

INFORMATION TRANSMISSION IN TRIGEMINAL
AFFERENTS FROM TEETH

by

T. J. Farnsworth

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

in the Department
of

Physiology

Faculty of Medicine

We accept this thesis as conform-
ing to the required standard

University of British Columbia

July, 1975

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study.

I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of

physiology

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date

Oct 6/1975

ABSTRACT

The purpose of this study was to measure the rate at which information is transmitted by periodontal mechanosensitive neurons. In addition, an attempt was made to develop a quantitative means for expressing their responsiveness, or sensitivity.

The experiments were carried out on cats anaesthetized with sodium pentobarbitol. The canine teeth were stimulated with computer-controlled forces, and the response of single first-order axons supplying the periodontal mechanoreceptors were recorded following microdissection of the inferior dental nerve. Thirty mechanosensitive units, selected from twelve animals, were used in the study.

For fifteen of these units a stimulus uncertainty of 5 bits, a stimulus duration of 500 msec, and an inter-stimulus interval of 500 msec was employed. Under these conditions, the mean channel capacity was found to be 3.2 bits/stimulus \pm S.E. = 0.09. The channel capacity for the other fifteen mechanosensitive units was measured using a 7 bit/stimulus uncertainty. This increased the channel capacity to a mean of 3.77 bits/stimulus, S.E. = \pm .14. It was found that this calculated rate of information transmission could be further significantly increased ($P > 0.0001$) to 4.33 bits/stimulus, S.E. = \pm 0.11, by treating the data so as to prevent over-sampling of parts of the response. When the

outputs of seven neurons innervating a single tooth were pooled, the maximum rate of information transmission for the group as a whole was found to be 6.84 bits/stimulus. Comparison of the pooled data with the results of psychophysical experiments suggests that the rates of encoding of information in primary afferents far exceeded the ability of the CNS to receive such data.

A computational method for the measurement of the sensitivity of the neurons was also developed. It is proposed that the sensitivity, or responsiveness, of neurons can be measured properly only by the use of information theory, and that it depends upon the particular method of coding chosen by the experimenter.

TABLE OF CONTENTS

	Page
ABSTRACT.....	14
LIST OF TABLES.....	vi
LIST OF FIGURES.....	vii
ACKNOWLEDGMENTS.....	viii
INTRODUCTION.....	1
I PREVIOUS STUDIES OF INTENSITY CODING BY MECHANORECEPTORS.....	3
II MATHEMATICAL DESCRIPTIONS OF INTENSITY CODING IN PRIMARY MECHANORECEPTIVE AFFERENTS AND ATTEMPTS TO RELATE THEM TO PSYCHO-PHYSICAL LAWS.....	5
(a) Weber-Fechner Law.....	5
(b) Stevens Power Function.....	6
(c) Other Descriptions of Input-Output Relations Not Involving Information Theory.....	6
III PHYSICAL PROCESSES INVOLVED IN ENCODING STIMULUS INTENSITY IN A FREQUENCY CODE.....	7
(a) Control of Mean Frequency.....	7
(b) Regularity of the Frequency Code.....	8
(i) Random Processes.....	8
(ii) Non-Random Variation in Frequency Code.....	11
IV INFORMATION THEORY MEASUREMENTS.....	12
(a) Factors Influencing Information.....	12
(b) How Information is Measured.....	13
(c) Information Flow.....	13
(d) Previous Studies of Neural Information Transmission.....	14
EXPERIMENTAL METHODS.....	24
Stimulation.....	26
Calibration of the Force Generator.....	28
Collection of Data.....	30
Analysis of Data.....	31
RESULTS.....	37
DISCUSSION.....	47
PART I.....	47

TABLE OF CONTENTS cont'd.

	Page
Receptor Supporting Structures.....	49
Coding Properties.....	50
PART II.....	50
PART III.....	54
Sensitivity.....	55
APPENDIX A.....	60
BIBLIOGRAPHY.....	69

LIST OF TABLES

	Page
TABLE I. Calibration Data for vibration generator.....	29
TABLE II. Information transmission rates (channel capacities) calculated for 15 slowly adapting periodontal neurons.....	40-42
TABLE III. Maximum information transmission rates (channel capacities) for 14 slowly adapting periodontal neurons.....	44
TABLE IV. Mean rates of information transmission (channel capacities) for Parts I, II and III of this experiment.....	45

LIST OF FIGURES

	Page
Figure 1. Diagram of preparation used for isolation and micro-dissection of nerve to canine tooth.....	25
Figure 2. Block diagram of the overall arrangement of apparatus used in this study.....	32
Figure 3. Matrix representation of the raw data from a typical neuron.....	33
Figure 4. Information transmission plotted vs. stimulus uncertainty for a single periodontal unit.....	35
Figure 5. Sequences of action potentials (upper traces) recorded from a single periodontal neuron in response to stimuli of increasing amplitude.....	38
Figure 6. Hypothetical stimulus response curve for neurons "A" and "B".....	56

ACKNOWLEDGMENTS

I would like to thank Dr. A. G. Hannam and Dr. P. Vaughn for their assistance in the preparation of this thesis. I would further like to express my appreciation of the excellent typing and proof reading skills of Mrs. E. Moore, and the help rendered by Mrs. Joyce Scott, Mr. Bob Paton, and Mr. Robert DeCou.

INTRODUCTION

The periodontal ligament consists of a matrix of collagen and reticulin fibres connecting the centre of the tooth to the lamina dura of the maxillary and mandibular bone. Within this matrix are located blood vessels, lymphatics and nerves. The nerve fibres run from the apical region of the tooth toward the gingival margin. These are joined by fibres entering laterally through the foramina in the alveolar process, (Simpson, 1966; Kizior et al, 1968). In the mandible, both these groups of fibres originate in the inferior alveolar branch of the trigeminal nerve, and in the case of the lower canine tooth of the cat, supply the apical part of this tooth from a discrete, identifiable terminal branch, (de Lange et al, 1968).

A number of studies indicate that periodontal mechanoreceptors terminate in encapsulated endings, (Simpson, 1966; Griffin and Spain, 1972; Griffin, 1972). It is believed that the fibres, upon entering the capsules, lose their myelin coating and break up into finer branches. Measurements of conduction velocity and direct observation indicate that mechanoreceptor fibre diameter varies between 4 and 14 μ , (Griffin, 1972; Pfaffman, 1939; de Lange et al, 1968). However, the observed terminal branches are somewhat smaller when observed using light or electron microscopy, (Griffin, 1972). Intracellular vesicles have been observed, (Canna and Ross, 1966; Munzer, 1965; Griffin and Harris, 1968), but the function of these vesicles is not known.

Electrophysiological studies have shown that the periodontal ligament is predominantly innervated by slowly adapting mechanosensitive afferent fibres, (Anderson et al., 1970). The general properties of these afferent fibres, in terms of their responses to dynamic and static loading of the teeth, appear to be similar to many of the mechanosensitive neurons supplying the skin, although the distribution of mechanoreceptors around the teeth imparts some unique directional properties to individual periodontal neurons, (Anderson, et al., 1970).

It is generally accepted that the periodontal innervation serves both somaesthetic and reflexogenic functions, in that it plays an important role in the perception of forces applied to the teeth, and in the reflex regulation of activity in the jaw-closing and jaw-opening musculature. As a result, periodontal mechanosensitive afferents provide a major source of information which is used in the discrimination and control of the high forces which can be developed between teeth during normal function. Although studies have been carried out on the responses of these neurons to changes in the magnitude of applied forces (Pfaffman, 1939; Ness, 1954; Hannam, 1969) no experiments appear to have been conducted to specifically measure their discriminability, a quality which would seem to require a high degree of definition in a sensitive control system. Information theory, the nature of which will be discussed in later sections, provides a means of assessing this property of discriminability.

The purpose of the experiments to be described was to measure the rate at which information about stimulus magnitude is transmitted by the periodontal

innervation. It was reasoned that not only would this provide an essential measure of the maximum signalling ability of the first order periodontal neurons per se, but, at the same time, provide a means for comparing the responsiveness of this afferent system with that of others, for example in skin or muscle. In this regard, it seems reasonable to expect that the different mechanical linkages between stimulus and mechanoreceptor in different sites would alter the flow of information through the various systems and therefore influence the discriminability of the particular neurons concerned.

It was also an objective of this study to examine the definition of sensitivity, or responsiveness, of primary afferent neurons in general, and to develop a better method of expressing sensitivity than that offered by the conventional stimulus-response curve.

I PREVIOUS STUDIES OF INTENSITY CODING BY MECHANORECEPTORS

When considering the relationship between input and output of systems, it becomes necessary to specify the input and output codes being used. A code may be defined as a system of symbols within which real events are represented. As far as the nervous system is concerned, it seems reasonable to assume that the output code of a first-order neuron is that part of its activity which is used by the second-order neuron in extracting a message. This creates the problem of deciding which characteristics of first-order neural activity should be chosen to represent its code.

Adrian and Zotterman (1926) were the first to record from a single primary afferent. Using a frog skeletal muscle stretch receptor preparation, they observed that the amplitude of the spike remained constant. At the same time, action potential frequency was observed to increase with increasing intensity of stimulation. Adrian (1928), followed by B.H.C. Matthews (1931), was the first to plot stimulus response curves. By plotting stimulus intensity against mean frequency of action potentials, early observers of primary afferent neuron behaviour chose a specific frequency code as the code of the mechanoreceptive neurons under consideration. Later studies have uniformly used the same class of code for slowly adapting mechanoreceptors.

In an experiment by Armet et al (1963) single phasic mechanoreceptors in the foot pad of the cat were stimulated. These units rarely produced more than one, and never more than three, action potentials in response to stimulation. At stimulus frequencies less than 1/sec., the firing frequency of the second-order neuron was totally uncorrelated with the intensity of the stimulus applied to the receptor. Thus no information about stimulus intensity was being transmitted. However, at frequencies greater than 10/sec., the correlation was excellent, indicating that intensity information was being transmitted. This demonstrates that even in systems with strongly phasic receptors, the signal may be decoded as a frequency code at second-order neurons in the brain.

However, as has been pointed out by Perkel and Bullock (1968) frequency constitutes a class of codes containing the following subclasses: interval

statistic codes; weighted average codes (of which the instantaneous frequency code is a special case); increment-above-background codes; rate of change codes; and frequency-of-firing, missing-at-fixed-interval codes, (this classification modified after Perkel and Bullock). A choice between these frequency codes is made extremely difficult by lack of knowledge of the exact characteristics of the synaptic interaction between the primary afferent and the second-order neuron. However, even in neurophysiological studies seemingly unrelated to information theory, such choices must be and are often made. This problem is treated further in the discussion.

II MATHEMATICAL DESCRIPTIONS OF INTENSITY CODING IN PRIMARY MECHANORECEPTIVE AFFERENTS AND ATTEMPTS TO RELATE THEM TO PSYCHO-PHYSICAL LAWS

(a) Weber-Fechner Law

Both Adrian (1928) and Matthews (1931) noted that the frequency of the action potentials was approximately proportional to the log of the intensity of the stimulus. This mimicked the Weber-Fechner law of psychophysics: sensation varies as the log of stimulus intensity. Thus, it appeared that the simple encoding properties of the primary afferent could explain some of the perceptual effects observed by psychologists. The Weber-Fechner law, however, has been strongly questioned almost from its inception (Hering, 1876; Boring, 1933, 1942;

Barlow, 1957; Stevens, 1961, 1966; and Rosner & Goff, 1967). Its weakness is that it produces poor predictions of the perceptual effects of alterations in stimulus intensity. This problem is caused by two factors. Weber's data was collected over only middle ranges of stimulus intensity, and Fechner fitted Weber's data with an inappropriate function.

(b) Stevens Power Function

The Stevens power function, which is of the form $R = K \cdot S^n$, produces a better fit than did the Weber-Fechner relation. Mountcastle, Poggio & Werner (1963), Werner & Mountcastle (1965), and Brown & Iggo (1967), have found that if proper constants are chosen, stimulus intensity and frequency of action potentials for some primary afferents are related to each other by this formula. The slope of this relation and of the stimulus response curve have been used as a measure of sensitivity. (Werner & Mountcastle, 1965; Mountcastle, 1968; Hensel, 1968).

(c) Other Descriptions of Input-Output Relations Not Involving Information Theory

Several forms of sinusoidal analysis have been used on primary afferents. Bode plots, graphs of phase lag, and amplitude versus frequency, have been employed to study the change in response produced by changes in stimulus frequency. This technique was used on the lobster stretch receptor by Terzuolo et al (1962), and Borsellino et al (1965).

Brown and Stein (1966) have used LaPlace analysis to achieve the same ends.

III PHYSICAL PROCESSES INVOLVED IN ENCODING STIMULUS INTENSITY IN A FREQUENCY CODE

A better understanding of the coding properties of neurons may be gained by examining some of the underlying physical processes involved in generating and regulating neuronal activity.

(a) Control of Mean Frequency

The primary response to stimulation of a receptor is an analogue signal (generator potential) (Katz, 1950) which is converted to a spike train by a specialized area in the axon (Edwards et al, 1959).

Hodgkin (1948) studied the initiation of repetitive discharge by constant currents in the isolated motor axon of *Carcinus maenas*. Current was delivered using external agar wick electrodes. From observations of potential change in the area where spike initiation occurs Hodgkin concluded that interspike interval is determined under certain conditions by (1) the rate of charge transfer from the generator area to the capacities of the spike initiating zone; (2) local graded responses which occur in the spike initiating zone; (3) refractoriness at high frequencies; and (4) a supranormal period, if any. Axons with pronounced supranormal periods tended to either fire at high frequencies (75-150/sec.) or not at all. These effects were verified for the slowly

adapting crustacean stretch receptor (Ezaguirre and Kuffler, 1955).

Thus it can be seen that a number of different phenomena control the frequency of output.

(b) Regularity of the Frequency Code

(i) Random Processes:

Any process which is uncorrelated with the stimulus, yet produces variation in the response, is called noise. Noise anywhere in the neuron will cause fluctuations in the neural output. But if the output fluctuates, for a constant input, it will be difficult for the observer of the output to tell whether a change in output was due to noise or to a change in signal. Therefore, increase in noise causes a decrease in the amount of information carried by the neuron, other things being equal.

Noise at the Encoder

Membrane noise has been extensively investigated by several workers. Blair and Erlanger (1933) excited two axons with the same stimulus. They noted that thresholds and latencies varied independently. They therefore concluded that the cause of the variation in response is found in the membrane. Pechier (1936, 1937, 1939) examined the statistical properties of these fluctuations and found them to be random.

Attempts to measure the amplitude of the noise voltage on the membrane began with the works of Blair and Erlanger (1933)

and Pechier (1939). A statistic called relative spread (RS), where

$$RS = \frac{\text{SD of membrane potential difference}}{\text{Potential difference for threshold}}$$

was employed. For axons of *Rana*, *Astacus* and *Sepia*, it was found that RS was related logarithmically to fibre diameter. A theoretical relation between RMS membrane noise and fibre diameter was developed by Fatt and Katz (1952). However, the RMS voltage fluctuations predicted by the work of Fatt and Katz (for expected thermal noise) were an order of magnitude too small to explain the relative spread data.

Verveen and Derksen (1965), using voltage clamp techniques, measured noise amplitudes and frequencies in a frog axon node of Ranvier. The noise was found to have: (1) a $1/F$ amplitude distribution (as opposed to a gaussian distribution expected for thermal noise) and (2) an amplitude about ten times greater than that predicted for thermal noise at threshold. The amplitude of the noise was related to the difference between the membrane potential and the potassium equilibrium potential and was reduced as the K^+ equilibrium potential was approached. It was therefore concluded that the cause of the noise was in the K^+ current.

Verveen and Derksen (1968) have examined the relationship of membrane noise to frequency coding. They note that neural noise manifests itself as fluctuations in latency (or interval) in spike trains. It has been found experimentally that different stimulus intensities produce latencies with different standard deviations. In

a mathematical model they studied the relationship between membrane noise, the final membrane potential reached due to the generator potential and local response (this final potential was called the ultimate depolarization), and the interval distribution. The type of interval distribution obtained for a given amplitude of noise, they assert, is related to (1) the size of the ultimate depolarization, and (2) the difference between the ultimate and the threshold potential. Relations are derived for various values of ultimate depolarization and threshold. Verveen and Derksen (1968) found that the above model, with the proper choice of constants, produced intervals which approximated measured distribution found by Buller et al (1953) and Hagiwara (1954) for the frog muscle spindle stretch receptor, and by Biscoe and Taylor (1963) for the cat carotid body chemoreceptor.

Noise in Generator Potential

Small miniature-like potentials called abortive spikes (Katz, 1950; Ito, 1969a, b) have been recorded in the frog muscle spindle. Their amplitude falls logarithmically with increasing distance from the first node of Ranvier. They occur mainly at low levels of generator potential and though they are not affected by preceding abortive spikes (in fact they sum), they do not occur for a considerable time after an action potential. Afferents supplying "single " muscle spindles produced abortive spikes of one or two amplitudes. Those supplying "tandem" arrangements of spindles

produced spikes of up to four distinct amplitudes. The amplitudes of the abortive spikes also decreased with increasing generator potential. The ionic mechanism causing these spikes is thought to be similar to the one causing normal spikes since they are abolished by tetrodotoxin. It is suggested by Ito that abortive spikes are all or none responses from numerous non-myelinated filaments which ramify to form the myelinated nerve. Quantal effects have been estimated by Katz and Miledi (1972) to produce charges of about 0.3 V each in the ACh receptor. One would expect to find both quantal and $1/F$ noise at the receptor.

(ii) Non-Random Variation in Frequency Code

Adaptation is the name given to the decay of action potential frequency in the face of a constant stimulus. It is a complex process which may be caused by (1) slippage of the mechanical linkage between the nervous tissue and the surrounding structures, and/or (2) a decline of the generator potential in spite of a maintained stimulus to the receptor and/or (3) a decline in the frequency of action potentials produced by a constant generator potential, (Husmark and Ottoson, 1971; Lowenstein and Skalak, 1966). Various mechanisms have been suggested for this latter process. Some of these are: (1) Na^+ inactivation (Vallbo, 1964); (2) membrane hyper-polarization due to an electrogenic Na^+ pump (Rang and Richie, 1969); and (3) failure of hyperpolarization (hyperpolarization helps restore inactivated Na^+ current) due to K^+ buildup or K^+ current inactivation, (Frankenhauser and Hodgkin, 1956; Nakajima

and Onodera, 1969). This variation may be viewed alternately as a carrier of information or as noise, depending on the decoding capabilities of the higher order neurons.

In summary, there appears to be ample evidence that noise levels differ for different primary mechanoreceptive afferents.

IV INFORMATION THEORY MEASUREMENTS

(a) Factors Influencing Information

The rate of information flow varies as a function of signal, noise, and reception time. Signal is important since, though an information channel may be able to convey 10^6 bits/second, if no signal is sent, no information is transmitted. Since information is that which allows us to distinguish between things, and since noise will cause confusion, it is therefore clear that noise will decrease the rate of information flow. Now say 10^6 bits/sec. of information is being carried by a neuron which was connected to another neuron which could only receive 1 bit/sec. It is obvious that the transmission rate is then no greater than 1 bit/sec. Thus the receiver is important. Finally, what is the role of time? Time is of great biological significance since the speed of occurrence of a reflex response can be of considerable importance to the survival of an organism. Thus one would expect selection to have worked in favour of animals which quickly received signals of danger. That is, without ever looking at the nervous system, we might, a priori, expect signals about strong noxious events to be transmitted more rapidly than signals having less immediate import.

(b) How Information is Measured

The amount of information needed to distinguish one element from all the other elements is a function of the probability of occurrence of that element. Suppose one has a series of weights with which a mechanoreceptor is to be stimulated. How much information does it take to distinguish between them? If there are only two weights, say, between 0 and 100 gms., it takes on the average an accuracy of only one significant figure to distinguish between them. Now if we greatly increased the number of weights within this same weight range, it would, on average, require a much more accurate measurement to distinguish them. Thus in some way, the amount of information necessary to tell two weights apart increases as the number (N) of different weights increases. Now one can always paint numbers on the weights and use the numbers to tell them apart. If we number them using a system to the base K then it will take numbers with $l = \log_K N$ digits in length to number any weight where l is the integral number of digits in each number. But since P , the probability of each element, $= 1/N$ for equi-probable elements, $\therefore l = -\log_K P$. (For Mathematical Treatment see Appendix A).

(c) Information Flow

Let us define the information carried by an event (or element) as $l = -\log_2 P$, the input set as X , and the output set as Y . Then the average information content of any element drawn from the set X is $H(X) = -\sum P_i \log_2 P_i$. Suppose we choose a receptor and stimulate it once a second with weights chosen at random. Thus on average we are feeding

into our neuron $H(X)$ bits of information per second, and we are interested in seeing how much of this information transfer can be calculated given only the probability of each input, the probability of each output state (e.g. frequency), and the probability of each of them occurring together. This is expressed mathematically as

$$R = H(X) + H(Y) - H(X, Y)$$

or

$$R = - \sum_i P_i \log P_i - \sum_j P_j \log P_j + \sum_{ij} P_{ij} \log P_{ij}$$

Where

R is the rate of information transmission in bits per second

$H(X)$ is the information content of the stimulus

set (stimulus uncertainty)

$H(Y)$ is the information content of the response

set (output certainty)

$H(X, Y)$ is the joint uncertainty.

The variable $C = \max (H(X) + H(Y) - H(X, Y))$ is found by increasing the number of different stimuli used until the amount of information obtained from the axon remains constant. This constant value is said to be channel capacity.

(d) Previous Studies of Neural Information Transmission

MacKay and McCulloch (1952) have attempted to deal mathematically with the theoretical maximum rate of information transfer across a synaptic link. It was estimated from data of Lloyd (1946) that the

time from the beginning of the rise of the generator potential to the peak of the spike would require 0.5 to 0.8 msec, (mean 0.65, SD = \pm 0.05). Since interspike intervals less than 1 msec are not seen and even those less than 2 msec are not long sustained, a maximum frequency after adaptation of 250 to 300 Hz was thus chosen (after observation by Galambos & Davis (1943)). Even so, the minimal interspike interval T_r was chosen as 1 - 4 msec, and it was argued that the system could be quantized by dividing the time scale into discrete segments T_r msec long. Such a time-quantized binary system would carry a maximum of $1/T_r$ bits/msec.

For interval coding, by allowing the time between successive pulses, T_s , to vary, where T_m is the maximum interval, the range of T_s possible is $T_m - T_r$. A change in T_s , ΔT , is assumed to be statistically significant if it is greater than the standard deviation. This is a rather inadequate step in the treatment since only about 60% of these intervals would fall in the proper slot. Thus if $n = (T_m - T_r)/\Delta T$, $\log_2 n$ bits of information may be transferred across the synaptic link per signal. Assuming all signals are equiprobable, the mean interval is $\frac{1}{2}(T_m + T_r)$ and the channel capacity $C = \frac{2}{T_m + T_r} \times \log_2 n$ (Eqn. 2).

Substituting $m\Delta T$ for T_m and $r\Delta T$ for T_r , one finds a maximum rate of transmission when $\frac{m+r}{n-r} = \ln(m-r)$. (Eqn. 3).

Solving (3), taking $T_r = 1$ msec. and $\Delta T = 0.5$ msec. and substituting therein eqn. (2), it is found that a maximum of 2.7 bits/msec could be attained (or about 4.3 bits/impulse at 670 impulses/sec.) for a pulse interval modulated system, versus 1 bit/msec for a binary linked system.

Mackay and McCulloch (1952) are quick to point out that such a mean impulse frequency is unrealistically high, and that their choice of ΔT might be far too low. They note that 50/sec. is a more reasonable figure for mean frequency. This mean pulse width of 20 msec, they estimate, could correspond to a range of 4 - 36 msec. Taking ΔT to be 0.05 msec gives a transmission rate of about 500 bits/sec. Successive intervals would probably be correlated and this would drastically reduce the rate of information transmission.

It is Stein's contention (Stein 1967) that the estimates produced by Mackay and McCulloch (1952) and by Rapaport and Horvath (1960) are much too high for several reasons. He suggests that mean frequency coding, not binary or interval coding, is the method of information transfer in the nervous system. On the other hand, Perkel & Bullock (1968) claim that the crayfish stretch receptor in the lower end of its frequency ranges is interval coded. Moreover, it is admitted by Stein that temporal coding must exist at least somewhere in the nervous system since very precise binaural localization of sound using a temporal method has been reported, (Hall 1964). Stein notes that Mackay and McCulloch (1952) underestimate the noise introduced at synapses by a factor of 10.

Jacobson (1950, 1951) has estimated the least channel capacity of retinal ganglion and auditory nerve fibres by dividing the best estimates for the number of bits per second of information that can be perceived for each of these senses by the number of ganglion cells involved in their transmission. He arrives at a figure of 5 bits/sec/fibre and 0.3 bits/sec/fibre for one eye and one ear respectively. This neglects color perception studies and the fact that many of the ganglion cells of the eye are not concerned with the high activity, cone-type vision of the fovea, for which the perceptual estimates are made. Further, it has been estimated by Stein (1967) that about 50% of the information carried by a fibre is lost in synaptic transfer. Now there are at least 2 synapses on the way to the cortex, and probably a large number more within it. Thus, there is reason to believe that either the rate at which Jacobson estimates the ganglion cells are carrying information is low, or that the information loss in these synapses is not nearly as great as that estimated by Stein (1967) and Walloe (1970) for the first synapse of muscle spindle afferents.

Stein (1967) developed some approximate computational methods to simplify the calculation of information transmission rates in neurons. Formulae were derived for regular and irregularly firing neurons, and serial correlation between impulses was taken into account. By successive approximations, optimal stimulus sets (i.e. those producing maximum information transfer) were produced for a model neuron assuming a frequency code. These maxima generally occurred when the stimuli were an equal number of standard deviation units apart. This work constitutes an important contribution to the methodology of information transmission calculations. Furthermore, a set of plots relating information carrying capacity to the time over which the

output is decoded were produced. From these results the channel capacity of regularly firing neurons would be expected to uniformly exceed that of randomly firing units for a given decoding time and a frequency code. Interval coding exceeded frequency coding for estimated transmission rates, and binary coding far exceeded both of the former two.

Werner and Mountcastle (1965) measured the rate of information transmission in primary afferent fibres ending in Iggo corpuscles in the monkey forepaw. Observation time for fibre output ranged from 20 to 1000 msec. Information transmission was found to reach a maximum between 200 and 400 msec. Stimulus uncertainties of 5 bits were used (i.e. a collection of 2^5 different stimulus intensities) and output categories were formed by lumping responses of 0 and 1 spike, 2 and 3 spikes, etc., into the same categories. This is called double binning. Each stimulus intensity was repeated ten times.

Calculations were performed by constructing and evaluating a stimulus response matrix. (see Methods). Rectangular pulse stimuli with evenly spaced intensities were delivered both sequentially and at random. It was found that randomizing the stimuli had little effect on the rate of transmission, indicating that serial correlations between stimuli were minimal. The portion of the response carrying most information was found to be the early "steady state". This period begins at about 100 msec after stimulus onset and ends after about 450 msec. It is characterized by great regularity. This finding is compatible with predictions made by Stein (1967). The mean rate of information transmission was found to be about 2.5 bits/sec. Werner and Mountcastle noted that this figure is similar to the 6 - 7

discriminable intensity steps observed in many psychophysical experiments.

Darian-Smith et al (1968), using a 4 bit (2^4 different input intensity levels) stimulus uncertainty, measured the rate of transmission of stimulus intensity information for first and second order mechanoreceptive neurons in the trigeminal nerve (lip, mental nerve). Rectangular pulse stimuli with evenly spaced stimulus intensities, lasting 1000 msec. were presented once/5 sec. Channel capacity was found to be $2.4 \text{ S.E.} = \pm 0.06 \text{ bits/stimulus}$ for first order afferents, $2.09 \text{ S.E.} = \pm 0.07 \text{ bits/stimulus}$ for second order neurons in nucleus oralis, and $1.07 \text{ S.E.} = \pm 0.07 \text{ bits/stimulus}$ for neurons in nucleus caudalis. Darian-Smith et al note that the observed difference in sensory deficit between the medial lemniscal and anterior lateral tracts in the spinal cord may be due to differences in their abilities to transmit information over the first synapse and seemed to be the case in the trigeminal sensory homologues of these tracts. They conclude: (1) that the second order neurons in the nucleus oralis receive information with little loss from the primary afferent neurons and (2) that the amount of information received here is responsible for the perceptual information levels seen in psychophysical experiments. Both of these conclusions are suspect. Since the diameter of the stylus used to deliver the stimulus was several mm., it is likely that numerous receptors were stimulated. Thus the information received by a second order neuron may constitute only a fraction of the total information which is being delivered to the neuron, and is not necessarily representative of the transmitting capacity of the region as a whole. It has been further noted by Kenton and Kruger (1971) that the similarity of the measured transmission values to the number of perceptual levels discriminable by the whole organism, as seen in both Mountcastle's experiment and Darian-Smith's, is probably the result of the experimental

design rather than the nature of the system.

Kenton and Kruger (1971), in a comprehensive and well-designed set of experiments, have studied mechanoreceptive afferents in the hind limb of the cat and caiman. Rectangular pulse mechanical displacements were applied to points of maximum sensitivities within cutaneous receptive fields. Up to one hundred eighty-eight (188) stimulus categories (amplitudes) were used with durations of 900 to 1000 msec., at a repetition rate of 6/minute. Responses to different stimulus distributions were measured by sampling each fibre with six sets of stimuli, graded at 50, 25, 10, 5, 2 and 1gm respectively, and then substituting the appropriate data in appropriate places in the information transmission matrix so as to obtain a degree homogeneity of sampling density. Only slowly adapting fibres were considered. Both Type I and Type II receptors were studied. Maximum rates of information transmission for for both Types I and II receptors on both glabrous and hairy skin of the cat fell between 4.0 and 5.4 bits. These values are considerably higher than those measured by Mountcastle et al, or Darian-Smith et al. Kenton et al (1971) claim that their findings are actually consistent with those of Werner and Mountcastle if one looks at a stimulus uncertainty of only 5 bits.

It was found that the information carrying capacity of the reptilian fibres was very nearly the same as that for mammals, and it was therefore suggested that if mammals are capable of finer tactile sense, it is due to central rather than peripheral differences.

Kenton and Kruger (1971), for reasons mentioned earlier in this section, believe that the similarity between the mean channel capacity found by Werner and Mountcastle for primary afferent data and psychological observations on the rate of perceptual information acquisition found in humans is fortuitous. They suggest that if a relation of this type does exist, the expected relation would be between the first order information transmission rates and the information transmission rates for just noticeable difference.

Kenton and Kruger (1971) felt that even their exhaustive methods were not sufficient to saturate these receptors and thus measure the true channel capacity. They also expressed doubt as to (1) the choice of output code, and (2) the ability of the errors of measurement of information carrying capacity to be adequately assessed. Further, they note that if high repetition rates are used, the variability is artificially increased, thus lowering information transmission rates, but if low repetition rates are employed only a small sample can be collected before the fibre dies.

Finally, Kenton and Kruger (1971) lament the lack of a method for calculating the placement of stimulus categories within the dynamic range. Stein (1967) has originated such a method. He has shown experimentally that the intensity of stimuli should be graded so that responses are an equal number of standard deviation units apart.

Walloe (1972), using data collected by Jansen et al (1966), has calculated the channel capacity for cat muscle spindle primary afferents, and 2nd order afferents, using an interval code. It was found that each muscle spindle afferent transmitted

about 5.9 bits/symbol. It is estimated that about 15 primary afferents converge on each second order neuron. Thus the 2nd order neuron should receive an amount of information equivalent to that transmitted by one neuron over 15 time periods. This Walloe calculates to be 7.8 bits/stimulus. He therefore concludes that about one-half the information is lost in synaptic transfer. This is in agreement with estimates of Stein (1967), and opposed to those of Werner and Mountcastle (1965) and Darian-Smith et al (1968). Formulae for the approximation of Stein's (1967) information transmission rates were tested by Walloe and the values which they produced were found to be in good agreement with those calculated by the exact method.

Stein and French (1972) have attempted to use coherence (cross correlation over a frequency band) between input and output frequencies (using band limited white noise as an input) in a neuron model. It was thus attempted to measure information transmission for different input frequencies. The results showed that information transmission falls off at high frequencies. Stein indicated that these results must be viewed with caution, since cross correlation is a linear measurement whereas neural input-output is basically non-linear.

Since different mechanical linkages can be expected to change the flow of information through a given afferent system, and since the rate of information flow through periodontal mechanoreceptive afferents has not been measured, it was decided to measure the rate of information transmission for these fibres. To make these values more readily comparable to those used by Werner and Mountcastle (1964) and Darian Smith et al, (1968), measurements were made originally using similar

input parameters, e.g., input uncertainties and non-random stimulus trains.

Methods were later modified to allow comparison to the results obtained by Kenton and Kreuger, (1971).

EXPERIMENTAL METHODS

Preparation of Animals

Cats were anesthetized with intra-peritoneal injection of Pentobarbital,¹ (64 mg + 13 mg/lb). A cannula was placed in the saphenous vein and further anesthetic was given whenever the animal withdrew its leg in response to a painful stimulus delivered to its toes. The skin beneath the jaw was cut, the muscle and periosteum cleared from the bone. Bone was chipped away with rongeurs exposing the inferior alveolar nerve, artery and vein. The cat's lower jaw was fixed by cementing its upper and lower teeth together with cold-cure acrylic resin. The acrylic was formed around a brass rod which was inserted into a clamp. The clamp was attached firmly to the table. Skin flaps around the wound were sewn to an F-shaped bar thereby forming a pool which was filled with light paraffin oil maintained at 37°C and which was deep enough to cover the nerve. With the aid of a dissecting microscope², bone chips and epineurium were removed and the nerve was cut and laid on a black acrylic dissecting platform. A schematic diagram of the surgical field is shown in Fig. 1. The nerve was teased into pseudo-single fibre preparations with surgical forceps and the strands were laid across a pair of fine platinum wire electrodes.

¹ Nembutal, Abbot Veterinary Labs, Montreal, Canada
² Olympus, Canada Ltd.

Stimulation

An 8K minicomputer³ was used to drive a power amplifier⁴ via a D to A converter⁵ from which it was isolated by a variable gain voltage-following configuration.⁶ The output of the voltage follower was displayed on a storage oscilloscope.⁷ The computer was programmed to give 10 stimuli of equal size, each 500 msec long, separated in some trials by 500 msec intervals, and in other trials by $4\frac{1}{2}$ sec intervals. The stimulus was varied in 32 equal steps in the first part of the experiment over a range of 0-50, 0-100 and 0-200 gms. In the second part of the experiment the stimulus was varied in 128 equal steps between 0 and 200 gms with the exception of one fibre which, because of its extremely low dynamic range, was stimulated with 128 equal steps between 0 and 100 gms. Only 500 msec intervals were used and with 5 trials at each stimulus intensity.

It is obvious that by a suitable choice of the distribution of stimulus strengths one can cause the measured channel capacity to assume any value between zero and the maximum possible for that neuron. The best choice for this distribution would be that actually experienced by the receptor in question under natural conditions. Though the measurement of this distribution is technically possible, it is beyond the scope of this thesis. Therefore, it was necessary to make certain assumptions. Chewing movements for which these receptors are important occur at about 1/sec. (Moller, 1966). Thus the 500 msec stimulus time and the 500 msec rest

³ HP 2114, Hewlett Packard Canada Ltd.

⁴ POA-1, Ling Industries

⁵ HP 2116-6198, Hewlett Packard Canada Ltd.

⁶ B194 Manifold, B100 OP AMP, Analog Devices, Canada Ltd.

⁷ 564B Oscilloscope, Tektronix, Canada Ltd.

time which were used are not unreasonable estimates of the actual conditions of the system. The stimulus magnitudes were placed evenly throughout the dynamic range so that this study could be compared with those of Werner and Mountcastle (1965) and Darian-Smith et al (1968).

In order to calculate information transmission, it is first necessary to measure a parameter of neural output which will adequately reflect the response of the neuron to stimulation. This raises the problem of which output code is the most probable. A frequency code is a likely candidate, but there is an infinite number of output frequency codes and, therefore, an infinite number of decoding algorithms, from which to choose. Two interesting choices might be: (1) the decoding algorithm which allows a maximum of information to be received by an ideal device; or (2) the decoding algorithm employed by the neuron(s) synaptically contacted by the neuron under consideration. However, neither of these algorithms are known.

The decoding system used by a second-order neuron(s) for a frequency code would likely depend, in part, on the duration of the EPSP. This, in turn, will depend on the electrical time constant of the membrane and the length of time during which the transmitter is active. An unweighted integrating period of 500 msec or longer, as employed by Werner and Mountcastle (1965), Darian-Smith et al (1968) and Kenton and Kruger (1971), does not seem a very reasonable candidate for an output code. However, this decoding algorithm, despite its faults, was chosen for the first two parts of this experiment because: (1) it made this work directly comparable to that done previously; and (2) it allowed one to get a feel

for the computational methods and the limits imposed by them. A stimulus duration of 500 msec has been shown by Werner and Mountcastle (1965) to be adequate to allow maximum rates of information transmission to be reached with observation times of between 200 and 400 msec. The stimuli were not randomized since Werner and Mountcastle (1965) claimed that this had little effect on the rate of transmission.

Calibration of the Force Generator

For the purposes of the experiment, it was necessary that the stimulating forces had a small probability of error, i.e., they should not overlap. It was also desired that the rise time of the stimuli remained nearly constant.

To judge the ability of the Ling Vibration Generator to meet these specifications, two calibrations were done:

- (1) 2 LVD (Linear Voltage Displacement) transducers⁸ and six springs were obtained. Each spring was mounted in series with a displacement transducer and weights were applied to it. From the resulting data, a graph of force versus displacement was produced for each spring. By preloading the displacement transducer-spring combination and then adding small weights the standard displacement vs. force and standard deviations were found. This procedure therefore defined the characteristics of the springs.
- (2) The Ling vibration generator was then coupled in turn to each spring via a 10 cm brass rod, and the voltage necessary to produce a given displacement was found in each case. By repeatedly applying the

⁸ Schaevitz Type 090 MS-L, Schaevitz Engineering, New Jersey, U.S.A.

TABLE I: Calibration Data for vibration generator. Mean Forces produced by various voltages were calculated by using calibrated springs. The S.D. of the displacement transducer output over 10 trials was calculated for each force. This value was converted to gms and the % deviation in gm % was calculated. Assuming a normal distribution of errors about the mean, the smallest voltage charge which could be applied to the vibration generator and still allow a probability of stimulus overlap $< 5\%$ was calculated. In the experiment, these limits were not approached.

Force (gms)	SD - Volts	SD - grams	% deviation Gram %	Volts - steps allow- able for 95% certainty of no overlap
13.5	0.042	0.058	$\pm .4$	$\pm 1\text{mv}$
27	0.05	.07	.2	$\pm 1\text{mv}$
54	0.1	.14	.2	$\pm 2\text{mv}$
108	0.14	.19	.16	$\pm 6.4\text{mv}$
217	0.26	.36	.16	$\pm 12\text{mv}$

same voltage to the Ling vibration generator, and measuring the displacement produced with an LVD transducer, the % deviation of the force was calculated. Table I lists applied force in grams, the standard deviations of the LVD output in volts, this value converted to grams and grams % of applied force, and the minimum voltage step for which the probability of no overlap of force categories will be greater than 95%.

The rise time of the force generator was examined by displaying the signal produced by the displacement transducers on a calibrated storage oscilloscope. The rise time was found to be 3 msec S.E. = ± 0.2 msec for forces ranging from 10 to 600 gms.

Collection of Data

The output of the nerve under examination was recorded by means of paired platinum electrodes (0.15 mm in diameter), amplified by a differential amplifier.⁹ The output of this amplifier was:

- (1) fed into a speaker amplifier for auditory feedback to the experimenter after isolation via an operational amplifier;
- (2) recorded on FM instrumentation recorder;¹⁰
- (3) displayed on the storage oscilloscope; and
- (4) fed into a pulse height discriminator.¹¹

The levels on the discriminator were set so as to reject all action potentials except those of the neuron under consideration. This was judged by spike heights and

⁹ Tektronix 3A9.

¹⁰ HP3960, Hewlett Packard.

¹¹ Frederick Haer Ltd., Boston, Mass., U.S.A.

regularity of firing. In case of certainty, the teased preparation was modified or discarded. The output of the pulse height discriminator was monitored on the oscilloscope and counted by a Multichannel Signal Analyzer.¹² The channel was advanced by the leading edge of the stimulus voltage produced by the computer. When one-half the total number of stimuli had been delivered, the data in the multichannel analyzer was read by the computer. The overall arrangement of the apparatus is summarized schematically in Figure 2. At the end of the experiment the contents of the analyzer were read again, and the data condensed and punched on paper tape by a high-speed punch. In this manner the integral number of impulses was counted and stored for each stimulus.

Analysis of Data

In the preliminary experiment, the data was analyzed immediately by means of the HP 2114 computer, using the formula

$$I(X) = H(X) + H(Y) - H(X, Y).$$

where $I(X)$ = rate of information transmission, $H(X)$ = the stimulus uncertainty, $H(Y)$ = the response uncertainty and $- H(X, Y)$ = the joint uncertainty. This was done as follows: The frequency of occurrence of a given number of spikes following each stimulus was represented in a 32×50 matrix. A typical raw data matrix is shown in Figure 3. Here only 36 of the 50 possible output categories are shown since the remaining categories were not filled.

¹² HP5480, Hewlett Packard.

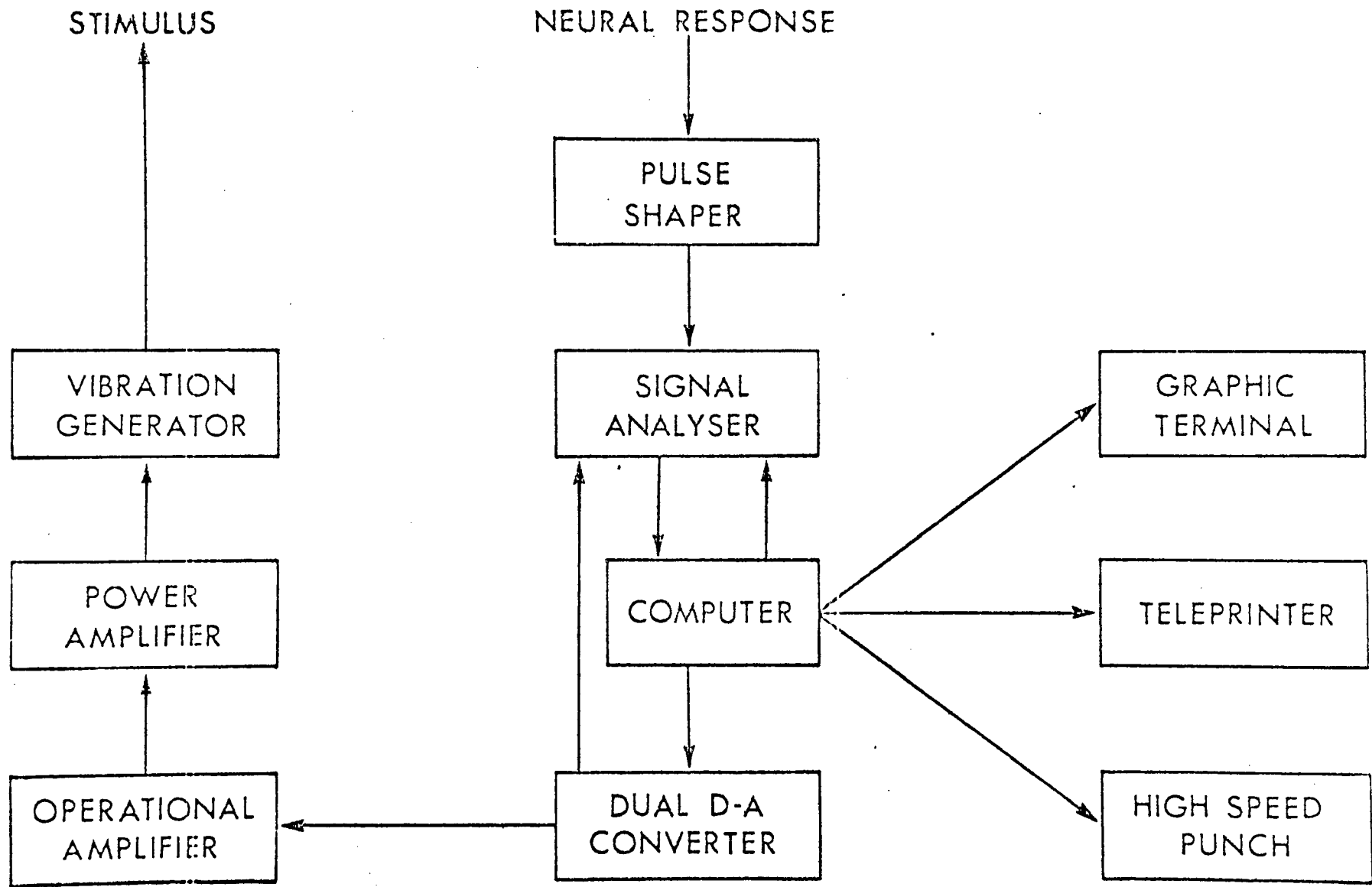


Fig. 2 Block diagram of the overall arrangement of apparatus used in this study.

The stimulus uncertainty $H(X)$ was calculated by the formula

$$H(X) = \log_2 N \text{ where } N = \text{number of categories.}$$

N was varied from 32 to 2 by successively deleting every 2nd, every 2nd and 3rd, every 2nd, 3rd and 4th...every 2nd, 3rd, 4th...16th row (stimulus). This allowed the stimulus uncertainty to be varied between 1 and 5 bits, and therefore allowed one to plot the rate of information transmission as a function of stimulus uncertainty.

The response uncertainty $H(Y)$ was calculated by the formula

$$H(Y) = \sum_{j=1}^N \frac{Y_j}{N \times 10} \log_2 \frac{Y_j}{10 \times N}$$

where Y_j = number of stimuli for which the number of spikes was j .

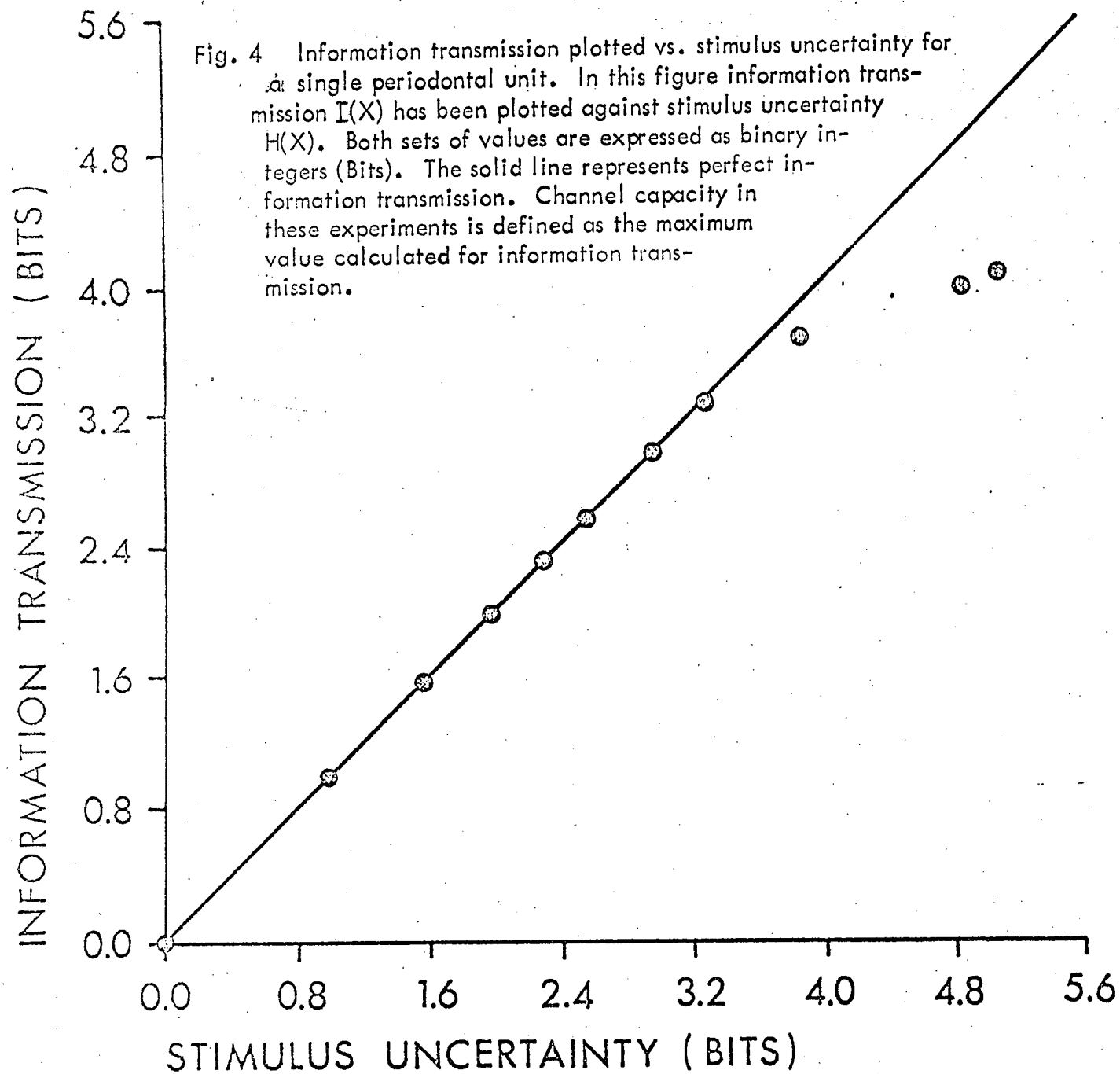
The joint uncertainty — $H(X, Y)$ was calculated by the formula

$$H(X, Y) = \sum_{ij=1}^N \frac{r_{ij}}{10 \times N} \log_2 \frac{r_{ij}}{10 \times N}$$

where r_{ij} = number of counts for the i^{th} stimulus in the j^{th} response category.

Information transmission was then plotted against stimulus uncertainty for each unit, and an example of one of these graphs is shown in Figure 4.

The second part of the experiment, which was carried out using 128 stimulus categories, was analyzed on the U.B.C. IBM 360 computer. The



computational methods employed were the same as those used for the preliminary investigations except that serial calculations of $I(X)$ were performed after deleting 1, 1 and 2, 1, 2 and 3 etc. stimulus categories from the matrix until the result had dropped 2.5% below the maximum previously found for $I(X)$. A similar procedure was carried out with the 128th, 128th and 127th...stimulus categories. In this way, the force range over which the unit transmitted 95% of its information output was found.

Finally, from the original data sets, stimulus categories were chosen such that the mean spike counts per stimulus increased by one or more with each increase in stimulus category. All others were discarded. This was done by finding the first and last category for which the mean number of spikes per stimulus category fell between two successive integers (e.g., between a mean of nine and ten spikes). Only stimulus categories equidistant from these first and last categories were then used in the calculations. This had the effect of matching the non-linearity of the stimulus response curve to the stimulus set and is hereafter referred to as delinearization. The channel capacity of units was then calculated as before.

RESULTS

The graded response of a typical slowly adapting periodontal mechanoreceptive unit to forces of increasing intensity is shown in Figure 5. The upper traces are sequences of action potentials elicited by 500 msec stimuli. The amplitude and duration of these stimuli are represented in the strain-gauge records shown in the lower traces. Stimulus intensity increases from A to D in this figure. At low intensities only a few action potentials were produced. At higher stimulus intensities, the response was maintained for a full 500 msec, but the frequency of action potentials could be seen to decrease during the application of a given stimulus as the length of time increased. When a series of oscilloscope traces for a given unit which had been stimulated with the same stimulus were superimposed, early action potentials were observed to occur at very nearly the same times. However, later action potentials in these spike trains did not superimpose.

Spontaneously active units were encountered occasionally during the experiments. A force applied to the tooth from a given direction increased the action potential frequency of these units in the same manner as described previously. Forces applied from the opposite direction decreased the activity of these units. Data from spontaneously active units was not included in this study.

Activity in a given unit generally ceased 15 to 30 minutes after the group of fibres containing it were placed across the electrodes. Thus, the duration

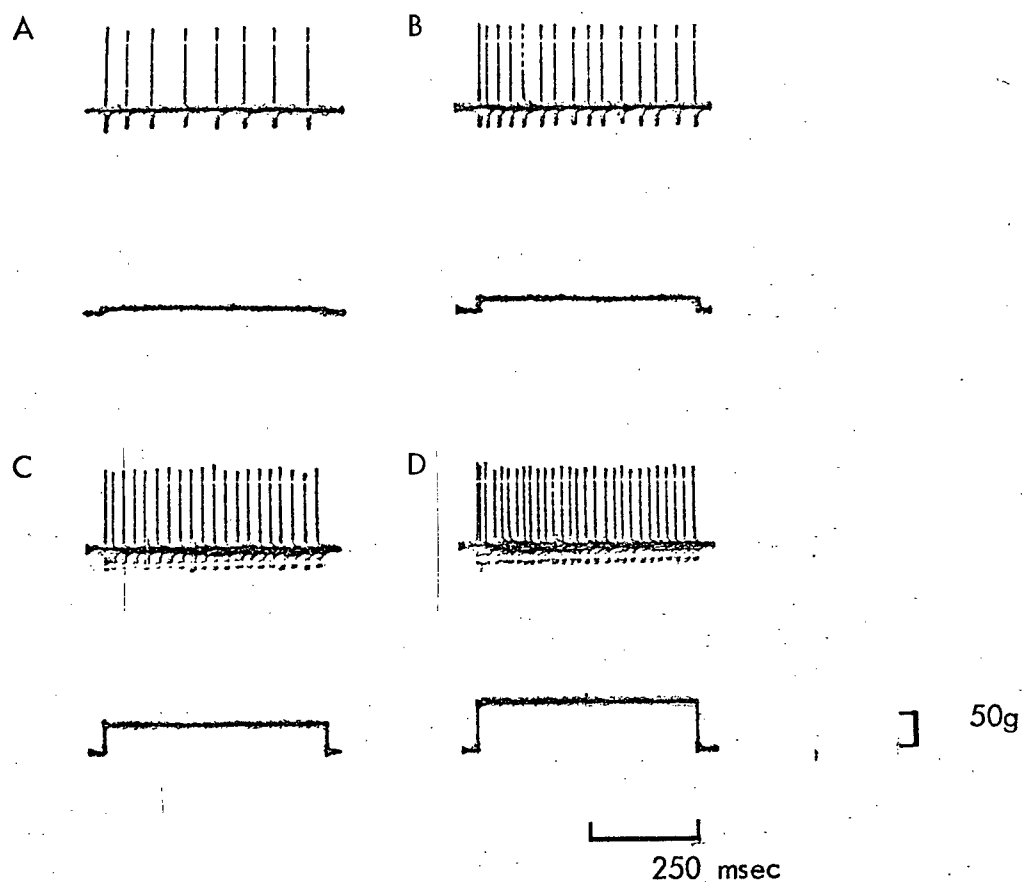


Fig. 5 Sequences of action potentials (upper traces) recorded from a single periodontal neuron in response to stimuli of increasing amplitude. In each case the upper trace represents the neural response, and the lower trace represents the strain-gauge record of applied force. The figure illustrates the grading of activity observed in these neurons during controlled increases in force.

of each stimulus, and the interstimulus intervals used were of extreme importance in determining the total number of responses which could be recorded reliably from a single fibre.

It was considered desirable to limit the interstimulus interval to 500 msec if possible, so as to match the experimental design set up by previous workers (see Introduction) and to permit successful completion of stimulation before the fibres were lost. However, it was felt that some effort should be made to determine whether 500 msec was sufficient time to allow the receptor system to recover after each stimulus. For this reason, experiments were carried out on 7 neurons in Part I of this study in order to compare their information transmission when interstimulus intervals of $\frac{1}{2}$ sec and $4\frac{1}{2}$ sec were used. The results of these experiments are presented in Table II and summarized in Table IV (A,B), where it is shown that the actual difference in transmission rates was relatively small, the $4\frac{1}{2}$ sec rest time providing a mean value of 2.82 bits/stimulus and the $\frac{1}{2}$ sec providing one of 2.91 bits/stimulus. It was therefore decided that 500 msec would be an adequate interstimulus interval in all future experiments.

Almost all neurons were tested with 32 stimulus intensities spread evenly within the 0 to 50 gm, 0 to 100 gm and 0 to 200 gm ranges. The results of these experiments are listed in Table II and summarized in Table IV (C). Over the 0-50 gm range the mean channel capacity was $2.55 \pm SE=.17$ bits/stimulus, over the 0-100 gm range $3.07 \pm SE=.18$ bits/stimulus, and over the 0-200 gm range $2.87 \pm SE=.2$ bits/stimulus. The mean value of the channel capacity for all fibres,

TABLE II: Information transmission rates (channel capacities) calculated for 15 slowly adapting periodontal neurons. 32 stimulus categories were used under various experimental conditions. In the experiments, the peak force varied between 50, 100 or 200 gm and the interstimulus interval between $4\frac{1}{2}$ and $\frac{1}{2}$ sec. The data provided here is summarized in Table IV. In the fibre codes R1, R2, etc. refer to measurement (runs) on the same fibre.

Fibre Code	Channel Capacity in Bits per Stimulus			
Stimulus Range	50g.	100g.	200g.	Inter-Stimulus Inter- val in Seconds
C2F1R1			2.772	$4\frac{1}{2}$
C2F1R2			3.82	$\frac{1}{2}$
C2F1R3		3.339		$\frac{1}{2}$
C2F2R1		2.469		$\frac{1}{2}$
C2F2R2	2.704			$\frac{1}{2}$
C2F2R3	2.521			$4\frac{1}{2}$
C3F2R4			1.782	$\frac{1}{2}$
C2F3R1		2.095		$\frac{1}{2}$
C2F3R2	2.090			$\frac{1}{2}$
C2F3R3	2.369			$4\frac{1}{2}$
C2F3R4			1.576	$\frac{1}{2}$
C3F2R1	2.552			$\frac{1}{2}$
C3F2R2		3.688		$\frac{1}{2}$
C3F2R3		3.427		$4\frac{1}{2}$

TABLE II (continued)

Fibre Code	Channel Capacity in Bits per Stimulus			Inter-Stimulus Inter- val in Seconds
	50g.	100g.	200g.	
C3F3R1		2.278		$\frac{1}{2}$
C3F3R2	1.471			$\frac{1}{2}$
C3F3R3		2.091		$4\frac{1}{2}$
C3F3R4			1.829	$\frac{1}{2}$
C3F4R1	1.906			$\frac{1}{2}$
C3F4R2		2.553		$\frac{1}{2}$
C3F4R3			2.736	$4\frac{1}{2}$
C3F4R4			3.136	$\frac{1}{2}$
C4F2R1		4.082		$\frac{1}{2}$
C4F2R2	3.584			$\frac{1}{2}$
C4F3R1	3.609			$\frac{1}{2}$
C4F3R2		3.485		$\frac{1}{2}$
C4F3R3	3.396			$4\frac{1}{2}$
C4F3R4			3.583	$\frac{1}{2}$
C4F5R1	2.468			$\frac{1}{2}$
C4F5R2		3.381		$\frac{1}{2}$
C4F5R3			3.066	$\frac{1}{2}$

TABLE II (concluded)

Fibre Code	Channel Capacity in Bits per Stimulus			Inter-Stimulus Inter- val in Seconds
	50g.	100 g.	200g.	
Stimulus Range				
C5F1R3	2.758			$\frac{1}{2}$
C5F1R2		3.675		$\frac{1}{2}$
C5F1R3			3.904	$\frac{1}{2}$
C5F2R1	3.355			$\frac{1}{2}$
C5F2R2		3.400		$\frac{1}{2}$
C5F2R3			3.074	$\frac{1}{2}$
C5F3R1	2.789			$\frac{1}{2}$
C5F3R2		3.613		$\frac{1}{2}$
C5F3R3			2.754	$\frac{1}{2}$
C5F4R1	2.326			$\frac{1}{2}$
C5F4R2		1.819		$\frac{1}{2}$
C5F4R3			2.216	$\frac{1}{2}$
C5F5R1	2.418			$\frac{1}{2}$
C5F5R2		3.476		$\frac{1}{2}$
C5F5R3			3.463	$\frac{1}{2}$
C5F6R1	1.202			$\frac{1}{2}$
C5F6R2		2.768		$\frac{1}{2}$
C5F6R3			2.779	$\frac{1}{2}$
C5F7R1	2.098			$\frac{1}{2}$
C5F7R2		3.112		$\frac{1}{2}$
C5F7R3			3.421	$\frac{1}{2}$

taken over the best range for each fibre, was $3.20 \text{ bits/stimulus} \pm 0.09$. All these values are uniformly larger than those found by Werner and Mountcastle.

During these studies it was noted that 32 stimulus categories were not consistently able to produce rates of information transmission approaching channel capacity, i.e., the graph of information output versus stimulus uncertainty usually maintained a positive slope (see Fig. 4). Further, the methods used in the first part of the experiment to choose the optimal operating range of the neuron were crude. Ranges between 0-50, 0-100 and 0-200 gm are not narrow enough, and the use of the best information transmission rate to select the best range over which to measure the information transmission rate is a statistically unsound procedure when one set of data is used for both purposes. Therefore, Part II of the study was carried out on 14 additional units and employed 128 stimuli separated by 500 msec intervals. Generally, in these experiments, a maximum force of 200 g was used.

The results of this part are presented in Table III and summarized in Table IV(D,E). It was found that the mean channel capacity for these units was $3.77 \pm 0.14 \text{ bits/stimulus}$, transmitted over a mean range of $159.1 \pm 7.9 \text{ gm}$. The mean value of the channel capacity found when the data was treated by matching the non-linearity of the stimulus set to the stimulus-response curve as explained in the methods section was $4.33 \text{ bits/stimulus S.E.} = \pm 0.11$. This procedure increased apparent rates of information transmission over those measured before matching in all cases. 95% of all information was transmitted over a mean range from threshold to $68.0 \text{ gms S.E.} = \pm 7.9$. Thus most of the information was transmitted over a

TABLE III: Maximum information transmission rates (channel capacities) for 14 slowly adapting periodontal neurons. 128 stimulus categories were used for these neurons, which are not the same fibres as those listed in Table II. This table presents channel capacities before and after stimulus delinearization (see text), the maximum forces used, the forces at which channel capacities were attained, the stimulus limits between which 95% of the information was transmitted, and the range between these limits. The data provided here is summarized in Table IV.

Fibre Number	Channel capacity in bits/stimulus before delinearization	Channel capacity in bits/stimulus after delinearization	Maximum force in gms.	Forces in gms between which 95% of the info. is transmitted gms.		Force at which channel capacity was reached	95% range in gms.
C6F1R1	3.24	3.78	200	4.4	58.6	125.9	54.2
C6F2R1	3.53	3.82	200	6.3	92.9	144.9	86.6
C6F3R2	3.73	4.16	200	25.2	74.0	143.3	49.2
C6F4R1	3.57	3.82	200	18.9	70.0	143.3	51.1
C6F5R1	3.99	4.33	200	31.5	75.5	151.2	44.0
C6F6R1	3.79	3.84	200	9.4	47.2	162.2	37.8
C7F1R1	4.09	4.33	100	15.7	64.6	89.8	48.9
C8F1R1	3.49	4.58	200	11.0	85.0	179.5	74.0
C8F2R1	3.68	4.62	200	15.7	141.7	193.7	126.0
C8F3R1	4.11	4.31	200	14.8	74.0	178.0	59.2
C8F4R1	3.88	4.12	200	9.4	122.8	198.0	113.2
C8F6R1	3.91	4.10	200	29.9	138.6	193.7	108.7
C8F8R1	3.86	3.23	200	28.38	85.0	144.9	56.7
C8F5R1	3.80	4.12	200	9.4	52.0	178.0	42.6

TABLE IV: Mean rates of information transmission (channel capacities) for Parts I, II and III of this experiment. Means for Part I are shown for rest times of $4\frac{1}{2}$ sec. and $\frac{1}{2}$ sec. (A,B). The difference is significant ($P = 0.005$). The mean for the most sensitive range of the fibres is 3.20 ± 0.09 bits/stimulus (C). Part II(D) gives the value obtained using 128 stimulus categories without delinearization. Part II(E) shows the mean after delinearization. The difference is significant ($P = 0.0001$). Part III gives the calculated channel capacity for pooled data taken from 7 neurons.

		Stim. Uncert. (Bits)	Length of Stim. (sec)	Rest Time (sec)	Estimate of C (bit/stim)	# of Fibres in Sample	Significance (P value) of difference between A & B
PART I	A	5	$\frac{1}{2}$	$4\frac{1}{2}$	2.82		
	B	5	$\frac{1}{2}$	$\frac{1}{2}$	2.91	7	< 0.05
	C	5	$\frac{1}{2}$	$\frac{1}{2}$	3.20 ± 0.09	15	
PART II *	D	7	$\frac{1}{2}$	$\frac{1}{2}$	3.77 ± 0.14	14	
	E	7	$\frac{1}{2}$	$\frac{1}{2}$	4.33 ± 0.11		< 0.0001
PART III		7	$\frac{1}{2}$	$\frac{1}{2}$	6.84	7	

*PART II 100% of information transmitted over mean range $159.1 \text{ gm} \pm 7.9$
 95% of information transmitted over mean range $68.0 \text{ gm} \pm 7.9$

relatively narrow range of forces.

Part III of the study was carried out in order to gain an estimate of the effect of pooled signals from several neurons. The outputs of seven neurons, all collected from stimulation of one tooth in one cat by the same forces acting in the same direction, were recorded, and their output was pooled by summing the spike output of the seven neurons recorded for each input category. From this data the pooled rate of information transmission under these conditions was calculated. The channel capacity for the group, when treated as an information transmitting unit, was found to be 6.84 bits/stimulus, (see Table IV).

This exercise was carried out in order to see how much information is theoretically available to second-order periodontal mechanoreceptive neurons. If all seven units behaved independently, and if a sufficiently finely graded stimulus set was available, second-order neurons should receive information at a maximum rate equal to the sum of the information transmission rates of all seven neurons. This should be over 20 bits/sec. In this case, the measured rate of information transmission for the seven fibres was probably limited to 6.84 bits/stimulus by the coarseness of the stimulus set.

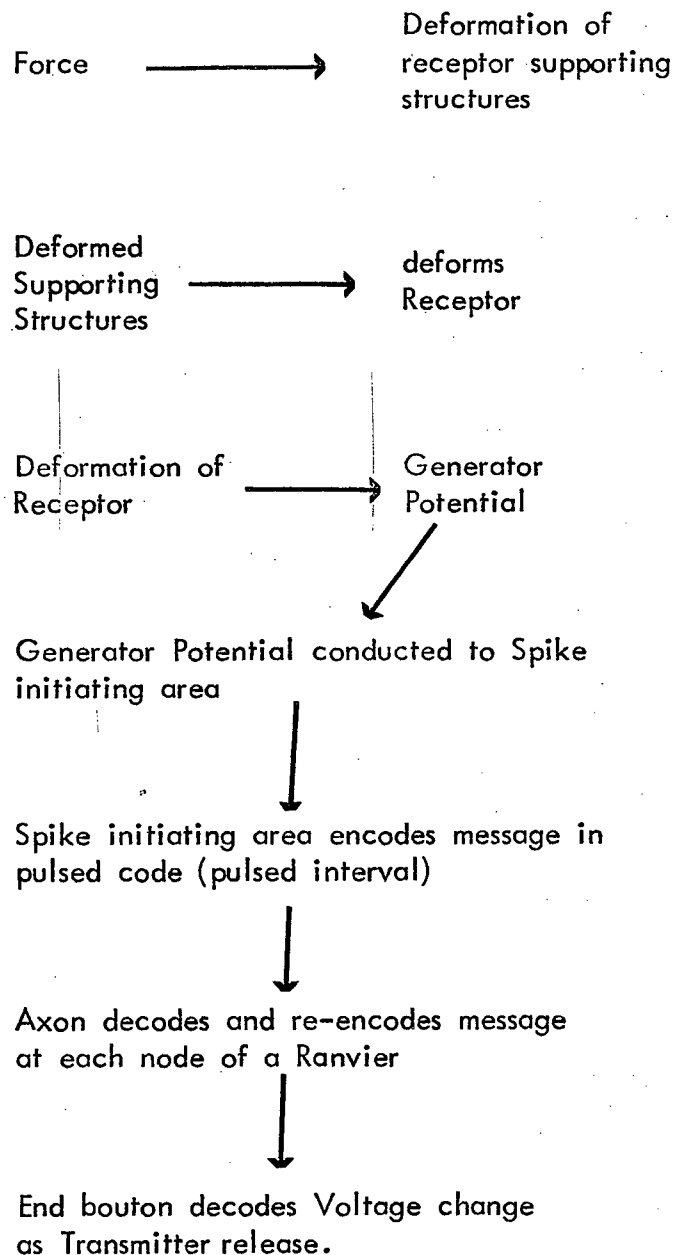
DISCUSSION

PART I

Slowly adapting periodontal mechanoreceptors, stimulated at a frequency of 1/sec, with 32 different stimulus intensities evenly spaced between 0 and 200 gms. produced a mean channel capacity of 3.2, S.E. = ± 0.09 bits per second. This exceeds values measured for cutaneous mechanoreceptive primary afferents which were assessed using similar putative input and output codes 2.5 bits/stimulus and 2.4 bits/stimulus respectively (Werner and Mountcastle, 1965; Darian-Smith et al, 1968). Thus, the periodontal mechanoreceptors would seem to be more sensitive than the primary afferents studied by Werner and Mountcastle and Darian-Smith et al. However, there are several possible explanations for these differences.

Since information output, for a particular decoding procedure, is equal to information input minus information lost due to noise, the increased channel capacity could be due to higher encoding rates at the receptor, reduced noise, or differences in decoding algorithms. The latter cause is by far the most likely. Werner and Mountcastle (1965) pooled response categories. Spike counts of 0 or 1, 2 or 3, 4 or 5 etc., were considered to be the same output. We did not carry out this procedure. Kenton and Kruger (1971) have shown that such treatment of the data causes a sizable drop in the measured channel capacity.

It should not be overlooked, however, that a number of other factors may cause one primary afferent neuron to be more sensitive than another. The general flow of events in a mechanoreceptor is as follows:



Noise

From the above argument, noise may be introduced in the following ways:

- (1) The same force produces different amounts of deformation in the supporting structure.

- (2) The same deformation in the supporting structure produces different amounts of deformation of the receptor in different trials.
- (3) The same deformation of the receptor produces different generator potentials.
- (4) The generator potential is noisy electrically.
- (5) The same generator potential produces a different amount of depolarization at the spike initiating area.
- (6) The spike generating area (encoder) is noisy.
- (7) The code output by the spike generating area is changed during transmission down the axon due, for example, to nodal noise.
- (8) Noise at the end bouton would not be of importance in this experiment.

In addition to random noise there exists accommodation and adaptation.

These processes may constitute either noise or part of the message depending upon the nature of the decoder.

Thus, noise may be added to the signal at numerous points along the neuron, in any of the processes which act to carry information. Less noise in any one of these processes in the periodontal mechanoreceptor, or a net decrease in noise over all these processes could explain the observed results. As pointed out in the Introduction, receptor noise is known to vary from structure to structure.

Receptor Supporting Structures

Differences between skin and periodontal ligament mechanoreceptor supporting structures could have contributed to differences in channel capacity.

It is possible that the supporting structure of the periodontal mechanoreceptor has a shorter mechanical recovery time than does that of a monkey hand or cat facial skin. A shorter recovery time would decrease correlation between successive outputs.

Coding Properties

Differences in the coding properties of the two structures may have caused the discrepancy between the results of Werner and Mountcastle's research (1965), and the findings of this study. It may be that 500 msec unweighted frequency codes chosen in these three experiments were closer to the optimal output code for the periodontal mechanoreceptor than for the receptors studied by Werner and Mountcastle and Darian-Smith. It might also be the case that the optimal input code for the periodontal mechanoreceptor more nearly fitted a rectangular distribution than did receptors studied by Mountcastle and Darian-Smith.

PART II

By increasing the stimulus uncertainty to 7 bits ($2^7 = 128$ stimulus categories) the mean channel capacity was increased to 3.77 bits/sec. S.E. = ± 0.14 . This was, of course, due to the fact that more finely graded stimuli allow more recognizable levels of response to occur.

In addition to this, by subsampling the original data set (see methods) the value calculated for the channel capacity was markedly increased to 4.33 bits/sec S.E. = ± 0.11 . All 14 channel capacities were increased after this procedure. The probability of this occurring by chance alone is $\frac{1}{2^{14}}$ (i.e. $P < 0.0001$). The reason for this phenomenon is intuitively obvious. If one has, for instance, a system which transmits information perfectly at a rate of 2 bits per second up to 4 grams and not at all above 4 grams, then more information will be transmitted when one restricts oneself to the 0 - 4 gram range than if one spends part of one's sampling time

unproductively sampling the 4 - 100 gram range which transmits no information.

For the same reason, taking many samples at forces for which transmission is relatively poor will severely reduce the overall rate of transmission and, therefore, the channel capacity. Thus, to accurately estimate the maximum possible channel capacity one must use:

- (1) an adequately fine stimulating and analysing system; and
- (2) a stimulus array in which the stimuli in the more sensitive parts of the dynamic range are more closely spaced than those in the less sensitive parts of the dynamic range.

It is worth pointing out that the reduced information transmission in the non-sampled data was due mainly to computational effects, and not to a lack of adequately refined measurement as implied by Kenton and Kruger (1971). Extremely large stimulus uncertainties such as the ones they employed (Kenton and Kruger, 1971) are probably unnecessary if one uses a sample distribution which is closely matched to the sensitivity of the receptor throughout its dynamic range. Reducing the stimulus uncertainty by a selective weeding out (albeit an approximate one) produced large increases in our measured channel capacity. In Part III of this discussion a system for more accurately achieving this end is considered.

Figures 3 and 4 present the raw data for a characteristic neuron in matrix form and the graph of input uncertainty versus information transmission respectively. The latter may be seen to round significantly as channel capacity is neared. This effect is due to the fact that, initially, doubling the number of input categories also doubles the number of categories which can be discriminated. However, later, when

there are already numerous stimulus categories, the added categories cannot all be discriminated from those previously present. For example, only information about some of the new categories can be transmitted. Thus the line deviates from its 45° climb and gradually flattens off. Now, if the receptor was stimulated with a non-linear set of force steps, the non-linearity of which matched the non-linearity of the receptor's sensitivity, the transition between slope = 45° and slope = 0° would be abrupt. This would occur since matching nonlinearities implies that all areas of the dynamic range reach their maximum discriminable sample density at about the same time. Therefore, either all added categories are discriminated, or none of them are. As the data of Kenton and Kruger (1971) did not behave in this manner, it must be assumed that their distribution was not optimum.

The outputs of seven neurons all collected from stimulation of one tooth in one cat by forces acting in one direction were summed (i.e. treated as if all terminated on one second order neuron and all were weighted equally) and decoded as before. This resulted in an information transmission rate of 6.84 bits per stimulus. Channel capacity had definitely not been reached as the slope of the plot of input uncertainty versus information transmission was still extremely steep at 7 bits of input uncertainty. This indicates that the sum of these seven inputs is capable of carrying far more information than our experimental setup allowed us to input.

However, from calculations made on the data of Soltis et al (1971) it was found that the perceptual information transmission for the human incisor was only

3.9 bits/stimulus. Thus, either the canines of cats are far better endowed with mechanoreceptors than are human incisors, or information is not carried undiminished between the first order neuron and the CNS. This conclusion is contrary to that of Werner and Mountcastle (1965) and Darian-Smith (1968). The experiment should be repeated on primates, since here studies of perception and primary afferent sensitivity could perhaps be done in the same animal. Also, to allow direct comparison, the data which is being compared to Mountcastle's should be recalculated using the same output code.

Soltis' data further indicates that following orthodontic treatment, patients are better able to discriminate forces applied to their teeth than they were originally, (Soltis et al, 1971). This surprising observation would seem to indicate that the limits of perceptual sensitivity are determined, at least in part, peripherally. Thus it would seem, considering the data of Kenton and Kruger, that either central or peripheral effects may limit perceptual ability.

Stimulus-response curves tend to flatten very slowly at the tops of their ranges. This makes it difficult to see exactly where the top of the dynamic range occurs. In this study it was hoped that by finding the part of the stimulus curve in which 100% and 95% of the information is transmitted, the important part of the dynamic range could be better defined. It was found that 100% of the information is transmitted over a mean range of $159.1 \text{ gm} \pm \text{S.E. } 7.9 (n = 14)$ with a mean rate of information transmission of $4.33 \text{ bits/stimulus S.E. } \pm 0.11$. Thus discriminable categories were separated on average by about 6.6 gms. On the other hand, 95% of the information ($4.09 \text{ bits/stimulus S.E. } \pm 0.10$) was transmitted over a mean range of 67.9 gms. $\text{S.E. } \pm 7.9$. Thus, over 95% of the range, the mean discriminable step

was 4.2 gms. It is of interest that the 95% range has a much higher % deviation than does the 100% range. This would seem to indicate that distribution of 100% of the dynamic range of a primary afferent ensemble may leave much unsaid about the range over which information is transmitted by the ensemble.

PART III

As mentioned in Part I, the choice of output code plays a very important part in determining the figure arrived at for information transmission. In our continuing studies of nerve neurons we are currently experimenting with an interval code because an interval code

- (1) more nearly approximates conditions at the first synapse;
- (2) allows the development of methods useful with continuous data in general; and
- (3) allows data collected to be re-evaluated for a large number of code types.

Preliminary estimates, on 5 neurons only, indicate that the first interval carries information at a rate of 4.5 bits/stimulus. Since the mean duration of the first interval was 12 msec, it might theoretically be possible to transmit information at rates up to 360 bits/sec. As has been pointed out by Stein (1967) such rates are probably not achievable in practice because of the high correlation expected between consecutive outputs at high rates. Nonetheless, it is also obvious that rates much higher than 5 bits/sec should be achievable.

Sensitivity

A method frequently used in the description of primary afferent behaviour is the stimulus response curve. Figure 6 shows the stimulus response curves for two hypothetical neurons. It can be seen that the dynamic range of neuron "A" lies approximately between m and n gms. Further, it is apparent by inspection that: (1) the slope of stimulus response curve "A" is greatest around "a"; (2) the slope of stimulus response curve "B" is greatest around "B" and (3) the slope at "a" is greater than the slope at "b". Such data has been used to infer that: (1) the sensitivity of neuron "A" is greater than the sensitivity of neuron "B"; (2) neuron "A" is most sensitive at "a"; and (3) neuron "B" is most sensitive at "b". (Brown, Engberg & Matthews, 1967; Mountcastle, 1968; Hensel, 1968; Werner and Mountcastle, 1965). Examination of these assertions in the light of information theory reveals some basic flaws.

A neuron tells us something about the state of the universe and often about changes of state. Its ability to discriminate between states is defined as its sensitivity. That is, the sensitivity of a sensor is great if it can distinguish small changes, and small if it can only distinguish large changes. But information is defined as the quality which allows us to distinguish two or more things. Thus, when we set out to measure sensitivity, we are actually trying to measure information. Information theory constitutes a method for measuring (1) the information content of a set, or (2) how much information we are obtaining on average about some set. It can be formally proved (see Appendix A) that

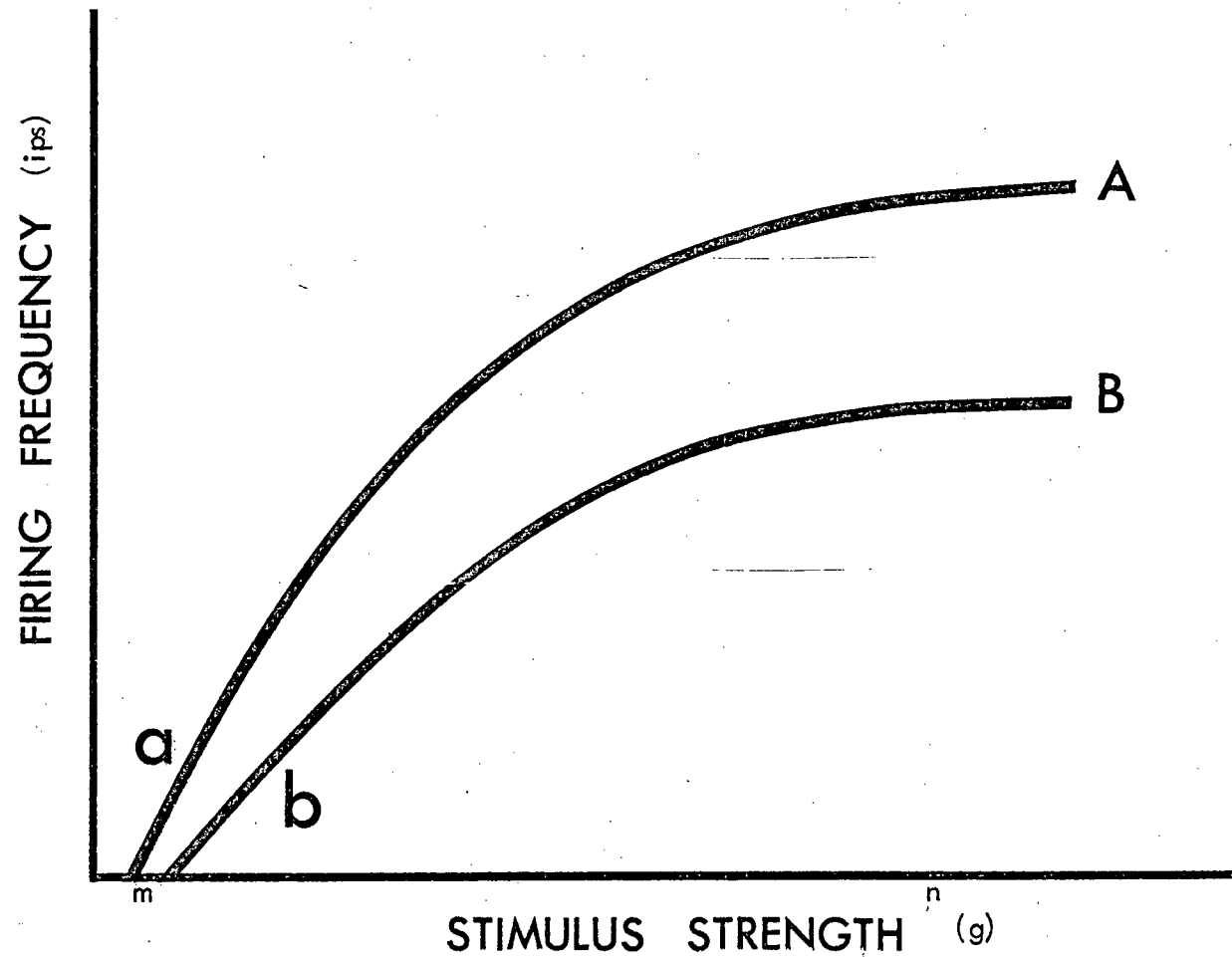


Fig. 6 Hypothetical stimulus response curves for neurons "A" and "B". The dynamic range of neuron "A" lies between m gms. and n gms. The slope of the stimulus response curve for "A" is greatest at a. The slope for "B" is greatest at b.

information may only be measured by information theory based measures. Thus, sensitivity, in the strict sense, may be measured only by the use of information theory-based metrics. Other methods for measuring sensitivity (e.g. slope of stimulus response curve, see Introduction) constitute approximations, the accuracy of which have never been appraised.

As discussed in the introduction, it seems that several types of noise are present and that their amplitudes will be correlated with the distribution of interspike intervals. What is more, the characteristics of this noise will vary with the tissue, geometry of the tissue, and the receptor potential. There is no reason to expect, a priori, that the noise in one receptor system is the same as in another. Thus, it is clear that two primary afferents which display the same stimulus response curves could have remarkably different sensitivities.

Behavioral experiments have shown that the low ranges of stimulus intensity are perceived most finely. This cannot be viewed as strong evidence for receptors being most sensitive in the low ranges as several alternate explanations are possible for these observations. Since many receptors are involved in psychophysically assessed phenomena, and since, as shown by our data for seven pooled fibres, very high rates of information transmission can be produced from pooled inputs, the proportionally greater loss of information from different parts of the dynamic range could produce this effect. Moreover, our data tends to indicate that much more information is encoded at the level of the primary afferent than is perceived at a conscious level. Thus, relatively high perceptual sensitivity in one part of the dynamic range could be due entirely to losses of

information from other parts of the dynamic range due to central effects.

In experiments still in progress, an attempt is being made to measure the relative sensitivity of a primary afferent neuron throughout its dynamic range. To measure the relative rates of information transmission in different parts of the dynamic range one would like to break the data matrix up into N matrices. This, however, cannot be done since removing one subset will have an effect on the apparent information transmitting abilities of its neighbours. But noise in an information matrix is expressed as the overlap of response categories filled by the results of different stimulus categories. When the overlap between response categories is everywhere the same, the same amount of information is carried by each stimulus category (given that the probabilities of the stimulus categories are made equal), i.e., each category then carries C/N bits of information where C is the channel capacity, and N is the number of stimulus categories. Now there is a theorem in information theory (Ash, 1965) which states that the information transmission through a system will be maximal when the code elements are equiprobable, and each carries the same amount of information. Stein (1967) found that the rate of information transmission was maximized for an artificial neuron when the stimulus categories were spaced in such a way that the responses occurred an equal number of standard deviation units apart.

Thus it seems that measurement of the sensitivity of a neuron throughout its dynamic range could be carried out in the following way:

- (1) Collect the data using forces such that each interval is $K \cdot S.D.$ units greater than the last, K being a convenient constant.
- (2) Then $\frac{C/N_{\max}}{F_m - F_{m-1}}$, for $m = 1$ to N_{\max} (where N_{\max} = the number of the largest force used and F_m = Force number m should give a measure of the regional sensitivity in bits/gm.

Further investigation of the information loss during decoding is warranted. This could be accomplished by measuring the rate of information transfer between a stimulus set and the generation potential, and between generator potential and the pulse code. In this way one might establish the most important factors limiting the rate of information coding.

APPENDIX A

The amount of information needed to distinguish an element is a function of the probability of that element. Suppose there is a box of stones. How much information does it take to distinguish between them? If there is only one stone, obviously no information at all is needed. Suppose we had ten stones, whose weights were distributed randomly, and decided to distinguish them by their weights. Only comparatively inaccurate measurements of weight would have to be used on average. But with 100,000 or 10^6 stones, with the same range and distribution of weights, one would have to measure extremely accurately to distinguish between them. Thus, on average, more information is needed to distinguish one element from another for a larger number of elements. But the probability of choosing a particular element at random = $1/N$, where N = the number of elements. Thus information is a function of probability.

Information is a function only of probability. What is the easiest way to distinguish between our stones? Intuitively, if we painted numbers on them, this would be about as easy a method as we could devise to tell them apart. How long do these numbers have to be? It is obvious that a binary number 2^I digits long will always be adequate to number N elements if $I = \log_2 N$ (where I is an integer). Thus, the length of the number which must be dealt with when trying to distinguish our stones, is a function only of the number of elements (stones) among which we must choose, and therefore is a function only of the probability of the elements' occurrence.

How much information is carried by an event whose probability is zero or an event whose probability is one? Imagine a fish living at the bottom of a deep sea trench. Say the fish has a water receptor and the only information this receptor encodes is yes - it is covered with water, or no - it is not covered with water. How much information is this receptor giving the fish about its environment assuming the fish never surfaces. Of course it is giving no information because we already know the fish can never surface. Similarly, a water receptor possessed by an animal that never touches water conveys no information. Thus events whose probability of occurrence equals zero or one carry no information.

How much information is carried by two independent events which occur together? This must, of course, be equal to the sum of the information carried by each event separately. Suppose event 1 has the probability of occurrence P_1 and event 2, P_2 . Then the probability of their occurring together $P_{1,2} = P_1 \cdot P_2$. Now let us define $I_i = F(P_i)$ as the information carried by event i whose probability of occurrence is P_i

$$\begin{aligned} \text{Then} \quad I_1 &= F(P_1) && \text{def.} \\ I_2 &= F(P_2) && \text{def.} \end{aligned} \quad (\text{Eqn. 1})$$

but the information carried by two uncorrelated events separately must be equal to the information carried by the two events together.

$$\begin{aligned} \therefore \quad I_1 + I_2 &= F(P_1 \cdot P_2) \\ &= F(P_1) + F(P_2) \end{aligned}$$

$$\text{or } F(P_1 \cdot P_2) = F(P_1) + F(P_2). \quad (\text{Eqn. 2})$$

It is obvious that $I_i = F(P_i) = -\log_k P_i$ satisfies equations 1 and 2
($K \neq 1, 0$). Let us assume $K = 2$.

Now let us prove that any function satisfying these equations must be equivalent to

$$-\log_2 P_i$$

Say we have a function $F(P_i)$ for which the above assertions are true.

Then, choosing C and r arbitrarily, we can make $F(C^r)$ equal to any number.

Let this number lie between two other numbers $F(M^K)$ and $F(M^{K+1})$

$$\therefore F(M^K) \leq F(C^r) \leq F(M^{K+1})$$

by induction on equation 2.

$$KF(M) \leq rF(C) \leq (K+1)F(M)$$

$$\text{or } K \leq \frac{rF(C)}{F(M)} \leq K+1$$

but we can also for these same numbers make the assertion

$$\log_2(M^K) \leq \log_2 C^r \leq \log_2(M^{K+1})$$

$$\text{or } K \leq \frac{r \log C}{\log_2 M} \leq K+1$$

$$\text{or } \frac{F(C)}{F(M)} - \frac{r \log C}{\log_2 M} \leq 1.$$

or
$$\frac{F(C)}{F(M)} - \frac{\log C}{\log_2 M} = \frac{1}{r}$$

but r can be any number. Assume $r = \infty$

then
$$\frac{F(C)}{F(M)} = \frac{\log C}{\log_2 M}$$

$$F(M) = \frac{\log(C)}{F(C)} \cdot \log_2(M)$$

or
$$F(M) = -K \log_2(M).$$

Q. E. D.

This is an extremely important theorem as it demonstrates that only information theory metrics can describe information flow, as defined in information theory.

Now suppose we start drawing stones from our box. If P_i is the probability of drawing the i^{th} stone, then each stone gives us $I_i = -\log_2 P_i$ bits of information. But if we draw these stones at a rate of 1/sec. for T sec., the i^{th} stone will be drawn $P_i \cdot T$ times. Or in T seconds we produce $I_{T,i} = -T \cdot P_i \log_2 P_i$ bits of information, or about $-P_i \log_2 P_i$ bits of information/sec. Let us define the source entropy

$$H(X) = - \sum_{i=1}^N P_i \log_2 P_i.$$

This is the average amount of information (in bits) that is being produced by our source (box of stones) per unit time.

Now let us drop our stone analogy and use the example of a person drawing weights and placing them on a mechanoreceptor. We wish to decide what weight was placed on the receptor from the output which we observe in the axon.

Now $2^{TH(X)}$ = the total number of input states in time T .

And if $q(j)$ is the probability of receiving from the axon a frequency indicating force j was the stimulus then the probability of any sequence is

$$P = \frac{N}{11} \prod_{j=1} q(j)^{N_j}$$

$$= \frac{N}{11} \prod_{j=1} q(j)^{Tq(j)}$$

where N_j = the number of action potential sequences per unit time (indicating force j) which were received, and T = the total sampling time with a rate of 1/sec.

Now the probability that the output frequency j is associated with the input force i

$$P_{i/j} = \frac{N}{11} \prod_{q_i(j)} P_i^{Tq_i(j)}$$

$$\therefore \log_2 P_{i/i} = \sum_{j=1}^M \sum_{i=1}^N P_j T_{ji}(i) \log_2 q_{ji}(i)$$

$$\text{define } H(X/Y) = - \sum_{j=1}^M \sum_{i=1}^N P_j q_{ji}(i) \log_2 q_{ji}(i)$$

$$\therefore \log_2 P_{i/i} = TH(X/Y)$$

$$\text{or } P_{i/i} = 2^{-TH(X/Y)}$$

$$\text{or } N_{i/i} = 2^{TH(X/Y)}$$

$$\therefore \text{The probability that an output is correct} = 2^{TH(X)} \cdot 2^{-TH(X/Y)}$$

now if R is the rate of information transmission in bits/sec

$$2^{TR} = 2^{TH(X)} \cdot 2^{-TH(X/Y)}$$

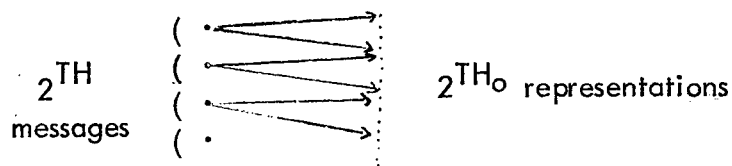
$$R = H(X) - H(X/Y)$$

Shannon Weaver Theorem

If there exists a channel with capacity = C, then for any source S with $H \leq C$ there exists a coding system by which S can be transmitted through the channel with an arbitrarily small probability of error.

Suppose we choose a source S_0 for which $H_0 < H$ and which contains S .

Define $C = \max (H_0(X) - H_0(X/Y))$



\therefore the probability that any of the 2^{TH_0} points is a message is

$$\frac{2^TH}{2^{TH_0}} = 2^{T(H - H_0)}$$

The probability that any one of the points is not the cause of the message is $(1 - 2^{T(H-H_0)})$

but each of the 2^{TH_0} points can be represented on average in $2^{TH_0(X/Y)}$ ways when transmitted through the channel.

\therefore there are $2^{TH_0(X/Y)} - 1$ possible causes for the message neglecting the real one.

\therefore the probability that the cause of the message was the correct corresponding input is

$$P = (1 - 2^{T(H_0 - H)}) 2^{TH(X/Y)} - 1$$

now $H - H_0 \leq H_0(X/Y)$ by def.

Let k be some positive number such that

$$H - H_0 = -H_0(X/Y) - k$$

$$\therefore P = (1 - 2^{(2^{-T(H(X/Y)+k)})} 2^{TH(X/Y)} - 1$$

but $(1 - a)^b \approx 1 - ab$.

$$\therefore P = 1 - 2^{-TK}$$

$$\text{or } P \rightarrow 1 \text{ as } T \rightarrow \infty$$

or for a long enough message we can produce an arbitrarily small probability of error.

Derivation of Computational Formula used to Calculate Information Transmission Rates.

$$\begin{aligned}
 H(X, Y) &= \sum_{i,j} P_{ij} \log P_{ij} \text{ (def)} \\
 &= \sum_{i,j} P_{ij} \log P_i \cdot R(X/Y) + \sum P_{ij} = P_i \cdot P_{i/i} \\
 &= \sum P_{ij} \log P_i + \sum P_{ij} \log P_{X/Y} \\
 &\quad \sum P_i \cdot P_i \log P_i + \sum P_i \cdot P_{i/i} \log P_{i/i} \\
 &= H(X) + H(X/Y)
 \end{aligned}$$

$$\therefore H(X/Y) = H(X, Y) - H(Y)$$

$$\text{or } R = H(X) - H(X/Y)$$

$$= H(X) + H(Y) - H(X/Y).$$

BIBLIOGRAPHY

- Adrian, (1928). The Basis of Sensation. New York : Norton.
- Adrian, E.D. (1931). The Messages in Sensory Nerve Fibres and Their Interpretation. Proc. Roy. Soc. London, Ser. B. 109, 1-18.
- Adrian, R.J., and Zotterman, Y. (1926). The Impulses Produced by Sensory Nerve Endings. J. Physiol., 61 : 151-171. 1926.
- Anderson, D.J., Hannam, A.G., and Mathews, B., (1970). Sensory Mechanisms in Mammalian Teeth and Their Supporting Structure. Physiol. Reviews, 50, 171-195.
- Armett, C.J., Gray, J.A.B., Hunsperger, R.W., and Lals (1962). The Transmission of Information in Primary Receptor Neurons and Second Order Neurons of Phasic Systems. J. Physiol., 1962, 164 : 395-421.
- Ash, R. (1965). Information Theory. John Wiley & Sons, Inc., New York.
- Barlow, H.B. (1963). (Kybernetik), 2 : 1. The Information Capacity of Nervous Transmission.
- Barlow, H.B. (1957). Increment Thresholds of Low Intensities Considered on Signal/Noise Discriminations. J. Physiol., 136 : 469-488, 1957.
- Biscoe, T.J., and Taylor, A. (1963). The Discharge Pattern Recorded in Chemo-Receptor Afferent Fibres from the Carotid Body of the Cat. J. Physiol. 1963, 168 : 332-344.
- Blair, E.A., and Erlanger, J. (1933). The Configuration of Axon and Simple Nerve Action Potentials. Am. J. Physiol. 106 : 524, 1933.
- Boring, E.G. (1933). Cutaneous Sensation After Nerve-Division. Quart. J. Exp. Physiol., 10 : 1-95, 196.
- (1942). Sensation and Perception in the History of Experimental Psychology. N.Y., Appleton-Century-Crofts. 1942.
- Borsellino, A., Fuortes, M.G.F., and Smith, T.G. (1965). The Visual Responses in Lumulus, Cold Spring Harbour Symposium of Quantitative Biology. 30, 429-443.
- Brown, M.C., Engberg, I., and Matthews, P.B.C. (1967). The Relative Sensitivity to Vibration of the Muscle Receptors of the Cat. J. Physiol. (Lond.) Vol. 192. pp 773-780.

- Brown, A.G., and Iggo, A. (1967). A Quantitative Study of Cutaneous Receptors and Afferent Fibres in the Cat and Rabbit. *J. Physiol.* 193 : 707-733, 1967.
- Brown, A.G., and Stein, R.B. (1966). Quantitative Studies on the Slowly Adapting Stretch Receptor of the Crayfish. *Kybernetik* 3 bd. Heft. 4, 1966.
- Buller, A.J., Nicholls, J.G., and Strom, G. (1953). Spontaneous Fluctuations in the Excitability of the Muscle Spindle of the Frog. *J. Physiol.* 1953, 122 : 409 - 418.
- Cauna, N. and Ross, L.C. (1966). The Fine Structure of Meisner's Touch Corpuscles in the Human Finger. *J. Biophys. Biochem. Cytol.* 8, 467-482.
- Darian-Smith, I., Rowe, M.J., Sessel, B.J. (1968). "Tactile" Stimulus Intensity: Information Transmission by Relay Neurons in the Trigeminal Nuclei. *Science.* Vol. 160, pp 791-794.
- De Lange, A., Hannam, A.G., and Matthews, B. (1969). The Diameter and Conduction Velocities of Fibres in the Terminal Branches of the Inferior Dental Nerve. *Archs. Oral. Biol.* Vol. 14, 513-519. 1969.
- Eccles, J. (1964). The Physiology of Synapses. Academic Press Inc., New York. (1964).
- Edwards, C., and Ottoson, D. (19). The Site of Impulse Initiation in the Nerve Cell of a Crustacean Stretch Receptor. *J. Physiol.* 143. pp 131-144.
- Erlanger, J., Blair, E.A., and Schoepfle: A Study of the Spontaneous Oscillations in the Excitability of Nerve Fibres with the Special Reference to the Action of Strychnine. *Amer. J. Physiol.* 134, 705. 1941.
- Eyzaguirre and Kuffler (1955). Process of Excitation in the Dendrites and the Source of a Single Isolated Sensory Nerve Cell of the Lobster and Crayfish. *J. Gen. Physiol.* 39, 87-119. (1955).
- Fatt, P., and Katz, B. (1952). Spontaneous Sub-Threshold Activity in Motor Nerve Endings. *J. Physiol.* 117, 109. 1952.
- Fechner, G.T. (1860). Elemente der Psychophysik, Leipzig. Breitkopf and Hartel. 1860.
- Frankenhaeuser, B., and Hodgkin, A.L. (1956). The After Effects of Impulses in the Giant Nerve Fibre of Loligo. *J. Physiol.* 131. p. 341. (1956).
- Fuortes, M.G.F., (1959). Initiation of Impulses in the Visual System of Limulus. *J. Physiol.* 148 : 14-28.
- Fuortes, (1958). Electric Activity in the Cells of the Eye of Limulus. *Americ. J. Opthal.* 46, 210-223 (1958).

- Fuortes, M.G.F., and Mantagazzini, F. (1962). Interpretation of the Repetitive Firing of Nerve Cells. *J. Gen. Physiol.* 45, 1163-1179 (1962).
- Galambos, R., and Davis, H. (1943). Responses of a Single Auditory Nerve Fibre. *J. Neurophysiol.* 6, 39-59.
- Griffin and Harris (1968). Unmyelinated Nerve Endings in the Periodontal Membrane of Human Teeth. *Arch. Oral Biol.* 13, 1207-1212.
- Griffin, C., and Spain, H. (1972). The Fine Structure of the Periodontal Nerve Plexus in Man. In: *A Symposium on the Maxillomandibular Apparatus. Proc. Int. Anat. Congr. Leningrad. 1970.*
- Griffin, C.J. (1972). The Fine Structure of End Rings in the Human Periodontal Ligament. *Arch. Oral Biol.* Vol. 17. pp 785-797.
- Hagiwara, S. (1954). Analysis of Interval Fluctuation of the Sensory Nerve Impulse. *Jap. J. Physiol.* 4, 234-240. 1954.
- Hall, R. (1964). Mass. Inst. Technol. Res. Lab. Electron. Tech. Rept. 416. as cited by Stein (1967).
- Hannam, A.G. (1969). The Response of Periodontal Mechanoreceptors to Controlled Loading of the Teeth. *Arch. Oral Biol.* 14: 781-791.
- Hannam, A.G. (1968). The Conduction Velocity of Nerve Impulses from Dental Mechanoreceptors in the Dog. *Arch. Oral Biol.* 13, 1385-1393.
- Harline, H.K., Coulter, N.A., Wagner, H.G. (1952). Effects of Electric Current on Responses of Single Photoreceptor Units in the Eye of Limulus. *Fed. Proc.* 11, 65-66 (1952).
- Hering, F. (1876). Zur Lehre de Beziehung Zwischen Lub und Seele. I. Uber Fechner's Psycho-physiisches Gesetz. *Sitzumb. Akad. d. Wessench., Math.-naturim.* K1 Wien, 72, Abth 3 : 310-348, 1876. - As cited by Solnsew (1972).
- Hensel, H. (1968). Electrophysiology of Cutaneous Thermoreceptors. In D.R. Kenshalo, ed. *The Skin Senses.* Springfield, Ill.: C.C. Thomas, 1968. pp 384-396.
- Hodgkin, A.L. (1948). The Local Electric Charges associated with Repetitive Action in the non-Myelinated Axon. *J. Physiol. (Lond.)* 165-181.
- Husmark, I., and Ottoson, D. (1971). The Contribution of Mechanical Factors to the Early Adaptation of the Spindle Response. *J. Physiol.* (1971) 212 : 557-597.

- Ito, F. (1969a). Abortive Spikes in Frog Muscle Spindles. Jap. J. Physiol. Vol. 17, pp 373-391.
- Ito, F. (1969b). Functional Properties of Tandem Muscle Spindles in Comparison with Those of the Muscle Spindle of the Frog. Jap. J. Physiol. Vol. 19. pp 641-651.
- Jacobsen, H. (1950). Information Capacity of the Human Ear. Science 112 : 143.
- Jacobsen, H. (1951). Information capacity of the Human Eye. Science 113 : 292.
- Jansen, J.K.S., Nicolaysen, K., and Rudyard, T. (1966). Discharge Pattern of Neuron of the DSCT activated by Static Extension of Primary Endings of Muscle Spindles. J. Neurophysiol. 29, 1061-1086, (1966).
- Katz, B. (1950). Depolarization of Sensory Terminals and the Initiation of Impulses in the Muscle Spindle. J. Physiol. (Lond.) 111, 261-282 (1950).
- Katz, B., and Miledi, R. (1972). The Statistical Nature of the Acetylcholine Potential and its Molecular Components. J. Physiol. (1972). 224, pp. 665-699.
- Kizior, J.E., Cuozzo, J.W., and Bowman, D.C. (1968). Functional and Histological Assignment of the Sensory Enervation of the Periodontal Ligament of the Cat. J. Dent. Res. 47 : 59-64. (1968).
- Kenton, B., and Kruger, L. (1971). Information Transmission in Slowly Adapting Mechanoreceptor Fibres. Exp. Neurology 31, 114-129 (1971).
- Krinjevic, K., and Van Gelder, N.M. (1961). Tension Changes in Crayfish Stretch Receptors. J. Physiol. (Lond.) 159. 310-325 (1961).
- Lloyd, F. (1946). Howell's Text of Physiol. pp. 140-141. Philadelphia and London. W. B. Saunders Co.
- Lowenstein, W.R., and Rathkamp, R. (1958). Sites of Mechanoelectrical Conversion in the Pacinian Corpuscle. J. Gen. Physiol. 41, 1245-1265 (1958).
- Lowenstein, W.R., (1959). The Generation of Electrical Activity in a Nerve Ending. Ann. N.Y. Acad. Sci. 81, 367-387 (1957).
- Lowenstein, W.R., and Mendelson, M. (1965). Components of Receptor Adaptation in a Pacinian Corpuscle. J. Physiol. (Lond.) 177, 377-397 (1965).
- Lowenstein, W.R., and Skalak, R. (1966). Mechanical Transmission in a Pacinian Corpuscle. J. Physiol. (Lond.) 182, 346-378 (1966).

- Lowenstein, W.R., Terzuolo, C.A., and Washigu, Y. (1963). Separation of Transducer and Impulse Generating Processes in the Pacinian Corpuscle. *Science* 142, 1180-1181. (1963).
- Lundberg, A. (1964). Ascending Spiral Hindlimb Pathways in the Cat. In J.C. Eccles and J.P. Schaeffer eds. *Physiol. of Spinal Neurons*. Proj. in Brain Res. 1963. Amsterdam : Elsevier.
- MacNichol (1956). Visual Receptors as Biological Transducers IN: *Molecular Structure and Functional Activity in Nerve Cells*. Eds. Grenell, K.G., Mullins, L.J., Wash. Amer. Inst. Biol. Sci. Pub.
- Mackay, D.M., and McCulloch, W.S. (1952). The Limiting Information Capacity of a Neuronal Link. *Bull. Math. Biophys.*, 14, 127-135. (1952).
- Mathews, B.H.C., (1931). The Response of a Single End Organ. *J. Physiol.* 71 : 64-110, 1931.
- Mountcastle, V.B. (1968). Quantitative Relations Between Mechanical Structure to the Skin and neural Responses Evoked by Them. IN D.R. Kenshalo, ed. *The Skin Senses*. Springfield, Ill. : C.C. Thomas, 1968 pp 112-138.
- Moller, E., 1966. The Chewing Apparatus. *Acta Physiol. Scand.* 60 : supp. 280.
- Mountcastle, V.B., Poggio, R., and Werner, G. (1963). The Relation of Thalamic Cell Response to Peripheral Stimuli varied Over an Intensity Continuum. *J. Neurophysiol.* 26 : 807-834. 1963.
- Munzer, B.L., (1965). The Intraepidermal Innervation of the Snout Skin of the Opposum. A Light and Electron Microscope Study with Observations of the Nature of Merkel's Tastzellen. *J. Cell. Biol.* 26, 79-97.
- Nakajima, S., and Onodera, K. (1969a). Membrane Properties of the Stretch Receptor Neurons of Crayfish with Particular Reference to the Mechanisms of Sensory Adaptation. *J. Physiol. (Lond.)* 200, 161-185 (1969a).
- _____. (1969b). Adaptation of the Generator Potential in the Crayfish Stretch Receptors under Constant Length and Constant Tension. *J. Physiol. (Lond.)* 200, 187-204. (1969b).
- Narahashi, T. (1964). Restoration of Action Potential by Anodal Polarization in the Lobster Giant Axons. *J. Cell. Comp. Physiol.* 64, 73-96 (1964).
- Ness, A.R. (1954). The Mechanoreceptor of the Rabbit Mandibular Nerve. *J. Physiol. (Lond.)* 126, 475-493.
- Ottoson, Dard, and Shepherd, G.M. (1970). Transducer Characteristics of the Muscle Spindle as Revealed by its Receptive Potential. *Acta Scand. Physiologica* 76, 49-63. (1970).

- Ottoson, D., and Shepherd, G.H. (1965). Receptor Potentials and Impulse Generation in the Isolated Spindle during Controlled Extension. Cold Spring Harb. Symp. Quart. Biol. 30, 105-113. (1965).
- Pecher, C. (1936). Etude Statistique des Variations Spontanees de l'excitabilite d'une fibre Nerveuse. CR. Soc. Biol. 1936, CXXII, 87.
- Pecher, C. (1937). Fluctuations Independents de l'excitabilite de Deux Fibre d'une Nerve. CR. Soc. Biol. 1937, CXXIV, 839.
- Pecher, C. (1939). La Fluctuation d'Excitabilite de la Fibre Nerveuse. Arch. Intern. Physiol. 49, 129-152 (1939). As cited by Verveen et al (1968).
- Perkel, D.H., and Bullock. (1968). A Report Based on an NRP Work Session. Neurosci. Res. Bull. 6 : 221-348. (1968).
- Pfaffman, C. (1939). Afferent Impulses from Teeth due to Pressure and Noxious Stimuli. J. Physiol. (Lond.) 180, 500-515.
- Quasler, H. (1956). In 3rd Symposium on Information Theory. C. Cherry, editor. Butterworth & Co. Ltd. London, 361. (1956).
- Rappaport, A. and Horvath, W.J. (1960). Maximum Information Capacity of Nerve Cells. Inform. and Control. 3 : 335.
- Rang, H.P., and Ritchie, J.M. (1968). On The Electrogenic Sodium Pump in Mammalian non-myelinated Nerve Fibres. J. Physiol. Vol. 196 pp 183-221.
- Shannon, C.E., and Weaver, W. (1949). The Mathematical Theory of Communication (Univ. Illinois Press. Urbana, Ill., 117 pp.) (1949).
- Simpson, H.E. The Enervation of the Periodontal Membrane. J. Periodontol. 37 : 374-376.
- Soltis, J.E., Nakfor, P.R., and Bowman, D.C. (1971). Changes in the Ability of Patients to Differentiate Intensity of Forces Applied to Maxillary Central Incisors During Orthodontic Treatment. J. Dent Res. 50 : 3 pp. 590-596. (1971).
- Somjen, J. (1972). Sensory Coding in the Mammalian Nervous System. Appleton-Century-Crofts. New York. N.Y.
- Stevens, S.S. (1951). Mathematical Measurements in Psycho-physics. In Stevens, S.S., ed., Handbook of Experimental Psychology. 1-47. New York, Wiley. (1951).

- Stevens, S.S. (1961). Transfer Functions of the Skin and Muscle Senses. In Rosenblith, W.A. ed. Sensory Communications Symposium, Endicott House, 1959, 1-33. Cambridge, Mass. MIT Press (1961).
- _____ (1966), and Knight, eds. Touch, Heat and Pain. Ciba Foundation Symposium. 3-17, Boston, Little, Brown & Co. (1966).
- _____ (1957). On the Psychophysical Law. Psychol. Rev. 64 : 153-181.
- Stein, R.B. (1967). The Information Capacity of Nerve Action Potentials Using a Frequency Code. Biophys. J. 7 : 797-826 (1967).
- Stein, R.B. (1970). The Role of Spike Trains in Transmitting and Distorting Sensory Signals. In the Neurosciences. Vol. 2. Rockefeller University Press. N.Y. (1970).
- Stein, R.B., and French, A.S. (1972). Models for the Transmission of Information by Nerve Cells. In Excitatory Synaptic Mechanisms. ed. Andersen, P., and Jansen, J. pp 247-257. Oslo. Oslo University Press.
- x Teorell, I. Electrokinetic Considerations of Mechano-Electrical Transduction. Ann. N.Y. Acad. Sci. 137. 950-966.
- Terzuolo, C.A., and Washigu, G. (1962). Relation Between Stimulus Strength, Generator Potential and Impulse Frequency in the Stretch Receptor of Crustacea. J. Neurophysiol. 25, 56-66. (1962).
- Vallbo, A.B. (1964). Accommodation in Relation to Inactivation of the Sodium Permeability in Single Myelinated Nerve Fibres from *Xenopus Laevis*. Acta Physiol. Scand. 1964. 61 : 429-444.
- Vallbo, A.B. (1964). Accommodation of Single Myelinated Nerve Fibres from *Xenopus Laevis* Related to Type of End Organ. Acta Physiol. Scand. 1964. 61 : 413-728.
- Verveen, A.A., and Derksen (1965). Fluctuations in the Membrane Potential of Axons and the Problem of Coding. Kybernetik, 2, 152-60 (1965).
- Verveen, A.A., and Derksen, H.E. (1968). Fluctuation Phenomena in Nerve Membrane. Proc. IEEE 56, 906-916. (1968).
- Walloe, L. (1972). Information Loss During Synaptic Transfer. In Excitatory Synaptic Mechanisms. Ed. Andersen, P., and Jensen, J. pp 258-261. Oslo : Oslo University Press.
- Walloe, L., Jansen, J.K.S., and Nygaard, K. (1969). A Computer Simulated Model of a Second Order Sensory Neuron. Kybernetik 6, 130-140. (1969).

Werner, G., and Mountcastle, V. B. (1965). Neural Activity in Mechanoreceptive Cutaneous Afferents: Stimulus Response Relations, Weber Functions, and Information Transmission. J. Neuro. 28, 1965.

Werner, G., and Mountcastle, V. B. (1968). Quantitative Relations Between Mechanical Stimulae to the Skin and Neural Responses Evoked by Them. pp 112-138. In "The Skin Senses". IN D.R. Kenshalo (ed.) Thomas, Springfield, Illinois, (1968).