CHAPTER 9

MODELING THE H-FUNCTION

It is time to pick up the thread of the argument that was begun in Chapter 2. The fundamental equation (2.6) was proposed,

$$F = kH \, , \qquad\qquad (2.6) \,/\, (9.1)$$

where F is a perceptually related variable such as subjective magnitude or impulse frequency in a sensory neuron, k is a constant that is greater than zero, and H is the entropy of a stimulus. Since H can be regarded as an uncertainty, or perhaps as a potential for receipt of information, F is a measure of this uncertainty or potential to receive information. Therefore, Equation (2.6)/(9.1) states that a quantity, F, which can be measured in the laboratory, is, in effect, a measure of uncertainty. In Chapter 2, it was suggested that the "denomination" of this uncertainty – that is, the quantity which is uncertain – is the exact value of the physical magnitude or intensity of the stimulus. This process of setting a measurable quantity proportional to an uncertainty no doubt strikes some people as enigmatic, to say the least. However, we may recall that we have already encountered this process in Chapter 6 in our survey of statistical mechanics. Boltzmann had discovered that

$$S = k_B \ln W = k_B H_I \, , \qquad\qquad (6.20) \,/\, (6.21)$$

where S is the thermodynamic entropy of a gas under equilibrium conditions, and k_B is now known as Boltzmann's constant. The denomination of the uncertainty, H, is the exact microstate corresponding to a given macrostate. A measurable quantity had been set equal to an uncertainty.

"Aha!" you may be thinking.

"So $F = kH$ has been modeled after $S = k_B H_I$!"

Would that I possessed such insight. There seems to be no direct link connecting physical entropy, S, with, for example, neural impulse rate as measured by F. There certainly have been studies involving the thermodynamic entropy of neural processes (e.g. Margineanu, 1972), but these studies have not, to my knowledge, been related to perceptual function. We shall, however, draw the equations $F = kH$ and $S = k_B H_I$ quite closely together in Chapter 15.

"Then what reasoning led you to postulate $F = kH$?" you may ask. A rather different sort of reasoning, closer to the philosophy of Berkeley than to the physics of Boltzmann.

Now is a good time to look back at Figure 1.2, which gives the order of exposition. We have proceeded to the right, across the top of the flow diagram, down the right-hand side, and partway across the bottom. We are now moving toward the block at the lower left corner. The center block (validation) will take us through a number of chapters. We shall then proceed up the left-hand side of the diagram to reach Berkeley, with whose concept of "relative perception" we really began.

At the level of Chapter 2, we had developed only the concept of entropy or information of events with discrete outcomes. So, in order to introduce Equation (9.1), it was necessary to postulate the existence of an organism that could perceive only discrete stimuli. That is, the organism could determine only that the stimulus assumed the values $I_1, I_2, ..., I_n$, but could not perceive a continuum of intensities[1]. Then, if the probabilities of occurrence of the stimuli were equal, $H = \log n$, and $F = kH = k \log n$: F was proportional to the logarithm of the number of equally probable stimulus intensities. This discrete representation certainly produces some conceptual problems, which I tried to mitigate, tentatively, in Note 6 to Chapter 2. The discrete representation, however, was intended primarily as a vehicle to introduce Equation (9.1). A more realistic interpretation of (9.1) requires that H be a measure of entropy in a system with continuous outcomes; that is, a system where intensity, I,

can vary continuously. For this reason, we took a long digression through Chapters 4, 5, 7 and 8. We should now be in possession of the conceptual apparatus necessary to proceed.

THE SENSORY RECEPTOR AS A QUANTUM DETECTOR

Many of the ideas developed here were put forward by the author in 1977.

The sensory stimulus belongs most properly to the microscopic or quantum world. Fewer than 10 molecules of an odorant may be detected by the olfactory receptor(s). About 15 photons may be required to see a flash of light (Barlow and Mollon, 1982, page 127). Certainly when operating at these near-liminal levels, the sensory detector can be regarded as a "quantum detector." However, the sensory afferent neuron issuing from these receptors, and most particularly the brain to which these neurons report, does not seem to be equipped to detect individual quanta (molecules, photons ...). Let me elaborate.

As stated before, we are concerned in this book primarily with constant stimuli of the intensity type; that is, with stimuli of light, sound, concentration, etc. that are presented in the form of step functions (Figure 1.1). However, to be more precise it should be said that we shall be dealing with stimuli whose *mean or average value* is constant, since most stimuli of the intensity type consist of a multitude of small particles or bundles of energy whose density is changing on a moment to moment basis (Figure 9.1). For example, if it is stated that the stimulus is an 0.1 M solution of sucrose, what is meant is that the bottle of stock solution contains 0.1 mole of sucrose per litre of solution. However, due to random movements of sucrose molecules in solution, a very small sample of this solution will seldom consist of exactly the stated concentration. The concentration of a very small sample will fluctuate about the value of 0.1 M. Therefore, the sensory receptors, which sample only very small quantities of the stimulus solution, *will detect variable concentrations of sucrose even when the stimulus is* (*macroscopically*) *constant* (Figure 9.1). Similarly, the photoreceptor, even when stimulated by a macroscopically constant light stimulus, will record a fluctuating photon density.

Not all the sensory stimuli that we deal with can be reduced to a density fluctuation diagram such as Figure 9.1, but the principle of fluctuation about the mean will still prevail. For example, if we apply a sudden, constant load to a tendon which activates a stretch receptor, there will be a period of transient oscillation, followed, probably by small, sustained fluctuations about the mean force, produced by muscle and elastic elements. Similarly, sound waves which activate the hearing mechanism consist of fluctuations in air pressure, etc.

The receptors are quantum receptors and, therefore, detect quantum fluctuations. However, the brain, which receives the report of sensory receptors, is not made aware of quantum fluctuations. The message the brain receives concerns the level of the steady, macroscopic stimulus, which is the mean of
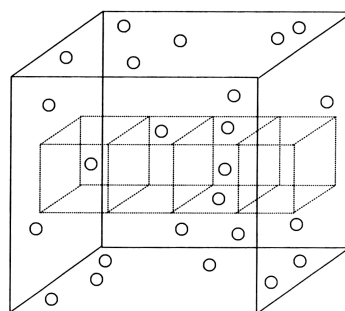

Figure 9.1   The large cube represents a large sample of a fluid containing 24 particles. Imagine the boundaries of this cube as impermeable so that, although particles are free to move about, they cannot escape from the large container. Particles move about randomly within this cube. The small curves represent small samples of the fluid such as might be obtained by a sensory receptor. The boundaries of the small cubes (in dotted lines) are just imagined boundaries, so that particles are free to move in and out. Due to random movements of the particles, one of the small cubes contains no particles, two of them contain one particle, and one contains three particles. In this way, a sensory receptor may obtain 0, 1, 2, 3, ... particles in any one process of sampling.

the signals detected by the receptors. The fluctuating environment of the sensory receptor is translated, presumably by a process of averaging, into a smooth report received at the brain. Since the message of the sensory receptors is usually relayed to the brain by means of a neural frequency code, it is expected that this frequency will also, usually, encode a smoothed or averaged signal (although, of course, we cannot be sure).[2]

SAMPLING THE STIMULUS POPULATION

The receptor will be regarded as drawing samples from the stimulus population which, in turn, will be regarded as an infinite population. Each sample will provide one value of the density of the stimulus. For concentration, density equals the number of molecules per sampled fluid volume; for light, density equals the number of photons per volume of light beam sampled. The samples are identically and independently distributed. Hence, if the original population of densities had mean, μ, and variance σ_S^2 ('S' designates 'Stimulus'), the mean of m samplings of the population will be a sample of size m, which, by the central limit theorem, will tend toward the normal distribution with mean, μ, and variance, σ_S^2/m, as m increases (see Chapter 8).

At what rate does a sensory receptor make samplings of its stimulus population? We do not know. We know only that such samplings must be occurring. In the absence of any definitive knowledge, I suggest a sort of null conjecture:[†] namely that the rate of sampling is constant, so that at time, t,

$$m = [t/t_s] \,, \tag{9.2}$$

where t_s is the time between samples, and $[t/t_s]$ designates the greatest integer not greater than t/t_s. In a continuous representation,

$$m = \alpha t \,, \qquad \alpha \text{ constant } > 0 \,. \tag{9.3}$$

$$\alpha = dm/dt \,, \tag{9,4}$$

or

$$\text{sampling rate} = dm/dt \,. \tag{9.5}$$

Clearly, for m small, the approximation of number of samplings, m, by the continuous variable, t, will weaken; and additionally, for small m the acceptability of the central limit theorem weakens. Therefore, the validity of the model for H as a function of t that we are developing will extend only beyond some minimum value for the variable, t.

The differential entropy of the distribution of means is

$$H_S = -\int p \ln p \, dx \,, \tag{9.6}$$

or explicitly (from Equation (8.18)),

$$H_S = \frac{1}{2} \ln(2\pi e \, \sigma_S^2/m) \,, \tag{9.7}$$

where p is the normal distribution with variance σ^2/m. We do not know the factors operating within the receptor system that limit the precision with which a mean may be measured. However, we understand from Chapter 7 that in the absence of such a "noise" distribution, no information can be received.

THE REFERENCE SIGNAL

Let us, therefore, assume the existence of "Gaussian white noise," which is just an interfering signal whose probability density is normal. Since the term "noise" is usually regarded in the pejorative, let us designate this signal as *reference* instead. It is a stimulus signal without which information would not be measurable. Accordingly, its variance will be designated by σ_R^2, and its differential entropy by

$$H_R = \frac{1}{2} \ln(2\pi e \, \sigma_R^2) \,. \tag{9.8}$$

The absolute entropy received per stimulus, H, is then given by

$$H = H_{S+R} - H_R = \tfrac{1}{2}\ln[2\pi e\,(\sigma_S^2/m + \sigma_R^2)] - \tfrac{1}{2}\ln(2\pi e\,\sigma_R^2)\,. \tag{9.9}$$

$$H = \tfrac{1}{2}\ln(1 + \sigma_S^2/m\sigma_R^2)\,. \tag{9.10}$$

In the theory that follows we shall take σ_R^2 to be constant for a given modality of sensation. Introducing Equation (9.3),

$$H = \tfrac{1}{2}\ln(1 + \sigma_S^2/\sigma_R^2\alpha t)\,. \tag{9.11}$$

Since $F = kH$ from Equation (9.1), we have

$$F = \tfrac{1}{2}k\,\ln(1 + \sigma_S^2/\sigma_R^2\alpha t)\,. \tag{9.12}$$

RELATIONSHIP BETWEEN VARIANCE AND MEAN

Equation (9.12), then, relates the perceptual variable, F, to σ_S^2, the variance of the stimulus population, and t, the duration of the stimulus. Unfortunately, in this form, the equation is still not readily applicable in the analysis of experiments, because σ_S^2 is not usually known. In the laboratory, we measure the mean stimulus value, μ, but usually not the variance. It is, therefore, necessary to introduce a relationship between mean and variance of the form

$$\sigma_S^2 = f(\mu)\,; \tag{9.13}$$

that is, σ_S^2 is some function of mean stimulus intensity, μ. However, in formulating such an equation it is necessary to bear in mind where σ_S^2 is being "measured," and where μ is being measured. σ_S^2 is the stimulus variance "as seen by" the sensory receptor; it is the variance obtained by the receptor's sampling of its stimulus environment. On the other hand, μ, the population mean, is estimated by the investigator as I, the intensity of the constant stimulus (that is, $\mu = I$), at some convenient location within the laboratory. For example, σ_S^2 for audition may be measured at the hair cells within the cochlea in the inner ear, while $\mu = I =$ sound intensity[††] is usually measured by a sound level meter at the outer ear. Hence, σ^2 incorporates fluctuation arising from both the physical stimulus and the biological receptor. Therefore, the character of the function, $f(\mu)$, is not obvious, although in the case of audition some idea may be obtained from the work of Rhode (1971). I selected for $f(\mu)$, empirically, a simple power function, so that

$$\sigma_S^2 \propto \mu^n, \qquad n\text{ constant} > 0\,, \tag{9.14}$$

or, since $\mu = I$,

$$\sigma_S^2 \propto I^n\,. \tag{9.14a}$$

The constant, n, may differ among the modalities of sensation.

I chose this function for several reasons. First, relationships between mean and variance within statistical physics often assume the form of Equation (9.14). For example (Jackson, 1968, or other textbooks on statistical mechanics), for a monoatomic ideal gas in thermal equilibrium with a thermal reservoir, and performing no mechanical work, the standard deviation of energy, $\sigma(E)$, is related to mean energy, μ, by the relation

$$\sigma(E)/\mu = 1/(3N/2)^{\frac{1}{2}}\,, \tag{9.15}$$

where N is the (constant) number of particles in the gas. That is,

$$\sigma(E)^2 \propto \mu^2\,. \tag{9.15a}$$

As another example, the number, N, of photons in a monochromatic beam emitted per unit time by a thermal source kept at constant temperature, will fluctuate. The general equation for $\sigma^2(N)$, the mean

square fluctuation, was found by Einstein. However, in the Wien approximation we have simply

$$\sigma(N)^2 = \mu . \tag{9.16}$$

Equations (9.15) and (9.16) are specific examples of mean-variance relationships of the form (9.14).

A second reason for selecting Equation (9.14) is that a similar relationship might be expected from any process wherein a total measurement, I, is the sum of a number of partial measurements. For example, a length, ma, is measured by laying down m times a measuring rod of length a. Suppose that each single measurement is regarded as a random variable whose standard deviation is $\sigma_1(a)$. Then, the variance of the sum of m measurements, $\sigma(ma)^2$, is equal to $m\sigma_1(a)^2$ by the rule of summing variances of independent random variables (for example, Freund and Walpole 1980, p 157 or Weatherburn 1961, p 27). That is,[3]

$$\sigma(ma)^2 \propto m = ma/a = I/a ,$$

or

$$\sigma(ma)^2 \propto I . \tag{9.17}$$

Total variance varies as the "intensity" (= length here) of the whole, which is just a way of saying that larger quantities are associated with greater fluctuations. The variance of measurements of the separation of two marked points in two different cities will be expected to exceed the variance of measurements of the separation of two marked points both within a given room.

None of the above examples demonstrate the validity of Equation (9.14); however, they illustrate the plausibility of such a relationship.

THE ENTROPY FUNCTION

Introducing Equation (9.14a) into (9.11),

$$H = \tfrac{1}{2}\ln[1 + (\beta'/\sigma_R^2)I^n/t)] , \tag{9.18}$$

where β' is a proportionality constant > 0. Finally, condensing the constant β'/σ_R^2 into a single constant, β,

$$\boxed{H = \tfrac{1}{2}\ln(1 + \beta I^n/t) , \qquad \text{[natural units]} .} \tag{9.19}$$

Of course, H [bits] is obtained by dividing H [natural units] by $\ln 2$. *Equation* (*9.19*), *then*, *gives the absolute entropy of m samplings by a sensory receptor taken over an interval of time*, t, *of a stationary stimulus population of mean intensity*, I. Since $F = kH$,

$$\boxed{F = \tfrac{1}{2}k\ln(1 + \beta I^n/t) .} \tag{9.20}$$

Before looking more deeply into the meaning of H as calculated from (9.19), let us review the primary assumptions that have been used to derive Equations (9.19) and (9.20). Six of them are listed below.

(1) There is a process of sampling by the receptor of its stimulus.

(2) The stimulus population approximates the normal distribution closely enough that the central limit theorem will apply to a sample of almost any size drawn from the stimulus population.

(3) Sensory neurons encode their state of certitude concerning the mean stimulus intensity.

(4) The rate of sampling of a stimulus by its receptor is constant.

(5) A "noise" or reference stimulus distribution is present. This reference stimulus is Gaussian with constant variance.

(6) Stimulus variance at the receptor is a simple power function of the mean stimulus intensity measured at a point external to the receptor.

Essentially, upon the above six assumptions, the entropy function was derived. In some applications of the *H*-function we shall not require one of either assumptions (4) or (6). In later chapters, we shall occasionally add an additional assumption, namely that ΔH, a small difference in H, is constant and plays the part of a threshold.

POPULATION VS. SAMPLE VARIANCE

In order to understand the H concept more clearly, and to understand the limitations of the modeled *H*-function given by Equation (9.19), we had best look at some of the intermediary stages of its development. Looking back at Equation (9.10), we see that for constant σ_R^2, the value of H is dependent upon σ_S^2/m, which is the variance of a *population* of means. However, the perceiving organism does not have access to the value of σ_S^2/m. Rather it must sample the stimulus population and obtain, at best, the value of the variance of a *sample* of the population means. That is, the organism measures s^2/m as an estimator of σ_S^2/m. In order to develop a more accurate H function, we should have to make allowance for the probability distribution of the sample variance, which we shall not do here.

MAXIMUM *H* AS POTENTIAL INFORMATION

We recall from Equation (9.10),

$$H = \tfrac{1}{2}\ln(1 + \sigma_S^2/m\sigma_R^2) \,. \tag{9.10}$$

For fixed *I*, *H* is seen to be maximum for $m = 1$. However, if the sensory receptors are sensing variances rather than means, no sample variance can be obtained for m less than 2. Nonetheless, for simplicity, let us take

$$H = H_{\max} \qquad \text{for } m = 1 \,. \tag{9.21}$$

Then $H_{\max}$, being an absolute entropy, represents the greatest amount of information *receivable* by the receptor per stimulus. Remember, though, that $H_{\max}$ is only *received* when the mean signal variance is reduced to zero, leaving only the noise or reference signal variance [recall the $(H_{SN} - H_N)$ idea of Equations (7.18) and (8.14)]. Looking at Equation (9.10), we see that this "extinction" of the mean signal variance, σ_S^2/m, will only occur for large values of *m*, the number of samplings made by the receptor – that is for large *t*. What transpires is a gradual reduction in the mean signal variance as *m* increases – the effects of Fisher's information felt here. We might look at $H_{\max}$, which we have termed an absolute entropy or an uncertainty, as the *potential* of the perceiving system to acquire information. *Potential information is transformed gradually into information as the perceptual act proceeds.*

To broaden the perspective, it is interesting to think back to the quantum physical picture outlined in Chapter 2. Perception of a quantum event occurs as the wave function (the square of whose modulus is a probability density function) collapses over time, Δt. That is, in time Δt, the variance of the quantum probability density decreases, with the resulting perception of the event. The physiological event occurs continuously, within the "paradigm" of classical physics; the wavefunction-mediated event occurs suddenly, within the paradigm of quantum physics.

THE SENSORY NEURON AS METACHANNEL

The concluding section of Chapter 6 should take on clearer meaning now. We see that the sensory receptor receives a "message" or "statement" from the so-called "outside world."[4] Within the context of the present theory, this statement is never made available to the brain, nor, presumably, to "consciousness." Since the sensory neuronal report, *F* is equal to *kH*, and *H* is the entropy or uncertainty or potential to receive information, the neuronal report is proportional to the uncertainty of
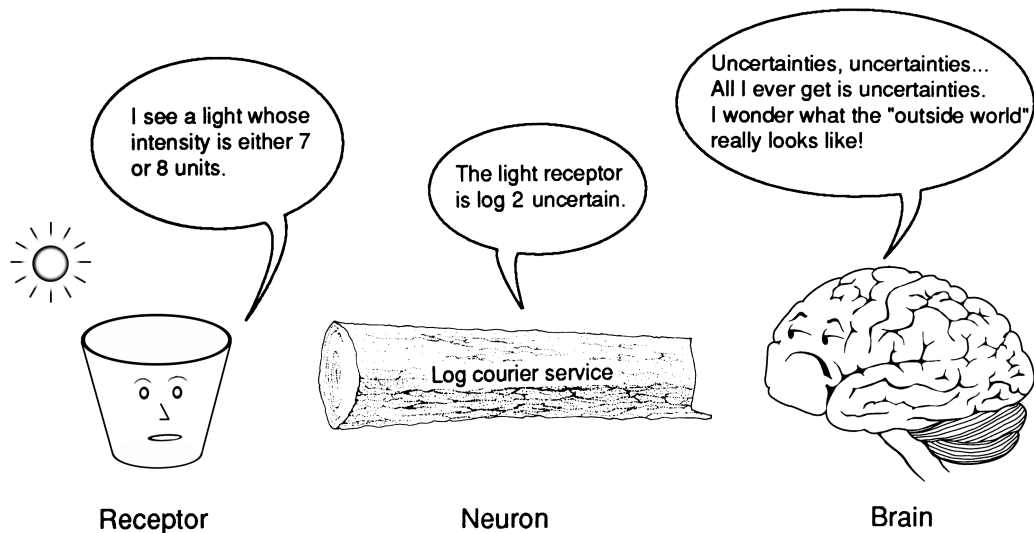
Figure 9.2 The different worlds of the receptor, sensory neuron and brain. The "outside world" is seen only by the receptor. The brain receives from the sensory neuron only a report concerning the "uncertainty" of the receptor about the state of the outside world. From these reports about uncertainty, the brain must synthesize its picture of the outside world. The cartoon depicts only the passive phase of perception: from the receptor to the brain. It omits the active phase of perception, which is the effect of the brain on the receptor.

the receptor about the statement. That is, the neuronal report is a *statement about a statement* or a *metastatement* (Norwich, 1983, 1984).

We see in cartoon fashion in Figure 9.2 a germ of the Berkeleian approach to sensory physiology: The substance of the world seems forever cut off from the mind that perceives it. We seem to be drawn ineluctably into the philosophical core; however, I am determined to delay it, so on to ...

F AT THE THRESHOLD

Let us consider the legitimacy of the use of the *H*-function, as obtained from Equation (9.19), for small values of *I*, when used in (9.1), $F = kH(I, t)$. It is well known that when *I* is less than some value, I_{thresh}, which may have to be determined with due attention to the statistical methods of signal detection theory, a stimulus is imperceptible. However, since *H* has been modeled as a continuous, monotone function of *I* for all values of $I \geq 0$, *F* will be greater than zero for all values of *I* greater than zero. When the perceptual variable, *F*, is interpreted as subjective magnitude, $F > 0$ for $I < I_{\text{thresh}}$ is not acceptable. Clearly, some correction is necessary in Equation (9.1) when *I* is near threshold, particularly when dealing with subjective magnitudes. Simplest would seem to be

$$F = k\left[H(I, t) - H(I_{\text{thresh}}, t)\right] . \tag{9.22}$$

This equation would be appropriate if I_{thresh} assumed the form of a masking stimulus. As we shall see in the next chapter, Equation (9.22) is in accord with the experimental findings of Lochner and Burger (1961).[5] But later we shall be treating thresholds in a manner quite different from the usual, by representing them in terms of the minimum quantities of information required to perceive a stimulus or to discriminate between stimuli; so it is, perhaps, premature to speculate about Equation (9.22). We shall continue to use Equation (9.20), but with caution or not at all in the proximity of I_{thresh}. So we have now issued caveats on the use of certain of our model equations both for very small *t* and for very small *I*.

INFORMATION-FREE-APPROACH

In Chapter 8 (Information) it was stated that we could, probably, proceed with this entropic interpretation of perception without using the concept of information at all. We can now see explicitly how that might be effected. Referring to Equation (9.9), we see that H can be split into its two components generically:

$$H = -\int p_{SR} \ln p_{SR}\, dx + \int p_R \ln p_R\, dx\,, \tag{9.23}$$

where p_{SR} is the summated probability density for stimulus and reference. The second integral on the right-hand side, representing the differential entropy for the reference stimulus, is constant, since p_R depends only on the constant, σ_R^2, which has been taken as constant. Therefore, we can write

$$H = [\text{Differential entropy for stimulus} + \text{reference}] + \lambda'\,, \tag{9.24}$$

where λ' is constant; or

$$F = kH = k\,[\text{Differential entropy for stimulus} + \text{reference}] + \lambda\,, \tag{9.25}$$

where λ is constant; or simply

$$F = kH_{\text{diff}} + \lambda\,. \tag{9.25a}$$

H_{diff} is, of course, decreasing with time due to its argument, σ_S^2/t. Written in the above form, the theory is, in principle, "information-free," since the differential entropy is not an information.

Equation (9.25) is of great interest to me because it parallels Boltzmann's equation

$$S = -k_B H_B \tag{9.26}$$

(cf. Equation (6.20)), where H_B is Boltzmann's H-function (akin to a differential entropy), k_B is Boltzmann's constant, and S is physical entropy. Equation (9.26), the equilibrium form for which has been derived in Chapter 6 [$S = k_B \ln W$ (Equation 6.21): see Tolman, 1979, page 135], states that, ignoring additive constants, physical entropy is proportional to Boltzmann's H-function *both in and out of equilibrium* (ter Haar, 1966, page 27). Boltzmann showed that, due to molecular collisions,

$$\frac{dH_B}{dt} \le 0\,, \tag{9.27}$$

thus demonstrating through the use of statistical mechanics that the thermodynamic entropy, S, is non-decreasing (recall the second law of thermodynamics). One cannot but note the parallel:

F [due to its dependence on $H(t)$] is necessarily non-increasing with increasing time.

S [due to its dependence on $H_B(t)$] is necessarily non-decreasing with increasing time.

We shall return to H_B in Chapter 15.

CONVERVATION LAWS AND UNIFICATION

The second part of Chapter 1 was devoted to a discussion of the nature of physical laws and their capacity to "unify" the observations of science. The distinction was made between laws of conservation and laws of mechanism, and I should like to pursue that distinction here in terms of sensory science. Chapter 3 consists of a survey of many of the empirical rules (or laws) of sensation and perception. These are rules that have been learned by experience; they are, essentially, codified measurements. Our job in the next few chapters will be to *derive* these laws theoretically from Equations (9.19) and (9.20). That is, *we shall try to derive ALL empirical laws that relate a single*, *constant stimulus of intensity* $I > I_{\text{thresh}}$, *applied for a duration of time* $t > t_o$, *to a selected perceptual variable*, F, *using the single theoretical law*, $F = kH$.

Megalomania, perhaps.

Now, it is important to realize that $F = kH$ is a theoretical law of the *conservation type*; it is *not* a mechanistic law. Please recall that the *H*-function was derived from considerations of uncertainties resident in the stimulus. That is, a model of the stimulus was formulated, and an expression for differential entropy was derived from it. We approached it from the point of view of fluctuations in the density of quanta. However, stimuli other than those of the "intensity" type may give rise to other models for *H*; the use of a fluctuating density of quanta may not always be appropriate. Some other quantity may be sampled by the sensory receptor.

We have seen how *H* measures the potential to deliver information. Neurophysiologically, the equation $F = kH$ depicts the transformation of that potential stimulus information to the sensory receptors. Consider a constant stimulus applied to a receptor. When m_1 samplings have been made, $H = H_1$. When m_2 samplings have been made, $H = H_2$. Let m_2 be greater than m_1. Then $F_1 = kH_1$, $F_2 = kH_2$, so that

$$(F_1 - F_2)/k = H_1 - H_2 . \tag{9.28}$$

We specify that the information $H_1 - H_2$ bits is transferred to the receptor, which acknowledges its receipt by signaling $F_1 - F_2$ units of information. When m_1 is very small and m_2 is very large, so that $H_1 \to H_{max}$ and $H_2 \to 0$, then

$$F_{max}/k = H_{max} . \tag{9.29}$$

That is, F_{max} is proportional to the total stimulus information. If we write, further,

$$\frac{F_1 - F_2}{F_1} = \frac{H_1 - H_2}{H_1} , \tag{9.30}$$

we just state that the difference between any two *F*-values, as a receptor continues to sample its stimulus, designates the fraction of the total available information transmitted to the receptor at any *m*-value or time. So we see that $F = kH$ can be viewed as a kind of conservation law that meters the flow of information from stimulus to receptor. No more than F_{max}/k units of information may be transmitted from stimulus to receptor.

We observe, though, that no *mechanism* for the transmitted information is specified; we don't know *how* the information passed from stimulus to receptor. We recall from Chapter 1 that laws of mechanism *complement* laws of conservation. The development and use of $F = kH$ in no way supplants the search for the principles of operation of sensory receptors – the electrical and mechanical properties of signal transducers. *A fortiori*, laws of the type $F = kH$ *need* laws of mechanism to provide a base for information transmittal. In the final analysis, knowledge of mechanism will enable us to formulate improved versions of the *H*-function, and, perhaps tailor the *H*-function to each specific type of receptor. The thrust of the next few chapters, however, will be to show how far we can go *without* knowledge of mechanism, using a common mathematical form of the *H*-function for all modalities of sensation.

APPENDIX: THE PARABLE OF THE SUPERMARKET

(*A Pedagogical Lamentation on the Elusiveness of Friendship*)

I suppose my motives might be suspect when I go to the supermarket with almost no money or plastic at my disposal but with a calculator capable of adding numbers as large as 10^{500}. I proceed to the fruit section where I note that each species of fruit is piled on its own special table: one table for grapes, one table for apples, etc. I go to one such table and proceed to weigh nearly every item of fruit on the table using the scale provided. I enter the weight into the memory on my calculator and then compute the average weight of a piece of fruit, which I duly record. Before leaving the table, I pick up one piece of fruit at random, re-weigh it, and record the weight. I then go on to a second table and repeat the above ritual. I am monopolizing the scales, so I make little grimaces and utter audible expression of annoyance to signify my dissatisfaction with the whole tableful of fruit, in the hopes of assuaging the anger of the woman behind me. What does she know of the methods of science?

I leave the store without buying anything, but brandishing my calculator, and I soon meet a friend on the street. He tells me how happy he is to meet me, but I suspect that he doesn't mean it. Nonetheless, I decide to pose the following problem to him. I provide him with the weights of the two pieces of fruit that I have selected at random, but I do not tell him what kinds of fruit they are. I ask him to guess, for each of the two types of fruit, what the *average weight* of the fruit on the table was, and to estimate the "uncertainty" in his guess of the average. That is, I am really asking him to estimate a sort of standard deviation of his guess of the average.

Well, the first piece of fruit weighed 5120 grams. My friend guesses that this fruit must have been a watermelon or pumpkin, estimates the average weight of the species as 5000 grams and his error as ± 2000 grams. The weight of the second piece of fruit was 45 grams. My friend guesses that it must be a berry of some kind, and estimates the average weight as 50 grams ± 20 grams. His uncertainties about the average weight are, then, 2000 grams and 20 grams. I suspect that he is burning with curiosity to know the true values of the means, so I kindly relieve his curiosity by telling him what they were. I wish to continue the dialogue by showing my friend what wisdom lay inherent in his choices, but he seems very anxious to go somewhere and I do not wish to detain him. He tells me that we really must get together some time, but I think he doesn't mean it. He leaves before I can teach him the two main lessons to be learned from the guessing game, but I'll tell him next time. Maybe I'll send him a letter ...

The first lesson is that when I provided him with the correct answer, his uncertainty "collapsed" (in the language of physics), and he received a quantity of information that depended on the magnitude of his prior uncertainty. We can, in fact, measure this information. The scale that I used in the supermarket had a resolution of about 5 grams; that is, measurements were accurate within a range of ± 5 grams. Consider, now, each of the two pieces of fruit which, were, in fact, a pumpkin and a strawberry.

(i)   Pumpkin:
estimated error = ± 2000 grams
"reference weight" or resolution of the scale = 5 grams
call 5 grams a "resolvable category"
information received = $\log_2$ (number of resolvable categories)
$$= \log_2 2000/5 = 8.64 \text{ bits.}$$
Perhaps more precisely, since his estimate of error did not take account of the error of the scale, we can combine errors using the rule for addition of variances. Then,
information received = $\frac{1}{2} \log_2[(5^2 + 2000^2)/5^2]$
$$= \frac{1}{2} \log_2(1 + 2000^2/5^2)$$
$$= 8.64 \text{ bits (Pumpkin).}$$

(ii)   Strawberry
estimated error = ± 20 grams
"reference weight" or resolution of the scale = 5 grams
again, call *5* grams a "resolvable category"
information received = $\log_2$ (number of resolvable categories)
$$= \log_2(20/5) = 2.00 \text{ bits}$$
Or, using again the addition of variances,
information received = $\frac{1}{2} \log_2[(5^2 + 20^2/5^2]$
$$= \frac{1}{2} \log_2(1 + 20^2/5^2)$$
$$= 2.04 \text{ bits (Strawberry).}$$

The second lesson, that I did not get the chance to teach my friend, is that he received a greater quantity of information when I provided him with the correct average weight of the pumpkin than when I provided him with the correct average weight of the strawberry. Larger measurements are associated with larger quantities of information. Intrinsically, he knew that when he ascribed a larger error to the pumpkin estimate than to the strawberry estimate. All this mathematical knowledge was actually present in my friend's mind all along. I just helped him to extract it. Socrates understood such matters long ago, as we learn in Plato's *The Meno*.

My friend is probably already regretting having treated me in such a distant manner. He is probably going to call me tonight to apologize. I'll have to remind him that, in this Parable of the Supermarket,

we are not forbidden to compare Equations (Pumpkin) and (Strawberry) to Equation (9.10) with $m = 1$:

$$H = H_S - H_R = \frac{1}{2}\ln(1 + \sigma_S^2/\sigma_R^2) .$$

He'll grow warm and friendly with the acquisition of this wisdom. I'm sure he will.

NOTES

1. Is this equivalent to the existence of an organism that possesses the ability to perceive a continuum of stimuli, immersed in a world in which stimuli always assume discrete values, like the eigenfunctions of a quantum mechanical system?

2. I know, when I look at a light signal, that I do not perceive quantal fluctuations. That is, I am not conscious of such fluctuations. But do the sensory neurons report such fluctuations anyway? I think not. However, even when a sensory neuron fires at a constant average rate, there are moment by moment fluctuations in the constant rate (stationary time series). I shall assume, in the absence of knowledge to the contrary, that these fluctuations in neuronal firing rate do not reflect quantal fluctuations in the stimulus density although, admittedly, future research may show otherwise. Recall, also, that we are confining our discussion, for the most part, to stimuli of the "intensity" type, so phenomena such as "phase locking" in pitch perception are not considered here.

3. This relation, $\sigma(na) \propto \sqrt{n}$, led Cattell (1893) to his well-known conjecture: "The algebraic sum of a number of variable errors tends to increase as the square root of the number ..." Recall also the generalization of the above by Guilford (1932) into the "Fullerton-Cattell law," discussed in Note 6, Chapter 2. However, it is most important that the reader make the following distinction. In the development of this entropy theory of perception, we are introducing a mean-variance relationship (Equation (9.14)) *at the level of the stimulus* (for example, a relationship between the mean and variance of the density of odorant molecules). We are *not* introducing any *a priori* relationships between the mean stimulus level and the differential threshold, ΔI, for purposes of obtaining Weber's law. We shall derive Weber's law from the general equation of entropy, (9.19). I am grateful to Professor Harry Lawless for drawing my attention to this situation.

4. Within this entropic view of perception, the meaning of the term "outside world" is not all that simple, as shown in Figure 9.2. Since a direct view of the "outside world" is denied to us, the best we can do is to formulate a model of this world. This model is one wherein magnitudes of outside-world events are inferred from the neuronal reports of uncertainties.

5. Lochner and Burger (1961) found for loudness that a function of the form

$$F = k(I^n - I_{\text{thresh}}^n)$$

was much more accurate in describing experimentally measured data than a function of the form

$$F = k(I - I_{\text{thresh}})^n .$$

We see in Chapter 10 that, in terms of H, Lochner and Burger's findings mean that

$$F = k[H(I, t) - H(I_{\text{thresh}}, t)]$$

is better than

$$F = k H(I - I_{\text{thresh}}, t) .$$

†. (2003 ed. note) The "null conjecture" regarding sampling rate by a receptor, introduced at Equation (9.2) was studied in more detail. The receptor, it seems, samples its stimulus environment more rapidly when the stimulus first appears, and then progressively more slowly (K. H. Norwich and W. Wong, "A Universal Model of Single-Unit Sensory Receptor Action", *Mathematical Biosciences*, **125**, 83-108, 1995). Using this model, one can account, in a quantitative manner, for neural activity during rapid onset and removal of stimuli (adaptation and de-adapatation), and for the responses to time-varying stimuli. Stimuli of this type are not embraced by the theory presented in this book.

††.   (2003 ed. note) The term *sound intensity*, as used in this book, will imply a scalar quantity that can be measured by the square of the amplitude of a sound wave. More properly used, *sound intensity* is a vector quantity discussed, for example, by Frank Fahy, *Sound Intensity*. E. and Fn. Spon, London, 1995.

REFERENCES

Barlow, H.B. and Mollon, J.D. 1982. Psychophysical measurements of visual performance. In: *The Senses*. H.B. Barlow and J.D. Mollon Eds. Cambridge University Press.

Cattell, J.M. 1893. On errors of observation. *American Journal of Psychology*, **5**, 285-293.

Freund, J.E. and Walpole, R.E. 1980. *Mathematical Statistics*, 3rd edition. Prentice-Hall, Englewood Cliffs , N.J.

Haar, D. ter. 1966. *Elements of Thermostatistics*. Holt, Rinehart and Winston.

Jackson, E.A. 1968. *Equilibrium Statistical Mechanics*. Prentice-Hall, Englewood Cliffs, N.J.

Lochner, J.P.A. and Burger, J.F. 1961. Form of the loudness function in the presence of masking noise. *Journal of the Acoustical Society of America*, **33**, 1705-1707.

Margineanu, D.-G. 1972. Entropy changes associated with a nerve impulse. *Kybernetic*, **11**, 73-76.

Norwich, K.H. 1977. On the information received by sensory receptors. *Bulletin of Mathematical Biology*, **39**, 453-461.

Norwich, K.H. 1983. To perceive is to doubt: The relativity of perception. *Journal of Theoretical Biology*, **102**, 175-190.

Norwich, K.H. 1984. Why the eye may be found to be a source of light. *Proceeding of the 6th International Congress of Cybernetics and Systems of the WOGSC*, Vol **2**, 831-836.

Rhode, W.S. 1971. Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique. *Journal of the Acoustical Society of America*, **49**, 1218-1231.

Tolman, R.C. 1979. *The Principles of Statistical Mechanics*. Dover, New York.

Weatherburn, C.E. 1961. *A First Course in Mathematical Statistics*. Cambridge University Press, London.