CHAPTER 11

SENSORY ADAPTATION

THE ENTROPIC INTERPRETATION OF ADAPTATION[1]

We continue, now, in our task of evaluating the fundamental entropy equation (9.20)/(10.1),

$$F = F(I, t) = kH = \frac{1}{2}k\ln(1 + \beta I^n/t)\ . \tag{11.1}$$

We recall from Chapter 3 ("The Weber Fraction: The Analogs") the general process that is employed in evaluating the function $F(I, t)$. In order to derive the law of sensation from $F(I, t)$, we set $t = t' =$ constant, and obtained Equation (10.3), the general equation embracing both logarithmic and power laws of sensation. Now, in order to explore the principle of adaptation using the same mathematical function, $F(I,t)$, we set $I = I' =$ constant. That is, we explore the behavior of the perceptual variable, F, and the entropy function, H, when a single, steady stimulus is applied to a sensory receptor for a period of time, t. When we deal with stimuli applied for "long" periods of time (for example, $t > 1$ s), as we do in this chapter, t may well refer to either the duration of the stimulus or to the time since onset of the stimulus, since generally these times will be equal. However, in Chapter 13, when we shall deal with very brief stimuli, t will refer unequivocally to the time since stimulus onset. For such brief stimuli, the ratio of time since onset to stimulus duration may exceed unity substantially. For the longer stimuli treated in this chapter the ratio is, effectively, unity.

Since stimulus intensity, I', is constant, we write[2]

$$\lambda = \beta(I')^n = \text{constant}, \tag{11.2}$$

analogous to (10.2).

Hence the equation

$$\boxed{F = kH = \tfrac{1}{2}k\ln(1 + \lambda/t)\ ,} \tag{11.3}$$

analogous to (10.3), describes the change in F with t for a constant stimulus. It is evident from Equation (11.3) that when t is small, F is large, and as t becomes larger, F declines. This behavior describes the process of sensory adaptation (see Chapter 3, Adaptation). It is well known that when F describes subjective magnitude, such as the sensation of taste, this magnitude will sometimes increase briefly after administration of the stimulus, and then decline. We deal with this early rise in sensation in a preliminary way in Chapter 14. Until then, we concern ourselves solely with the declining phase of the adaptation process.

There are both mathematical and physiological limitations governing the range of values taken by the time variable, t. We recall from Equation (9.3) that

$$t = m/\alpha\ ,$$

where α is a constant, greater than zero, and m is an integer, greater than zero, representing the number of samples taken by a receptor of its stimulus population. It would seem that when $m = 1$, t takes on its smallest value[3] of $1/\alpha$. At the other extreme, there is almost certainly a maximum value for m that characterizes each type of receptor. Maximum m would represent the greatest number of sampling-values (intensity values) that a receptor could retain in its local "memory." Although we do not know from anatomical or physiological considerations exactly how great this number is, we do

know that it must exist. For example, it is not likely that your pressure receptors are currently storing the values of pressures applied 15 minutes ago. That is, the receptor's memory is less than 15 minutes. Some minimum and maximum for m must exist, but we cannot, at the moment, provide values for these extremes.†

If the maximum value for m (and hence, t) is great enough, then, from Equation (11.3), F will tend toward $\ln(1) = 0$ — the case of complete adaptation, or adaptation to extinction. If, however, the maximum value for m (and hence for t) is somewhat smaller, F will fall to some level greater than zero, and will not decrease further, corresponding to the case of incomplete adaptation. *The entropy equation* (11.3) *is valid only to the time when F first reaches its minimum value*. Beyond this equilibration point, we cannot use the entropy equation in its present form.

The theoretical graph of F vs. t is governed, then, by the ln-function (11.3). For small t, the function falls steeply, and for large t, it falls gently toward zero. The characteristics of the curve differ markedly from those of the exponential function, which is often used, empirically, to describe such adaptation phenomena. The exponential function does not rise as sharply for small t, and descends more steeply for larger t.

"Adaptation" is often discussed together with "fatigue," some writers making a distinction between the terms. Adaptation, as seen from the entropy perspective, has nothing whatever to do with fatigue.

Since $F = kH$, F is a "mirror" of H. As H does, so F does. The entropic view of adaptation may be inferred from the preliminary discussion given in Chapter 8 (Central Limit Theorem) and Chapter 9 (Maximum H as Potential Information). Referring to Equation (11.3) — When t is small (m small), receptor uncertainty is great and the *potential* to receive information is high; when t is large (m large), receptor uncertainty is reduced, and thus the potential to receive information is low or absent. *Adaptation*, *therefore*, *refers to the progressive acquisition of information* (*eradication of uncertainty*). When the maximum value of t for a given receptor is insufficient to reduce the H-function to zero, adaptation will be incomplete. A receptor which has adapted to its fullest extent cannot receive further information. When a receptor has adapted completely, it retains no further uncertainty about its stimulus magnitude, leading to a rather dramatic conclusion: *A receptor cannot perceive a certainty*. If the outcome to an event is completely certain, the receptor cannot perceive it.

The reader will notice that here, as elsewhere, I seem to have anthropomorphized the receptor, relegating to it the capacity to be certain or uncertain, as if it possessed a mind. This view is, however, incomplete. A receptor, as an isolated unit, cannot reasonably be expected to possess the potential for certainty or uncertainty. I shall return later to this matter, but even then, I am afraid, my views may be considered audacious.

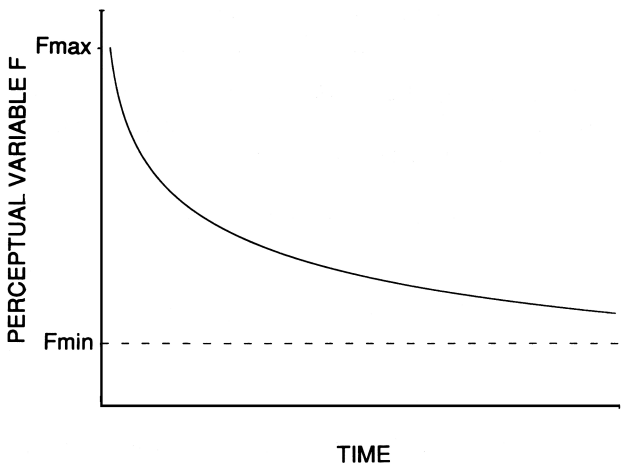

Figure 11.1 Schematic diagram of an adaptation curve. The variable, F, represents either impulse frequency in a sensory neuron or subjective magnitude of a stimulus. F_{max} is the maximum value obtainable by F. Sometimes one finds an early portion of the curve where F rises with time (not shown here). This early rising phase can be detected neurophysiologically and, occasionally, psychophysically, particularly with the chemical senses. F_{min}, the minimum value obtainable by F, is often not obtained during the course of the experiment, but rather, must be estimated by means of an asymptote to the curve, as shown in the diagram. When F_{min} is equal to zero, adaptation is said to go to completion.

The fall of F with t is illustrated schematically by the curves in Figure 3.2 and Figure 11.1. Again, F, the perceptual variable, will be interpreted either as subjective magnitude (magnitude estimates) or as impulse rate in a sensory neuron. However, we shall see in Chapter 13 that this dual interpretation of F is, at best, an approximation. That is, F (subjective magnitude) is not precisely synchronous with F (impulse rate). In fact the two types of curve are not even necessarily of exactly the same shape, as discussed in note 1 of Chapter 13.

INFORMATION, RECEPTORS AND CATEGORICAL JUDGMENTS

From Equation (11.3),

$$\boxed{H = \tfrac{1}{2}\ln(1 + \lambda / t)\,.} \tag{11.4}$$

We have established, now, the existence of a minimum and maximum value of t, which can be designated t_0 and $t_{\max}$, respectively.[4] Since $H(t)$ is a monotone decreasing function of t, therefore,

$$H_{\max} = \tfrac{1}{2}\ln(1 + \lambda / t_0) \tag{11.5}$$

and

$$H_{\min} = \tfrac{1}{2}\ln(1 + \lambda / t_{\max})\,. \tag{11.6}$$

Let

$$\Delta H = H_{\max} - H_{\min}\,. \tag{11.7}$$

Then ΔH is the difference between the receptor uncertainty at the beginning and end of the process of adaptation. It is the reduction in potential information. *That is*, ΔH *equals the quantity of information received by the receptor during the process of adaptation*. Since $F = kH$, therefore,

$$\Delta H = \Delta F / k\,. \tag{11.8}$$

That is, the *excursion* of the adaptation curve divided by the scaling constant, k, is equal to the total information received by the receptor during the adaptation process. Moreover, both quantities k and ΔF, on the right-hand side of (11.8) can be measured. Since experimental adaptation data can be fitted to Equation (11.3), we can obtain a value for the constant, k; since

$$\Delta F = F_{\max} - F_{\min}\,, \tag{11.9}$$

we can measure ΔF by taking the difference between the highest and lowest values on the adaptation curve. We can, thereby, measure ΔH using Equation (11.8) to obtain a value in natural units of information per stimulus. Examples of this technique will follow.

We might just note at this point that there is an approximation to Equation (11.7) which not only permits one to estimate the value of ΔH from visual inspection of an adaptation curve, but seems to provide some insight into ΔH as well (cf. Chapter 8, The Central Limit Theorem). Let us assume (rightly or wrongly) that we can use the "Fechner approximation," $\lambda/t_{\max} >> 1$. Admittedly, use of this approximation will sometimes lead us astray. Since $t_o < t_{\max}$, it will also be true that $\lambda / t_o >> 1$. Introducing these approximations into Equations (11.5) and (11.6),

$$H_{\max} \simeq \tfrac{1}{2}\ln(\lambda / t_o)$$

and

$$H_{\min} \simeq \tfrac{1}{2}\ln(\lambda / t_{\max})\,,$$

so that, from Equation (11.7),

$$\Delta H \simeq \tfrac{1}{2}\ln(\lambda / t_o) - \tfrac{1}{2}\ln(\lambda / t_{\max})\,,$$

$$\Delta H = \tfrac{1}{2}\ln(t_{\max} / t_o)\,,$$

or

$$\boxed{\Delta H = \ln \sqrt{t_{\max} / t_o}\ .} \tag{11.10}$$

This equation was demonstrated by the author (1981, 1984). It is useful because the right-hand side can be approximately evaluated by simple visual inspection of an adaptation curve. For example, from a curve such as that illustrated in Fig (11.1) one can estimate (sometimes) the value of t_o at which the curve is maximum, and of $t_{\max}$ at which the curve is minimum. Sometimes the curve seems to continue falling slowly throughout its duration, so I usually suggest taking $t_{\max}$ as the time when the curve has fallen through, say, 90% of its total excursion. The final result of the calculation is not all that sensitive to small changes in $t_{\max}$, but is exquisitely sensitive to changes in the estimate of t_o. Anyway, a glance at the adaptation curve provides an estimate of the ratio of $t_{\max} / t_o$, and the log of the square root of this ratio equals the information transmitted (or an approximation thereto).

If, moreover, we introduce Equation (9.3) into (11.10), then

$$\Delta H = \tfrac{1}{2} \ln(m_{\max} / m_o)\ ,$$

where m is the number of samplings made by the receptor. But since m_o is the minimum number of samplings = 1 (say), therefore

$$\Delta H \simeq \tfrac{1}{2} \ln m_{\max}\ . \tag{11.11}$$

That is, information received by the receptor is approximately equal to one-half the logarithm of the greatest number of samplings that can be contained in the memory of the receptor.

The reader will recall from Chapters 4 and 5 that one can also calculate an information per stimulus by means of an experiment on categorical judgments. Quite a lot of labor was expended to derive the quantity $\mathscr{I}(X|Y)$ from the confusion matrix. It was shown that for all modalities of sensation, a maximum of about 2.5 bits of information was transmitted per stimulus, the so-called channel capacity. This 2.5 bits or 1.75 natural units of information corresponded to $2^{2.5} = e^{1.75} \simeq 6$ categories. That is, a human being is capable of distinguishing, without error, "the equivalent of" about 6 categories of light or sound intensity, concentration of solutions, etc. We now explore the question of the relationship between this quantity of information, $\mathscr{I}(X|Y)$, obtained from measurements of categorical judgments, and the information ΔH, obtained using Equation (11.7) from measurements made on adaptation curves.

The factors limiting the maximum value of $\mathscr{I}(X|Y)$ are usually assumed to reside in the brain. Perhaps the main reason for this assumption is that the number of jnd's distinguishable (range of 20 – 350) is much greater than the number of absolute categorical judgments (about 6). Since the senses can distinguish many more than 6 jnd's, the limitation in making an absolute judgment about an unknown stimulus must, surely, lie in the brain. And so it may be. But I would like to suggest an alternative. I suggest that the limitation on the quantity of information available for an absolute judgment may be, in the final analysis, due to the limitation in the amount of information provided by the sensory receptors. That is, if the sensory receptors can receive a maximum of ΔH bits of information per stimulus, then no more than ΔH bits can be used to make an absolute judgment. Implicit in this suggestion is that, to a degree of approximation, the sensory receptors operate in parallel; for example, the amount of information received by the brain from n olfactory receptors is the same as that received from only one receptor. That is, redundancy (numerosity) in receptors may be necessary to insure receipt of a stimulus, but once received, the information from the large number is the same as the information from only one receptor. Some psychologists have expressed dissatisfaction with this hypothesis. Ward (1991) has put forward an alternative hypothesis: N receptors acting in parallel and sampling at rate α samples/s may produce the effect of a single adaptation process with sampling at the rate $N\alpha$. Such may be the case, but further work must be done on it.

I submit, anyway, by way of conjecture, that

$$\Delta H = \mathscr{I}(X|Y)\ . \tag{11.12}$$

The truth of Equation (11.12) is not vital to the integrity of the entropy theory of perception, but we shall see presently that (11.12) does seem to verify when tested on experimental data — which we now proceed to do.

EXPERIMENTAL TESTS OF THE ENTROPIC THEORY OF ADAPTATION

Let us now test the entropic theory, to the level that we have developed it hitherto, using published data. Gent and McBurney (1978) measured the change in magnitude estimate, F, with time, for the sense of taste. Let us select from their paper a medium (0.32 M) solution of sodium chloride. I digitized, as carefully as possible, the data from Gent and McBurney's Figure 1. When Equation (11.3) was fitted to the data using a least squares technique, the following parameter values were obtained:

$$k = 7.634 ,$$
$$\lambda = 46.04 \text{ s}.$$

The fitted equation was, therefore,

$$F = (7.634/2)\ln(1 + 46.04/t) . \tag{11.13}$$

The results are graphed in Figure 11.2. The fitted curve, while perhaps tolerable, seems to fall a little too slowly at larger times, a feature characteristic of all data in this paper.[5] Gent and McBurney fitted the curves to monoexponential functions.

Using the above parameter values, we can measure the information transmitted per sodium chloride stimulus. From Equation (11.9),

$$\Delta F = F_{\max} - F_{\min} = 8.57 - 1.00 = 7.57 , \tag{11.14}$$

if we consider the adaptation process to be complete at 90 seconds. The excursion, ΔF, would equal 8.57 – 0 = 8.57, if we suspect that the adaptation process would proceed to extinction. Therefore, from Equation (11.8), using the smaller of the two values for ΔF, the information transmitted by this solution of medium intensity

$$\Delta H = \Delta F / k = 7.57 / 7.634 = 0.992 \text{ natural units.}$$

Dividing by $\ln 2$,

$$\Delta H = \Delta F / (k \ln 2) = 1.431 \text{ bits}, \tag{11.15}$$

a value slightly less than the channel capacity for taste (1.7 – 2.0 bits). Or, using the larger of the two

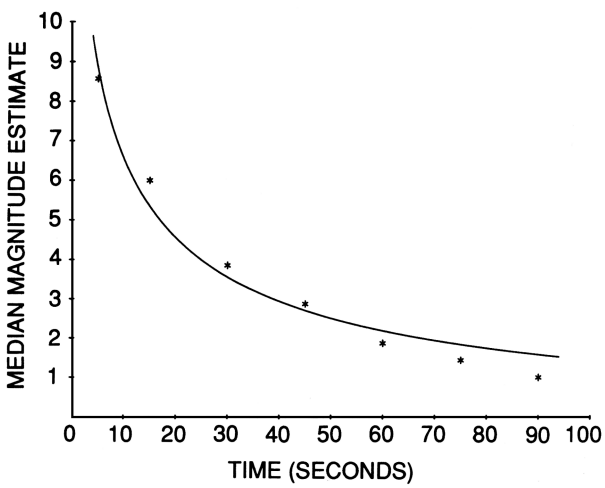

Figure 11.2 Data of Gent and McBurney (1978) illustrating psychophysical adaptation. Mean magnitude estimate of 0.32 M solution of sodium chloride plotted against time. The smooth curve is given by the entropy equation (11.13).

Table 11.1 Data of Matthews (1931), Figure 3: Two-Gram Weight Applied to Muscle of Frog.

Time [s]	Impulse frequency [s^{-1}] fitted	Impulse frequency [s^{-1}] measured
0.228	117.47	118.98
0.411	101.68	97.79
0.826	83.28	84.63
1.277	72.11	74.80
1.658	65.56	65.72
2.059	60.24	59.86
3.073	50.76	50.73
4.009	44.78	44.82
5.068	39.74	39.36
6.048	36.13	34.40
7.052	33.11	32.51
8.091	30.53	29.44
9.056	28.50	27.80
10.030	26.74	26.56
11.080	25.07	24.62
12.050	23.71	24.62
13.060	22.45	23.63
14.050	21.36	23.33

Note: Data were fitted to the entropic adaptation equation (11.3) using a simplex procedure with the least squares criterion (see Figure 11.3). The number of digits used for a given entry does not reflect the number of significant figures in the measurement, a quantity that is not really known to the author. Function fitted is $F = 0.5 \times (\text{Parameter } 1) \times \ln(1 + (\text{Parameter } 2)/t)$. Value of Parameter 1 = 54.51. Value of Parameter 2 = 16.71. Sum of squares of residuals = 38.22

values for ΔF,

$$\Delta H = 8.57/(7.634 \ln 2) = 1.620 \text{ bits.} \tag{11.16}$$

Moving now to an example in which F is measured as impulse frequency in a sensory neuron, we examine the data of B.H.C. Matthews (1931). Matthews investigated the response to stretch receptors in the muscles of a frog. He provided stretch or tension stimuli to a small muscle in the upper, outer side of the middle toe of the frog, and recorded from the lateral branch of the peroneal nerve. A 2-gram load was applied to the muscle tendon, which was immersed in Ringer's solution, and the resulting impulse frequency plotted against time is shown in the Matthews' Figure 3. The entropy equation for adaptation, (11.3), was fitted to the data digitized from this graph.[6] The parameter values obtained were as follows (Table 11.1):

$$k = 54.51 \,,$$
$$\lambda = 16.71 \text{ s.}$$

The fitted equation was, therefore,

$$F = (\tfrac{1}{2})(54.51) \ln(1 + 16.71/t) \,. \tag{11.17}$$

From Equation (11.9)

$$\Delta F = F_{\max} - F_{\min} = 119 - 23 = 96 \text{ impulses/s.} \tag{11.18}$$

From Equation (11.8),

$$\Delta H = \Delta F/(k \ln 2) = 96/(54.51 \ln 2) = 2.541 \text{ bits.} \tag{11.19}$$

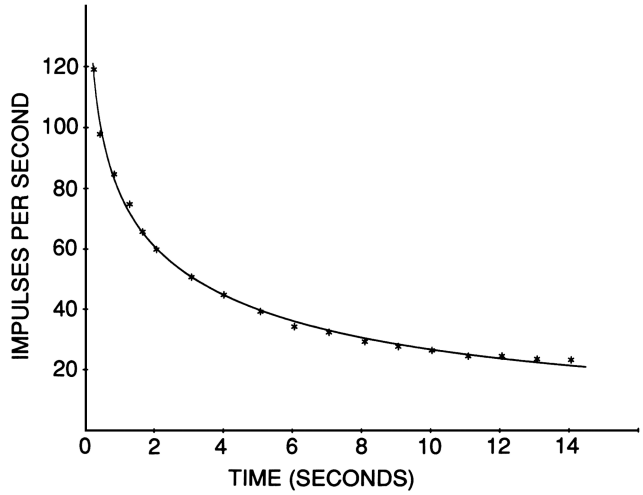

Figure 11.3 Data of Matthews (1931), Figure 3 illustrating neural adaptation. Two-gram load applied to a muscle tendon of the frog. The smooth curve is given by the entropy equation (11.17).

Again we find that the value for information transmitted per stimulus is close the the "global" value of about 2.5 bits. Since $2^{2.541} = 5.82$, we see that we remain close to the "magical number 6" categories. Matthews' data and its fitted curve are shown in Figure 11.3. Using the very approximate Equation (11.10) with data from Table 11.1, we have

$$\Delta H \simeq \tfrac{1}{2}\ln(14.05/0.228)/\ln 2 = 2.97 \text{ bits.} \tag{11.19a}$$

Continuing with examples where F is measured neurophysiologically, we consider adaptation in the guinea pig auditory nerve, as reported by Yates, Robertson and Johnstone (1985) (see Chapter 3, Adaptation). The responses in auditory nerve ganglion cells to 100 ms tone bursts were measured by these investigators. Adaptation was incomplete; that is, firing rate did not descend to zero. Judging from results of experiment GP53/08:4, the decline in firing rate was complete by 25 to 30 ms following onset[††] of the stimulus. To use the graphs in this paper, probability density was interpreted as firing rate, as suggested by the authors. These data (Yates *et al.*, Figure 5) were digitized and fitted to the entropy equation (11.3). The values of the parameters were as follows:

$$k = 157.7\,,$$
$$\lambda = 37.64 \text{ ms.}$$

Thus, the fitted equation was

$$F = (\tfrac{1}{2})(157.7)\ln(1 + 37.64/t)\,. \tag{11.20}$$

From Equation (11.9)

$$\Delta F = F_{\max} - F_{\min} = 284 - 67 = 217\,. \tag{11.21}$$

From Equation (11.8)

$$\Delta H = \Delta F/(k\ln 2) = 217/(157.7\ln 2) = 1.99 \text{ bits.} \tag{11.22}$$

The data and fitted curve are shown in Figure 11.4.

In passing, may I draw your attention to the older neurophysiological data on audition by Galambos and Davis (1943), also introduced in Chapter 3. Pure tones were delivered to the ears of cats, and the frequency of impulses was measured in what were believed originally to be single auditory fibers (1943 paper). However, in a note of correction (1948), the authors amended the interpretation of their measurements, and attributed them to single unit activity of second-order neurons, rather than first-order neurons, as originally expected. The data of Galambos and Davis were digitized from their graph (Figure 3), and published in Table 2 of Norwich (1981). The entropic equation (11.3) was fitted

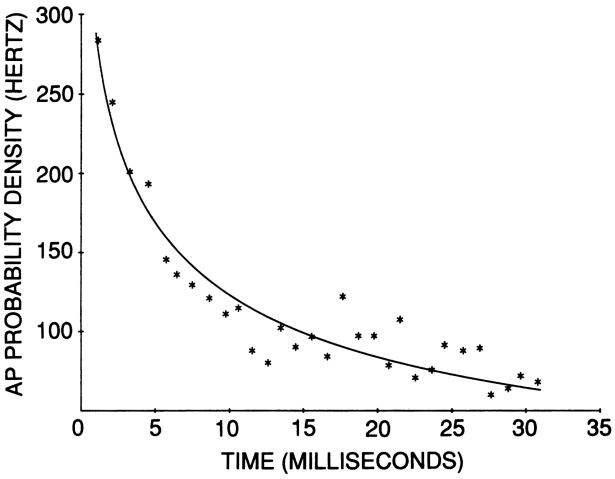

Figure 11.4 Data of Yates *et al.* (1985). Neural adaptation in the auditory ganglion cell of a guinea pig to a 100 ms tone burst. F is measured as action potential probability density [Hz], similar to impulse frequency. The smooth curve is given by the entropy function (11.20). To estimate the information transmitted by this adaptation process using Equation (11.10), we observe that $t_{\max} \simeq 25$ ms and $t_o \simeq 1$ ms. Thus, $\Delta H \simeq \ln\sqrt{25} / \ln 2 = 2.3$ bits of information per stimulus.

as closely as possible to their data. The following parameter values were obtained:

$$k = 120.31\,,$$
$$\lambda = 4.00 \text{ s.}$$

Hence

$$\Delta F = 400 - 80 = 320 \text{ Hz} \tag{11.23}$$

$$\Delta H = 320/(120.31 \ln 2) = 3.84 \text{ bits.} \tag{11.24}$$

Analogous to auditory nerve fiber response to acoustic stimuli, we examine now optic nerve response to visual stimuli. Cleland and Enroth-Cugell (1968) applied square-wave inputs of light to the

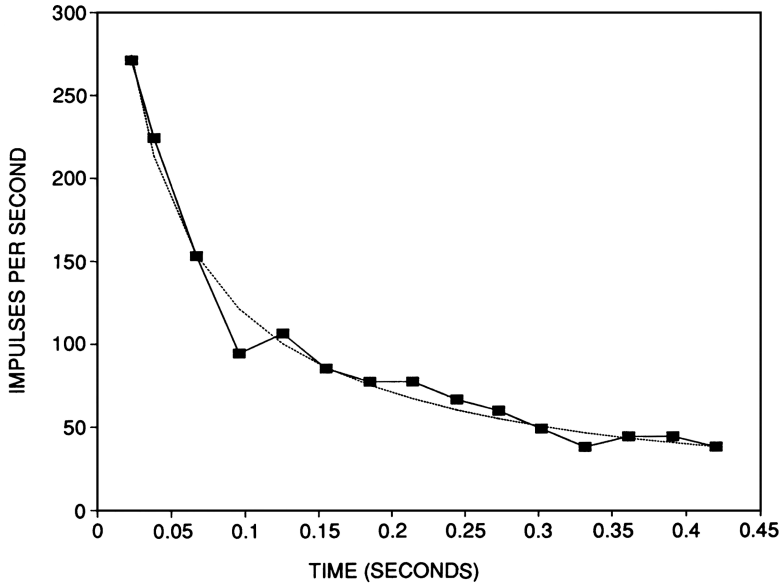

Figure 11.5 Data of Cleland and Enroth-Cugell (1968). Neural adaptation in the on-center ganglion cells of the cat to square-wave inputs of light to the retina. The smooth curve is given by the entropy equation (11.3) with $k = 302.5$ and $\lambda = 0.1175$ s.

retina of cats, and recorded the impulse frequency of on-center ganglion cells in the cat retina from optic tract fibers. A number of their tracings for different light intensities were provided in the authors' Figure 2, showing the effects of adaptation. Their tracing #2 has been digitized, as well as we were able, and the entropy equation for adaptation (11.3) was fitted to the data (Figure 11.5). The exact position of $t = 0$ for these data was hard to determine. I took the liberty here of shifting the origin back and forth by a few milliseconds in order to produce the best curve fit. The values for the parameters were as follows:

$$k = 302.5\,,$$
$$\lambda = 0.1175 \text{ s.}$$
$$\Delta F \geq 271.0 - 37.6 = 233.4 \tag{11.25}$$
$$\Delta H \geq 233.4/(302.5 \ln 2) = 1.11 \text{ bits.} \tag{11.26}$$

I use ">" because the adaptation process is continuing at 0.42 seconds, where the graph in Figure 11.5 terminates. Interpretation of these results is not straightforward. We are decidedly *not* dealing here with independent sensory receptors (see, for example, Dowling, 1987, page 35-36). Moreover, the input was not a single step function but a square wave, so that receptors were not totally unadapted at $t = 0$, the time of stimulus onset.

EXPERIMENTAL TESTS OF THE ENTROPIC THEORY: COMBINED STIMULUS INTENSITY AND ADAPTATION

In Chapter 10, we subjected the entropy equation with t = constant [Equation (10.3)] to various experimental tests, and we confirmed the fact that the WF and PBS laws merge and are contained within the entropy equation. In the preceding section of the present chapter we subjected the entropy equation with I = constant [Equation (11.3)] to a number of experimental tests. We confirmed the adequacy of the equation to fit adaptation data, although in some instances, particularly with magnitude estimates, one would be happier with a closer fit.[5] We also observed that the values of information transmitted per stimulus calculated from adaptation data were compatible with corresponding values calculated from experiments on categorical judgments. Henceforth, it will be assumed that the "channel capacity," whether determined by entropic or category analysis, will be about 2.5 bits of information per stimulus.

Let us proceed, now, and subject the *general* equation of entropy, (9.20)/(10.1)/(11.1), to experimental test. That is, we let *both I and t* vary,[7] and study the *surface*,

$$F = F(I, t)\,. \tag{11.27}$$

That is, we may regard I and t as two independent variables and F as a dependent variable. The graph of F as a function of I and t will describe a surface. However, for simplicity in presenting the data, we shall take "slices" through the surface for $I = I_1$, I_2, I_3, and present a series of two-dimensional graphs of F vs. t.

Let us consider first the older work of Matthews (1931) on stretch receptors, whose experiments were described above. Instead of dealing with only one experiment in which a 2-gram force was applied, we now consider, together, three experiments in which forces of 1, 2, and 5 grams were applied to the tendon. Each force gave rise to its own, unique adaptation curve. The three sets of data were fitted simultaneously to the general entropy equation (11.1), each set being assigned equal weighting. The total sum of squares was minimized to give values for the three parameters. Data analyzed were taken from Matthews' Figure 15B (in Ringer's solution) which, I am assuming, were all obtained from the same preparation (or, at least, are governed by the same set of parameters). The parameter values are:

$k = 32.24$,
$\beta = 17.19$ (cf. $\beta = 54.51$ for the preparation analyzed in Figure 11.3)
$n = 0.9496$.

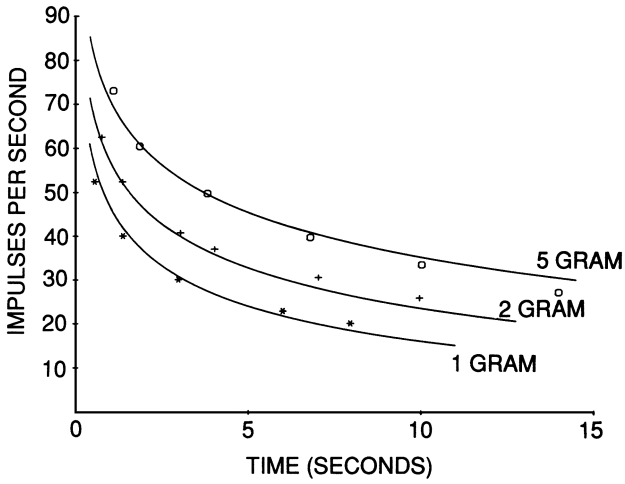

Figure 11.6 Data of Matthews (1931), Figure 15B. Frog stretch receptor. All three curves were fitted simultaneously to the same set of 3 parameters (given in the text). That is, only 3 adjustable parameters were used, in all, to curve-fit all 3 curves, or a ratio of one parameter per curve. The fitted entropy function is given by Equation (11.28).

The results of the three simultaneous curve fits are shown in Figure 11.6. Remember that a *total* of 3 parameters were used to curve-fit the 3 sets of data, or a ratio of one parameter per curve. Equation (11.1) is represented explicitly by

$$F = (\tfrac{1}{2})(32.24)\ln(1 + 17.20\, I^{0.9496})\ . \tag{11.28}$$

Calculating the channel capacity using the 5-gram data set,

$$\Delta F = 72.9 - 27.2 = 45.7 \text{ impulses per second.} \tag{11.29}$$

$$(\Delta H = 45.7/(32.34\ln 2) = 2.04 \text{ bits per stimulus.} \tag{11.30}$$

The three data sets and the corresponding parameter values are given in Table 11.2.

As a second example of a surface-fit to Equation (11.1), and the first example of insect sensation, we consider the sugar receptor of the blowfly. Dethier and Bowdan (1984) stimulated this receptor with sucrose solutions of 3 different concentrations: 1.0 M, 0.1 M and 0.01 M. They measured the frequency of impulses in bipolar neurons in tarsal hairs. In this way, 3 sets of data were obtained showing the adaptation of the sugar receptor. The 3 sets of data were fitted simultaneously to the general entropy equation (11.1), with weightings in the ratio of 1:3:5 (1.0 M, 0.1 M, 0.01 M). With parameters evaluated, Equation (11.1) became

$$F = (\tfrac{1}{2})(121)\ln(1 + 1.15\, I^{0.585}/t)\ . \tag{11.31}$$

The data and fitted surface (curves) are shown in Figure 11.7. Again we note that *all three curves* have been fitted with the *same* 3 parameter values, or, the average ratio of one parameter per curve. Maximum information per stimulus was calculated to be 2.9 bits per stimulus. More details about these curves are given by Norwich and Valter-McConville (1991). In this same paper the reader may find more extensive mathematical tests of the adaptation principle and a simultaneous curve-fit to adaptation data from the slit sense organ on the walking leg of the Hunting Spider.

We can now understand the near-parallel straight lines of Schmidt (1988, page 88), introduced in Chapter 3. The logarithm of the neural response of a pressure receptor plotted against the logarithm of the stimulus pressure (for t constant) gives a straight line (power law of sensation). As the state of adaptation increases (t increases), the straight line shifts downward on the graph. The result is a series of nearly parallel straight lines (Figure 3.3). We proceed by expanding the right-hand side of Equation (11.1) in a Taylor series (where t is large enough to make the expansion legitimate), as in Equation (10.7). Retaining only the first order term,

$$F = \tfrac{1}{2} k\beta\, I^n/t\ . \tag{11.32}$$

Table 11.2 Data of Matthews (1931), Figure 15B. Three Experiments in Which a 1-gram, 2-gram and 5-gram Weight Were Applied, Respectively, to the Muscle of a Frog.

t [s]	Frequency Fitted	Frequency Measured	Data set
			1 gram
0.548	56.06	52.47	
1.373	41.98	40.06	
2.975	30.85	30.12	
5.992	21.81	22.95	
7.955	18.55	20.18	
			2 gram
0.746	61.55	62.48	
1.354	52.22	52.43	
3.036	39.97	40.82	
4.021	35.87	37.10	
7.030	28.12	30.61	
9.944	23.66	26.02	
			5 gram
1.077	69.51	72.91	
1.845	60.98	60.32	
3.808	49.69	49.73	
6.787	40.94	39.64	
10.01	35.27	33.46	
13.97	30.60	27.20	

Note. The three data sets were curve-fitted *simultaneously* to Equation (9.20)/(10.1)/(11.1) *using the same three parameters for all three curves* (Figure 11.7). The equation is

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

Intensity, I, was set equal to 1, 2 and 5 respectively in each of the three curve-fits. The above equation was then fitted to the F vs. t data. Function fitted is $F = 0.5 \times$ (Parameter 1) $\times \ln(1 +$ (Parameter 2) $\times$ I ^ (Parameter 3)$/t$). Value of Parameter 1 = 32.24. Value of Parameter 2 = 17.19. Value of Parameter 3 = 0.9496. Sum of squares of residuals = 64.50.

The receptor can be represented in different states of adaptation by setting $t = t_i$ = constant. Taking logs of both sides of (11.32),

$$\log F = n\log I + \log(\tfrac{1}{2}k\beta) - \log t_i\ . \tag{11.33}$$

That is, plotting $\log F$ vs. $\log I$ gives a straight line whose slope is n and whose intercept, $\log(\frac{1}{2}k\beta) - \log t_i$, slides down as t_i increases, exactly as shown in Figure 3.3. As an exercise, the reader might like to check the theoretical values of the spacing between the straight lines in Figure 3.3. Take I as constant in Equation (11.33), and calculate the expected displacement of each line from the uppermost line using the t-values on the right-hand side of the graph. Compare with the observed displacement.

I think that the point is amply illustrated: Equation (9.20)/(10.1)/(11.1) is capable of allowing for both adaptation and stimulus magnitude changes using the same set of parameter values, when F is measured neurophysiologically. The information transmitted is usually close to the nominal value of 2.5 bits per stimulus.

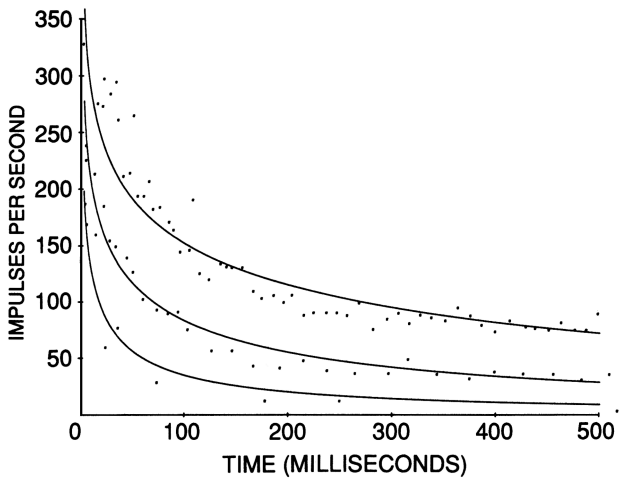

Figure 11.7 Data of Dethier and Bowdan (1984) for the sugar receptor of the blowfly. For the upper curve, the concentration of the sugar solution was 1.0 M (i.e. $I = 1.0$); for the middle curve, 0.1 M ($I = 0.1$); and for the lowest curve, 0.01 M ($I = 0.01$). Again, all three curves were fitted simultaneously to the same set of 3 parameters; that is, only 3 adjustable parameters were used, in all, to fit all 3 curves, or a ratio of one parameter per curve. The fitted entropy function is given by Equation (11.31).

MAGICAL NUMBERS FROM THE LAW OF SENSATION

You will recall from our studies in Chapter 10 that it was not always possible to obtain three robust parameters from Equation (10.3) using measurements of F vs. I (t constant), but that we did manage to do so in a number of cases. I suggested that there was magic in the values obtained for the parameter, k, and we are now in a position to be appropriately enchanted.

For the sodium chloride data of S. S. Stevens (Figure 10.2), we obtained the value $k = 41.31$. The values of $F_{\max}$ and $F_{\min}$ can be read directly from the graph: $F_{\max} = e^{3.91} = 50.0$ and $F_{\min} = e^{-0.288} = 0.75$. Hence we can calculate ΔF from this experiment in the same manner as for an adaptation curve. As in Equation (11.9),

$$\Delta F = 50.0 - 0.75 = 49.25 . \tag{11.34}$$

Using Equation (11.8)

$$\Delta H = 49.25/(41.31 \ln 2) = 1.72 \text{ bits.} \tag{11.35}$$

We observe that this value is quite close to the value of 1.62 bits of information per stimulus obtained by analysis of the taste adaptation data of Gent and McBurney for sodium chloride. It is also close to the value of 1.70 bits per stimulus for $\mathscr{I}(X|Y)$ (sodium chloride) obtained from an experiment on categorical judgments (Beebe-Center *et al.*, 1955).

For the auditory data of Luce and Mo (Figure 10.3), $k = 113.1$.

$$F_{\max} = e^{4.96} = 142.6 , \qquad F_{\min} = e^{3.18} = 24.0 .$$

$$\Delta F = 118.6 . \tag{11.36}$$

$$\Delta H = 118.6/(113.1 \ln 2) = 1.51 \text{ bits.} \tag{11.37}$$

As much as one would like to supply limits of variability for such calculations, I know of no way to calculate these limits.

From Equation (11.22), the information transmitted in the guinea pig ganglion cell was 1.99 bits. Category experiments on human audition will give values between 1.62 and 2.51 bits per stimulus (Garner 1953).

For the weight-lifting data of Luce and Mo (Figure 10.3), $k = 1040$.

$$F_{\max} = e^{7.25} = 1408\,, \qquad F_{\min} = e^{3.22} = 25.0\,,$$

$$\Delta F = 1408.0 - 25.0 = 1383\,. \tag{11.38}$$

$$\Delta H = 1383/(1040\ln 2) = 1.92 \text{ bits.} \tag{11.39}$$

I have no corresponding values for $\mathscr{I}(X|Y)$ for comparison but, again, the value is close to the expected global value for channel capacity.

Finally, from Chapter 10, the sweetness-of-sucrose data of Moskowitz gave

$$k = 24.6\,, \qquad F_{\max} = 45.0\,, \qquad F_{\min} = 0.80\,.$$

$$\Delta F = 44.2\,, \tag{11.40}$$

$$\Delta H = 44.2/(24.6\ln 2) = 2.59 \text{ bits,} \tag{11.41}$$

a little high for taste but, again, these data were quite approximate.

Again, the reader is reminded that data measured to study the law of sensation usually do not permit estimation of 3 robust parameters, and, therefore, do not usually permit calculation of channel capacities.

NUMERO, NUMERO, WHEREFORE ART THOU NUMERO?[8]

What function, you might wonder, do numbers have in science?

What is the difference, I put to you, between a physicist and an engineer?

Well, you may be inclined to answer, the physicist deals with pure science and the engineer with applied science.

While that is true, I believe that there is another more salient distinction referable to their respective uses of mathematics. The engineer uses mathematics to obtain (often) a numerical result: the weight tolerance of a bridge, the characteristics of a circuit. The physicist uses mathematics to obtain (often) a "verbal" or non-numerical result: energy and mass are, in principle, the same; the universe began from an explosion at a single point; etc. As I stated before, I am writing this material with the Weltanschauung of the physicist. Therefore, to me, the pages of numbers given above are not end-values in themselves. They are useful only in helping to confirm a position which can be stated in words, namely that a single principle of entropy, *F is a measure of H* ($F = kH$), seems to account for all stimulus-magnitude and all adaptation experiments in which a single, constant stimulus is used. As we move forward in this book, we shall see that the scope of this principle is even broader.

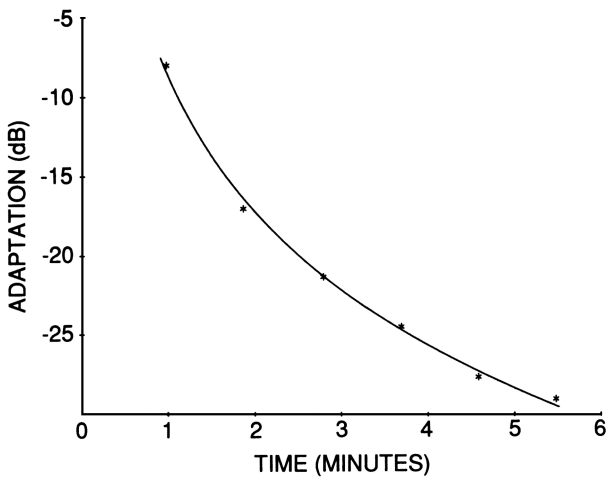

Figure 11.8 Data of Small and Minifie (1961), Figure 3(d), 50 seconds off and 10 seconds on: simultaneous dichotic loudness balance. Data are plotted here in the usual way: decibels of adaptation are plotted against time. The data fall on a curve.

AUDITION: THE TECHNIQUE OF SIMULTANEOUS DICHOTIC LOUDNESS BALANCE (SDLB)

This technique was described in Chapter 3, Adaptation. A tone of constant intensity, I_a, is presented to the adapting ear, beginning at $t = 0$. An intermittent tone of variable intensity, I, is presented to the opposite or test ear for total duration, t'. The subject must adjust the intensity of the intermittent, test tone until its loudness matches that of the steady, adapting tone.

The application of the entropy principle to the SDLB technique is very simple (if not perfectly accurate). Let us assume that both ears of a given subject have identical values for the parameters k, β and n. If the loudness of the test tone is equal to that of the adapting tone then

$$\tfrac{1}{2}k\ln(1+\beta I^n/t') = \tfrac{1}{2}k\ln(1+\beta I_a^n/t). \tag{11.42}$$

Hence

$$\begin{aligned} I^n/t' &= I_a^n/t \\ n\log_{10}I - n\log_{10}I_a &= \log_{10}t' - \log_{10}t \\ 10\log_{10}(I/I_a) &= (10/n)\log_{10}t' - (10/n)\log_{10}t\,. \end{aligned} \tag{11.43}$$

But

$$10\log_{10}(I/I_a) = \text{ decibels of adaptation,}$$

so that

$$\text{decibels of adaptation} = (10/n)\log_{10}t' - (10/n)\log_{10}t\,. \tag{11.44}$$

Therefore, if we plot dB of adaptation against $\log_{10}t$, we expect to obtain a straight line whose slope equals $-10/n$, where n is the power function exponent from the power law of sensation.

In the various papers describing experiments of the SDLB type, dB of adaptation is plotted against t (not log t), and is seen to produce a curve, as shown in Figure 11.8 (data from Figure 3d, 50 seconds off, 10 seconds on, of Small and Minifie, l961). If, however, we re-plot the data with a $\log_{10}t$–scale instead of a t-scale, the data are seen to lie nearly along a straight line, as shown in Figure 11.9. Fitting

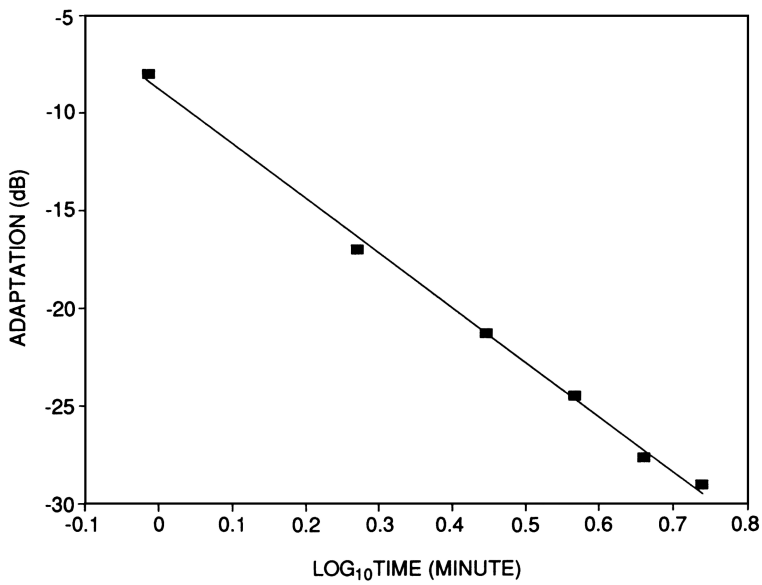

Figure 11.9 Same data as in Figure 11.8. However, decibels of adaptation are now plotted against the log of time. The data tend to lie on a straight line. Then −10/slope of line gives the power function exponent. The intercept with the line log $t = 0$ gives the "on time," in principle.

a straight line to these data, we obtain

$$\text{decibels of adaptation} = -8.816 - 27.98 \log_{10} t \,. \tag{11.45}$$

Comparing (11.44) with (11.45) we have

$$(10/n) = 27.98 \,, \qquad n = 0.36 \,. \tag{11.46}$$

This value for n is quite close to the expected value for the power function exponent for sound intensity at 4000 Hz, the frequency used by Small and Minifie. We are not quite so accurate, however, when we test the intercept. Again comparing (11.44) with (11.45),

$$(10/n) \log_{10} t' = -8.816 \,. \tag{11.47}$$

Inserting the value of n from (11.46),

$$\log_{10} t' = (-8.816)(0.36)/10 = -0.317$$
$$t' = 0.482 \text{ minutes } = 29 \text{ seconds.} \tag{11.48}$$

However, the experimental "n time" for this graph was only 10 seconds. Other graphs obtained from SDLB experiments gave somewhat closer estimates of t'. But one must remember that Equation (11.42), from which these computations were made, assumed that the test ear was always completely unadapted, which was not the case experimentally. I think that the predictions made by the theory are reasonable.

Loudness adaptation does not occur, or is not prominent, monaurally, as discussed by Scharf (1978, p. 219). Yet adaptation of one ear with respect to another is prominent. Such an effect can be explained if k or β for audition were slowly increasing functions of time. For example, β/t might tend to remain constant. We may observe, though, that Equation (11.43) can be obtained from Equation (11.42) even if k and β are time-varying parameters, provided that β[left ear] = β[right ear], etc.

The reader is encouraged to carry out more of these SDLB analyses, and perhaps to refine Equation (11.42). The mathematical analyses are quite straightforward, requiring only a regression line. Don't forget: "decibels of adaptation" are negative numbers as we calculate them.

CONCLUDING REMARKS

I have presented rather a new view of the process of sensory adaptation. Eschewing the more traditional ideas such as the association of adaptation with fatigue, I have represented adaptation as the process of acquisition of information. As adaptation proceeds, information is gained progressively. When adaptation is complete, we cease to sensate[9] because we have received all possible information from a particular stimulus. We might extend this idea beyond simple stimuli of the intensity type. When an image is fixed in constant position on the retina, the image fades or "grays out" (Troxler's effect). Perhaps this is another example of the principle: *To perceive is to receive information*; *in the absence of new information*, *sensation desists*. Troxler's effect has, recently, been simulated by a "silicon retina" as an adaptation process (Mahowald and Mead, 1991).

Another view of adaptation as gain in information was presented by Keidel *et al*. (1961).

In the process of deriving Equation (11.3), we have held I constant, and set $\lambda = \beta I^n$ = constant in Equation (11.2). In so doing, we have removed assumption (6) from the set of six assumptions listed in Chapter 9.

Equation (11.3) was descended from its ancestor,

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

You will recall that stimulus population variance, σ_S^2, was divided by m, the number of samplings made by the receptor since the time of introduction of the stimulus, to provide a reduced variance, σ_S^2/m, by virtue of the central limit theorem. That is, the receptor was modeled as an averaging device. Hence adaptation as reduction of uncertainty. However, mathematically, we obtain the same Equation

(9.10) if we associate each sampling with the population variance, σ_S^2, and regard the reference variance, σ_R^2, as increasing with the number of samplings. For example, by the rule of addition of variances, we may obtain m times the error squared after m measurements: hence $m\sigma_R^2$. Perhaps the reader prefers to model the sampling process in this manner, but I find it less comprehensible.

Our fundamental equation of adaptation, (11.3), is in logarithmic form, but for all except the smallest values of t, the F-function can be expanded in a Taylor series [cf. Equation (11.32)], in the usual way, to give a power function. Retaining only the first term of the series, we have

$$F = \tfrac{1}{2} k\lambda\, t^{-b}\,, \tag{11.49}$$

where $b = -1$. It is of interest that in recent years, insect physiologists have been curve-fitting adaptation curves by power functions of the form of (11.49), but usually with fractional values for the constant b, such as $b = -0.31$ (Chapman and Smith, 1963; Thorson and Biederman-Thorson, 1974; Mann and Chapman, 1975). In re-analyzing some of their data, our research group has found that better curve-fits are sometimes obtained using $b = -1$. See, for example, the analysis of the data of Bohnenberger (1981) as presented by Norwich and McConville (1991). You will recall that our value, $b = -1$, arose originally from the "null conjecture" of linear sampling rate, expressed by Equation (9.3). There is, of course, no reason why b should not be retained as an additional parameter; it's just that there is, perhaps, not sufficient evidence to support abrogation of the null conjecture.

Finally, a remark about the synchrony of the process of adaptation with the time interval of exposure to the stimulus. My original conjecture (Norwich 1981) was that the neurophysiological adaptation process (which we have analyzed above) was synchronous with the exposure of the receptor to the stimulus environment. That is, as the receptor sampled, so the neuron fired. Nature, however, was not to be this simple. Experiments carried out by L. Ward (1991) have shown that the sampling process (at least for vision) is completed exceedingly rapidly. In the order of one millisecond, human perceivers have obtained nearly the full complement of information from a flash stimulus, as measured later in an experiment on categorical judgments. In contrast, however, the neurophysiological adaptation curve may *begin* at about one millisecond, and proceed for many milliseconds, as seen from the guinea pig data of Cleland and Enroth-Cugell (1968). It is as if the retina takes a "photograph" of the flash using a very rapid "shutter," and this photograph is later scanned by the more slowly-reacting nervous system. The scanning comprises the sampling process which is responsible for adaptation to the stimulus. But the above is pure conjecture.

NOTES

1. We shall confine the discussion of adaptation to changes in the perceptual variable, F, with time. That is, we shall not be concerned here with changes in threshold that accompany adaptation.

2. I am resisting the temptation to represent the constant βI^n by the Greek letter τ, even though it has the dimensions of time. τ is too easily misread as t.

3. It is arguable that m may not take on values less than 2 in Equation (11.3). Remember that, fundamentally, Equation (11.3) comes from Equation (9.12). *If* the receptor is utilizing sample variances as estimators of σ_S^2, then at least 2 samples are needed. Then $m \geq 2$.

4. We shall be using t_o fairly frequently, so I used "t_o" rather than the longer "$t_{\min}$."

5. When one allows explicitly for a threshold value of ΔH, as given later by Equation (14.4), the rate decline of the theoretical curve is found to match the observed data much more closely.

6. The digitization performed here and elsewhere in this book was usually carried out using a digitizing site mounted on a Hewlett Packard plotter (which was also used to draw many of the graphs). A computer program was written which facilitated the process and permitted digitization from log or linear scales.

7. We actually carried out a similar exercise at the end of Chapter 10, when we studied the total number of impulses recorded over a period of time.

8. *Wherefore*? means *why*? or *for what reason*? Juliet was not searching for Romeo from her balcony! Refer to commentary on Shakespeare's play.

9. The Oxford English Dictionary recognizes the intransitive verb, *sensate*, as an obsolete form. However, I recommend its revitalization for use in the field of sensation and perception. *He ceased to sense while under anesthesia* seems less satisfactory than *He ceased to sensate while under anesthesia*.

——————— 2003 ed. notes:

†. The Universal model however, does obviate the problem of maximum *m* by establishing, effectively, a "forgeting function": the receptor memory can retain only a certain number of bits of information.

††. Probably not. Judging from the data of Kiang *et al*. (Research Monograph 35, The MIT Press, Cambridge Mass., 1965.), the adaptation would continue at a slow rate for some time, perhaps in excess of one minute.

REFERENCES

Beebe-Center, J.G., Rogers, M.S. and O'Connell, D.N. 1955. Transmission of information about sucrose and saline solutions through the sense of taste. *Journal of Psychology*, **39**, 157-160.

Chapman, K.M. and Smith, R.S. 1963. A linear transfer function underlying impulse frequency modulation in a cockroach mechanoreceptor. *Nature*, **197**, 699-700.

Cleland, B.G. and Enroth-Cugell, C. 1968. Quantitative aspects of sensitivity and summation in the cat retina. *Journal of Physiology*, **198**, 17-38.

Dethier, V.G. and Bowdan, E. 1984. Relations between differential threshold and sugar receptor mechanisms in the blowfly. *Behavioral Neuroscience*, **98**, 791-803.

Dowling, J.E. 1987. *The Retina: An approachable Part of the Brain*. The Belknap Press of Harvard University Press, Cambridge, Mass.

Galambos, R. and Davis, H. 1943. The response of single auditory-nerve fibers to acoustic stimulation. *Journal of Neurophysiology*, **6**, 39-57.

Galambos, R. and Davis, H. 1948. Action potentials from single auditory-nerve fibers? *Science*, **108**, 513.

Garner, W.R. 1953. An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology*, **46**, 373-380.

Gent, J.F. and McBurney, D.H. 1978. Time course of gustatory adaptation. *Perception and Psychophysics*, **23**, 171-175.

Keidel, W.D., Keidel, U.O. and Wigand, M.E. 1961. Adaptation: Loss or gain of sensory information? In: *Sensory Communication*, W.A. Rosenblith, Ed. MIT Press and Wiley, New York.

Mahowald, M.A. and Mead, C. 1991. The silicon retina. *Scientific American*, **May** 76-82.

Mann, D.W. and Chapman, K.M. 1975. Component mechanisms of sensitivity and adaptation in an insect mechanoreceptor. *Brain Research*, **97**, 331-336.

Matthews, B.H.C. 1931. The response of a single end organ. *Journal of Physiology*, **71**, 64-110.

Norwich, K.H. 1981. The magical number seven: Making a "bit" of "sense". *Perception and Psychophysics*, **29**, 409-422.

Norwich, K.H. 1984. The psychophysics of taste from the entropy of the stimulus. *Perception and Psychophysics*, **35**, 269-278.

Norwich, K.H., and Valter McConville, K.M. 1991. An informational approach to sensory adaptation. *Journal of Comparative Physiology A*, **168**, 151-157.

Scharf, B. 1978. Loudness. In: *Handbook of Perception*, Vol. IV, Hearing. E.C. Carterette and M.P. Friedman, Eds. Academic Press, New York.

Schmidt, R.F. 1978. Somatovisceral sensibility. In: *Fundamentals of Sensory Physiology* (R. F. Schmidt, ed.). Springer-Verlag, New York.

Small, Jr., A.M., and Minifie, F.D. 1961. Effect of matching time on perstimulatory adaptation. *Journal of the Acoustical Society of America*, **33**, 1028-1033.

Thorson, J. and Biederman-Thorson, M. 1974. Distributed relaxation processes in sensory adaptation. *Science*, **183**, 161-172.

Ward, L.M. 1991. Informational and neural adaptation curves are asynchronous. *Perception and Psychophysics*, **50**, 117-128.

Yates, G.K., Robertson, D. and Johnstone, B.M. 1985. Very rapid adaptation in the guinea pig auditory nerve. *Hearing Research*, **17**, 1-12.