The bumps that follow result from decreases in the current entering the rod outer segment. The variability in the size of the responses follows the expected Poisson statistics. (From Baylor *et al.* 1979.)

From: Baylor, D. A., Lamb, T. D., & Yau, K. W. (1979). Responses of retinal rods to single photons. Journal of Physiology, Lond., 288, 613-634.

Preview: Computational Theory for Light Discrimination

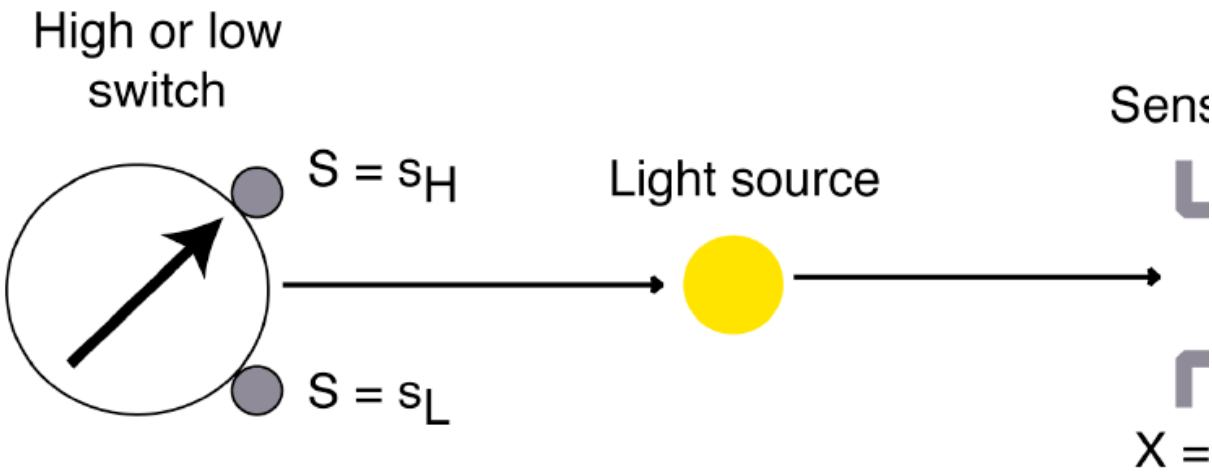
Despite all the fascinating facts that crop up when considering Hecht et al.'s experiment, we are now going to single-mindedly focus on just one aspect, the psychometric function. From this starting point, we'll develop signal detection theory, the theory of ideal observers, and from there a framework to understand perception generally as a process of statistical inference.

Why was the subject response curve sigmoidal, rather than a step function? Variability may be due to the statistics of photon emission and absorption, human responses, and neural transmission. Signal detection theory was developed 1940's and 1950's to describe the ability to detect, discrimination, and estimate signals in the presence of variability or "noise", without carrying much about exactly where the noise comes from, just about its properties and affects on sensory decisions. Let's return to the light detection example, but generalize it to light discrimination.

Above, our generative model for light transmission in essence had a knob for which the light intensity could be varied continuously. But the observer was asked to make a discrete decision. Our theoretical development will proceed better if we match the signal space to the decision space. We could ask an observer to estimate the amount of light (e.g. on a scale of 1 to 100, and psychophysicsts sometimes do this). But psychophysicsts usually prefer to make the signal space simple, and discrete.

A generative model for light intensity discrimination

■ Schematic of set-up for ideal photon counter



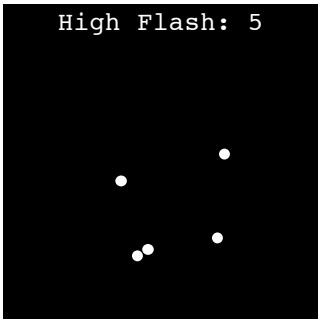
■ Physical model for light, two switch settings

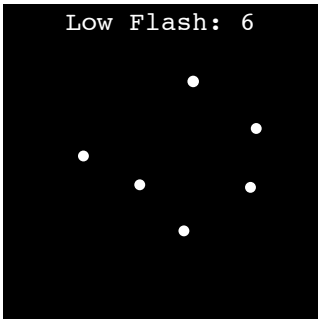
As above, let's define a Poisson distribution with a mean of **mean**, with a function to draw a sample from this distribution:

```
numberofphotons[mean_] := Random[PoissonDistribution[mean]];
```

Let highmean = 7 for the "high" setting, and darkmean =5 for the "dark" setting. Now, turn the switch to "high", draw a sample, and then to "low" and draw a sample. Let's plot up the two:

```
highmean = 7; numhighsample = numberofphotons[highmean];
darkmean = 5; numdarksample = numberofphotons[darkmean];
highsample = Table[{Random[], Random[]}, {numhighsample}];
darksample = Table[{Random[], Random[]}, {numdarksample}];

highg = Graphics[{PointSize[0.04], Point /@ highsample},
   AspectRatio → 1, Frame → False, FrameTicks → None,
   Background → GrayLevel[0.0], PlotRange → {{-.2, 1.2}, {-.2, 1.2}},
   PlotLabel → "High Flash: " <> ToString[numhighsample]];
darkg = Graphics[{PointSize[0.04], Point /@ darksample},
  AspectRatio → 1, Frame → False, FrameTicks → None,
  Background → GrayLevel[0.0], PlotRange → {{-.2, 1.2}, {-.2, 1.2}},
  PlotLabel → "Low Flash: " <> ToString[numdarksample]];
Show[GraphicsArray[{highg, darkg}, GraphicsSpacing → .3],
 Frame → False];
```





■ Histograms for high and low photon counts:

Suppose we compile a histogram of photon counts conditional on the switch setting:

Define high and low models:

```
highpdist = PoissonDistribution[7];
lowpdist = PoissonDistribution[4];
PDF[lowpdist, x]
```

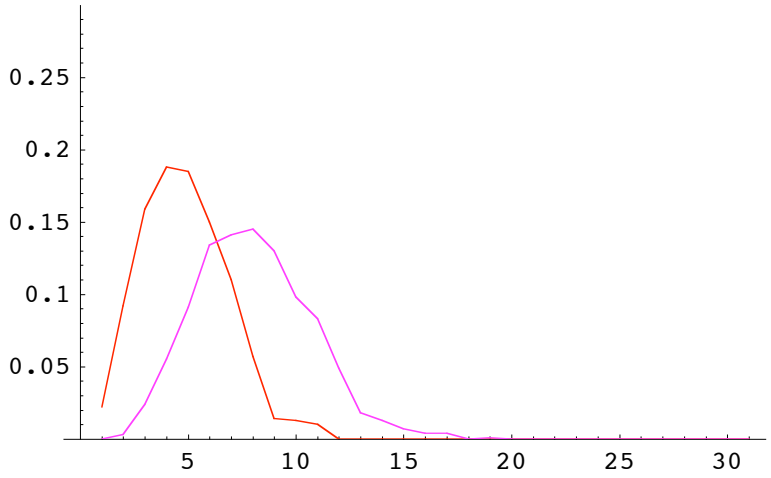
$$\frac{4^x}{e^4\, x!}$$

```
sample[ntimes_] :=
    Table[Random[lowpdist],{ntimes}];
    z = sample[1000];
domain = Range[0,30];
Freq = Map[Count[z,#]&,domain];

lowg=ListPlot[N[Freq/Length[z]],PlotJoined→True,
      PlotStyle→{RGBColor[1,0,0]},DisplayFunction→Identity];
```

General::spell1 :
Possible spelling error: new symbol name "domain" is similar to existing symbol "Domain". More…

```
sample[ntimes_] :=
    Table[Random[highpdist],{ntimes}];
    z = sample[1000];
domain = Range[0,30];
Freq = Map[Count[z,#]&,domain];

highg=ListPlot[N[Freq/Length[z]],PlotJoined→True,
      PlotStyle→{RGBColor[1,0,1]},DisplayFunction→Identity];
```

```
Show[{lowg, highg}, PlotRange → {0, .3}, DisplayFunction → $DisplayFunction];
```



The point is: photon detection is inherently statistical, and in fact because of the independence of photon absorption, the histogram can be modeled as a Poisson distribution. More on this next time.

Discrimination as inference

How do quantum fluctuations limit discrimination?

Inference is choosing an hypothesis based on data. In statistical inference, the decision is based on a model of the probabilities of the hypotheses, H and the data, k (or x above). What are our hypotheses?

Think of the task in terms of signal transmission. The sender wishes to set the intensity of the light switch to one of two positions, corresponding to a high and low light. The hypothesis space is simple: H = S_L (low), or H = S_H (high). The idea is that we have two hypotheses that each determine a conditional probability distribution.

Thus, given a measurement of x photons, there is an inherent ambiguity in determining the cause or signal--S_L or S_H? Next time, we will study the signal detection theory, and then see how to extend the fundamental principles of SDT to general perceptual inference for image patterns and visual tasks.

References

Barlow, H. B. (1962). A method of determining the overall quantum efficiency of visual discriminations. Journal of Physiology (London), 160, 155-168.

Barlow, H. B., & Mollon, J. D. (1982). The Senses. Cambridge: Cambridge University Press.

Baylor, D. A., Lamb, T. D., & Yau, K. W. (1979). Responses of retinal rods to single photons. Journal of Physiology, Lond., 288, 613-634.

Cornsweet, T. N. (1970). Visual Perception . New York: Academic Press.

Green, D. M., & Swets, J. A. (1974). Signal Detection Theory and Psychophysics . Huntington, New York: Robert E. Krieger Publishing Company.

Hecht, S., Shlaer, S., & Pirenne, M. H. (1942). Energy, quanta, and vision. Journal of General Physiology, 25, 819-840.

Rose, A. (1948). The sensitivity performance of the human eye on an absolute scale. Journal of the Optical Society of America, 38, 196-208.

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