



PHYSICS/NEURO 141. THE PHYSICS OF SENSORY SYSTEMS IN BIOLOGY

Prof. Aravi Samuel, Department of Physics.

Dr. Alina Vrabioiu, Department of Physics.

Ariana-Dalia Vlad, Department of Physics.

Living organisms use sensory systems to inform themselves of the sights, sounds, and smells of their surrounding environments. Sensory systems are physical measuring devices, and are therefore subject to the laws and limits of physics. Here, we will consider the physics of sensory measurement and perception, and study ways that biological systems have solved their underlying physical problems. We will discuss specific cases in vision, olfaction, and hearing from a physicist's point of view. **N.B.:** This is one course that is taught as both Neuro 141 and Physics 141. Whether a student enrolls in Neuro 141 or Physics 141 is flexible, to best suit each student's academic planning and concentration requirements.

ARAVI SAMUEL grew up in upstate New York, and graduated with a BA in physics and PhD in biophysics from Harvard. For his PhD, he studied the biophysics of bacterial chemotaxis with Howard Berg, a pioneering biophysicist at Harvard. Berg established our mechanistic understanding of sensory perception in bacterial chemotaxis, one of the best understood systems in biology ¹ (Figs. 2, 3). Aravi thinks about the biophysics of circuits and behavior in bacteria, worms, and flies in much the same mechanistic way as his mentor. To find out more about Aravi's work, see [his website](#).

Email: samuel@physics.harvard.edu

ALINA VRABIOIU was a postdoc with Howard Berg, and is continuing her postdoctoral work with Aravi on the bacterial flagellar motor.

Email: alina_vrabioiu@fas.harvard.edu

ARIANA-DALIA VLAD was a student in this course in 2022, and is currently a senior studying quantitative biology and physics.

Email: avlad@college.harvard.edu

Updated: September 28, 2023

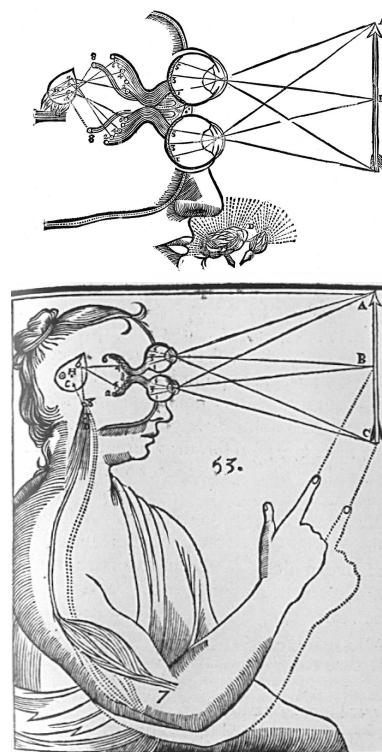


Figure 1: René Descartes (1596-1650) depicts sensory integration between two of Aristotle's five senses, vision and smell (above) and a sensorimotor pathway that mediates a behavioral response (below). Descartes was an early believer that sensory perception and animal behavior could be reduced to physical processes and the mechanical flow of information through an organism, the conceptual framework for modern neuroscience and this course. Descartes did make some mistakes. In these drawings, sensory perception and behavior are mediated by the pineal gland, not the brain.

¹ H C Berg. A physicist looks at bacterial chemotaxis. *Cold Spring Harbor Symposia on Quantitative Biology*, 53 Pt 1: 1-9, 1988. ISSN 0091-7451

LECTURES (Tu, Th 9-10:15 EST in NW255) will teach fundamental physics and math that illuminates sensory neuroscience. We will also discuss classic papers about vision, hearing, and chemical sensing that can be unlocked with physical reasoning.

SECTIONS will occur at the same time and same place as lectures, occasionally taking the place of an ordinary lecture but where the teaching fellows will go over problem-solving that is relevant to problem sets.

OFFICE HOURS will be held by Aravi, Alina, and Ariana by appointment.

COURSE MATERIALS will be distributed as this main PDF course packet, containing hyperlinks to directly download all other additional required and recommended reading material (chapters of other books and primary papers). There is no required course textbook. **Primary papers** that we cover each week will draw on both classic and current studies in vision, hearing, and chemical sensation.

GRADES will be based on Problem sets and coding assignments, 60%; Final student presentation (20%); Class participation (20%).

- **Problem Sets** We will have occasional problem sets, roughly every third week, that will include questions about the physics and biology of sensory systems. The problem sets will be longer than the typical weekly problem set given in most classes, so it is advisable to start early. In section, you will go over solving similar problems with the help of the TFs. We will occasionally include small coding assignments involving simulations or data analysis.
- **Final presentation** At the end of the course, students will make a final presentation. Students will
 1. Write their own review-style paper
 2. Present a short talk that describes the importance of their chosen paper and communicates any essential physics or math needed as background
 3. Construct their own coding project (e.g., data analysis or simulation) that illuminates their paper

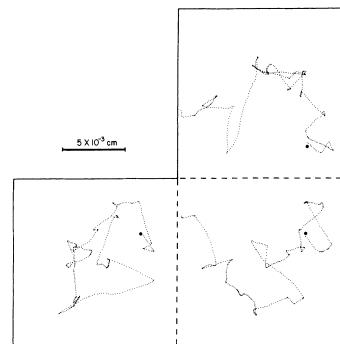


Figure 2: Bacteria perform chemotaxis by a random walk. The three-dimensional track of a single swimming bacteria viewed in xy, yz, and xz projections. The movement can be characterized as an alternating sequence of runs (periods of forward movement) and tumbles (periods of erratic rotational movement). When the bacteria is pointed in a direction it wants to go, runs get longer. The random walk becomes biased towards preferred environments.

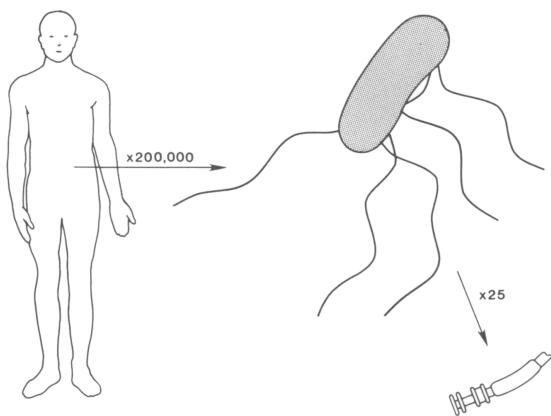


Figure 3: Man, *E. coli*, and its flagellar motor, a study in scale. Physics is different at different scales. No sensory or behavioral system is better understood than that of bacterial chemotaxis, where behavior can be reduced to the rotation of individual bacterial flagella, and every event from perception of sensory inputs (chemoreceptors binding to receptors) to motor control has been determined.

Contents

CALENDAR	5
WHY STUDY SENSORY SYSTEMS?	6
HOW BIOLOGISTS STUDY SENSORY SYSTEMS	16
SOME STATISTICAL MECHANICS	33
COLOR VISION AND THE DIMENSIONALITY OF PERCEPTION	39
THE STATISTICS OF PHOTON ABSORPTION BY PHOTORECEPTORS	48
HOW MANY PHOTONS CREATE VISION	54

CALENDAR

Class Meeting	Lecture Topic	Slides
Sep 5	Why Study Sensory Systems?	Slides
Sep 7	How biologists study sensory systems	Slides
Sep 12	Some Statistical Mechanics	Slides
Sep 14	Vision	Slides
Sep 19	Vision	Zoom
Sep 21	Review Section	Slides
Sep 26	Vision	Problem Set One Due
Sep 28	Vision	Slides
Oct 3	Vision	Slides
Oct 5	Vision	Slides
Oct 10	Review Section	Slides
Oct 12	Smell	Problem Set Two Due
Oct 17	Smell	
Oct 19	Smell	
Oct 24	Smell	
Oct 26	Review Section	
Oct 31	Hearing	Problem Set Three Due
Nov 2	Hearing	
Nov 7	Hearing	
Nov 9	Hearing	
Nov 14	Review Section	
Nov 16	Other senses	Problem Set Four Due
Nov 28	Student Presentations	
Nov 30	Student Presentations	
Dec 5	Prof. Samuel's Birthday	Something fun

WHY STUDY SENSORY SYSTEMS?

OUR BEHAVIORS begin with sensory perception. We gather information about the sights, sounds, smells, tastes, and textures of our worlds using our sensory systems. We respond to this sensory information using our brains and motor systems. We have wondered about our capacity for sensory perception as long as we have wondered about ourselves. In his treatise *de Anima*, Aristotle focused on sensory perception and motility, arguing that these were the most essential biological processes in making us alive. He thought sensory perception and motility would be windows to the soul.

We might argue with Aristotle about the “soul”, but a modern neuroscientist would accept that our sensory and motor systems are credible windows into the brain. Understanding how the human brain creates cognition and behavior is our major challenge in 21st-century science. To make progress, we can make a strong case to start with sensory systems. Every neuron is an “information processing unit” with the task of mapping incoming signals to outgoing signals. Incoming signals might come from other neurons (e.g., by synaptic communication) or from the rest of the body (e.g., by non-synaptic chemical communication). Outgoing signals are sent to other neurons, muscle cells, or the rest of the body through synapses and chemical messengers. The experimental advantage with sensory neurons is that these cells provide the scientist with a direct handle on their most salient incoming signals, environmental stimuli that can be controlled in the laboratory, like photons for photoreceptors or molecules for chemoreceptors. If we can understand the principles by which one sensory neuron maps an incoming signals to outgoing signals, we can illuminate the principles by which any neuron maps inputs into outputs.

Evolutionary conservation and homology allows us to leverage sensory neurons to understand common features of any neuron. “*Nothing in biology makes sense except in the light of evolution.*” – Dobzhansky. We will learn that sensory neurons share general computational principles in the mapping of incoming signals to outgoing signals that can be recognized throughout biology, from single-celled microorganisms to the photon-counting rod cells in the human retina. that abstract operating principles with other neurons, More concretely, any given sensory neuron in an animal can share homologous molecules and mechanisms, with other sensory neurons across Aristotle’s five sensory modalities and beyond. The molecular basis of olfaction, where olfactory neurons smell molecules from the environment (Fig. 5) is closely linked to the molecular basis of synaptic



Figure 4: Aristotle’s Five Senses. Vision, Hearing, Taste, Smell, Touch. Hearing and touch are both mediated by mechanosensory neurons. Smell and taste are both mediated by chemosensory neurons.

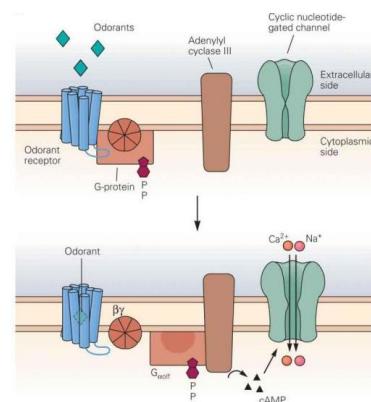


Figure 5: Odorant receptors. Binding an odorant causes an odorant receptor to interact with a G-protein signal transduction cascade that increases cAMP concentration. Elevated [cAMP] opens cyclic nucleotide-gated cation channels, causing a change in membrane potential, the neural signal for olfactory detection. Here, we illustrate a “metabotropic” olfactory receptor where stimulus energy is indirectly converted to neural activity through biochemistry that activates separate ion channels. Some olfactory receptors are metabotropic, but some, like insect olfactory receptors, are ionotropic (like the neurotransmitter receptor shown in Fig. 6 where the receptor contains its own ion channel).

communication, where postsynaptic neurons essentially “smell” neurotransmitters released by presynaptic neurons at synaptic clefts (Fig. 6).

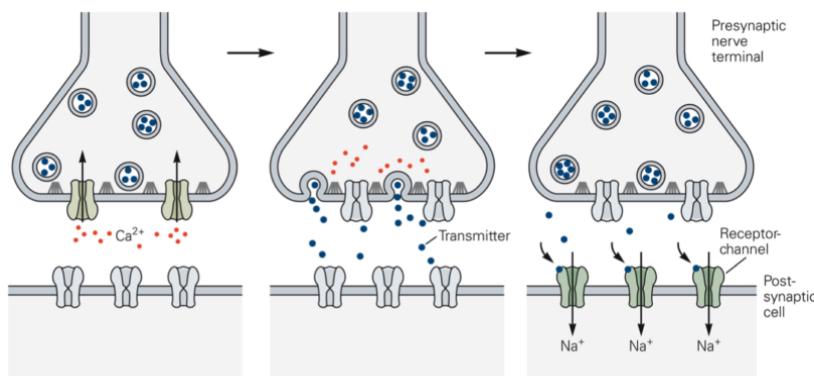


Figure 6: Synaptic transmission as olfaction. An action potential arriving at the terminal of a presynaptic neuron causes voltage-gated Ca^{++} channels at the active zone to open. Channel opening and the rise in intracellular calcium levels causes vesicles containing neurotransmitter to fuse with the cell membrane and release their contents. Released neurotransmitter molecules diffuse across the synaptic cleft and are “smelled” by specific receptors on the postsynaptic membrane. Here, we illustrate an “ionotropic” receptor where neurotransmitter binding is directly converted to neural activity by opening an ion channel in the receptor itself. Some neurotransmitter receptors are ionotropic, but some are metabotropic (in the same way as the vertebrate olfactory receptor shown in Fig. 5).

Evolution creates bigger brains with added functions, but always by reusing and adapting a relatively small set of especially pliant information-processing mechanisms. The first recognizable nervous systems probably belonged to animals where only one or two cells carried out every information-processing step from perception to action. Larger nervous systems separate the many information-processing steps that were carried out by single primordial neurons, dividing functions among different neurons across multicellular circuits and brains, always by repeating, reusing, and adapting homologous building blocks. If we can understand how one sensory neuron responds to the outside world, we are well equipped to understand how any neuron responds to the inside world of the brain.

Another reason that physicists might want to study sensory systems is that sensory phenomena are particularly amenable to identifying and characterizing the primary physical triggers that turn environmental stimuli (photons, chemicals, mechanical force) into neural activity. A pattern of stimulus energy is translated into a pattern of neural activity that is then used by the brain to assess its surrounding world. Sensory systems are highly evolved, often working at the physical limits of operation. We see individual photons. Many organisms smell individual molecules. Understanding the physical challenges in counting photons and molecules requires understanding the physics of light, thermal noise, molecular diffusion, fluid mechanics, and statistical mechanics. Understanding sensory systems is a motivation for learning many areas of fundamental physics in the context of concrete biological application. We will learn this physics from the “ground up” in a way that is directly applied to biological

questions. Not only might we learn physics more deeply by *using* it, we will gain practice in using physics to build models of the natural world, an important skill beyond neuroscience.

Beyond Aristotle

THE FIVE SENSES that Aristotle cataloged in *de Anima* remain the five senses of conventional wisdom. But if the essence of sensory perception is gathering *information*, there are countless legitimate sensory systems throughout biology from microorganisms to man. Most animals sense pain, called nociception. Many animals sense the position and posture of their own bodies, called proprioception. Many animals sense internal states like visceral information about sickness or hunger that is communicated to the brain, called interoception. Many animals have sensory modalities that we do not. Weakly electric fish sense the perturbations of self-generated electric fields as a type of radar (Fig. 8). Bats use sound and hearing as another type of radar.

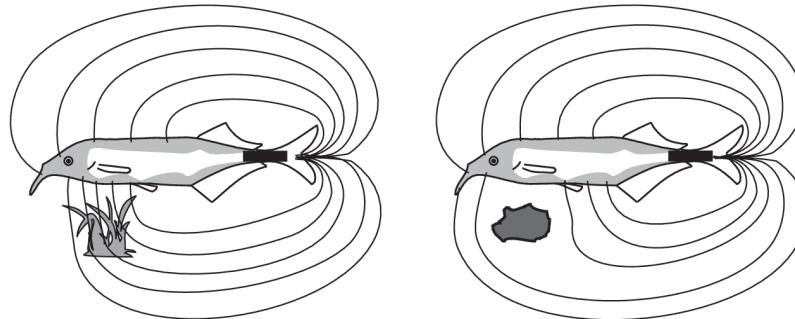


Figure 7: **Thermal imaging cameras.**
The pit organ of vipers is an exquisitely sensitive thermal imaging camera, albeit with lower resolution than this one formed by a man-made camera.

Sensory Receptors

We know how many environmental stimuli activate their corresponding receptor molecules. One large, highly conserved family of G-protein coupled receptors are common to all animals. GPCRs have proved especially pliant receptors across evolutionary history, having been adapted to detect photons (rhodopsin molecules in our retina), chemicals (olfactory receptors in our noses), and countless neurotransmitter and hormonal receptors that effectively “smell” molecules in our brains and bodies. GPCRs detect photons by virtue of a pigment (retinal) that changes shape upon photon absorption, triggering the conformational change in the surrounding protein that modulates its activity (Fig 38). GPCRs detect molecules when ligand binding

Figure 8: **Imaging with electric fields.**
We create visual maps of our surroundings our eyes, but other animals do the same with other sense. Schematic two-dimensional drawings of the electric fields of a *Gnathonemus petersii* distorted by a water plant (good conductor, left) and a stone (isolator, right). The fish is viewed from the side. Electrical field lines are drawn as thin lines. The electrorreceptive body surface of the fish is shown in grey.

directly triggers a conformational change that modulates protein activity.

The conformational change that accompanies the activation of an olfactory receptor has now been visualized. Unlike mammalian olfactory receptors which are GPCRs and *metabotropic* (meaning that they activate an intracellular signaling pathway that eventually changes the electrical activity of the cell) insect olfactory receptors are *ionotropic* (meaning that odorant binding directly opens an ion channel in the receptor itself). **Cryo-electron microscopy** allowed direct visualization of this structural change – opening and closing the ion channel within a receptor– by comparing bound and unbound states (Fig. 10). The evolutionary divergence between insect ionotropic olfactory receptors and vertebrate metabotropic receptors highlights the ancientness of a critical sensory modality that has evolved separately in different phyla for hundreds of millions of years. Even older are the olfactory receptors and signaling pathways that allow bacteria to ‘smell’ their environments, using “metabotropic” mechanisms without shared evolutionary history to animals.

Sometimes we know what receptors sense a particular stimulus, but not how they do it. Most organisms sense temperature. Temperature sensing is critical for small animals that do not regulate their own body temperatures – called poikilotherms or ectotherms or cold-blooded (if they have blood). Ectotherms rely on environmental temperature to regulate their own body temperatures by moving from place to place, think of reptiles sunning themselves to warm up. Thermosensation can also be used to gather information, like the pit organs of certain snakes that act like image-forming *pinhole cameras*, seeing warm objects by focusing infrared radiation onto highly sensitive thermoreceptors (Fig. 7). Thus, the most sensitive known thermoreceptors in biology can sense temperature changes as small as $0.001\text{ }^{\circ}\text{C/sec}$! Many thermoreceptors have been identified – like certain insect gustatory receptors that evolved to sense warming and cooling. However, the precise biophysical mechanism for temperature sensing remains poorly understood. While we know what receptors are needed, we do not know how thermal stimuli as small as millidegrees drive the specific changes in receptors that toggle their activities.

Receptor diversity within sensory modalities

Evolutionary divergence of receptors within a sensory modality can increase the sophistication and dimensionality of sensory coding. We have *trichromatic* vision because our retinas have three types of *cone cells* that are sensitive to long, medium, or short wavelengths of

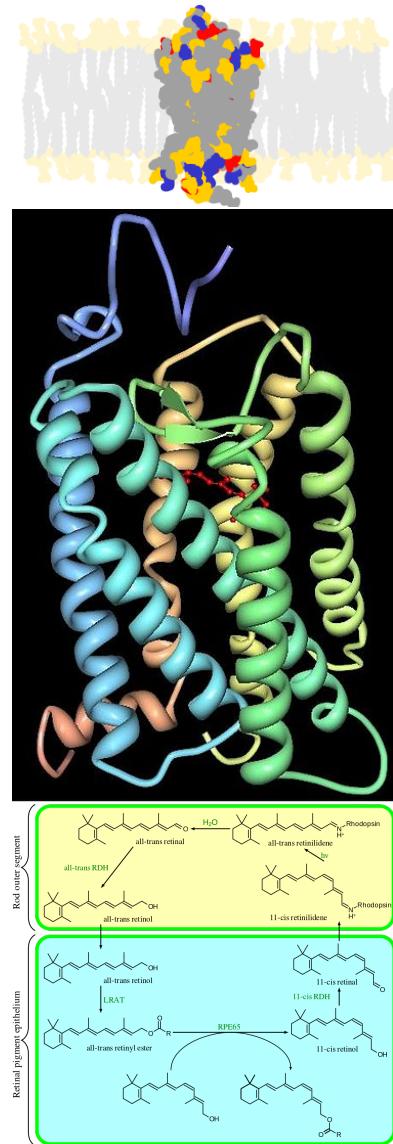


Figure 9: **Rhodopsin.** Three dimensional structure of bovine rhodopsin, a membrane bound protein. The chromophore, retinal, is embedded within the protein. Absorption of light energy, $h\nu$, causes a conformational change – 11-cis-retinal becomes all-trans-retinal – which alters protein structure to activate G-protein-coupled signal transduction. Biochemistry in the rod cell restores the cis configuration.

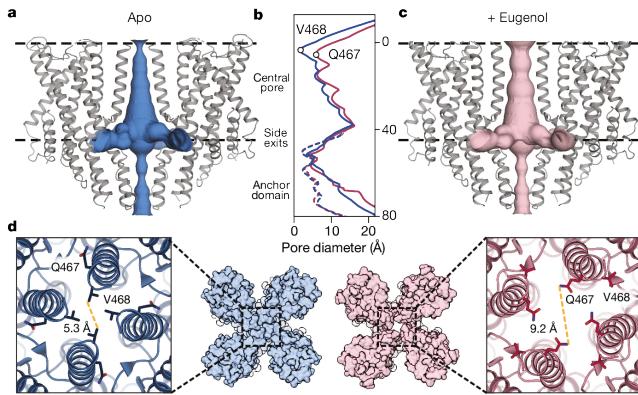


Figure 10: Odorant-evoked opening of an ionotropic olfactory receptor. a, c, The channel pores of unbound (a, blue) and eugenol-bound (c, pink). Black dashed lines, membrane boundaries. b, The diameter of the ion conduction pathway (solid lines) and along the anchor domain (dashed lines). d, Close-up view of the pore from the extracellular side. Marmol et al. (2021).

bright light by virtue of three slightly different rhodopsin molecules (Fig. 14). We have one type of rod photoreceptor that is sensitive to dim light. Because we only have one type of rod photoreceptor with one type of wavelength sensitivity, vision in dim light is *monochromatic*. But unlike the cone cells, our rod cells can detect single photons, a level of sensitivity that is needed to see the dimmest stars in the night sky.

Sensory discrimination is more complicated when different stimulus types must be separated and identified by different receptors. In our noses, diverse *olfactory receptors* in the nasal epithelium are tuned to different volatile chemicals. Mammalian noses contain many different types of olfactory receptor neurons, each with its own molecular olfactory receptor (Fig. 11). When we whiff a scent, blends of different types and concentrations of molecules enter the nasal cavity. These molecules bind to different olfactory receptors with different chemical specificity. Odorant molecules can exhibit relatively small differences in their chemical structure and properties (Fig. 12). Thus, one receptor can bind (and be activated) by many different kinds of molecule. One molecule can bind (and activate) many different kinds of receptors. Because typical smelly objects will emit many different odorant molecules, patterns of olfactory receptor activity can be expected to be complicated. A *combinatorial code* might be needed to discriminate odor molecules and identify smells (Fig. 13).

When many receptors in many cells are used to collect information along the different dimensions of a sensory modality (different colors in vision, different chemicals in smell), the information must be collected and analyzed by downstream circuits. Information travels from sensory receptors to our brains along *afferent* neuronal processes. Our modern circuit-level understanding of sensory processing

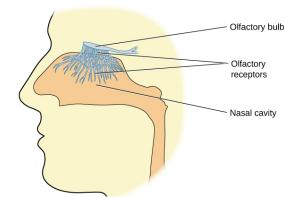


Figure 11: Olfactory receptor neurons innervate the olfactory epithelium in the human nasal cavity.

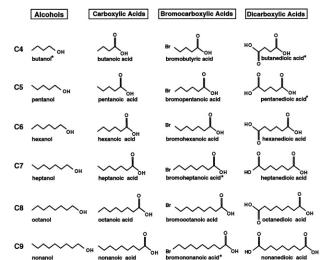


Figure 12: Volatile odorant molecules among the many thousands of chemicals that we can smell.

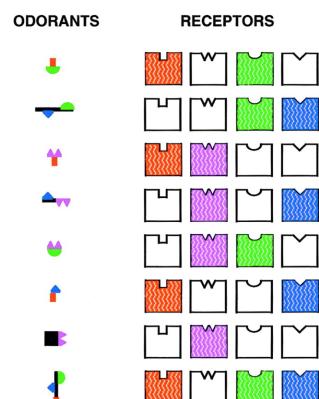
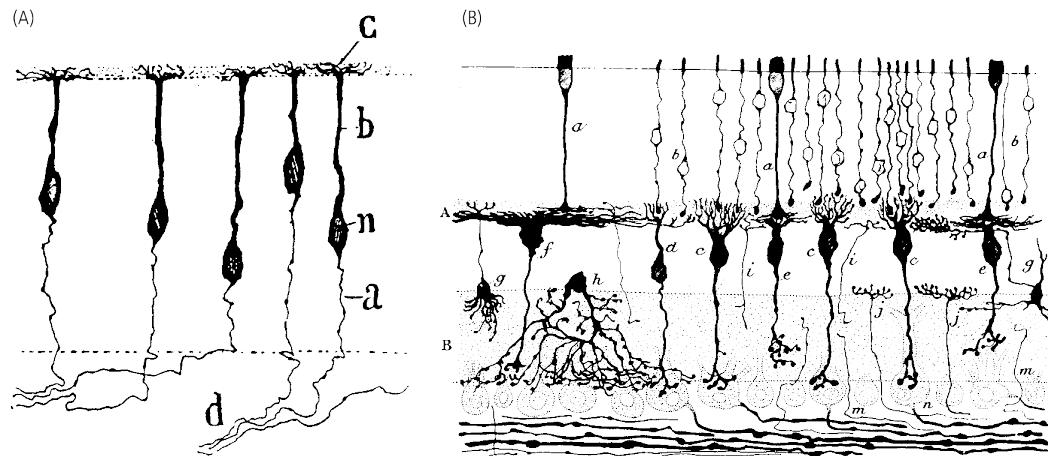


Figure 13: Cartoon of the combinatorial code where differently shaped molecules can bind to olfactory receptors with differently shaped binding pockets. Mammals can have 100s of different olfactory receptors that presumably detect many 1000s of different odorant molecules.

began with the anatomical studies of **Santiago Ramón y Cajal** (1852–1934) who visualized the detailed structure of many types of sensory neurons, as well as the wires and intercellular contacts (synapses) that carried information to the brain (Fig. 14). Cajal did this by perfecting Golgi's method for sparsely staining neural tissues with silver salts. The Golgi staining method made individual neurons visible at random. Each neuron could be categorized and reconstructed using light microscopy. Collecting neuronal structures across samples, Cajal worked out the basic structure and connectivity of many neural circuits, revealing many principles of information processing that remain valid today.



The physics in sensory perception

STIMULUS ENERGY is the initiating event of all sensory perception. Photoreceptors transduce the energy of photons. Chemoreceptors transduce chemical binding energy. Mechanoreceptors transduce mechanical stretch or movement. Information is generated by some ‘energy absorbing’ primary event in sensory detection – a photoreceptor molecule absorbs a photon, a chemoreceptor binds a molecule, a mechanoreceptor is tugged open.

Sensory information eventually reaches circuits that shape organism behavior. For this to happen, the primary sensory signal – the energy from absorbing single photons or binding single molecules – must be filtered, amplified, and transmitted through intervening pathways. These different stages of sensory perception can be inter-

Figure 14: Santiago Ramón y Cajal has claim to be the father of modern neuroscience through his detailed anatomical studies of neural tissue. (A) Bipolar sensory neurons from mammalian olfactory mucosa. a, Axon; b, peripheral process; c, sensory dendrites; d, axon; n, nucleus. (B) Section of retina of an adult dog. A, Outer plexiform (synaptic) layer; B, inner plexiform (synaptic) layer; a, cone fiber; b, rod cell body and fiber; c, rod bipolar cell with vertical dendrites; d, cone bipolar cell with vertical dendrites; e, cone bipolar cell with flattened dendrites; f, giant bipolar cell with flattened dendrites; g, special cells stained very rarely (perhaps inter-plexiform cells); h, diffuse amacrine cell; i, ascendant nerve fibers (probably processes of cell not well stained); j, centrifugal fibers coming from central nervous system; m, nerve fiber (probably again of poorly stained cell); n, ganglion cell. (from Fain, Chapter 1)

preted in engineering terms (Fig.). Every sensory system has some sort of ‘antenna’ that serve to detect primary signals. At the cellular level, the antenna for vision, for example, might be interpreted to be the photoreceptor cell. At the molecular level, the antenna might be molecular rhodopsin which directly absorbs each photon.

The primary sensory signal brings an amount of energy to the antenna that must be amplified. The energy of a single photon is calculated using Einstein’s relation $E = hc/\lambda$. The energy of one blue-green photon ($\lambda=500 \text{ nm}$) is $4 \times 10^{-19} \text{ J}$.

After detection, primary signals must be filtered or analyzed to create useful information. Filtering is needed to separate true signals from noise. Thermal energy constitutes a constant and unavoidable source of random noise that contaminates the sensory antenna. One of the most useful results from **Statistical Mechanics** is the *Equipartition Theorem* that states that all bodies in thermal equilibrium have $k_B T/2$ of average energy in every quadratic degree of freedom. For example, a gas molecule has an average kinetic energy of $k_B T/2$ along each axis of motion. Kinetic energy is quadratic in velocity components, three degrees of freedom in three-dimensional space:

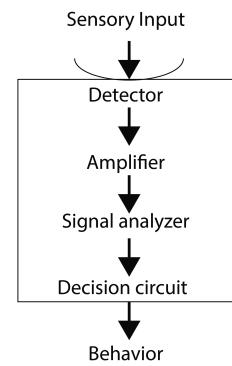
$$\left\langle \frac{mv_x^2}{2} \right\rangle = \left\langle \frac{mv_y^2}{2} \right\rangle = \left\langle \frac{mv_z^2}{2} \right\rangle = \frac{k_B T}{2} \quad (1)$$

Thermal energy at room temperature ($\sim 25^\circ\text{C}$), roughly the temperature of most biology, is $4 \times 10^{-21} \text{ J}$. By comparison, the energy of one blue photon is $\sim 100 \times$ thermal energy. Although thermal energy fluctuates about $k_B T/2$, because a single photon is so much more energetic, its corresponding signal is much larger than this fluctuating noise. If the energy threshold for single photon detection is set to be much larger than $k_B T/2$ and below $\sim 100k_B T/2$, the energy of ‘true signals’ caused by single photons can be reliably filtered from ‘false signals’ caused by thermal energy fluctuations.

THE BOLTZMANN DISTRIBUTION predicts that any detection threshold can be broached, albeit perhaps rarely, by thermal noise. For a body at thermal equilibrium, the probability of a thermal fluctuation with energy E_c is exponentially distributed:

$$P(E_c) \propto e^{-E_c/k_B T} \quad (2)$$

This means that the probability that a rhodopsin molecule reaches energies comparable to visible photons by thermal fluctuations, although vanishingly small, is also non-zero. Our eyes have hundreds of millions of photoreceptor cells (Fig. 21). Each photoreceptor cell is filled with billions of rhodopsin molecules. Although the likelihood



$$h = 6.6 \times 10^{-34} \text{ J s}$$

$$k_B = 1.4 \times 10^{-23} \text{ J K}^{-1}$$

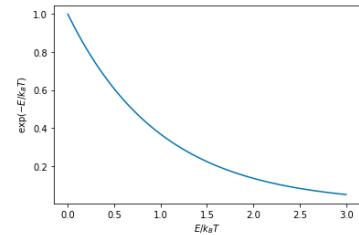


Figure 15: The Boltzmann Distribution

that a single rhodopsin molecule reaches detection threshold because of temperature might be small, so many rhodopsin molecules mean that ‘dark noise’ events – the spontaneous activation of a rhodopsin molecule by temperature, not by photon absorption – can occur with significant frequency. The activation of rhodopsins by true photons have to be discriminated from the static noise of spontaneous rhodopsin activation by temperature. It is impossible to say whether a single given rhodopsin activation was triggered by a photon and not by thermal fluctuation. Multiple simultaneous rhodopsin activations are required to discriminate a true flash of light from a dark noise event.

Seeing single photons

WHAT IS THE SMALLEST NUMBER OF PHOTONS that a human observer can reliably detect? In the early history of the photon, Lorentz realized that a ‘just detectable’ flash of light delivered ~ 100 photons to the cornea. Most photons that reach the cornea do not reach the retina. The deeper question is: What is the threshold number of photons that is absorbed by photoreceptor cells to produce ‘seeing’? Hecht, Shlaer, and Pirenne (1942) did the classic experiment that established that single photons absorbed by single rod cells could be integrated into a perception of a flash of light. The threshold number of photons individually absorbed by a group of photoreceptor cells at the ‘threshold of seeing’ was 5–7 photons (Fig. 16).

Smelling single molecules

BACTERIAL CHEMOTAXIS is driven by smell. Bacteria are covered with chemoreceptors for molecules like amino acids that signify food. They use these chemoreceptors to assess surroundings and swim to favorable places. *E. coli* swims by rotating helical flagellar filaments attached to its $\sim 1 \mu\text{m}$ size body (Fig.). When all flagella rotate counterclockwise, as viewed from outside the cell, a bundle forms that pushes the cell forward at $\sim 25 \mu\text{m}/\text{s}$. These ‘runs’ typically last $\sim 1 \text{ s}$. Occasionally, one or more flagellar motors switch from CCW to CW rotation. This ends the run by disrupting the flagellar bundle. The cell stays in place and ‘tumbles’ until all flagella return to CCW rotation, and the cell starts a new run in a new direction.

Bacterial chemotaxis works by counting molecules. If the bacteria counts more attractant molecules over time during a run, it postpones the next tumble. Runs in favorable directions are thus longer than runs in unfavorable directions. Although each tumble randomly

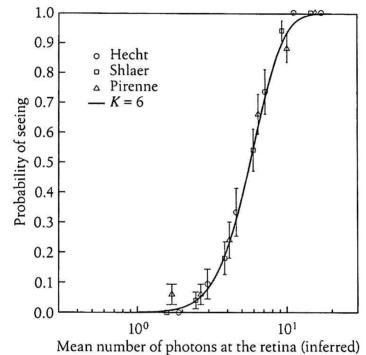


Figure 16: **The threshold of seeing** as the probability of seeing a flash plotted against the logarithm of the number of photons estimated to be absorbed by the retina at different flash strengths. Open symbols represent different measurements from three observers, the authors of the experiment.

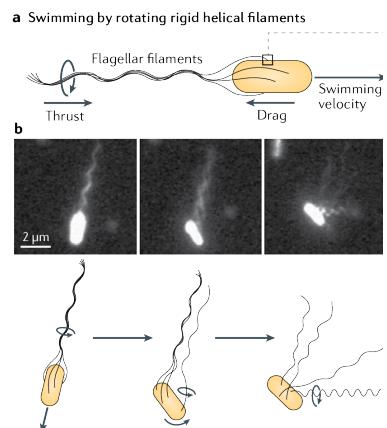


Figure 17: **Bacteria swim with a random walk** alternating periods of forward movement (counterclockwise rotation of flagellar filaments) with tumbles (clockwise rotation of one or more flagellar filaments).

reorients each run, a *biased random walk* ensues that inexorably drives the bacteria where it wants to go.

Noise in counting molecules during chemotaxis has a thermal origin, but in another sense than thermal activation of unbound receptors. Consider a bacterium-sized volume, $L=1 \mu\text{m}$ on each side. If the cell instantaneously counted all molecules inside this volume, it would count $\sim 600,000$ molecules if the mean concentration was 1 mM , 600 molecules if the concentration was $1 \mu\text{M}$, and 60 molecules if the concentration was 10^{-7} M . But molecules constantly move in and out of the measurement volume, and so the number in an instantaneous count will fluctuate in a way governed by **Poisson statistics** (Fig. 18). The standard deviation in the number of counted molecules will be the square root of the mean number of molecules. And so the relative error in estimating molecular concentration based on a single count within the $1 \mu\text{m}^3$ -sized measurement volume will be $600,000 \pm 800$ at 1 mM ($\sim 1.3\%$ error), 600 ± 24 at $1 \mu\text{M}$ ($\sim 4\%$ error), and 60 ± 8 molecules at 10^{-7} M ($\sim 13\%$ error). Berg and Purcell (1977) showed that bacteria do better than this by integrating measurements over time. During bacterial chemotaxis, measurable changes in bacterial behavior are caused by changing the occupancy of single receptors.

All biology is sensing molecules

Understanding the biophysics of chemoreception is not idiosyncratic to bacterial chemotaxis. Virtually every process in intracellular and intercellular biological signaling involves the diffusive movement of molecules. Intracellular biochemical pathways involve the binding of ligands and enzymes (Fig. 19). Synaptic communication between nerve cells involves the diffusion of neurotransmitters from a presynaptic cell to the receptors on a postsynaptic cell. Understanding the biophysics of chemoreception is fundamental to understanding life.

RECOMMENDED READING

- S M Block. Biophysical principles of sensory transduction. *Society of General Physiologists*, 47:1–17, 1992. ISSN 0094-7733 [Download paper](#)
- H C Berg. A physicist looks at bacterial chemotaxis. *Cold Spring Harbor Symposia on Quantitative Biology*, 53 Pt 1:1–9, 1988. ISSN 0091-7451 [Download paper](#)
- Chapter One. Gordon L Fain. *Sensory Transduction*. Sinauer Associates, Sunderland, Mass., 2003. ISBN 0878931716 [Download paper](#)

ADDITIONAL READING

- Josefina del Marmol, Mackenzie A. Yedlin, and Vanessa Ruta. The structural basis of odorant recognition in insect olfactory receptors. *Nature*, 597(7874):126–131, 2021. ISSN 0028-0836 [Download paper](#)

Updated: September 28, 2023

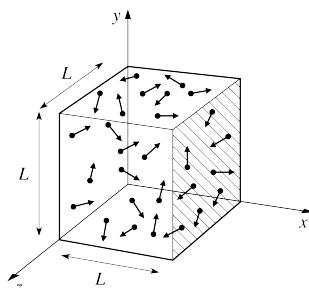


Figure 18: **Diffusion** is driven by the random and incessant motion of particles in and out of sampling volumes. At any point in time, the number of particles in a sampling volume fluctuates about a mean concentration owing to these ‘Brownian movements’.

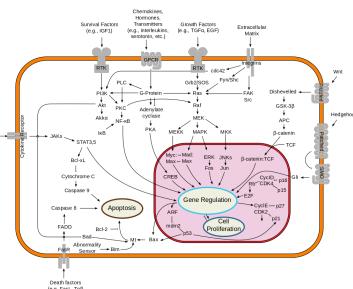


Figure 19: **Eukaryotic signal processing** is dominated by diffusion and the random binding and unbinding of molecules to receptors.

- S Hecht, S Shlaer, and M H Pirenne. Energy, quanta, and vision. *The Journal of General Physiology*, 25(6): 819–840, 1942. ISSN 0022-1295 [Download paper](#)

HOW BIOLOGISTS STUDY SENSORY SYSTEMS

Psychophysics

THE QUANTITATIVE STUDY of sensory perception, *psychophysics*, began with 19th century experimental psychology. Conscious human responses to well-controlled stimuli began to be measured and analyzed. With a human subject, a scientist can obtain meaningful perceptual measurements simply by asking questions. Did the subject detect a stimulus – did she or didn't she see the flash of dim light? Can the subject tell the difference (discriminate) between two stimuli – which is heavier, block A or block B? For the answers to such questions to be meaningful, one should use well-controlled stimuli and ask simple questions with clear answers. But even with rigorously designed experiments and simple questions, interpreting the results of psychophysical experiments can be subtle. *Psychometric curves*, mathematical functions that interrelate sensory response and stimulus intensity, must be carefully interpreted (Fig. 20).

Psychometric curves are shaped by the activities of underlying molecular and cellular sensors. An animal perceives a quantitative change in a given stimulus because some sensory molecules or cells undergo a quantitative change in activity. A stimulus might be perceived as stronger with an increase in stimulus intensity – brighter light, louder sound, higher odorant concentration – for different reasons. One possibility is that more sensors of a given type become activated – our low-light vision is mediated by one type of photoreceptor, when more rods within the retinal field of a dim object are activated, object appears brighter. Another possibility is that individual sensors become more strongly activated – the mechanosensory hair cells of the inner ear detect cochlear vibrations, the larger the vibration amplitude, the louder the sound. A third possibility is that different sensors with different intensity thresholds for the same stimulus become activated – our noses contain many different olfactory receptors with different activation thresholds for a given odor molecule, so the set of olfactory receptors that are activated by a given smell can contain information about the identity and intensity of odor molecules. Our goal as neuroscientists is to explain animal perception in terms of behavior with underlying mechanisms in terms of molecules and cells. To do this, sensory neuroscientists can study the relationships between psychometric curves that describe behavior and “dose-response” curves that describe molecular and cellular activity.

When studying sensory systems, it is useful to design experiments

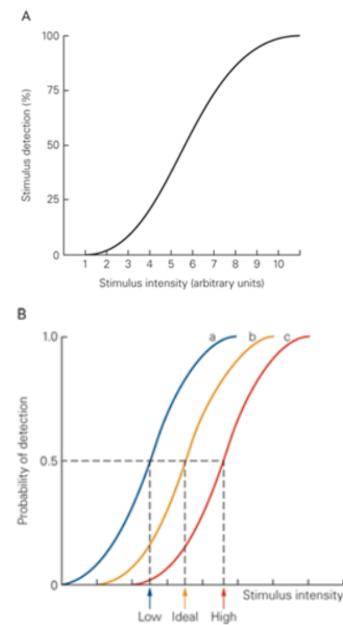


Figure 20: **The psychometric curve.** A. The psychometric function plots the percentage of stimuli detected by a human observer as a function of stimulus magnitude. Threshold is defined as the stimulus intensity detected on 50% of the trials. B. Detection and discrimination thresholds depend on the criteria used by individual subjects. Where an ‘ideal’ observer correctly detects the presence and absence of stimuli at the response threshold with equal probability (curve b), an observer who is told to respond to the slightest indication of a stimulus may report many false positives when no stimuli occur and has a lower response threshold (curve a). An observer who is told to respond only when very certain that a stimulus has occurred reports more hits than false positives and has a higher response thresholds (curve c). From *Principles of Neural Science*.

above *response thresholds* but below saturation, using stimulus intensities where the sensory system switches between ‘inactive’ and ‘active’ states with changes in stimulus intensity. One good reason is to study sensory systems where they are most relevant to natural behavior.

All biological systems *evolved* by natural selection. Sensory systems evolved to provide information about naturally occurring stimuli with certain characteristics in the real world. One learns less from sensory systems with stimuli far below threshold (i.e., too small to be detected) or far above saturation (i.e., too large to be discriminated).

Above response thresholds and below saturation, sensory systems have the most discriminatory power and provide the most information to the behaving organism. But in a regime where responses to changes in stimulus intensity are graded, sensory perception is not ‘all’ or ‘none’, binary performance that has the virtue of being easy for the experimenter to characterize on the basis of ‘yes’ and ‘no’ questions. Instead, sensory and behavioral responses will change in a gradual and typically probabilistic manner with changes in stimulus intensity, requiring more subtle (but also more informative!) analyses.

Let’s consider a concrete example. The rod cells of the human eye are specialized low-light (also called scotopic) vision. A good example of a naturally occurring stimulus that requires rod vision is the night sky. The dimmest visible stars will shower hundreds to thousands of photons per second on the dark-adapted pupil. Because most photons that reach the cornea will be scattered or absorbed before reaching the retina, rod cells must be able to detect small numbers of photons. The extraordinary sensitivity of the rod cell as photon counters was established at the turn of the last century. The astronomer Samuel Langley invented his *bolometer* in 1878, a device that allowed him to measure radiant energy with extraordinary precision – he could detect the infrared radiation from a cow at a quarter mile, quite a feat for a 19th century physicist. The bolometer allowed precise calibration of the radiant energy in visual stimuli that were just barely visible. With the realization that electromagnetic radiation is quantized as photons, Lorentz was able to estimate the “threshold for seeing” in terms of numbers of photons: just ~ 100 photons reaching the cornea could be reliably “seen”. Given photon losses due to reflection from the cornea or absorption within the eye, Lorentz’s estimate meant that single photons could trigger the activity of individual rod cells and induce a psychometric effect.

Indeed, the anatomy of the rod cell seems expressly engineered to catch photons (Fig. 21). A photon traveling along the long axis of a rod cell has to run a gamut of parallel disks loaded with visual pigment (rhodopsin molecules), but still only has a $\sim 2/3$ chance of being absorbed. Understanding the results of any photon-counting

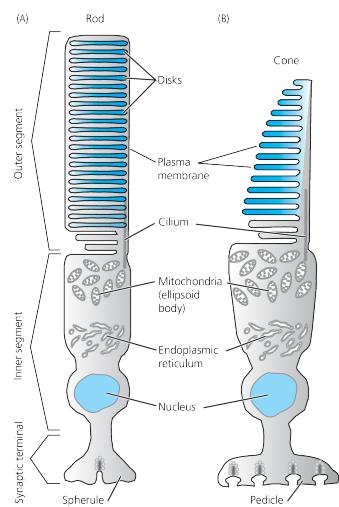


Figure 21: Vertebrate rods and cones. Principal structural features of vertebrate photoreceptors. (A) Rod. The outer segment is composed of disks detached from external plasma membrane. (B) Cone. The outer segment has membrane infoldings or lamellae instead of disks. The total area of sensory membrane is increased by these disk and lamellar structures, which are loaded with visual pigment. In a mammalian rod cell that is 0.025 cm long, roughly 2/3 of light is absorbed. (from Fain, Chapter 9).

experiment, whether at the level of animal perception or rod cell activity, requires thinking about probability and statistics. Not only is there intrinsic randomness in sensory response to a given stimulus, there is often intrinsic randomness in the delivery of any stimulus. Quantum mechanics does not allow a device that reliably deliver single photons – one at a time in a stream of identical pellets – to the eye of an observer or to a photoreceptor cell. The best we can do is understand the many sources of randomness that shape stimulus-response relationships, and develop probabilistic models that are consistent with behavioral and physiological measurements.

To make it easier to quantitatively analyze experimental results, studies are asked to make *comparative judgments* of a stimulus property, such as stimulus amplitude or frequency. To do this, we often use *two-alternative forced-choice protocol* with two observation intervals and a pair of stimuli. Subjects might be asked whether the second stimulus is stronger or weaker, higher or lower, faster or slower, same or different than the first stimulus. Or the subject might simply be asked whether the stimulus occurred in each interval. With binary questions and answers, there are only four outcomes – true positive, false positive, true negative, and false negative – making it easy to quantify and tabulate data (Fig. 22).

BAYES' THEOREM is naturally useful for interpreting psychophysical experiments in terms of *conditional probabilities*. Whether a stimulus is delivered in each trial has a well-defined probability determined by the experimenter. Whether a stimulus is detected has a probability that can be measured in the course of each experiment. Bayes' theorem is stated mathematically as the following:

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)}$$

where A and B are events and $P(B) \neq 0$. Here, the events are stimuli and responses.

- $P(A | B)$ is a *conditional probability*: the probability of event A given event B .
- $P(B | A)$ is also a conditional probability: the probability of event B given event A .
- $P(A)$ and $P(B)$ are the probabilities of events A and B respectively without conditions on other variables; they are also known as the marginal probability or prior probability.

In the experiment shown in Fig. 22, a stimulus event is true (A) if a red flash is delivered and false (\bar{A}) is a blue flash is delivered. In

		Response		Total stimuli
		Yes Red	No Blue	
Stimulus	Red	Hits (65)	Misses (35)	100
	Blue	False positives (20)	Correct rejections (80)	100
		85	115	200

Figure 22: **Two-alternative forced choice tasks.** The stimulus-response matrix for a stimulus detection task (yes-no) or a categorical judgment task (red-blue). Although there are two possible stimuli and two possible responses, the data represent conditional probabilities in which the experimenter controls the stimuli and measures the subject's responses. The numbers provide examples of behavioral data obtained from a strict observer who responds "yes" less often than the actual frequency of occurrence of the stimulus.

the same experiment, a response event is true (B) if a red flash is seen and false (\bar{B}) if a blue flash is seen. Red and blue flashes are delivered with equal probabilities so that $P(A) = P(\bar{A}) = 0.5$. However, a red flash is not always seen as red – of 100 red flashes, 65 are properly seen as red (hits or true positives) but 35 are falsely seen as blue (misses or false negatives). In terms of conditional probabilities, we write $P(B | A) = 0.65$ and $P(\bar{B} | A) = 0.35$. At the same time, a blue flash is not always seen as blue. Of 100 blue flashes, 20 are falsely seen as red (false positives) and 80 are properly seen as blue (true negatives). In terms of conditional probabilities, we write $P(B | \bar{A}) = 0.20$ and $P(\bar{B} | \bar{A}) = 0.80$. The probability of seeing a red flash is the sum of true positives and false positives: $P(B) = P(B | A)P(A) + P(B | \bar{A})P(\bar{A}) = 0.425$. The probability of seeing a blue flash is the sum of true negatives and false negatives: $P(\bar{B}) = P(\bar{B} | A) + P(\bar{B} | \bar{A})P(\bar{A}) = 0.575$.

The probability that a human subject sees a red or blue flash is an assessment of perceptual accuracy in the context of an experiment, but has less to do with perceptual accuracy in the real world. When we see something, we want to know the likelihood that the ‘something’ occurred. Using the data shown in Fig. 22, we can calculate the answer to this inverse question with the help of Bayes: what is the probability that a photon that is seen as red is actually red: $P(A | B)$.

Bayes' Theorem

BAYES' THEOREM MAY BE DERIVED from the definition of conditional probability:

$$P(A | B) = \frac{P(A \cap B)}{P(B)}, \text{ if } P(B) \neq 0$$

where $P(A \cap B)$ is the probability of both A and B being true. Similarly,

$$P(B | A) = \frac{P(B \cap A)}{P(A)}, \text{ if } P(A) \neq 0$$

Solving for $P(A \cap B)$ and substituting into the above expression for $P(A | B)$ yields Bayes' theorem:

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)} \text{ if } P(B) \neq 0$$

Interesting Bayesian results

Surprising results can emanate from Bayes' theorem. A classic example is testing for disease in a population, a pertinent modern example. Suppose, a particular test for whether someone has COVID-19 is 90% sensitive, meaning that the test is correctly positive result for 90% of infected persons (and incorrectly negative for the other 10%). The test is also 80% specific, meaning that it is correctly negative for 80% of uninfected persons (and incorrectly positive for the other 20%). Assuming 5% of the population is infected, what is the probability that a random person who tests positive is really infected? This conditional probability is: $P(\text{Infected} | \text{Positive})$.

$$\begin{aligned} P(\text{Infected} | \text{Positive}) &= \frac{P(\text{Positive} | \text{Infected})P(\text{Infected})}{P(\text{Positive})} \\ &= \frac{P(\text{Positive} | \text{Infected})P(\text{Infected})}{P(\text{Positive} | \text{Infected})P(\text{Infected}) + P(\text{Positive} | \text{Uninfected})P(\text{Uninfected})} \\ &= \frac{0.90 \times 0.05}{0.90 \times 0.05 + 0.20 \times 0.95} = \frac{0.045}{0.045 + 0.19} \approx 19\% \end{aligned}$$

In other words, even if someone tests positive, the probability that they are infected is only 19% - this is because in this group, only 5% of people are infected, and most positives are false positives coming from the remaining 95%.

THE MONTY HALL PROBLEM can also be solved by Bayes' Theorem.

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

Stimulus variability and thresholds

How might the same stimulus give rise to a true positive or false negative? One reason is that any stimulus is rarely well-characterized as a binary variable, but is usually drawn from a probability distribution over a continuum of possible values. The observer applies a threshold to each stimulus to perform binary classification: ‘seen’ or ‘unseen’; ‘true’ or ‘false’; or ‘red’ or ‘blue’ as in Fig. 22. An ‘ideal’ observer would make the fewest mistakes, such as the sum of false positives and false negatives. But most observers might have different thresholds. An optimist (or lax observer) might be more likely to classify a stimulus as ‘true’ than the ideal observer, perhaps making more false positives but fewer false negatives. A pessimist (or strict observer) might be more likely to classify a stimulus as ‘false’, perhaps making fewer false positives but also more false negatives. The psychometric curves for different observers with different thresholds would accordingly shift along the stimulus axis (Fig. 20).

To model the effects of sensory threshold and stimulus variability on psychometric curves and error rates, we need to characterize the variability of the stimulus itself. Two stimuli might exist on a continuous range of measurable value, on the basis of which they must be discriminated. If the probability distributions of the measurable values for different stimuli exhibit overlap, it is not possible to make error-free judgment. Wherever a threshold can be set to delineate ‘true’ events from ‘false’ events, there is always a finite possibility for a ‘false’ event to exceed the threshold (leading to a false positive) or for a ‘true’ event to fall below threshold (leading to a false negative).

Stimulus variability around a mean measured value is most often characterized using Gaussian distributions. For example, whether a subject sees a ‘red’ flash (A , or true event) or blue flash (\bar{A} , or false event) is made on the basis of a continuously measured variable y . The mean value of y will be different for true and false events, $\langle y_A \rangle$ and $\langle y_{\bar{A}} \rangle$, respectively. We can then write the conditional probabilities of different values of y using Gaussian distributions with different means but (for simplicity) the same variance, as illustrated in (Fig. 23):

$$P(y | A) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(y - \langle y_A \rangle)^2}{2\sigma^2} \right]$$

$$P(y | \bar{A}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(y - \langle y_{\bar{A}} \rangle)^2}{2\sigma^2} \right]$$

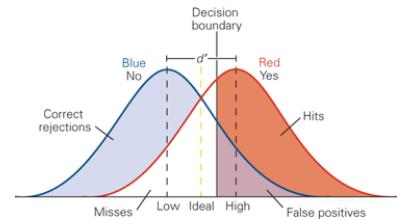


Figure 23: **Gaussian stimulus magnitudes.** Stimulus magnitudes can be represented by Gaussian curves with standard deviations that measure the fluctuation in sensations from trial to trial. The discriminability of a pair of stimuli is correlated with the distance between the two curves. When two stimuli are similar in magnitude, the two Gaussian curves overlap and no single criterion allows error-free responses. The frequency of true and false positives (and true and false negatives) is determined by the criteria used in the decision task. An ideal observer maximizes the number of correct responses and minimizes the total errors, setting the decision boundary at the intersection of the two curves. A strict observer minimizes the number of false positives but also reduces the total hits, setting the decision boundary to the right (solid line). A lax observer maximizes the number of hits but also increases the total false positives, setting the decision boundary to the left of the ideal subject.

Why do we use Gaussian distributions to characterize stimulus variability? A principled reason is the **Central Limit Theorem**. Suppose that a large sample of observations is obtained, each observation being randomly produced in a way that does not depend on the values of the other observations, and that the average (arithmetic mean) of the observed values is computed. If this procedure is performed many times, resulting in a collection of observed averages, the central limit theorem says that the probability distribution of these averages will converge to a normal distribution. An unprincipled reason is that the Gaussian distribution is simple-to-characterize with two just parameters (mean and variance) and an elegant mathematically function that facilitates calculation. Even when variability is non-Gaussian, the Gaussian is often used anyway to get useful approximate results without too much work or without too much danger.

Sensory Anatomy

FORM AND FUNCTION are complementary in biology, just like any in area of man-made engineering – designing electrical circuits, machinery, automobiles, or buildings – where structure dictates performance. The difference is that ‘engineering’ in biology was done by evolution and natural selection. The study of ‘form’ in biology occurs at many levels from the anatomy of body parts to cells to molecules.

The functional study of sensory systems began with gross anatomy. Long before magnifying glasses and microscopes were invented, analysis was limited to manual dissection and unaided human observation. For example, Galen’s early studies of the eye marveled at its unique and specialized structures. But without a proper understanding of physics of optics, the ‘model’ that Galen built was inherently flawed. Fascinated by the lens, an object like none other in the animal body, Galen made it the central structure in the vision mechanism (Fig. 24). When physical optics was discovered in the Renaissance, the properties of refraction became understood. On the practical side, optics led to useful inventions like the magnifying glass, telescope, and microscope. On the conceptual side, physical optics also clarified the role of the lens in eye. The lens is only a device for focuses images onto the retina at the back of the eye. The retina, a thin sheet of brain tissue, is the central mechanism for detecting images and relaying information to the brain.

Galen’s model is a significant distortion of eye anatomy, and it is hard to imagine how it held sway for a millennium. Scientists needed to understand the physics before they could properly see the parts of

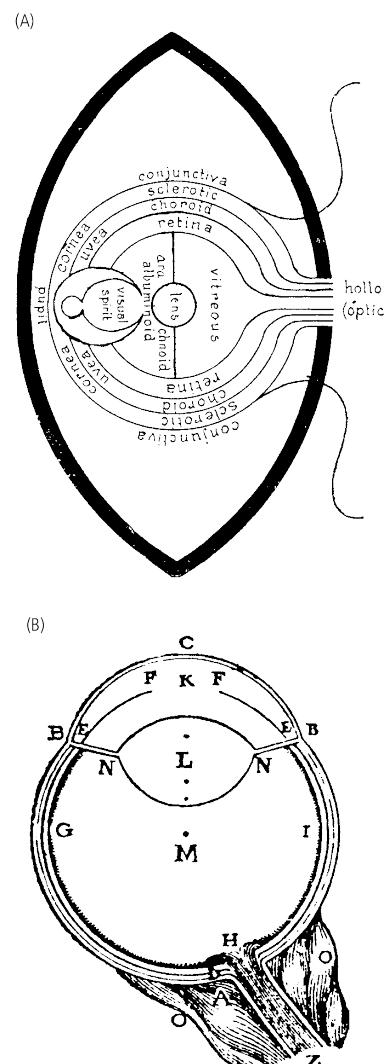


Figure 24: **Structure of the eye.** (A) Diagram of the eye from a ninth-century ad translation of Galen. (B) More anatomically correct diagram of cross-section of the eye made by Descartes. ABCB, Cornea and sclera; EF, iris (in actual fact closer to the lens than shown in Descartes' diagram); K, aqueous humor; L, lens; EN, zonyl fibers; M, vitreous humor; GHI, retina; H, optic nerve head; O, ocular muscles; and Z, optic nerve. From Fain, Chapter 1.

the eye in their proper places. Seeing is believing. Believing is also seeing. Descartes' model put the lens closer to its actual position and identified the muscles that change the shape of the lens during focal accommodation.

Microscopy

THE INVENTION OF THE OPTICAL MICROSCOPE led to anatomical investigations at the cellular level. Most of the sensory systems that we will study are in animals, vertebrate hearing, smell, and vision. The primary sensory receptors are specialized neurons for detecting sound, scent, and photons. The basic conceptual framework of neural circuit organization was largely established by Santiago Ramón y Cajal (1852-1934), the Spanish neuroscientist who was primarily a cellular-level anatomist (Fig. 25). Based on careful systematic analyses of sparsely labeled neurons in many brain tissues across animals and across developmental stages, Cajal identified basic principles about the organization of neural circuits in general and sensory circuits in particular. First, individual neurons interact with other neurons via contact or contiguity, synaptic contacts in today's language. Second, information flows with directionality through circuits, with dendrites and cell bodies on the input side and axons on the output side. These basic principles have been confirmed through anatomical investigations with higher spatial resolution using **electron microscopy**.

THE GOLGI STAINING METHOD was critical to Ramón y Cajal's success in resolving individual neurons. Neurons can be densely packed in brain tissue. Cell bodies are on the scale of micrometers. Synapses and nerve fiber thicknesses are on the scale of tens of nanometers. A theoretical limit to the resolution of any imaging system is the **Abbe diffraction limit**. Most optical systems, from light microscopes to electron microscopes to the human eye, suffer a spatial limit to resolution comparable to the wavelength due to diffraction. Even with the best compound microscopes, Cajal would not have been able to resolve neighboring structures closer than $\sim 300 - 500$ nm (corresponding to the wavelengths of visible light). But by individually labeling neurons that were separated by much greater distances, he could still perform exquisite anatomical reconstructions of each neuron, one cell at a time.

The Fly Eye

The wave nature of light is an important factor in determining the spatial limit of optical resolution, not just of microscopes but in vision. The "pixel sizes" of our retinas corresponds to the diameters of rod and cone photoreceptor cells, roughly 500 nm in the fovea where

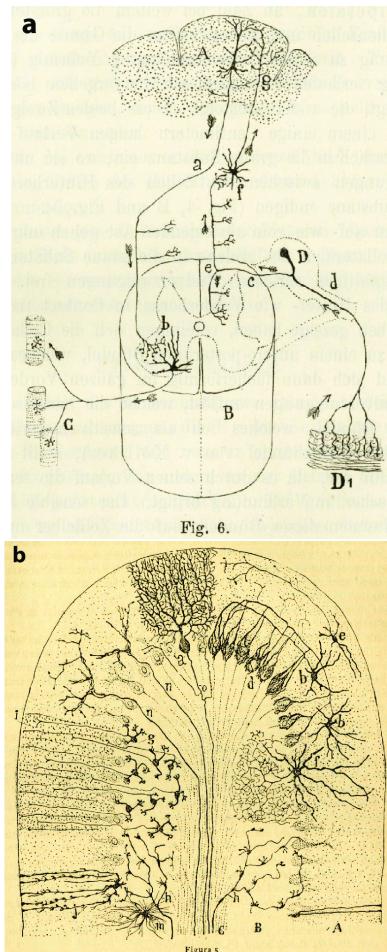


Figure 25: Cellular level anatomical analyses. (a) A nervous system-wide diagram of reflex and voluntary control of behavior by Cajal. Sensory information from the skin (D) is transmitted by dorsal root ganglion cells (d) to spinal cord (B) gray matter and to pyramidal neurons in the cerebral cortex (A), which in turn transmit impulses to motor neurons (b) in the spinal cord. For clarity, an interneuron between the spinal ending of (c) and an ipsilateral motoneuron are not shown. In this diagram, function (arrows) is predicted from structure. Neuron types in a gray matter region, the cerebellum. From Swanson and Lichtman (2016).

we obtain our highest spatial acuity. The visual angle of our highest spatial acuity, corresponding to the smallest letters on the eye chart, is about 1 minute of arc, spanning about 1-2 cone diameters in the fovea.

Insects have compound eyes, where each “pixel” of the field of view is provided by an ommatidium that points in a different direction, each gathering visual information from one angular field of view (Fig. 26). Feynman (the physicist) and Barlow (the neuroscientist) thought deeply about the anatomy of the compound eye. Take the insect eye to be a sphere of radius r divided into ommatidia (Fig. 27).

The larger the diameter of the ommatidium, the less angular resolution. This geometrical estimate of angular resolution is simply:

$$\Delta\theta_g = \frac{\delta}{r}$$

Shrinking δ would increase visual acuity, but diffraction puts a limit to how small δ can be. Light traveling through a thin slit will diffract. The thinner the slit, the wider the central peak of the diffracted wavefront (Fig. 28). Thus, if δ is too small, the ommatidium will “see” light at angles far from its axis:

$$\Delta\theta_d = \frac{\lambda}{\delta}$$

If δ is too large, enough ommatidia will see light because of poor angular resolution. So we adjust the δ to minimize poor angular resolution at excessively small or large values. We add the two effects and find the δ where the sum is minimized:

$$\frac{d(\Delta\theta_d + \Delta\theta_g)}{d\delta} = -\frac{\lambda}{\delta^2} + \frac{1}{r} = 0$$

This gives us an estimate for the optimum size of the ommatidium (Fig. 29):

$$\delta = \sqrt{\lambda r}$$

The diameter of the ommatidium should increase with the square root of the size of the eye. Barlow tested this idea by measuring insect eyes from 27 species of Hymenoptera (sawflies, wasps, bees, and ants), and thus experimentally verified his prediction.

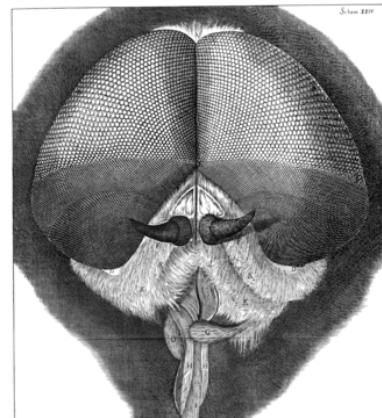


Figure 26: The compound eye.

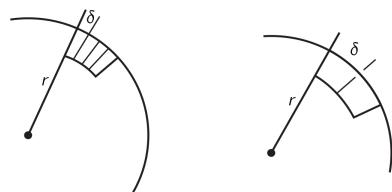


Figure 27: Schematic view of packing of ommatidia in insect eye.

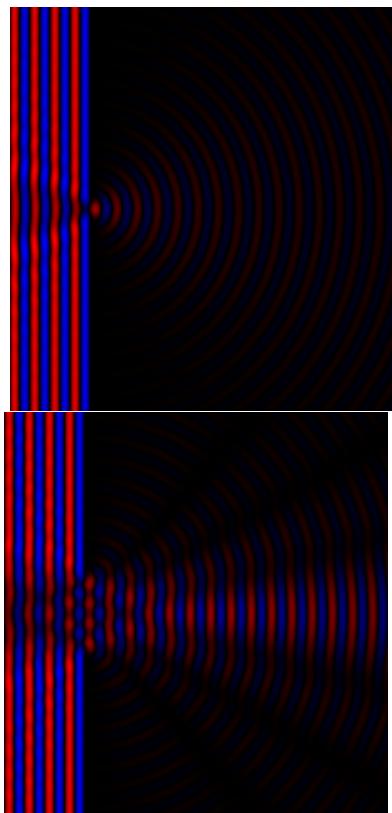


Figure 28: Diffraction.

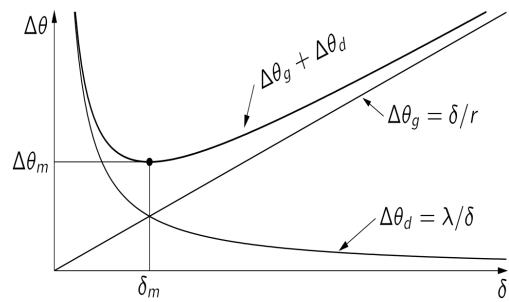


Figure 29: Optimum ommatidia diameters

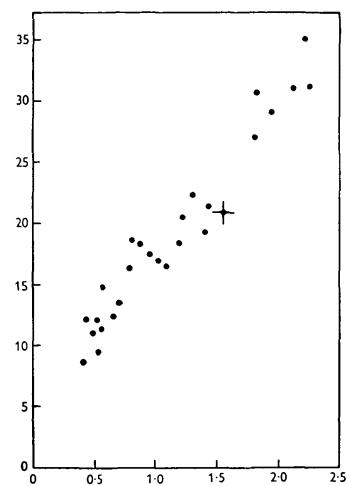


Figure 30: Actual ommatidia diameters
The ordinate axis is ommatidia diameter in micrometers. The abscissa is the square root of eye in millimeters.

ELECTRON MICROSCOPES have greater spatial resolution because the electron wavelength, defined by quantum mechanics by the de Broglie equation:

$$\lambda = \frac{h}{p}$$

where h is Planck's constant and $p = mv$ is momentum, is so much smaller in scanning or transmission electron microscopes. In typical microscopes, electron velocities reach 20%-70% the speed of light. The electron wavelength reaches ~ 12 picometers in a 10 kV SEM and ~ 2 picometers in a 200 kV TEM.

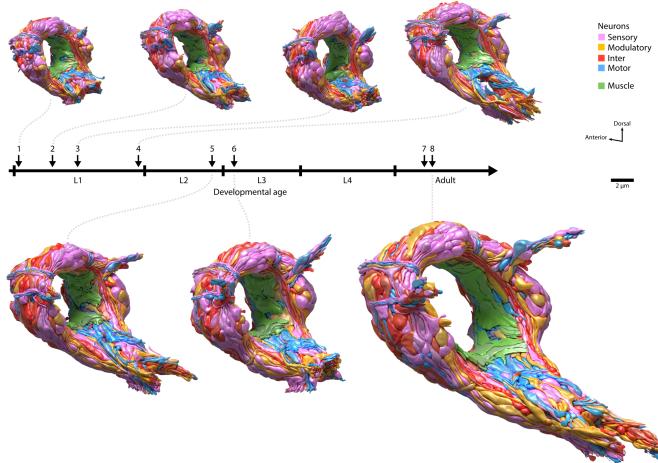


Figure 31: *C. elegans* *C. elegans* was the first animal to have its entire nervous system mapped by electron microscopy by John White and co-workers in the 1980s. More recently, eight *C. elegans* brains were reconstructed from birth to adulthood, mapping and comparing every chemical synapse and neuronal shape across whole-brain connectomes. From Witvliet et al. 2021

Electron microscopy imaging of brain sections has led to the characterization of synapses and neurons in diverse tissues. The main disadvantage of electron microscopy is that it requires ultrathin ($1\text{ }\mu\text{m}$) histological sections, which means that a single section is never adequate to generate full structural analysis of any cell. To overcome the problem of thin sections, scientists have used serial sectioning, in which each brain section is part of a sequence that transects a volume into hundreds or even tens of thousands of sections. Tracing objects from one section to the next reconstructs the geometry of the neurons and can also be used to identify the sites and cellular participants of each synapse. Connectomics has been used to reconstruct the entire nervous systems of small animals like nematodes and fruit flies (Fig. 31), and small parts of the nervous system of vertebrates (Fig. 32).

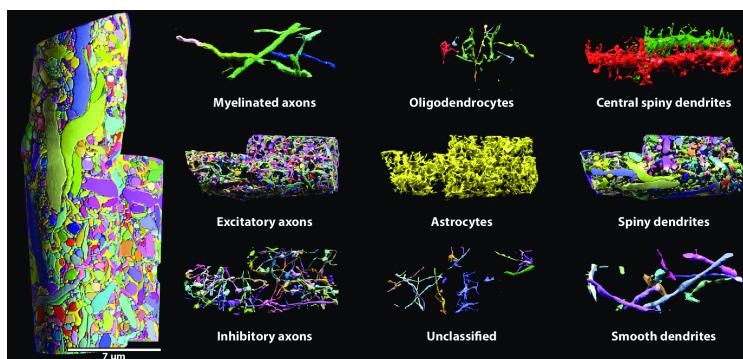


Figure 32: Reconstructions using modern, serial electron-microscopy connectomics approaches. On the far left is a fully reconstructed volume surrounding two apical dendrite segments of layer 5 pyramidal neurons in somatosensory cortex. Each colored object is a separate neuronal or glial cell process or extension. These processes are categorized in the images to the right. From Swanson and Lichtman, 2016.

ANATOMICAL ANALYSIS AT THE LEVEL OF INDIVIDUAL SENSORY RECEPTORS requires atomic resolution that greatly exceeds what is used for connectomics. X-ray crystallography has long been used to determine atomic-level structures of biomolecules. Crystallography only works for molecules that can be crystallized. Most sensory receptors are embedded in cell membranes, making them hard to crystallize in their native environments. Electron microscopy has atomic resolution, in principle, but sample degradation and poor contrast has long made it difficult to apply to molecules. Technical breakthroughs have led to cryo-electron microscopy (cryo-eM), which has allowed scientists to image many biomolecules for the first time, including membrane-bound, multimolecular sensory receptors. Cryo-EM, combined with the computer-assisted approach of obtaining numerous images of the molecule of interest in different orientations, followed by algorithmic reconstruction, enables 3D structures of biomolecules at Angstrom-level resolution, such as that of Piezo, a major mechanosensory channel in animals (Fig. 33).

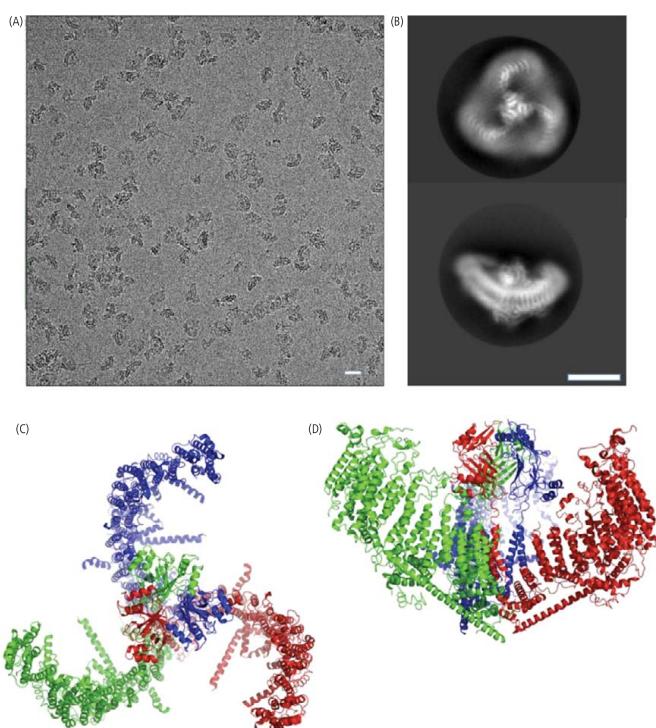


Figure 33: CryoEM of Piezo1 The gene for the Piezo1 protein was expressed in a cell line and purified. The protein was suspended on an electron-microscope grid and rapidly frozen by plunging the grid first into liquid ethane and then into liquid nitrogen. (A) Representative raw micrograph of the protein on the grid; scale bar is 20 nM. (B) Protein images like those in (A) were separated into groups according to their orientation, first manually to produce templates and then automatically under computer control. Images in each class were averaged, and representative averaged classes are shown viewed from the top (upper image) and side (lower image); scale bar is 100 Å. (C, D) Atomic model of the trimeric channel at an overall resolution of 3.7 Å shown as a ribbon diagram, viewed from the top (C) and side (D). The three subunits have been given different colors. From Fain, Chapter 1.

Sensory Physiology

ANIMAL SUBJECTS can be asked whether they perceived a stimulus, whether directly (if a human) or by experimental design (quantitative analysis of behavioral responses in monkey, mouse, or smaller organisms). Knowing whether internal mechanisms – molecules, neurons, or brain circuits – perceived a stimulus requires *physiological analysis*. We will mostly discuss perception in animals, where sensory mechanisms are located in neurons. Neuronal activity can be measured in terms of electrical signals that travel along their ‘wires’ in a way that is mediated and amplified by ion channels in their cell membranes. Sensory cells have specialized ion channels that are directly (in the case of ionotropic receptors) or indirectly (in the case of metabotropic receptors) activated by the environmental stimulus itself.

E. D. Adrian recorded some of the first *extracellular recordings* from sensory neurons by placing the axons of touch receptor cells near wire electrodes. Skin pressure changed the frequency of *action potentials* in different ways depending on the stimulus. Thus, action potential firing patterns are a mechanism for encoding and communicating sensory information to the brain. The first photosensory responses were made by Hartline in the horseshoe crab, where action potential patterns also depended on the intensity and duration of light stimulation. Patterns of neuronal activity in sensory neurons encode the incoming stimulus.

THE BRAIN SOLVES AN INVERSE PROBLEM. It must reconstruct a past incoming stimulus based on the present pattern of neuronal activity. As before, when we considered the statistics of stimulus and response at the level of organism behavior, Bayesian analysis is needed to think about the statistics of stimulus and response at the level of neuronal activity. Because neuron activity patterns can be intrinsically randomness, a physiologist must measure the *probability* of a specific neuronal response that might be conditioned on a specific stimulus: $P(\text{response} \mid \text{stimulus})$. The neuronal activity pattern is what the brain “knows” about the external environment. The animal must then perform a Bayesian inference, estimating the probability that a specific stimulus occurred in the external world conditioned on a specific neuronal response in its brain: $P(\text{stimulus} \mid \text{response})$.

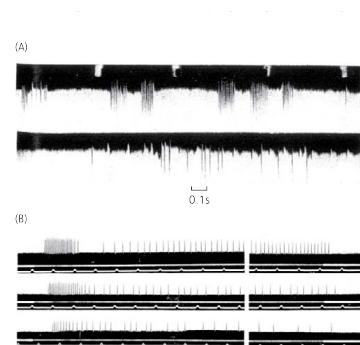
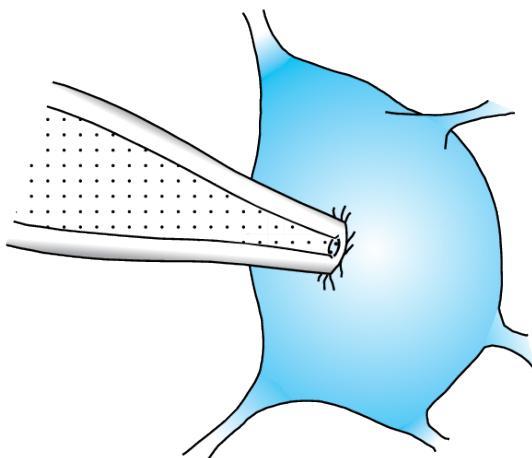


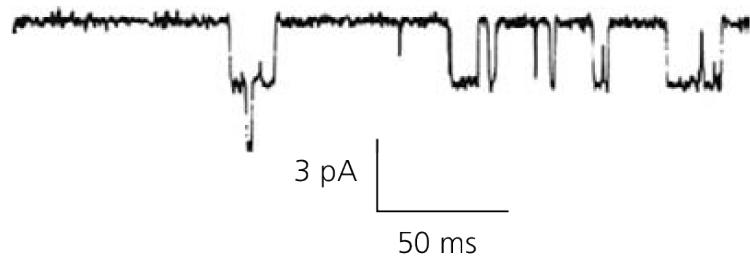
Figure 34: **Early electrical recordings of sensory responses.** (A) Action potentials recorded from single axons dissected from the cutaneous nerve of a frog. (B) Action potentials from the lateral eye of the horseshoe crab *Limulus*. Each trace gives the response to a different light intensity, which was systematically increased by an additional factor of ten from dimmest (bottom) to brightest (top). From Fain, Chapter One.

INTRACELLULAR RECORDINGS made it possible to record the detailed electrical responses of single neurons (Fig. 35). Electrodes are inserted into glass tubes that are melted and pulled to a fine point and filled with salt solution. Sharp electrode recordings can be made by penetrating individual cells. *Patch electrode recordings* can be made by polishing the tip of the glass electrode to be very smooth, such that when it is pressed against a cell membrane and a slight suction is applied, a very tight seal is formed, sometimes called a gigaseal with typical electrical resistances of $10\text{-}100\text{ G}\Omega$. High seal resistance ensures that most of the electrical current in each recording travels through the cell membrane that is being tested, not leaking through the seal.

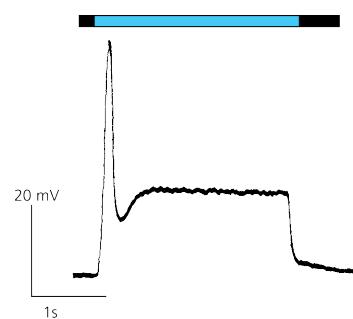
(A)



(B)



(A)



(B)

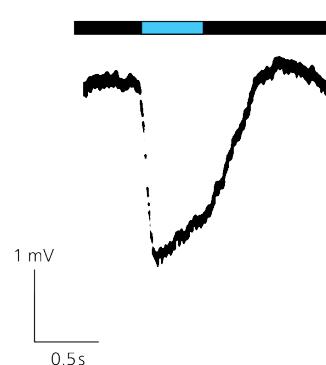


Figure 35: Intracellular recordings from sensory receptors. Bars above recordings show timing and duration of light flashes. (A) Depolarizing voltage response from photoreceptor of *Limulus* ventral eye. (B) Hyperpolarizing voltage response from photoreceptor (cone) of a fish. This is the first published recording of the response of a vertebrate photoreceptor. From Fain, Chapter One.

Figure 36: Patch-clamp recording from single channels. (A) The tip of a patch pipette is pushed against the cell body of a cell and slight suction is applied to form a seal. (B) Single-channel currents recorded from muscle acetylcholine receptors. The pipette contained $0.3\text{ }\mu\text{M}$ acetylcholine. Downward deflections indicate channel opening. At least two channels were present in this membrane patch. From Fain, Chapter One.

Molecular Biology

THE REVOLUTION IN MOLECULAR BIOLOGY deepened the analysis of sensory mechanisms (Fig. 37). Most sensory receptors are integral membrane proteins. Most known receptors were identified either by protein purification and sequencing, or by genetic analysis (finding a mutant that lacks a sensory modality, and working out the missing gene that was responsible for phenotype). Modern techniques to find sensory receptors include single-cell transcriptional and whole-genome sequencing and analysis. The genetic sequence of a putative receptor has characteristics that can betray its identity. Integral membrane proteins must have extensive sequences within the hydrophobic interior of the lipid bilayer. From amino acid sequences, which amino acids lie within the membrane and which face the cytoplasmic or extracellular solution can be inferred. Some amino acids (such as valine and isoleucine) are hydrophobic. Some amino acids (such as aspartate and lysine) are hydrophilic. Hydropathy analysis can be used to estimate the rough structure of putative receptors.

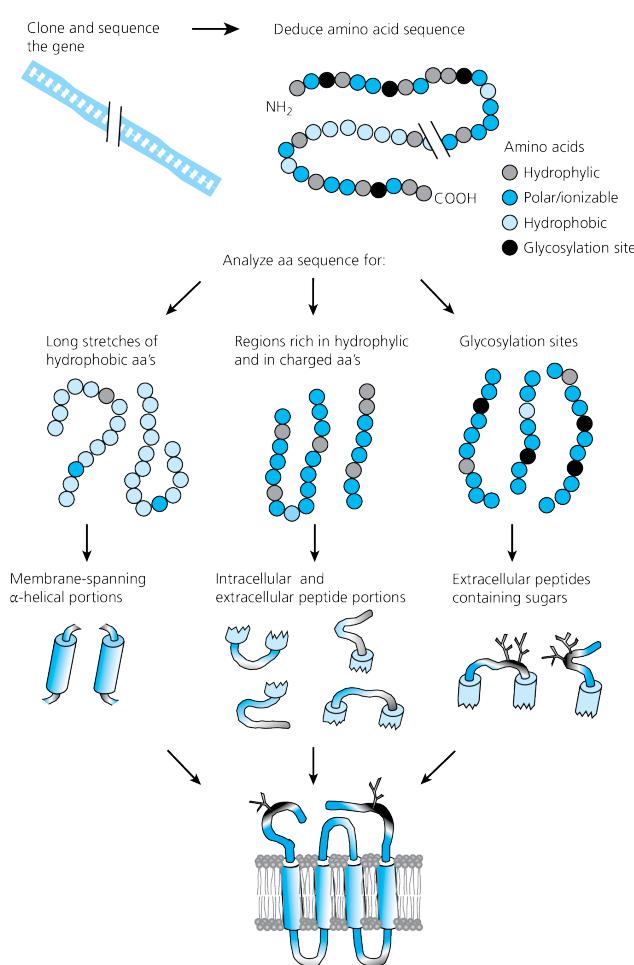


Figure 37: Analysis of hydropathy and the folding of membrane proteins. The amino acid sequence of a membrane protein can be used to make inferences about protein structure. From Fain Chapter One.

RECOMMENDED READING

- Eric R. Kandel, James H. Schwartz, and Thomas M. Jessell, editors. *Principles of Neural Science*. Elsevier, New York, third edition, 1991 [Download](#)
- Chapter One. Gordon L Fain. *Sensory Transduction*. Sinauer Associates, Sunderland, Mass., 2003. ISBN 0878931716 [Download paper](#)
- Larry W. Swanson and Jeff W. Lichtman. From Cajal to Connectome and Beyond. *Annual Review of Neuroscience*, 39(1):197–216, 2016 [Download](#)

ADDITIONAL READING

- Barlow's explanation of ommatidial size across insects.
 - H. B. Barlow. The size of ommatidia in apposition eyes. *Journal of Experimental Biology*, 29(4):667–674, 1952. ISSN 0022-0949 [Download](#)
- This paper describes the modern use of high-throughput connectomics to measure the developmental dynamics of an entire animal brain from birth to adulthood.
 - Daniel Witvliet, Ben Mulcahy, James K. Mitchell, Yaron Meirovitch, Daniel R. Berger, Yuelong Wu, Yufang Liu, Wan Xian Koh, Rajeev Parvathala, Douglas Holmyard, Richard L. Schalek, Nir Shavit, Andrew D. Chisholm, Jeff W. Lichtman, Aravinthan D. T. Samuel, and Mei Zhen. Connectomes across development reveal principles of brain maturation. *Nature*, 596(7871):257–261, 2021. ISSN 0028-0836 [Download](#)
- The discovery of the Piezo mechanosensory channels involved a remarkable integration of modern techniques in molecular biology and electrophysiology to yield one of the most elusive sets of proteins in sensory perception.
 - Bertrand Coste, Jayanti Mathur, Manuela Schmidt, Taryn J Earley, Sanjeev Ranade, Matt J Petrus, Adrienne E Dubin, and Ardem Patapoutian. Piezo1 and piezo2 are essential components of distinct mechanically activated cation channels. *Science*, 330(6000):55–60, 2010 [Download](#)

SOME STATISTICAL MECHANICS

Rhodopsin

RHODOPSIN, also called visual purple, is the light-sensitive receptor protein in the retina (Fig. 39). Rhodopsin is also a membrane-bound photoswitchable G-protein-coupled receptor (GPCR) with both a protein component and covalently-bound cofactor, a photoreactive chromophore called *retinal*. Isomerization of 11-cis-retinal into all-trans-retinal by light triggers a conformational change that causes rhodopsin to activate another G protein called *transducin*. Transducin activation begins a signal transduction cascade that eventually modulates cyclic guanosine monophosphate (cGMP) levels. Changing cGMP levels change the electrical activity of the rod cell through cGMP-gated ion channels in the cell membrane of the photoreceptor cell. Changes in the electrical activity of the photoreceptor cell membrane lead to changes in their chemical synaptic outputs that are communicated to the rest of the retina as visual events.

THE ENERGY of one blue-green photon ($\lambda=500 \text{ nm}$) is $4 \times 10^{-19} \text{ J}$. Because rhodopsin is in thermal equilibrium with its environment ($\sim 25^\circ\text{C}$), the chromophore will have, on average, an amount of thermal energy around $4 \times 10^{-21} \text{ J}$. The energy of a blue photon is much higher than thermal energy. On average, the thermal energy fluctuations of the chromophore can be expected to be much lower than a threshold needed for its activation. But how frequently will the chromophore exceed the threshold of activation by thermal fluctuations alone? Answering this question requires knowing about probability distribution of energy levels of an object in thermal equilibrium, not just its mean thermal energy.

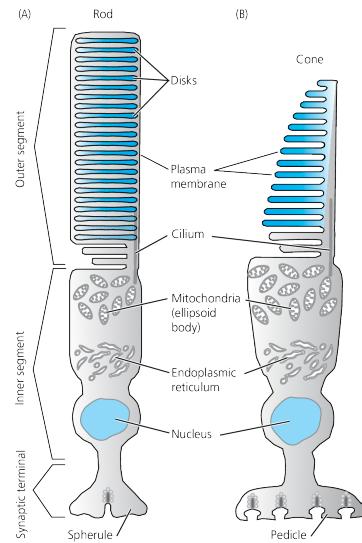


Figure 39: Rods and Cones.

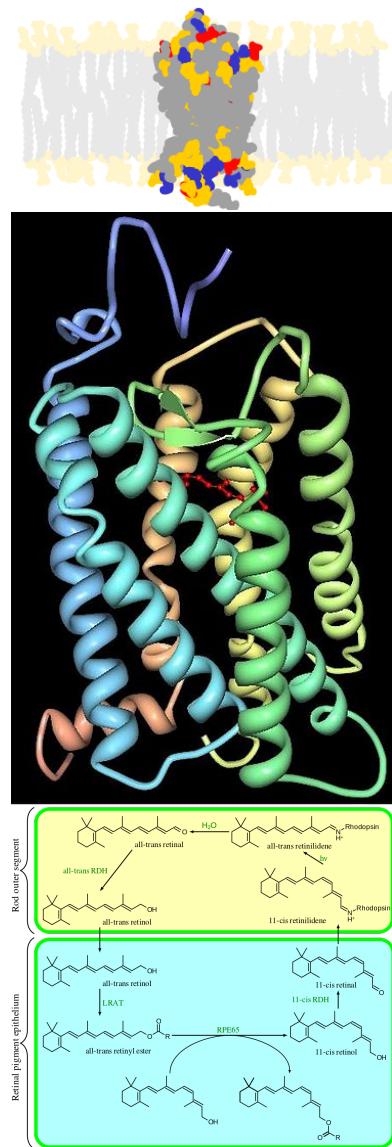


Figure 38: Rhodopsin. Three dimensional structure of bovine rhodopsin, a membrane bound protein. The chromophore, retinal, is embedded within the protein. Rhodopsin is embedded in detached discs in the outer segment of the rod photoreceptor cell and embedded in infolded lamellae in cones. Absorption of light energy, $h\nu$, causes a conformational change – 11-cis-retinal becomes all-trans-retinal – which thereby alters protein structure to activate G-protein-coupled signal transduction. Biochemistry in the rod cell restores the cis configuration.

The Boltzmann Distribution

IN STATISTICAL MECHANICS, a Boltzmann distribution is a probability distribution that gives the probability that a system will be in a certain state as a function of state energy and temperature:

$$p_i \propto e^{-\varepsilon_i/(kT)}$$

where p_i is the probability of the system being in state i , ε_i is the energy of that state, and a constant kT of the Boltzmann constant and temperature T . The symbol \propto denotes proportionality.

System can have broad meaning; an macroscopic ensemble of components (a molecule) or a single atom. What matters is that the system has a *degree of freedom* that allows it to enter different measurable states, and that energy is freely exchanged into and out of the system. An exponential distribution means that states with lower energy have a higher probability of being occupied (Fig. 40).

THE RATIO OF PROBABILITIES of two states is the Boltzmann factor and depends on their energy difference:

$$\frac{p_i}{p_j} = e^{(\varepsilon_j - \varepsilon_i)/(kT)}$$

If a system has M possible states, and the sum of the probabilities of being in each state is normalized, $\sum_{j=1}^M p_i = 1$, one can compute the probability of being in each state:

$$p_i = \frac{1}{Q} e^{-\varepsilon_i/(kT)} = \frac{e^{-\varepsilon_i/(kT)}}{\sum_{j=1}^M e^{-\varepsilon_j/(kT)}}$$

THE EXPONENTIAL ATMOSPHERE is a classic example of a continuous Boltzmann distribution (Fig. 41). The altitude of gas molecules in Earth's atmosphere is one spatial degree of freedom. The altitude of each gas molecule is associated with a gravitational potential energy: $E = mgh$. Assuming a uniform atmospheric temperature T , the probability that a molecule is at height h is:

$$p(h) \propto e^{-mgh/kT}$$

For this probability density to be properly normalized:

$$p(h) = \frac{e^{-mgh/kT}}{\int_{h=0}^{\infty} e^{-mgh/kT} dh}$$

The expectation value for the altitude of a gas molecule depends on its mass: $\langle h \rangle = \frac{kT}{mg}$.

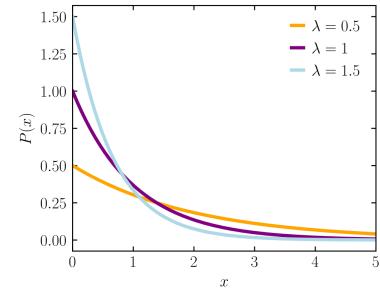


Figure 40: **Exponential Distributions.** Plot of the probability density function of the exponential distribution ($P(x) = \frac{e^{-\lambda x}}{\lambda}$) for rates $\lambda = 0.5, 1$ or 1.5 .

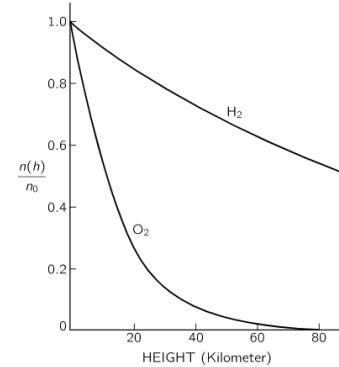


Figure 41: **Exponential atmosphere.** From Feynman's *Lectures in Physics*. The mass of one molecule of O_2 is 5×10^{-23} g. The mass of one molecule of H_2 is 3×10^{-24} g. Does Feynman's drawing make sense?

THE BOLTZMANN DISTRIBUTION governs the likelihood that a sensory receptor like rhodopsin has different states with different thermal energies. We describe the different conformational states of a molecule using a “reaction coordinate”, and assign an energy to each point along the reaction coordinate (Fig. 42).

Using Boltzmann factors, we can infer the relative likelihood of being in different states. Going from one stable state A to another state B might require transit through an activation state. The energy of this activation state creates a barrier that slows the reaction. At equilibrium, the relative likelihood of any two states is:

$$\frac{p_A}{p_B} = e^{-\frac{\Delta E}{kT}}$$

To estimate the reaction rate from A to B , we need the relative likelihood of being at the top of the activation barrier compared to state A . This estimates the fraction of particles in state A that are able to move over the barrier thanks to fluctuations in thermal energy:

$$k_{A \rightarrow B} \propto e^{-\frac{E_{act}}{k_B T}}$$

The Boltzmann distribution gives us insight into the fluctuating energies of sensory receptors before the arrival of stimulus energy, as well as the speed of the reactions that characterize sensory transduction.

Deriving the Boltzmann distribution by counting

Say that you have N particles that are given a total amount of energy E . These particles are able to freely exchange energy among them, but are constrained to quantal energy levels. How many particles can you expect to find at each energy level? In other words, what is the probability of a particle having a certain amount of energy? To find the distribution of the particles over their possible energy states, we enumerate the states ($s = 1, 2, \dots$), associate each state with a discrete energy ($\varepsilon_1, \varepsilon_2, \dots$), and assign a number of particles to each state (n_1, n_2, \dots). The Boltzmann distribution should tell us how many particles, n_s , of the N total particles that we can expect to find in the s state with energy ε_s .

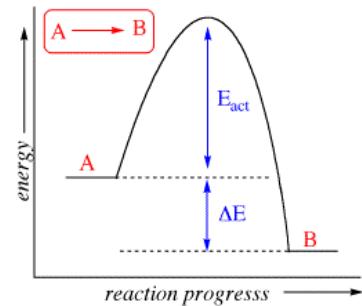


Figure 42: Reaction coordinate. From <http://butane.chem.uiuc.edu/pshapley/genchem2>

State	Energy	Number
1	ε_1	n_1
2	ε_2	n_2
3	ε_3	n_3
.	.	.
.	.	.
s	ε_s	n_s
.	.	.
.	.	.

Mass Conservation:
 $\sum_s n_s = N$.

Energy Conservation:
 $\sum_s n_s \varepsilon_s = E$.

The probability of any particle being in a given state s is n_s/N . Thus, the average energy of each particle is:

$$\langle \epsilon \rangle = \frac{\sum_s n_s \epsilon_s}{\sum_s n_s}$$

What is the probability of observing a specific distribution with say n_1 particles in state 1, n_2 in state 2, and so on? Our central assumption is that every possible distinct arrangement that satisfies conservation laws of mass and energy is equally possible. The probability of observing a particular distribution of n_s is thus proportional to the number of distinct arrangements that can be achieved with the N particles. The number of distinct particle arrangements that corresponds to the same distribution of particles among energy levels is a measure of the probability of that distribution. The number of ways a given distribution can be formed is a combinatorial problem:

$$W = \frac{N!}{n_1! n_2! n_3! \dots}$$

To find the distribution where W is largest, we need to maximize W with respect to n_s and with respect to the conservation laws. To do this, we need a convenient analytical expression for the factorial problem and we need to use Lagrange Multipliers.

We choose to maximize $\log W$ subject to the constraints of conservation of energy and mass. Thus,

$$d \left(\log W - \alpha \sum_s n_s - \beta \sum_s \epsilon_s n_s \right) = 0$$

The α and β are the Lagrange multipliers. Varying with respect to n_s and incorporating Stirling's Formula gives:

$$-\sum_s d n_s (\log n_s + \alpha + \beta \epsilon_s) = 0$$

Because this must hold true for every δn_s , every term in the sum must vanish. The values of n_s which do this are

$$\log n_s + \alpha + \beta \epsilon_s = 0$$

We can thus conclude that the occupancy of state s depends exponentially on its energy:

$$n_s \propto e^{-\beta \epsilon_s}$$

Actually, we have shown that the exponential dependence of state occupancy on energy is true only at the most probable W . We have not shown that every other arrangement can be ignored. The probability

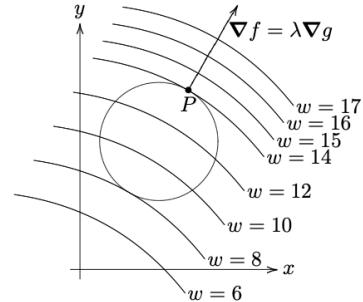


Figure 43: **Boltzmann's Tomb.** When you take Statistical Mechanics you will learn that $\log W = S/k_B$, the equation for entropy etched on his tomb.

Lagrange multipliers

Here, we sketch a geometric proof that builds intuition without worrying about rigor. For the function $w = f(x, y, z)$ constrained by $g(x, y, z) = c$, the maxima and minima are those points where ∇f is parallel to ∇g :

$$\nabla f - \lambda \nabla g = 0$$



For concreteness, we've drawn the constraint curve, $g(x, y) = c$, as a circle and some level curves for $w = f(x, y) = c$ with explicit (made up) values. Geometrically, we are looking for the point on the circle where w takes its maximum or minimum values.

Start at the level curve with $w = 17$, which has no points on the circle. Clearly, the maximum value of w on the constraint circle is less than 17. Move down the level curves until they first touch the circle when $w = 14$. Call the point where they first touch P. It is clear that P gives a local maximum for w on $g = c$, because if you move away from P in either direction on the circle you'll be on a level curve with a smaller value.

Since the circle is a level curve for g , we know ∇g is perpendicular to it. We also know ∇f is perpendicular to the level curve $w = 14$, since the curves themselves are tangent, these two gradients must be parallel. Q.E.D.

of alternative arrangements becomes negligible when N is very large, but this is hard to show and we will skip it.

We also have not shown that β , introduced here as a Lagrange multiplier, is related to temperature. For our purposes, we define $\beta = 1/k_B T$. When you take thermodynamics, you will learn that β behaves like $1/k_B T$ for various thermodynamic relationships and can have no other meaning. We are going to declare victory with the result that energy levels are exponentially distributed at thermal equilibrium.

Deriving the Boltzmann distribution Using Information Theory

CLAUDE SHANNON invented *information theory* starting with his own notion of the *entropy* of a probability distribution. Let X be a discrete random variable that can have different possible values, x . The probability density function, $p(x)$, is the likelihood of having different values.

SHANNON DEFINED HIS ENTROPY $H(x)$ of a discrete random variable X as:

$$H(X) = - \sum_x p(x) \log_2 p(x)$$

Say the random variable is the outcome of tossing a fair coin. X can either be heads ($p = 0.5$) or tails ($p = 0.5$). In this case, $H = 1$. Shannon's entropy is the amount of information in *bits* that you need to characterize the outcome of the toss.

What if the coin was biased and always came up heads? In this case, $H = 0$. You don't need any information to characterize the outcome of the coin toss (i.e., 0 bits), because the outcome is guaranteed. What if we vary p ? The entropy is minimum at $p = 0$ or $p = 1$ and maximum at $p = 0.5$ (Figure 44).

Roll a fair die. In this case, every outcome has $p = 1/6$ and $H = \log_2 6 \approx 2.5$. In Dungeons and Dragons, we have 8-sided die, and the number of bits needed to characterize one roll is 3.

The maximum entropy distribution is the one that minimizes the amount of prior information that is built into the distribution. If you know nothing about a probability distribution except the range of outcomes or the mean of the distribution, these constraints can be used to calculate the distribution that maximizes entropy using methods like Lagrange multipliers. This does not mean that the maximum entropy distribution is the right distribution, but it is a safe place to start. If the maximum entropy distribution isn't the distribution that correctly describes the system, then you lack prior

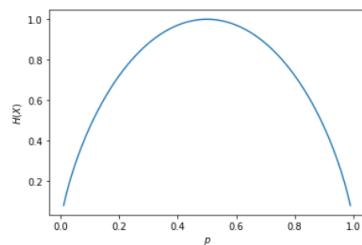


Figure 44: Entropy of a coin toss

and salient constraints about the system that would lead to different maximum entropy distribution.

Reconsider the problem of N particles exchanging a total amount of energy E among them. Each particle is constrained to discrete energy levels $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_s$

What is the probability distribution that governs the energies of each particle? This problem is fully described. We are not missing any salient facts needed to calculate a distribution. The correct probability distribution must be the one that maximizes entropy. We have phrased another Lagrange multiplier problem:

$$\delta \left(H - \alpha \sum_s p_s - \beta \sum_s \varepsilon_s p_s \right) = 0$$

We conclude that $p_s \propto e^{-\varepsilon_s \beta}$. If the distribution were anything else, we would have required additional salient constraints on the physical problem. But because we set up the problem without any other constraints, only the exponential distribution of energies, the Boltzmann distribution, is possible. *Q.E.D.*

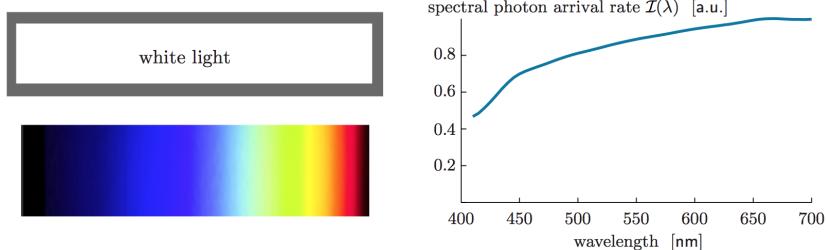
REFERENCES

- This is Shannon's original paper where he invented Information Theory.
 - C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(4):623–656, 1948. ISSN 0005-8580
[Download](#)
- Mehran Kardar. *Statistical physics of particles*. Cambridge University Press, Cambridge ; New York, 2007. ISBN 9780521873420

COLOR VISION AND THE DIMENSIONALITY OF PERCEPTION

VERTEBRATES HAVE TWO TYPES OF PHOTORECEPTORS, rods and cones. Our human eyes contain ~ 130 million rod cells for scotopic (low light) vision and ~ 7 million cone cells for color vision. All rod cells have one peak wavelength sensitivity, whereas our three cone sub-types are tuned to long, medium, and short wavelengths (Fig. 50). Thus, rod vision is monochromatic and cone vision is trichromatic. Trichromacy means that any color that we can see can be built from three different wavelengths of visible light such as red/green/blue or cyan/magenta/yellow. Thus, the stimulus space of human color vision has three dimensions. Let's make the idea of stimulus dimensionality more precise.

Human color vision provided early clues about the physical nature of light. Newton used a prism to separate white light into a spectrum of different colors (Fig. 46). He used a second prism to recombine spectrally-separated light to recreate white light. Somehow, light of one color (white) can be formed by summing light from different colors.



In principle, electromagnetic radiation of one wavelength can carry information that can be completely separated from electromagnetic radiation of another wavelength. This is how frequency-modulated (FM) radio transmission works: each radio station sends information in a different electromagnetic 'color'. Stimulus information that is carried on two different orthogonal axes (electromagnetic signals with different wavelengths) correspond to a 2-dimensional stimulus space. One can imagine a visual system that more fully exploits the human visual spectrum – from 400 nm (blue) to 700 nm (red) – by disentangling the visual information that is separately carried on a very large number of different wavelengths. But to read N -dimensional color information that is carried on N different wavelengths, the receiver needs at least N receptors that are separately tuned to different wavelengths. With only three cone sub-types, our

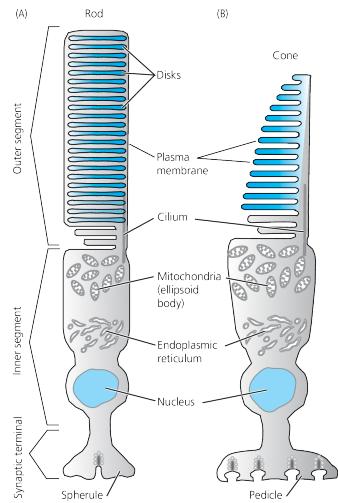


Figure 45: Rods and Cones.

Figure 46: Sunlight consists of a broad spectrum of photon energies. When passed through a prism, it separates into a continuous distribution of colors. We can represent the spectrum by a photon arrival rate function that is nearly constant over the range of visible light.

trichromatic retinas project all color information that might be carried by many different wavelengths onto a three-dimensional space. We discard any color information that might be contained in a high-dimensional stimulus space onto our much smaller three-dimensional internal representations of color. Most animals have fewer dimensions of color perception than we do (horses have two cone types, corresponding to dichromatic vision) but some more (mantis shrimps have 16 cone types, and might have the greatest power of color discrimination on the planet).

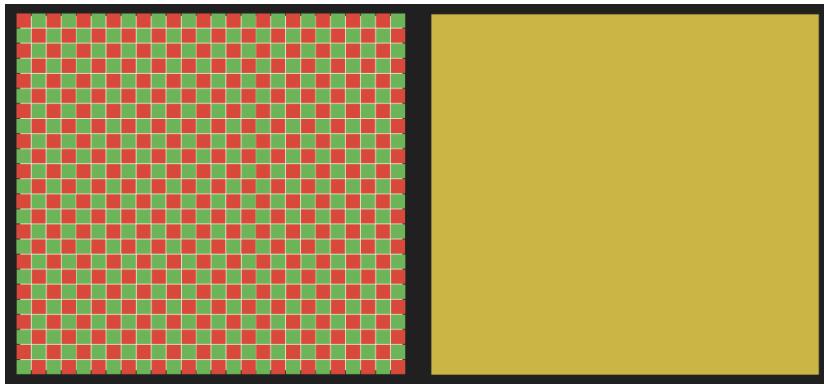


Figure 47: **Color illusion.** When viewed up close, the left box is seen to consist of small red and green squares. When viewed from afar, each photoreceptor receives light from both red and green squares, whose spectra merge. The resulting percept is closer to the color in the right box than red or green. Red + Green \sim Yellow.

We intuitively know that trichromatic vision discards spectral information. A mixture of spectrally pure red and green light looks yellow (Fig. 47). *Metamers* refer to physically different light spectra that lead to the same color perception. Long before biologists knew the molecular and cellular basis of color perception – the existence of three cone cells that express three different rhodopsin molecules – optical scientists used perceptual ‘color matching experiments’ to postulate trichromacy.

In color matching experiments, different sets of *basis lights* to illuminate a viewing screen (Fig. 48). Each basis light corresponds to a distinct spectrum of light wavelengths. Scientists discovered that only three basis lights, when varied in relative intensity, were sufficient to create a perceptual match to the full range of human color perception from red to violet. Moreover, three basis lights with different spectra can be chosen, as long as three spectra are different from one another and fall within the spectrum of visible light. The practical consequence of trichromatic vision when building computer or television screens is that the full gamut of color vision can be achieved with only three types of pixels (red, green, and blue). By mixing a relatively small palette of primary colors, a painter can endlessly create different hues. Chemical analysis of all of Vermeer’s paintings has revealed only twenty distinct pigments.

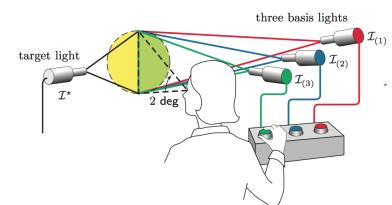


Figure 48: **Color matching.** A target light is projected onto the left half of a screen. A subject attempts to obtain a perceptual match by adjusting the intensities of three otherwise fixed basis lights that converge on the right half of the screen.

A quantitative model of color matching.

Suppose that we choose three basis lights with different spectra, meaning that they have different photon arrival rates as a function of wavelength: $\mathcal{I}_1(\lambda)$, $\mathcal{I}_2(\lambda)$, and $\mathcal{I}_3(\lambda)$. The units of these spectral light intensities are photons $s^{-1} nm^{-1}$, because light energy is distributed over time and over different wavelengths. To calculate the total photon arrival rate for each basis light (Φ_i), one must integrate over all wavelengths:

$$\Phi_i = \int \mathcal{I}_i(\lambda) d\lambda$$

In a color matching experiment, we must deliver these basis lights with adjustable intensities. To do this, we might fix the mean photon delivery rate of all basis lights to one baseline value, $\Phi = \Phi_1 = \Phi_2 = \Phi_3$, and use an adjustable scaling factor to individually change the intensities of each light: ζ_1 , ζ_2 , and ζ_3 . We give a human subject knobs to adjust the three ζ scaling factors to create a perceptual match to a monochromatic target (Fig. 48).

For simplicity, we can choose this monochromatic target to have a spectrum that is sharply peaked at one wavelength λ^* and fix the photon arrival rate at the same baseline as all basis lights, Φ . In the experiment corresponding to Fig. 49, the selected basis lights were also sharply peaked at single wavelengths – $\lambda_1=645$ nm, $\lambda_2=526$ nm, and $\lambda_3=444$ nm. In this experiment, a different combination of ζ_1 , ζ_2 , and ζ_3 were needed to create an experimentally observed perceptual match to the target light. Different target lights were chosen throughout the visible spectrum, each demanding a different set of ζ_1 , ζ_2 , and ζ_3 to create perceptual matches.

An interesting and subtle point for the experiment shown in Fig. 49 is the discovery of a range of target wavelengths that *cannot* be matched with three positive ζ values. There is a range of target wavelengths that require *negative* values for ζ_1 . Since negative light intensities are physically meaningless, how were experimental measurements made in this seemingly impossible range? When red basis light with negative intensity was needed to match the target wavelength, the experimenters simply added red basis light to the target light. Then, instead of matching the target light to the sum of red, green, and blue basis lights, the sum of the target and red basis light was matched to the sum of the green and blue basis lights. This works in a linear model of color perception, where perception follows from the simply addition of light at all wavelengths. If the stimulus is not too strong (such that cones are saturated), a linear model can be sufficient and is usually a good place to start.

In any case, every color in the spectrum of human perception can

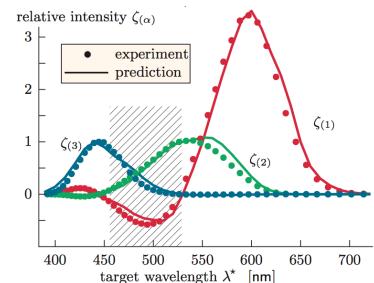


Figure 49: **Quantitative model of color matching.** Dots: experimental measurement of the relative intensities of basis lights needed to replicate target monochromatic wavelengths λ^* and results predicted from the measured spectral sensitivities of human cone cells using three basis lights with fixed wavelengths $\lambda_1=645$ nm, $\lambda_2=526$ nm, and $\lambda_3=444$ nm. It is impossible to create perceptual matches for monochromatic target lights in the hatched region that require negative values of long wavelength illumination (\mathcal{I}_1) in a linear model of color matching.

be represented with a stimulus space where every color is uniquely specified with only three free variables. A three-dimensional stimulus space saturates the full range of human color perception. .

The three-dimensional space of color representation.

The discovery that using three stimulus dimensions is sufficient to match any color led to the prediction that our internal representation of color also uses three dimensions. A three-dimensional internal representation of color would be formed at the beginning of ‘seeing’ if we used three different classes of photoreceptor for color – each with its own spectral sensitivity – to capture visual images. Now that we know that humans have three types of cone photoreceptor, we know that this is true, but let us build a quantitative model. First, we assign a spectral sensitivity function to all neurons of each photoreceptor class. This spectral sensitivity function is proportional to the probability that a photon of a given wavelength triggers the activity of the corresponding photoreceptor. The activity of each photoreceptor effectively counts all incoming and absorbed photons. But once a photon is counted, its color information (i.e., wavelength) is discarded. The rhodopsin molecule that absorbed the photon undergoes a binary change, from inactive to active, but no longer encodes the specific wavelength of the photon that elicited this change.

The three photoreceptor classes can be ordered by peak wavelength of their corresponding sensitivity functions (S, short, blue; M, medium, green; L, long, red): $S_L(\lambda)$, $S_M(\lambda)$, and $S_S(\lambda)$. Although the color of incoming photons is lost in the activity of an individual photoreceptor, color information is preserved and encoded in the *relative* activities of the three photoreceptor types.

In each small wavelength range, $\Delta\lambda$, photons arrive at the retina with a mean rate $\mathcal{I}\Delta\lambda$. Photons stimulate the activity of each photoreceptor in proportion to the rate of photon arrival and its sensitivity function: $S_i\mathcal{I}\Delta\lambda$. Collecting all wavelengths of light, the total number of photon absorptions by each photoreceptor is:

$$\begin{aligned}\beta_S &= \int d\lambda S_S(\lambda) \mathcal{I} \\ \beta_M &= \int d\lambda S_M(\lambda) \mathcal{I} \\ \beta_L &= \int d\lambda S_L(\lambda) \mathcal{I}\end{aligned}$$

The brain concludes that two colors match when their spectra, or spectral arrival rates, cannot be distinguished – $\mathcal{I} \sim \mathcal{I}'$). The brain only has access to the activity of its three photoreceptor classes. Thus, the brain would conclude that two colors match when both colors

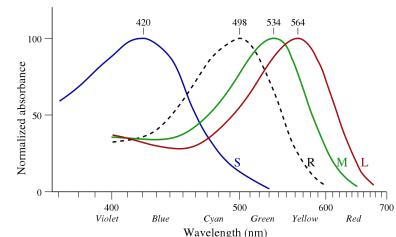


Figure 50: **Color vision. Normalized human photoreceptor absorbances for different wavelengths of light.**

evoke the same photon absorption rates in the three photoreceptors – $\beta_L, \beta_M, \beta_S$.

Given a target light spectrum (with intensity \mathcal{I}^*), how do we adjust the relative amounts of three basis lights with spectral intensities \mathcal{I}_i by setting values of ζ_i . What is the amount of each basis light in a sum that matches a target light when measured by photoreceptor activation? We need to solve the following equation for ζ_i :

$$\mathcal{I}^* \sim \zeta_1 \mathcal{I}_1 + \zeta_2 \mathcal{I}_2 + \zeta_3 \mathcal{I}_3$$

For the target light, the number of photon absorptions by the i th photoreceptor is by:

$$\begin{aligned}\beta_S^* &= \int d\lambda S_S(\lambda) \mathcal{I}^*(\lambda) \\ \beta_M^* &= \int d\lambda S_M(\lambda) \mathcal{I}^*(\lambda) \\ \beta_L^* &= \int d\lambda S_L(\lambda) \mathcal{I}^*(\lambda)\end{aligned}$$

Define $3 \times 3 = 9$ numbers, that describe the photon absorptions rates in each photoreceptor (S,M,L) evoked by each basis light ($i = 1, 2, 3$):

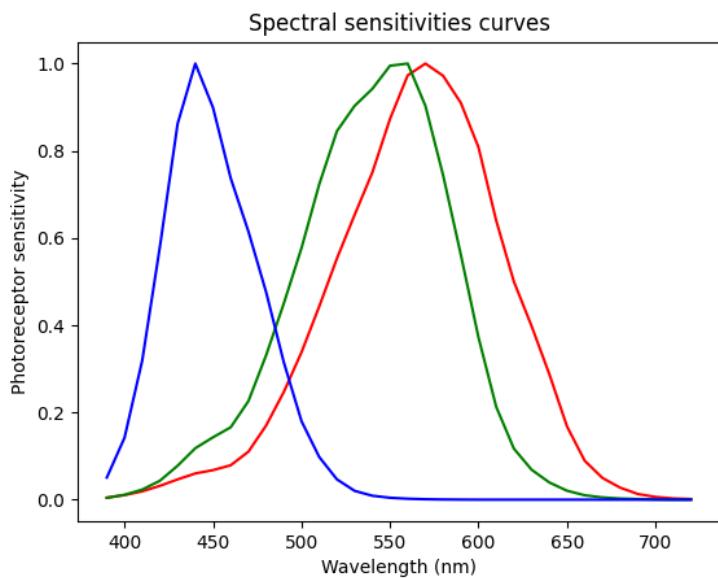
$$\begin{aligned}B_{i,S} &= \int d\lambda S_S(\lambda) \mathcal{I}_i(\lambda) \\ B_{i,M} &= \int d\lambda S_M(\lambda) \mathcal{I}_i(\lambda) \\ B_{i,L} &= \int d\lambda S_L(\lambda) \mathcal{I}_i(\lambda)\end{aligned}$$

To calculate the ζ_i that achieves color matching, we need to solve three linear equations that describe the activities of each photoreceptor in response to the target light and the to the summed basis lights. Rewriting the requirement for color matching in vector-matrix form :

$$\begin{bmatrix} \beta_S^* \\ \beta_M^* \\ \beta_L^* \end{bmatrix} = \begin{bmatrix} \zeta_1 B_{1,S} + \zeta_2 B_{2,S} + \zeta_3 B_{3,S} \\ \zeta_1 B_{1,M} + \zeta_2 B_{2,M} + \zeta_3 B_{3,M} \\ \zeta_1 B_{1,L} + \zeta_2 B_{2,L} + \zeta_3 B_{3,L} \end{bmatrix}$$

As homework, we will use the measured spectral sensitivities of human photoreceptors to calculate the basis lights needed to achieve color matching with different target lights.

Tabulated spectral sensitivities of photoreceptors.



[Download CSV file](#) corresponding to the measured spectral sensitivities of human photoreceptors.

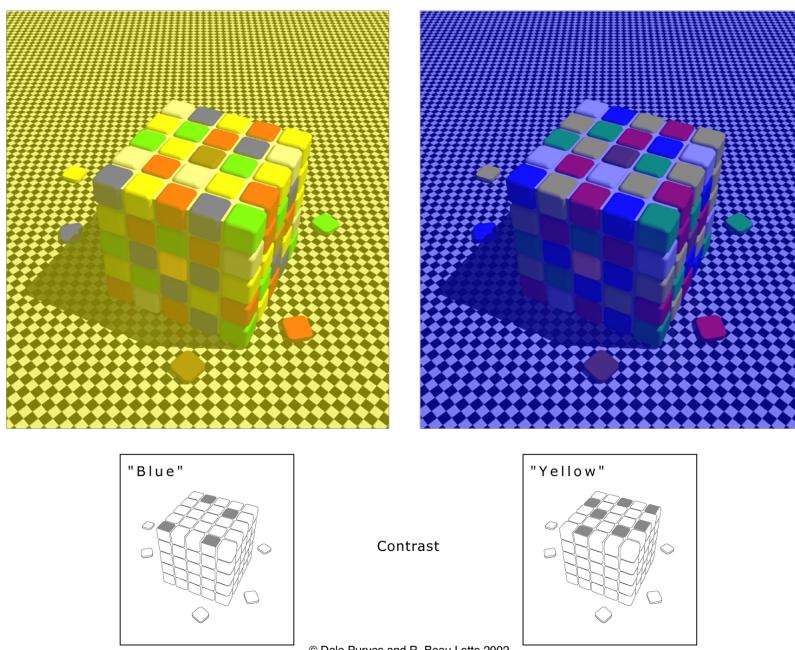
Wavelength	$\mathcal{S}_L(\lambda)$	$\mathcal{S}_M(\lambda)$	$\mathcal{S}_S(\lambda)$
390	0.00442	0.00440	0.0507
400	0.0105	0.0111	0.142
410	0.019	0.0231	0.319
420	0.0317	0.0434	0.580
430	0.0465	0.0778	0.862
440	0.0601	0.118	1
450	0.0677	0.143	0.899
460	0.0789	0.166	0.737
470	0.110	0.226	0.615
480	0.17	0.331	0.476
490	0.247	0.450	0.315
500	0.338	0.577	0.18
510	0.443	0.723	0.0979
520	0.553	0.845	0.0466
530	0.653	0.903	0.0202
540	0.75	0.942	0.00899
550	0.873	0.995	0.00438
560	0.973	1	0.00215
570	1	0.903	0.000969
580	0.972	0.744	0.000402
590	0.91	0.562	0.000157
600	0.808	0.372	0
610	0.64	0.213	0
620	0.500	0.117	0
630	0.398	0.0687	0
640	0.288	0.0395	0
650	0.168	0.0205	0
660	0.0893	0.0101	0
670	0.0499	0.00514	0
680	0.0276	0.00264	0
690	0.013	0.00133	0
700	0.00627	0.000626	0
710	0.00283	0	0
720	0.00134	0	0

Mysteries about color vision

In reality, color perception is much more than pixel-wise analysis of cone photoreceptor activity patterns. We perceive color by integrating information throughout the visual field. This is clearly demonstrated by various visual illusions about color.

One visual illusion is demonstrated when two targets with the same spectral intensities are surrounded by backgrounds with different spectral sensitivities. When this happens, the color of the target can appear different, even though the spectral composition of the light that is being absorbed by cone photoreceptors is exactly the same.

A more dramatic example of the context-dependence of color perception was devised when yellowish or bluish illumination was provided to a Rubik's cube tiled with different colors. Tiles that are spectrally gray when viewed in isolation can become blue when viewed with yellowish illumination or can become yellow when viewed with bluish illumination.



These examples suggest that measurements of human trichromatic color vision is strongly contingent on the experimental setup, involving the matching of large, homogeneous, context-free blocks of color. Color perception in the real world is a more subtle problem.

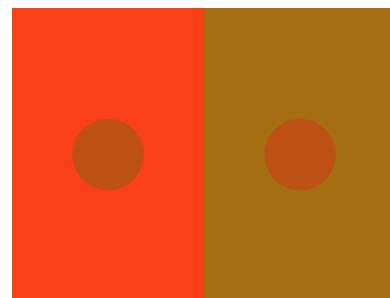


Figure 51: Spectrally identical patches can look differently colored when placed in spectrally different surrounds. The two central targets here are identical, as can be seen by masking out the surround.

Figure 52: Upper images show the cubes as if in yellowish (top left) or bluish (top right) illumination. The lower images show specific tiles of interest in the absence of these contexts. The yellow-looking tiles depicted as if under blue light and blue-looking tiles depicted as if under yellow light are actually a gray on their own.

REFERENCES

- Philip Charles Nelson. *From photon to neuron : light, imaging, vision*. Princeton University Press, Princeton, New Jersey, 2017. ISBN 9780691175188
- Dale Purves, R Beau Lotto, and Surajit Nundy. Why we see what we do. *American scientist*, 90(3):236–243, 2002. ISSN 0003-0996 [Download paper](#)

THE STATISTICS OF PHOTON ABSORPTION BY PHOTORECEPTORS

The anatomy of rod cells is specialized for low-light detection. Each rod cell is ~ 2 microns in diameter, but ~ 100 microns long. Incoming photons travel along the long axis, the better the chance of photon capture by visual pigment. Some animals have a shiny tissue layer behind the retina that acts as a retroreflector (Fig. 53). This *tapetum lucidum* reflects photons not captured by rods in a first pass, adding a chance of absorption in a second pass. Thus, the *tapetum lucidum* effectively doubles rod length without doubling visual pigment.

Starlight has photon fluxes of $<10^{-2}$ photons $\mu\text{m}^{-2} \text{ sec}^{-1}$. Bright sunlight has photon fluxes of $>10^8$ photons $\mu\text{m}^{-2} \text{ sec}^{-1}$. Cones mediate vision over the upper 7-8 log units of this range. Rods mediate vision at rates of photons per second. Rods are engineered to detect single photons. For rods to also *count* photons, they must reliably amplify the signal associated with each photon to be able to distinguish one photon from zero or two.

In photoreceptor cells, the visual pigments and signal transduction molecules that convert photon detection into electrical activity are in the outer segments. In rods, the visual pigments are rhodopsin molecules that are loaded into the membranes of packed disks (Fig. 54). Typical human rods have ~ 1000 disks. Each rod has about 10^5 rhodopsin molecules in each disk.

When a rod cell absorbs a photon, it creates and communicates a signal by changing synaptic output to downstream horizontal cells and bipolar cells in deeper retinal layers. Rods and cones lack the voltage-gated Na^+ needed to create fast all-or-none action potentials. Instead, photoreceptor cells have relatively slow, graded changes in membrane potential that modulate the rate of synaptic release. A biochemical signal transduction cascade connects changes in photon absorption to changes in membrane potential.

ROD PHOTORECEPTOR CELLS are designed to capture photons, but not *every* photon. The longer the rod, the more likely that an entering photon (at $x = 0$) is absorbed before exiting (at $x = L$). What is the probability, P , that a rod of length L captures a photon? This total probability is the sum over the probabilities that the photon is absorbed in each segment (between x and $x + dx$) along rod length.

Consider a thin segment of rod with thickness Δx . A photon that enters this segment has a small finite probability of being absorbed. This probability is proportional to both rhodopsin concentration, C , and an absorption cross-section, σ :

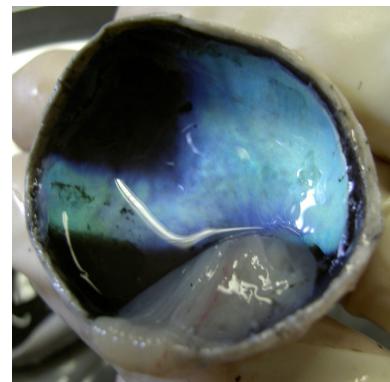


Figure 53: **Tapetum Lucidum**. Choroid dissected from a calf's eye appearing iridescent blue.

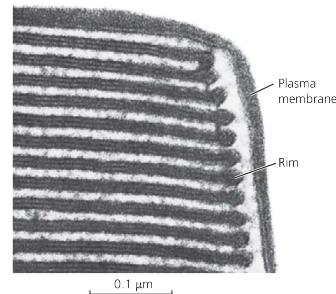
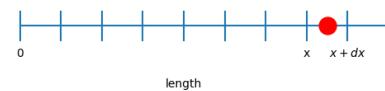


Figure 54: **Rod disks**. Low power electron micrograph of rod outer segment from Fain, Chapter 9.



$$p = \sigma C \Delta x$$

For a photon to be absorbed in the segment between x and $x + \Delta x$, it must have *not* been absorbed in any prior segment between 0 and x . What is the probability that the photon is not absorbed in any prior segment? Divide the distance from 0 to x into N segments. The probability that a photon entering any prior segment is absorbed by that segment is $\sigma C x / N$. So the probability that a photon is not absorbed by each prior segment is $q = 1 - \frac{\sigma C x}{N}$. These probabilities must be multiplied to achieve the total probability of not being absorbed by all N prior segments before reaching x :

$$\left(1 - \frac{\sigma C x}{N}\right)^N$$

Hence, the probability that a photon that enters the rod at $x = 0$ is *first* absorbed between x and $x + dx$ is:

$$dP = \left(1 - \frac{\sigma C x}{N}\right)^N \sigma C dx$$

In the limit of large N , this differential probability is $dP = e^{-\sigma C x} \sigma C dx$. The total probability that an entering photon is absorbed by a rod within its length L is the sum of the probabilities that the photon is first absorbed between 0 and L :

$$P = \int_0^L dP = 1 - e^{-\sigma C L}$$

A definition of the exponential function in terms of a limit:

$$\lim_{N \rightarrow \infty} \left(1 - \frac{x}{N}\right)^N = e^{-x}$$

AS A ROD CELL INCREASES IN LENGTH the probability that each photon is absorbed will exponentially converge to 1. Longer rod cells require more rhodopsin molecules. Each rhodopsin molecule has a finite probability of spontaneous thermal isomerization. These spontaneous isomerizations create “dark noise” events. There must be a trade-off between the amount of signal (true absorbed photons) and the amount of noise (spontaneous isomerizations) when increasing rod length.

If the rate of spontaneous isomerization is r_{dark} , the number of dark-noise events in a given interval of time, τ , is proportional to the total number of rhodopsin molecules in the cell, N_{rh} : $\langle n_{dark} \rangle = r_{dark}\tau N_{rh}$. If the number of spontaneous isomerizations were exactly the average number of spontaneous isomerizations, the cell could subtract this number from the total number of isomerizations to calculate the fraction due to real photons. But spontaneous isomerizations occur randomly. A true photon is seen when its signal is larger than the fluctuating noise due to spontaneous isomerizations.

Binomial Statistics

Say that each rhodopsin molecule has a probability, p , of spontaneous isomerization in each trial. Say that we have N rhodopsin molecules. The mean number of ‘successful’ isomerizations in each trial is pN . What are the fluctuations around this mean? The answer requires binomial statistics in the limit of small p and large N . Before flipping rhodopsin molecules, let’s flip coins.

Start with a biased coin, which gives heads with probability p and tails with probability q . What is the probability of k heads from N flips? The probability that a single flip is heads is p . Since each flip is independent, I multiply together the probabilities of results from each set of flips: p^k for the heads, and q^{n-k} for the tails. I have to account for all permutations —HTTHT and HHTTT are distinct permutations with two heads, where each permutation has equal probability . The probability of k heads from N flips is:

$$P(k; N, p) = \binom{N}{k} p^k q^{N-k}, \quad (3)$$

$$\binom{N}{k} = \frac{N!}{k!(N-k)!}. \quad (4)$$

This is the **binomial distribution**. Is the binomial distribution normalized? In algebra, we use the binomial distribution to expand $(a + b)^N$. So:

$$\sum_{k=0}^N P(k; N, p) = \sum_{k=0}^N \binom{N}{k} p^k q^{N-k} \quad (5)$$

$$= (p + q)^N = 1^N = 1. \quad (6)$$

What is the mean of the binomial distribution. In other words, what is the expected value of k , the number of successes?

$$\langle k \rangle = \sum_{k=0}^N k P(k; N, p) \quad (7)$$

$$= \sum_{k=0}^N k \frac{N!}{k!(N-k)!} p^k q^{N-k} \quad (8)$$

$$= \sum_{k=1}^N \frac{N!}{(k-1)!(N-k)!} p^k q^{N-k} \quad (9)$$

$$= Np \sum_{k=1}^N \frac{(N-1)!}{(k-1)!(N-k)!} p^{k-1} q^{N-k}. \quad (10)$$

A change of variables to $m = N - 1$ and $s = k - 1$ makes it clear that the sum is the binomial distribution, which we know is normalized. Therefore,

$$\langle k \rangle = Np \sum_{s=0}^m \frac{(m)!}{(s)!(m-s)!} p^m q^{m-s} \quad (11)$$

$$= Np. \quad (12)$$

The expected number of successes is the probability of success multiplied by the number of trials. However, the mean is only one parameter of a probability distribution. Another parameter is **variance**, a measure of the spread of a distribution about the mean. Variance is defined as:

$$\text{Var}(k) = \sigma_k^2 = \langle k^2 \rangle - \langle k \rangle^2. \quad (13)$$

The square root of variance, σ_k , is the **standard deviation**. To solve for the variance of the binomial distribution, we must compute another expected value, $\langle k^2 \rangle$.

$$\langle k^2 \rangle = \sum_{k=0}^N k^2 P(k; N, p) \quad (14)$$

$$= Np \sum_{k=1}^N k \frac{(N-1)!}{(k-1)!(N-k)!} p^{k-1} q^{N-k} \quad (15)$$

$$= Np \sum_{s=0}^m (s+1) \frac{(m)!}{(s)!(m-s)!} p^m q^{m-s} \quad (16)$$

$$= Np (\langle s \rangle + 1) \quad (17)$$

$$= Np ((N-1)p + 1) \quad (18)$$

$$= N^2 p^2 + Np(1-p) \quad (19)$$

$$= N^2 p^2 + Npq. \quad (20)$$

$$(21)$$

So the variance of the binomial distribution is:

$$\sigma_k^2 = \langle k^2 \rangle - \langle k \rangle^2 \quad (22)$$

$$= N^2 p^2 + Npq - (Np)^2 \quad (23)$$

$$= Npq. \quad (24)$$

Poisson Statistics

Let us now try to take the limit of the binomial distribution as $N \rightarrow \infty$, but $\mu = Np$ is constant. Then:

$$P(k; N, p) = \frac{N!}{k!(N-k)!} p^k (1-p)^{N-k} \quad (25)$$

$$= \frac{N(N-1)(N-2)\dots(N-k+1)}{k!} \left(\frac{\mu}{N}\right)^k \left(1 - \frac{\mu}{N}\right)^{N-k} \quad (26)$$

$$= \frac{\mu^k}{k!} \frac{N(N-1)(N-2)\dots(N-k+1)}{N^k} \left(1 - \frac{\mu}{N}\right)^N \left(1 - \frac{\mu}{N}\right)^{-k}. \quad (27)$$

Taking the limit of each of the three N -dependent terms as $N \rightarrow \infty$, we get 1, $e^{-\mu}$, and 1, respectively. So:

$$P(k, \mu) = \frac{\mu^k}{k!} e^{-\mu}. \quad (28)$$

Since in the limit of small $p, q \rightarrow 1$, the variance $\sigma^2 = Np = \mu$. Poisson processes often show up in biology, where many phenomena are variations on the counting problem (binding processes, photon detection, random excitation, molecule synthesis etc.). Like the binomial distribution, the Poisson distribution is a discrete distribution. Note that the probability of having no events is:

$$P(0, \mu) = e^{-\mu}. \quad (29)$$

It follows that the probability of having at least one success is:

$$P(k \geq 1, \mu) = 1 - e^{-\mu}. \quad (30)$$

The Poisson distribution often shows up describing what is known as a Poisson process. A Poisson process is one where an event occurs with some constant probability per unit time λ , such that over some time t the mean number of events $\mu = \lambda t$. Thus, the probability that k events occur in time t is:

$$P(k, \lambda) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}. \quad (31)$$

Optimal rod length

We are now ready to calculate the optimal length of a rod cell in terms of signal to noise. If the mean number of spontaneous isomerizations in a cell is $\langle n_{dark} \rangle = r_{dark} \tau N_{rh}$, the standard deviations of fluctuations in this number will be:

$$\delta = \sqrt{r_{dark}\tau N_{rh}}$$

A fraction of photons in each flash of light that enters a rod cell are absorbed by that rod cell. This constitutes the true signal:

$$N_{flash} (1 - e^{-\sigma CL})$$

The total number of rhodopsin molecules in the rod cell is a function of rhodopsin concentration, C , and rod volume, AL . The ratio of signal to noise is thus:

$$SNR = \frac{N_{flash} (1 - e^{-\sigma CL})}{\sqrt{CALr_{dark}\tau}}$$

This function has a maximum at an intermediate value of L between 0 and ∞ . Its maximum is reached when $CL \sim 1.26/\sigma$. This means that the probability of an incident photon not being absorbed when signal to noise is maximum is:

$$1 - P = e^{-CL\sigma} \sim e^{-1.26} \sim 0.28$$

Thus, to maximize signal-to-noise ratio, nearly 30% of photons should pass through the rod without being absorbed.

REFERENCES

- Chapter Nine. Gordon L Fain. *Sensory Transduction*. Sinauer Associates, Sunderland, Mass., 2003. ISBN 0878931716 [Download paper](#)
- F Rieke and DA Baylor. Single-photon detection by rod cells of the retina. *Reviews of Modern Physics*, 70 (3):1027–1036, 1998. ISSN 0034-6861 [Download paper](#)

HOW MANY PHOTONS CREATE VISION

RODS DETECT DIM LIGHTS, but how many photons are needed to create light that can be seen. The human eye has so many rod cells (~ 100 million) and each rod cell has so many rhodopsin molecules (~ 100 million) that even if spontaneous isomerizations were exceedingly rare per rhodopsin, they would occur with appreciable rate in the retina as a whole. Unless you knew you were in an absolutely dark room, the brain has no way of telling the difference between a rod cell that is activated by thermal isomerization from a rod cell that is activated by single photon absorption. To see a flash of photons, its signal must be larger than baseline activity caused by thermal isomerizations.

Scientists were interested in the smallest see-able light before the photon was discovered. In 1881, Langley reported a *bolometer*, a device capable of measuring temperature changes as small as 0.00001°C . The bolometer detected the temperature-induced change in electrical resistance of a metal conductor. Increasing temperature increases metal resistance. Langley's bolometer compared the tiny differences in resistance of an illuminated and un-illuminated reference metal. His bolometer could detect thermal radiation from a cow from a quarter mile. Langley estimated the minimum energy of a see-able flash: 3×10^{-16} Joules.

By 1905 when Einstein explained the photoelectric effect in terms of photons, the minimal see-able energy for vision had been reduced another 10-fold. Lorentz used the new equation for the energy of one photon, $E = h\nu$, to estimate the minimum number of *photons* needed to see. He estimated that at least ~ 100 photons must be delivered to the cornea to see a flash. Different experiments in different conditions made similar estimates. Hecht, Shlaer, and Pirenne (1934) noted the most reliable:

Hecht et al. knew that rhodopsin (then called visual purple) was the molecular photosensor. They asked a new question: how many rhodopsins must be activated to be seen? This lower bound might be estimated using the measurable corneal reflectance, scattering of the vitreous humor, and rhodopsin absorption. They guessed that the number of photons that might be absorbed by rhodopsin is about one-tenth the number that reach the cornea. They sought a direct measure.

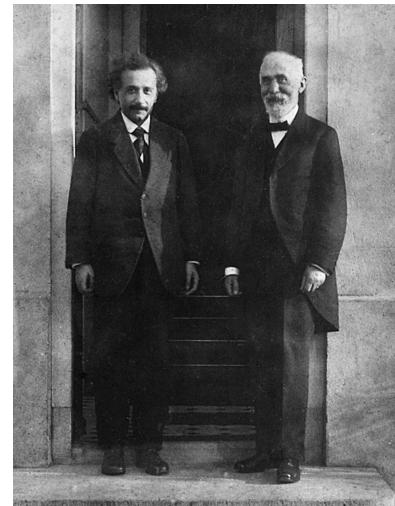


Figure 55: **Einstein and Lorentz.** In 1905, Einstein (left) published his paper on the photoelectric effect, the emission of electrons when electromagnetic radiation, such as light, hits a material. In classical electricity and magnetism, continuous light waves transfer energy to electrons, which would then be emitted when they accumulate enough energy. Einstein explained that the kinetic energy of emitted electrons was the difference in energy of single photons and a threshold voltage. There was no dependence on the number of photons, only the energy of single photons. The photoelectric effect could be fully explained to be a function of the frequency (energy) of single photons, quanta of light. His friend Lorentz (right) used the new equation for the energy of single photons to make the first estimate of the smallest see-able number of photons.

Wavelength	No. of quanta	Source
505	17-30	Chariton and Lea, 1929
507	34-68	von Kries and Eyster, 1907
530	40-90	Barnes and Czerny, 1932

TO SEEK THE SMALLEST SEE-ABLE LIGHT, Hecht, Shlaer, and Pirenne maximized the chance to see a given flash by optimizing experimental conditions. First, they measured scotopic (or rod-dominated) vision which occurs after complete dark adaptation. In the dark, pupils dilate to allow more incoming photons to reach the retina. This happens in a few seconds and increases sensitivity by about ten-fold. In bright light, all rhodopsin molecules in a rod cell will be “bleached” and non-absorptive in the *all-trans* configuration of retinal, the visual pigment. In bright light, all rod cells are hyperpolarized with low rates of synaptic release. The light sensitivity of scotopic vision is reached after biochemistry restores all visual pigment to the *cis* configuration. This takes about 30 minutes.

Rods and cones are not evenly distributed in the retina (Fig. 56). Foveal vision with highest-spatial resolution, at the center-point of our visual field, is mediated only by cones. Rods contribute only to peripheral vision. To see a dim light, don’t look straight at it.

Photons delivered to the retina are spread over space and time. Empirically, a larger test area that is illuminated weakly is seen as well as a smaller test area that is illuminated strongly. But this reciprocal relationship between intensity and area is not perfect. There is a maximum sensitivity corresponding to a circular retinal area that spans about 500 rod cells.

Small numbers of photons are more easily seen when they arrive in one visual ‘moment’ when all photons can be counted at once. The visual ‘moment’ in both rod and cone photoreceptor cells is surprisingly long. Metabotropic receptors (like rhodopsin) require a biochemical signal transduction cascade to open and close ion channels. So metabotropic receptors are often slower than ionotropic receptors (like mechanosensory channels) where stimulus energy directly opens and closes the ion channel. This is why the *flicker fusion threshold*, the frequency of visual stimuli when an intermittent stimulus appears steady, is slow. Both rods and cones lump together photons that arrive within 0.01–0.1 s. The plateau sensitivity of human rod-mediated vision to flicker is 15 Hz. Cones are faster, with plateau sensitivity for bright illumination near 60 Hz. This is why televisions operate at 30 or 60 frames per second. Any slower would create motion jitter between frames. Any faster would not improve the smoothness of movement in video. The reciprocal relationship between intensity and exposure time in scotopic vision holds is perfect for stimuli <0.01 sec. Hecht, Shlaer, and Pirenne used 0.001 s flashes.

Finally, the color of the visual stimulus should be optimized for scotopic vision. The wavelength sensitivity of purified rhodopsin and

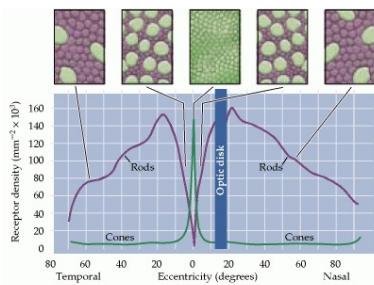


Figure 56: **Rod distribution.** Distribution of cones and rods in a typical human retina

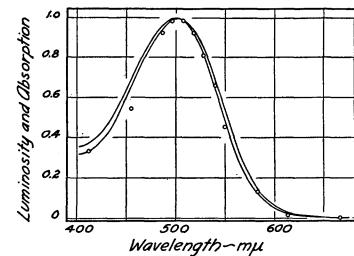


Figure 57: **Rod absorption.** Comparison of scotopic luminosity at the retina with visual purple absorption. The curves are the percentage absorption spectra of visual purple; the upper curve represents 20 per cent maximal absorption, and the lower one 5 per cent maximal absorption. All curves have been made equal to 1 at the maximum, 500 nm, for ease in comparison.

scotopic vision peak near 510 nm (Fig. 57).

THE HSP EXPERIMENT involved a human observer who triggered the release of a flash of light that would fall on their retina in an area spanning ~ 500 rods at about 20° from the center of vision. Another person manipulated the filters and wedges that controlled the flash intensity. Because the observer triggered each flash, the observer also knew when to pay attention. This aspect of the experimental design, by maximizing the readiness and receptivity of the observer, might have increased the sensitivity of their measurements to minimal flashes. Honesty mattered. The observer had to be truthful about admitting to not seeing a flash of light that they themselves had delivered to their own retinas but failed to see.

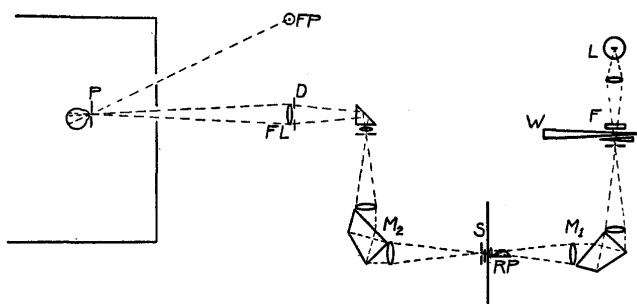


Figure 58: **Apparatus** for measuring minimum energies necessary for vision. The eye at the pupil P fixates the red point FP and observes the test field formed by the lens FL and the diaphragm D . The light for this field comes from the lamp L through the neutral filter F and wedge W , through the double monochromator $M_1 M_2$ and is controlled by the shutter S .

A meaningful threshold for seeing is when an observer successfully perceives a majority of stimuli. The HSP experiment did not allow *false positives* – the observer cannot see a flash that was not triggered. But every failure to see a flash is a *false negative*, where the observer must admit to failing to see a self-administered light. Valid statistics require truthfulness of the observer, who must not pretend to see flashes, even when they knew that they occurred. HSP reduced the ‘threshold for seeing’ to one simple probability, when an observer performed at a rate of 60% true positive/40% false negative.

HECHT, SHLAER, AND PIRENNE obtained remarkably consistent results for a threshold for seeing, the number of photons that hit the cornea, that are successfully seen 60% of the time. Seven different subjects saw flashes of light that delivered between 54 and 148 blue-green photons to the cornea. These numbers are comparable to the earlier ‘most reliable’ measurements.

Observer	Quanta
S.H.	126
	135
	107
	87
	79
	123
S.S.	148
	79
	54
	56
	62
	96
C.D.H.	99
	104
	65
	76
M.S.	58
	58
S.R.F.	81
	112
A.F.B.	120
M.H.P.	83
A.F.B.	79
	83
M.H.P.	138

A DEEPER QUESTION is how many photons are needed to activate rod cells to see? How many individual rods need to be activated for the retina to ‘see’ a flash? One approach is to estimate the fraction of photons that arrive at the cornea that reach the retina and are absorbed by rhodopsin. The cornea reflects ~4% of incident photons. The vitreous humor absorbs ~50% of transmitted photons. Even rods of optimal length will only absorb ~70% of photons that reach the retina. The retina absorbs a small fraction of the photons that arrive at the cornea. If the number of photons that arrive at the cornea is N , the number of absorbed photons is $a = \alpha N$, where α is unknown and takes into account all losses by reflection and absorption.

Say that flashes result, on average, in one absorbed photon at the retina: $a = 1$. Not every flash will result in one photon absorption. Some will result in zero absorptions. Some

will result in two or more absorptions. Poisson statistics describes the complete probability distribution of the number of absorbed photons in each trial, $P(k)$

$$P(k) = \frac{a^k}{k!} e^{-a}$$

, where a is the mean arrival number of photons across trials. Why Poisson statistics? No light source reliably delivers a fixed number of photons to the retina in every trial. Each photon that is released by a light source has a small probability of making its way to the retina and being absorbed. Large numbers of photons at the source and small probabilities that each photon might eventually be absorbed results in Poisson statistics. The probability of seeing curve can be plotted as a function of the mean stimulus size, but at each stimulus size, there is unavoidable trial-to-trial variability in the size of the stimulus that will affect the probability of seeing.

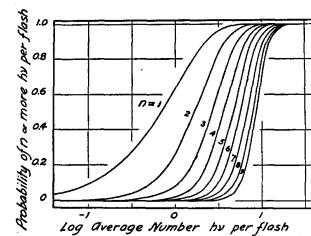


Figure 59: **Probability of seeing, the theory.** For any average number of quanta per flash, the y-axis gives the probabilities that the flash will deliver n or more quanta to the retina, with different values assumed for n .

If the threshold of seeing is one absorbed photon at the retina, then the 'probability of seeing' curve will be a smooth sigmoidal function. This sigmoidal curve can be calculated using the Poisson distribution and summing the probabilities of absorption of one, two, three photons and so on (Fig. 59). If the threshold of seeing is two absorbed photons, then Poisson distributions must be summed from the absorption of two photons and up. For a threshold of n photons, the probability of seeing curve is:

$$P_{see} = \sum_{k=n}^{\infty} \frac{a^k}{k!} e^{-a}$$

As the threshold increases, the sigmoidal curve shifts towards more photons being absorbed. As the curve shifts towards larger stimuli, it becomes steeper when plotted as probability versus the *logarithm* of stimulus size (Fig. 59).

There is merit in plotting P_{see} against the logarithm of stimulus size. Plot P_{see} as a function of the logarithm of the mean number of photons that arrive at the retina, $\log a$. Plot P_{see} as a function of the logarithm of the mean number of photons that arrive at the cornea, $\log N$. Because $a = \alpha N$, $\log a = \log \alpha + \log N$. The two P_{see} curves will have identical shapes except for a horizontal shift of $\log a$.

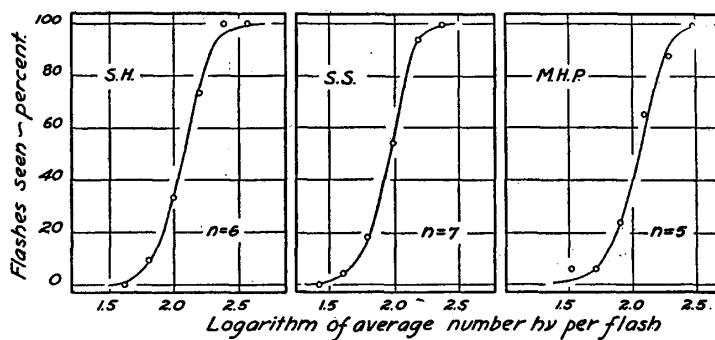


Figure 60: **Probability of seeing, the experiment.** Relation between the average energy content of a flash of light (in number of photons) and the frequency with which it is seen by three observers. Each point represents 50 flashes, except for S.H. where the number is 35. The curves are cumulative Poisson distributions with different thresholds.

Another merit in plotting against the logarithm of stimulus size (whether stimulus size is measured at the cornea or retina) is that the slope of the curve is a direct measure of the threshold:

$$\frac{dP_{see}}{d \log N} \approx \sqrt{n}$$

To derive this square root dependence, one needs to differentiate the probability of seeing, use Stirling's formula, use the chain rule, and evaluate the slope near the inflection where $a \approx n$. As n increases, the steepness of the probability of seeing curve will increase. Whether the probability of seeing curve is plotted as the logarithm of photons at the cornea or the logarithm of photons at the retina does not matter to this steepness. Steepness is always a function of n , the threshold number of photons needed to see.

The steepness of the probability of seeing curve resembles the **signal-to-noise ratio** of sensory perception. The size of the signal – the numerator of the signal-to-noise ratio – is the number of absorbed photons. Because this signal obeys Poisson statistics, the standard deviation in stimulus size – the noise in the denominator of the signal-to-noise ratio – is its square root: $\sigma = \sqrt{n}$. For the three subjects for whom they had the most data – with initials S.S., S.H., and M.H.P. – HSP derived thresholds of 5,6, and 7 photons from the steepness of their respective probability of seeing curves.

WHY ARE MULTIPLE PHOTONS NEEDED TO SEE? HSP argued that probability of seeing curves depended on the statistical variability of any flash stimulus. But the probability of seeing curve can also depend on the statistical variability of internal noise. A weak signal corresponding to 5-8 photons must be distinguished from a background of spurious signals that will occur in the retina in total darkness. The retina cannot know whether a given rhodopsin activation was due to thermal activation or photon activation. The retina can only count rhodopsin activations. The threshold decision – whether a subject sees a flash of light in a given trial – must be based on the sum of weak and spurious signals. If the weak signal is ‘seen’, it is because this summed signal is larger than typical spurious signals in darkness.

The threshold for a perceptual decision is linked to the reliability of the response. If the threshold is lowered, then response reliability will be lowered. This is because more spurious signals can be interpreted as ‘seeing’, increasing the rate of false positives. If the threshold is raised, some weak signals will not be large enough to be distinguished from spurious signals alone, increasing the rate of false negatives. The probability of seeing and response reliability will be functions of signal size, noise, and the internal detection threshold used by the observer.

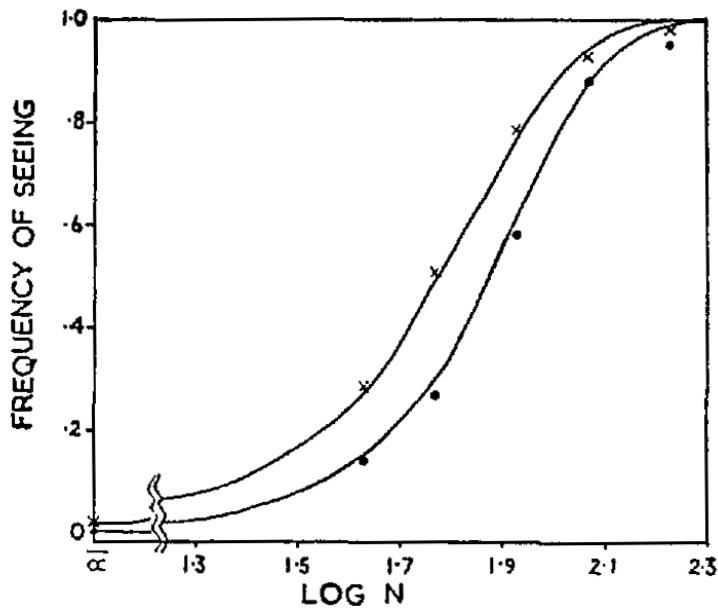


Figure 61: **Poisson probability distributions.** For any average number of quanta per flash, the y-axis gives the probabilities that the flash will deliver n or more quanta to the retina, with different values assumed for n .

HORACE BARLOW wanted to test the effect of response reliability on detecting weak signals. Human observers are conscious of the certainty of their own perceptions and the reliability of their own decisions. Barlow took advantage of this perceptual awareness to test the responses of one human observer, a fellow named Roy Rumble, at two different detection thresholds. Barlow delivered flashes of different intensities, including blanks where no flash was given, to Rumble. Rumble was encouraged to signal when a flash was “seen” and also when a flash was “possible”. Thus, the threshold number of absorbed photons “possible” would be lower than the threshold for “seen”. Because Barlow knew when blanks were delivered, he knew

when Rumble reported “false positives” and “false negatives”.

Indeed, the probability of seeing curve for “seen” events had a higher threshold (shifted to higher flash intensities) and higher steepness (required more photons) than probability of seeing for “possible” plus “seen” events. Response reliability differed for the two curves. The subject never said a blank was “seen”, but reported 3 of the 300 blanks as “possible”. Lowering the response threshold increased the rate of false positives to 1%. From these curves, Barlow extracted somewhat different parameters from his experiment than those of HSP.

- N , average number of photons at the cornea
- n , average number of photon absorptions, i.e., *bona fide* rod excitations
- x , average number of confusable events, i.e., spurious rod excitations
- $a = x + n$, total average of events (real photons plus noise)
- c , threshold number of events to “see”

Just as for HSP, the probability of seeing is a cumulative probability distribution involving these parameters.

$$P_{\text{see}} = \sum_{k=c}^{\infty} \frac{a^k}{k!} e^{-a}$$

The HSP experiment used the frequency of seeing curve as a function of the logarithm of flash intensity to estimate two values, 1) the fraction of incident photons at the cornea that activated rod cells, and 2) the threshold number of activated rod cells needed to see. These two values can be inferred from one frequency of seeing curve, because this sigmoidal curve has two quantifiable parameters, its horizontal shift and steepness. Barlow’s experiment had more parameters. Like the HSP experiment, a fraction of incident photons would activate rods (one unknown value). But two frequency of seeing curves in Barlow’s experiment meant two separate thresholds (two more unknown value). Lastly, every flash stimulus would be accompanied by a certain number of spurious rod activations (a fourth unknown, noise that has the same average amplitude in every trial). Barlow’s experimental measurement of two frequency of seeing curves contained enough information to calculate the four unknowns.

The steepness of the probability of seeing curve in Barlow’s formulation is somewhat different than the steepness in HSP’s formulation:

$$\frac{dP_{\text{see}}}{d \log N} \approx \frac{c - x}{\sqrt{c}}$$

As before, the steepness evokes the **signal-to-noise ratio**. In this case, the size of the signal in the numerator is the distance of the threshold c from confusable events x , and the size of the noise in the denominator is the expected fluctuations in that signal given by Poisson statistics, \sqrt{c} .

Parameters	Best fits
n/N	0.14
x	8.9
c for Possible or Seen	17
False positive rate for Possible or Seen	0.01
c for Seen	19
False positive rate for Seen	0.002

For “seen” events, meeting the threshold requires 10 events above the number of confusable events (spontaneous rhodopsin activation). In terms of signal-to-noise, events corresponding to “seen” *bona fide* flashes are ~ 2.3 standard deviations from the confusable events. Events corresponding to “possible” plus “seen” flashes are ~ 2 standard deviations from confusable events.

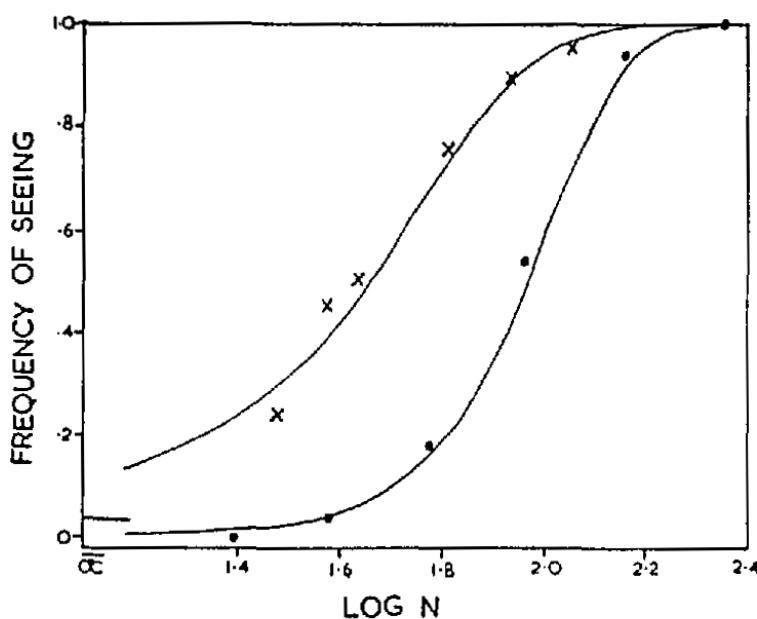


Figure 62: **Other experiments.** Data from Hecht et al. (dots) and van der Velden (crosses) fitted using Barlow's theoretical curves and parameters.

Is Barlow's formulation consistent with earlier measurements? From one probability of seeing curve in HSP (or one probability of seeing curve in another measurement by van der Velden), it is impossible to extract all the parameters that Barlow used. Earlier measurements did estimate the number of photons at the cornea, N , from which they inferred the fraction of photons that were lost and the threshold number of rhodopsin activations needed to see. If the number of spontaneous rhodopsin activations in earlier measurements was similar to that measured by Barlow, $x = 8.9$, these numbers could be inferred (Fig. 62). Fitting Barlow's parameters to HSP's experiments revealed roughly consistent numbers for the fraction of absorbed photons ($n/N = 0.13$) and the threshold of rhodopsin activations ($c = 21$). Fitting Barlow's parameters to another experiment by van der Velden revealed an unrealistically large fraction of absorbed photons ($n/N = 0.9$) and lower threshold of rhodopsin activations ($c = 15$).

REFERENCES

- S Hecht, S Shlaer, and M H Pirenne. Energy, quanta, and vision. *The Journal of General Physiology*, 25(6): 819–840, 1942. ISSN 0022-1295 [Download paper](#)
- HB Barlow. Retinal noise and absolute threshold. *Journal of the Optical Society of America*, 46(8):634–639, 1956. ISSN 0030-3941 [Download paper](#)