ON THE TRANSMISSION OF INFORMATION THROUGH SENSORY NEURONS

L. WALLØE

From the Institute of Physiology, University of Oslo, Oslo, Norway

ABSTRACT Information about muscle length is transmitted to the cerebellum from muscle spindle receptors through the dorsal spinocerebellar tract (DSCT). The "transinformation" about muscle length in single DSCT fibers was calculated from steady-state spike trains by two different methods, assuming that the decoding mechanisms use a frequency code. By the first method, the number of distinguishable muscle lengths (and thus the transinformation) was determined from the rate of convergence of the mean frequency of firing (with increasing number of intervals). The observation time necessary to estimate the mean frequency of the impulse train with a certain accuracy was independent of the stretch level, even though the number of intervals necessary to make this estimate was different at high and low levels of stretch. By the second method an input frequency-output frequency matrix was calculated. The transinformations and the rate of transinformation was then calculated from this matrix. There was an acceptable agreement in the estimates of transinformation by the two methods. The rates of transinformation are significantly increased by the particular time structure of the discharge patterns of the nerve cells. Consequently, the loss of information due to the synaptic coupling is appreciably reduced.

INTRODUCTION

The main functions of nerve cells are to transmit and to process information. Information, which is obtained from the surroundings through sense organs, is transmitted to the central nervous system through chains of nerve cells. Information is also conveyed from one part of the nervous system to another and from the central nervous system out to the effector organs through nerve cells. In the nervous system the information is coded in trains of action potentials. The pattern in these spike trains is altered as the signals traverse the different synapses in the chain.

The present paper tries to elucidate the first step of the transformations of the information content of the nervous signals which take place in a chain of neurons.

The main difficulty in investigations of this kind is obtaining a measure of the amount of information. Different approaches have been attempted. Werner and Mountcastle have tried two different methods (Mountcastle, Poggio, and Werner, 1962; Werner and Mountcastle, 1963; Werner and Mountcastle, 1965). (*a*) An estimator was established for the discriminability between two different trains of impulses, based upon measurements of each individual interspike interval in the trains. This statistical test was applied on signals from thalamic joint neurons. The reading times necessary to reach "decision" could be used to establish a measure of the information content in the spike trains. (*b*) Some concepts of information theory were applied to a stimulus-response matrix obtained from first-order "mechanoreceptive" fibers, and the transmitted information was calculated on the assumption that the signal is read in a frequency code. Darian-Smith, Rowe, and Sessle (1968) have used a similar method on relay neurons in trigeminal nuclei.

Grüsser, also applying information theory, estimated the maximum information transmission rate in visual systems assuming a frequency code (Grüsser, Hellner, and Grüsser-Cornehls, 1962; Grüsser, 1962).

Other investigators have attempted a more theoretical application of information theory to the study of nerve cells. Different coding schemes have been assumed by MacKay and McCulloch (1952), Rapoport and Howarth (1960), and Stein (1967). All these investigations, and especially the last, which was published while this work was in progress, are highly relevant for the present problems.

In the present study, data from dorsal spinocerebellar tract (DSCT) cells which are monosynaptically activated from primary endings of muscle spindles, have been used (Jansen, Nicolaysen, and Rudjord, 1966). These data are well suited for such an analysis for a number of reasons.

(*a*) The response of the input elements (primary fibers) to muscle stretch is well known.

(*b*) The input elements are slowly adapting so that approximately steady-state spike trains can be produced.

(*c*) Information about the length of the muscle is the content of the input signals from primary endings which are statically extended.

Information about the muscle length can be considered as the content of the signal from muscle spindle receptors statically extended (Matthews, 1964). This also holds true for the second-order neuron under the present experimental conditions. In each experiment one muscle was dissected free and stretched without affecting other muscles. Some possible estimates of the amount of information about muscle lengths that is conveyed to the cerebellum through the DSCT neurons are therefore developed.

The transinformations are calculated by two different methods assuming that the decoding mechanism uses a frequency code. It is also assumed that all muscle lengths from a slack muscle (0 mm) up to full extension (14 mm) are equiprobable. Other (physiological?) stimulus distributions in the same length interval will, of course, give smaller transinformations. The spike pattern in the DSCT neurons have been investigated earlier (Jansen et al., 1966; Walløe, Jansen, and Nygaard, 1969). There is a strong negative serial dependency between neighboring intervals in these spike trains. Both calculation methods were applied on the original sequence of interimpulse intervals as generated by the DSCT neuron and on a random permutation of the same intervals in order to study the effect of the interimpulse dependency on the transinformation. The results from these two methods will be compared with each other and with Werner and Mountcastle and Stein's results in the Discussion.

METHODS

The data were obtained during the series of experiments reported earlier by Jansen et al. (1966). A full description of the experimental arrangement is given in their paper. The muscles used were deprived of motor innervation. The action potentials of DSCT neurons were recorded on a magnetic tape and later played back through an electric counter, which counted the number of unitary time intervals (usually of 1.6 msec duration) that occurred between the action potentials. These series of numbers were transferred to a Univac 1107 computer via perforated paper tapes.

Three cells were selected on the basis of a large body of recordings from different cells. These neurons were chosen because it had been possible to obtain records at many different levels of activity. Altogether, spike trains corresponding to 79 different muscle lengths were obtained from these three cells, 42 of them from one cell only. Most of the samples had a duration of more than 10 sec corresponding to some 500–1000 intervals. One cell was activated by stretch of the tibialis anterior-extensor digitorum longus muscles, while the two others were activated by stretch of the gastrocnemius-soleus muscle.

Investigations on the statistical properties of the spike trains from the same three cells have recently been published (Walløe et al., 1969).

RESULTS

To obtain a quantitative measurement of the transinformation conveyed by the DSCT neuronal spike sequences, two different methods have been employed. The first utilizes the increasing accuracy of the estimation of the mean frequency of an irregular train of spikes with increasing observation time. The second depends on the estimation of contingent probabilities between input and output signals, and is in principle similar to the traditional method of information theory for analysis of a "noisy channel". The relevant formulas can be found in Reza (1961).

Method I: Convergence of Mean Frequency of Firing

By this method the number of distinguishable muscle lengths is determined as a function of the observation time. The amount of information about muscle length which is conveyed to the cerebellum is then calculated. To estimate the number of distinguishable muscle lengths, the following procedure is a possible approach.

Assume that the decoder has no knowledge of the muscle length at a certain moment. From that time on the decoder receives a train of impulses and makes estimates about the mean frequency. The best estimate of the mean frequency after reading two spikes is the reciprocal of the interimpulse interval. After reading *n* in-

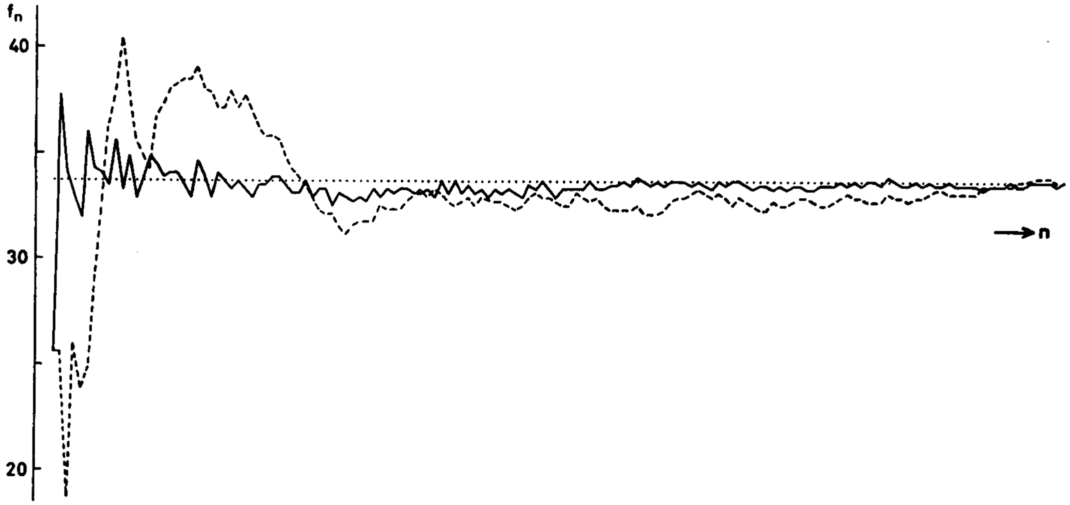

FIGURE 1 The mean frequency f_n of n interimpulse intervals as a function of n. Stretch about 2 mm. Solid line: intervals in normal order. Broken line: the same 150 intervals randomly shuffled. Dotted line: mean frequency after 150 intervals. The first interval is the same in both sequences.

tervals each with length t_i, a possible estimate of the mean frequency is

$$f_n = \frac{n}{\sum_{i=1}^{n} t_i}. \tag{1}$$

f_n will converge towards the mean frequency when $n \rightarrow \infty$. Fig. 1 shows an example of f_n as a function of n. The solid line shows f_n as a function of n when the intervals are read in original order; the broken line shows f_n when the intervals are read in a shuffled order. f_n seems to converge much faster when the intervals are read in normal order. This was a regular finding.

This hypothesis was tested further. 150 interimpulse intervals from a steady-state registration were read by the computer. f_n was calculated for all n from 1 to 150 for the sequence of intervals in original order. Call this frequency $f_{n(\mathrm{orig})}$. The computer then made a random permutation of all intervals t_i from $i = 2$ to $i = 150$, and calculated f_n of the shuffled sequence. Call this frequency $f_{n(\mathrm{shuff})}$. t_1 in both sequences were thus the same, and $f_{1(\mathrm{orig})} = f_{1(\mathrm{shuff})}$. Further $f_{150(\mathrm{orig})} = f_{150(\mathrm{shuff})}$, since the shuffled sequence is composed of exactly the same 150 intervals.

The usual definition of convergence can be stated as follows: an infinite sequence f_n converges towards f_∞ if to every positive Δf there exists one N, such that

$$|f_n - f_\infty| < \Delta f \text{ for all } n > N. \tag{2}$$

It is possible to make a measure of the rate of convergence in the sequence f_n from 1 to 150 analogous to the condition of equation 2. Let $N_{\min}$ be the smallest number N that fulfills the condition

$$|f_n - f_{150}| < \Delta f \text{ when } n \geqq N \tag{3}$$

To each positive value of Δf there corresponds one value of $N_{\min}$ in the interval [1, 150]. It is reasonable to say that $f_{n(\text{orig})}$ converges faster than $f_{n(\text{shuff})}$ when $N_{\min}$ is smaller for the original sequence than for the shuffled sequence with the same Δf. The meaning of some of these definitions is illustrated in Fig. 2.

Let T be defined by equation 4

$$T = \sum_{i=1}^{N_{\min}} t_i . \tag{4}$$

To avoid influence from adaptation in the receptor, the first 0.5 sec of the recordings from a static stretch was disregarded. $N_{\min}$ and T were calculated by the computer for the next 150 intervals. The same calculations were carried out with the intervals from t_{50} to t_{200}, from t_{100} to t_{250}, and so on. This procedure was carried out through all the registered intervals from one stretch level. Each sequence of 150 intervals was thus overlapping the preceding sequence by 100 intervals.

The null hypothesis that $N_{\min}$ from both sequences were equal, was rejected at the 1 % level of significance by the signed-ranks test.

Similar results were found in all spike sequences of sufficiently long duration.

$N_{\min}$ is dependent upon Δf. A small Δf gives a large $N_{\min}$ and vice versa. For medium values of Δf, about 6 imp/sec, ("imp" for impulses), the difference between the sequences $f_{n(\text{orig})}$ and $f_{n(\text{shuff})}$ is significant at the 5 % level, except in the extremely short series (with less than 3–400 imp). Thus, judged from the values of $N_{\min}$, the mean frequency seems to converge faster when the intervals are read in original order than when the intervals are shuffled.

The mean $N_{\min}$ was calculated for all stretch levels in three cells with Δf = 4 imp/sec. The mean $N_{\min}$ was dependent upon the stretch level. In series with short average

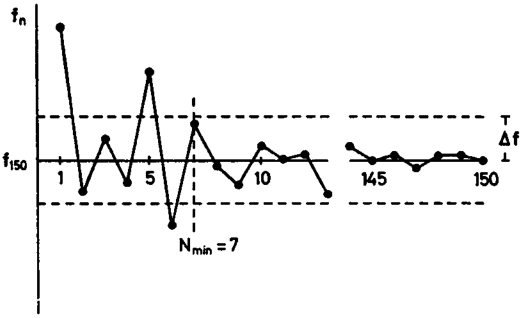

FIGURE 2 The mean frequency f_n as a function of n. This figure illustrates the meaning of Δf and $N_{\min}$.

interimpulse intervals the mean $N_{\min}$ was generally large, in series with long intervals, $N_{\min}$ was generally small. However, Fig. 3 reveals that the mean of the "observation time", T, is independent of the mean interval length in the series. Each stretch level in this cell is represented by two points in this figure, an open circle, corresponding to the shuffled series, and filled circle, corresponding to the original series.

Thus, the observation time necessary to estimate the mean frequency of the impulse train with a certain accuracy is independent of the stretch level. Notice that the number of impulses necessary to make this estimate is different in high and low levels of stretch. Instead of estimating the mean frequency with a certain accuracy, the decoder might try to estimate the mean interval length with a certain accuracy. In this case, neither the number of impulses, nor the observation time that is necessary, would be independent of the stretch level.

Similar results were found for spike sequences of the two other cells, that were examined, except at very low rates of firing, when the firing pattern is dominated by the regular background activity of the cells (Walløe, 1968).

The mean observation time $\bar{T}$ is a function of Δf. A large Δf gives a small $\bar{T}$ and

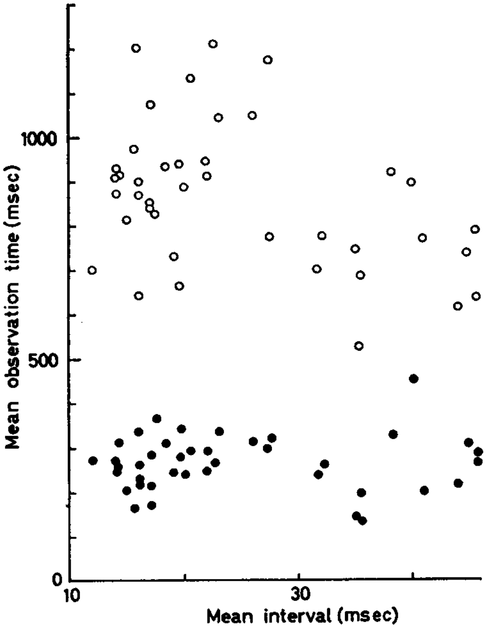

FIGURE 3

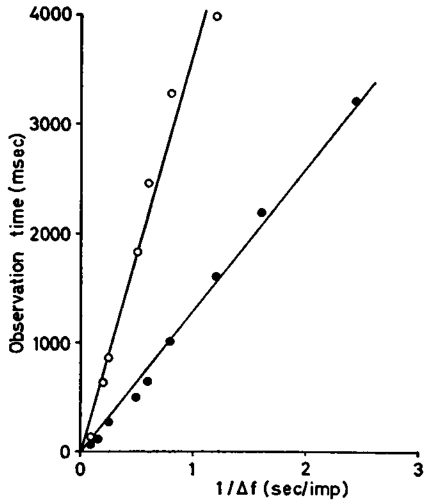

FIGURE 4

FIGURE 3 The mean observation time $\bar{T}$ as a function of the mean interimpulse interval $\bar{I}$ in each stretch level. $\Delta f = 4$ imp/sec. Data from one cell. Filled circles: intervals read in normal order. Open circles: intervals read in shuffled order. Mean $\bar{T}$ in the two populations are 265 msec and 865 msec, respectively.

FIGURE 4 The mean of the mean observation times $\bar{T}$ as a function of Δf. Data from one cell. Filled circles: intervals read in normal order. Open circles: intervals read in shuffled order.

vice versa. Fig. 4 reveals that the relationship is simple and can be expressed by the formula

$$\bar{T} = a\frac{1}{\Delta f}, \tag{5}$$

where a is a constant, a is different for the shuffled and original sequences. Each pair of points in Fig. 4 is calculated from sets of data similar to those presented in Fig. 3.

As mentioned earlier, the impulse frequency in the second-order neuron is a linear function of the muscle length. Let I be the difference in impulse frequency in the neuron between a fully extended muscle and the unstretched muscle. The problem is now to estimate the number of distinguishable muscle lengths. It is possible to use

$$n = \frac{I}{2\cdot\Delta f} \tag{6}$$

as a measure of this number. The number of distinguishable muscle lengths is a function of the observation time, and Δf can be calculated from $\bar{T}$ by equation 5.

In other words, the total frequency range is divided into parts each $2\cdot\Delta f$ wide. The number of such parts is a possible guess of the number of distinguishable lengths in the observation time $\bar{T}$. This procedure is possible because Δf corresponding to a given $\bar{T}$ is independent of the muscle length.

$\bar{T}$ is the mean of the observation times calculated by equation 4. In approximately half of the cases the mean frequency is within $\pm\Delta f$ of the estimate based on the intervals read in the time $\bar{T}$. It is, of course, possible to use stronger criteria to determine the number of distinguishable lengths. One of the upper fractiles in the distributions of T (with constant Δf) could be used instead of the mean. The calculations show that the distribution histograms of T are unimodal and approximately symmetric. Fig. 5 shows that the standard deviation in the distribution increases proportionally with $\bar{T}$. The slope of the line is similar in all three cells both in the original and the shuffled sequences. If one of the upper fractiles in the distribution of T is used instead of the mean, the constant a in equation 5 will thus have a larger value, but no other change will be introduced. I will return to this point later.

From equations 5 and 6

$$n = \frac{I\bar{T}}{2a}. \tag{7}$$

Assuming that all muscle lengths from a slack muscle (0 mm) up to full stretch (14 mm) are equiprobable, the information H can be expressed as

$$H = \log_2\frac{I\bar{T}}{2a}. \tag{8}$$

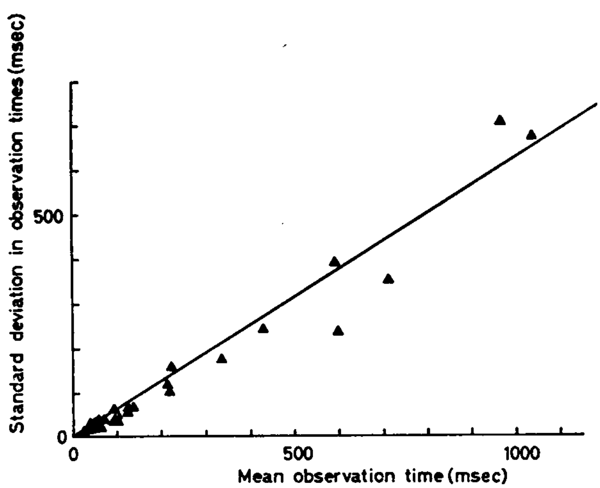

FIGURE 5 Standard deviations (SD) in the distribution of observation times T as a function of the mean observation time. Data from one cell. Intervals read in normal order. SD $\approx$ 0.63 T.

I and a are constants for a certain cell and can be determined experimentally. The information that is transmitted, can thus be calculated as a function of the observation time, assuming that the decoder uses a frequency code.

a has been determined to 1.3 imp for the normal sequences and to 3.6 imp for the permutated sequences from Fig. 4 and similar figures. I is somewhat more uncertain. One cell had a frequency range from about 20 imp/sec (corresponding to muscle not stretched) to about 80 imp/sec (corresponding to fully extended muscle) in our experiments. Another cell had a frequency range from 10 imp/sec to 65 imp/sec, and the third cell from 30 imp/sec to 110 imp/sec. The corresponding values of I are thus 60, 55, and 80 imp/sec, respectively. The value $I = 60$ imp/sec will be used in the calculations. But this is certainly an underestimate of the range of these neurons, since they can be expected to fire at appreciably higher rates under physiological conditions with intact fusimotor supply. Eide, Fedina, Jansen, Lundberg, and Vyklický, (1969) have furthermore found a linear increase in firing frequency with increasing transmembrane current up to more than 200 imp/sec.

Figs. 6 and 7 show the information as a function of the observation time calculated from these assumptions.

Equation 8 and the figures reveal that the amount of information increases as a logarithmic function of the observation time. When the observation time is doubled, the information has increased by 1 bit.

The figures also show that the rate of information has a maximum near observation times of 115 msec in the original sequences and 330 msec in the shuffled sequences. The corresponding information rates are 12.35 bits/sec and 4.40 bits/sec, respectively.

As mentioned, it is possible to use stronger criteria than $\bar{T}$ for the determination of

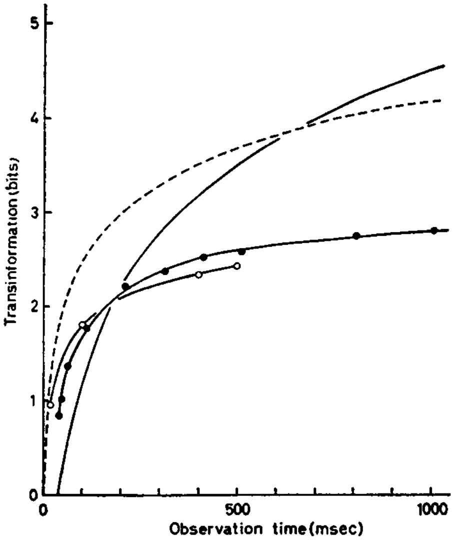

FIGURE 6

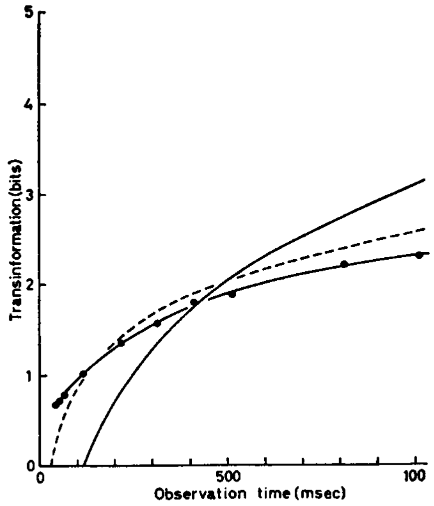

FIGURE 7

FIGURE 6 Transinformations calculated by different methods as functions of the observation time. Intervals read in normal order. ———: transinformations calculated by Method I. ●——●: transinformation calculated by Method II. - - - - - - -: transinformations calculated by Stein's formula. ○——○: Mountcastle and Werner's results from first-order touch fibers (1965).

FIGURE 7 Transinformations calculated by different methods as functions of the observation time. Intervals read in shuffled order. ———: transinformations calculated by Method I. ●——●: transinformations calculated by Method II. - - - - - - -: transinformations calculated by Stein's formula.

the number of distinguishable muscle lengths. Fig. 5 reveals that the standard deviation in the distribution of T is a linear function of $\bar{T}$:

$$\text{SD} \approx 0.63\,\bar{T}. \qquad (9)$$

97.5 % of the estimates based on the intervals read in the time $\bar{T} + 2\text{SD}$ are within $\pm\,\Delta f$ of the mean frequency if the distribution is approximately normal. Call this observation time $T_{2.5}$. The equation analogous to equation 14, when these stronger criteria are used, will be

$$T_{2.5} = 2.26a\,\frac{1}{\Delta f}, \qquad (10)$$

and the information that is transmitted will decrease by $\log_2 2.26$ for all observation times. The corresponding curves in Fig. 6 and 7 will thus be lowered by 1.2 bits. The corresponding changes in the information rates will be a lowering of maximum rates to about 5.4 bits/sec in the original sequences and 2.0 bits/sec in the shuffled sequences, and an increase in optimum observation times to 250 msec and 800 msec, respectively.

The frequency range I was found to be about 60 imp/sec in one of the cells. This might be an underestimate of the actual range. If I is increased from 60 to 120 imp/sec, the number of distinguishable muscle lengths will be doubled. The curves in Fig. 6 and 7 will thus be elevated 1 bit for all observation times. The maximum rate will then be about 8.9 bits/sec for observation times near 170 msec for the shuffled sequences and 25.6 bits/sec for observation times near 50 msec for the original sequences.

Neither of these two extreme changes in the assumptions changes the order of magnitude of the results. The peak in information rate is higher, and the corresponding observation time shorter, when the intervals are read in original order, than when they are shuffled.

For the original sequences the maximum rate 12 bits/sec occurring with observation times of 115 msec, at least indicates the order of magnitude of the possible real value.

Method II: Stimulus-Response Matrix

An input-output matrix was determined, and the transinformation and the rate of transinformation were calculated from the matrix.

The starting point in these calculations is an assumption about a discrete number n of possible muscle lengths. Later I shall show that the results are independent of n as long as the number is not too small. These possible muscle lengths are assumed to be equiprobable and represent the input to the channel. The mean impulse frequency in a certain observation time is the output of the system. The frequencies belong to one of a discrete number m of possible frequency groups. Again, I shall show that the results can be extrapolated from m possible outputs to an infinite number, that is to a continuous output.

n steady-state recordings of the activity in one second-order neuron were selected. Ideally, the recordings should correspond to n equidistant muscle lengths and cover the total length range of a muscle. The actual recordings had mean frequencies that differed somewhat from the desired values. In each of the recordings all intervals t were therefore multiplied by a correction factor. An example follows: in one calculation a recording with mean frequency 31.000 imp/sec was wanted, corresponding to a mean t of 32.26 msec; the best recording available had a mean t of 32.41 msec; each t in the recording was multiplied by $32.26/32.41 = 0.995$. All the correction factors used were in the interval from 0.9 to 1.1.

From one of the n recordings, intervals were counted and added by the computer until the accumulated interimpulse time was equal to or just greater than a preset reading time d. Let d_k be the accumulated intervals. The mean impulse frequency in the time d_k was calculated by the computer and arrayed in the appropriate frequency group F_l. The computer repeated this procedure until all intervals in the recording were used. The number of elements in the different frequency groups F_l were nor-

TABLE I

STIMULUS-RESPONSE MATRIX CALCULATED FROM ONE CELL

Number of muscle lengths $n = 4$, frequency group size 3.0 imp/sec, reading time $d = 100$ msec. Intervals read in normal order. The numbers in the matrix are relative frequencies multiplied by 10,000.

	Frequency group number																		
n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	p_k
1	0	1	89	1701	689	13	0	0	0	0	0	0	0	0	0	0	0	0	2500
2	0	0	0	14	541	1676	230	41	0	0	0	0	0	0	0	0	0	0	2500
3	0	0	0	0	0	111	1327	575	376	66	44	0	0	0	0	0	0	0	2500
4	0	0	0	0	0	0	13	13	278	370	582	529	410	172	53	40	13	26	2500
q_l	0	0	89	1714	1239	1799	1570	629	654	437	626	529	410	172	53	40	13	26	10000

n	Number of observations	$\bar{d}_k$
1	197	108.47 msec
2	185	110.23 msec
3	113	112.98 msec
4	189	123.14 msec

$H(X) = 2.00000$ bits
$H(Y) = 3.31149$ bits
$H(X, Y) = 3.76609$ bits
$I(X, Y) = 1.54540$ bits

Observation time $\bar{d}_k = 113.70$ msec

malized corresponding to the condition $p_k = 1/n$. The mean observation time $\bar{d}_k$ was also calculated. The mean observation time $\bar{d}_k$ is, of course, greater than the reading time d, which is a parameter in the calculations.

This procedure was repeated in all recordings, and the result was an input-output matrix as shown in Table I. The entropies of the input, $H(X)$, of the output, $H(Y)$, of the matrix, $H(X, Y)$, and the transinformation $I(X, Y)$ were calculated. The rate of transinformation $I(X, Y)/\bar{d}_k$ was also calculated.

The transinformation and the rate of transinformation are functions of the number of muscle lengths, n, the number of frequency groups m, and the reading time d. Each of these functional relationships were studied in both the original sequences and the shuffled sequences.

(*a*) The reading time d and the number of muscle lengths n were kept constant, and the number m of frequency groups was varied. Figure 8 shows the results from one set of calculations with $d = 100$ msec and $n = 7$. The transinformation is shown as a function of the size of frequency groups F_l. The transinformation increases as the frequency groups decrease in size. When the size of each group F_l decreases, the number of groups increases. The calculations are carried out with group sizes ranging from 0.4 imp/sec to 3.5 imp/sec. The regression line through the calculated points is extrapolated from group size 0.4 imp/sec to 0 imp/sec. The validity of this extrapolation is, of course, dubious. It is equivalent to a change from discrete to continuous output. However, from the theory it is known that $I(X, Y)$ generally remains finite even when $H(X)$, H(Y) are infinite, and the figure really suggests this

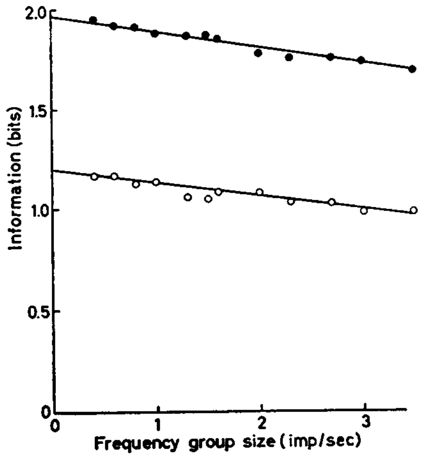

FIGURE 8 The transinformation as a function of the size of the frequency groups. Filled circles: intervals read in normal order. Open circles: intervals read in shuffled order. Preset reading time d = 100 msec, real reading time $\bar{d}_k$ = 115 msec, number of different muscle lengths n = 7.

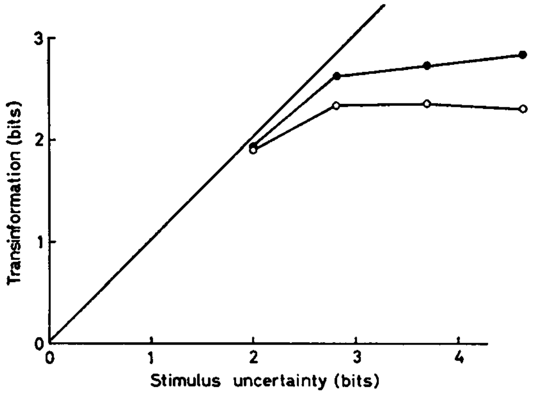

FIGURE 9 The transinformation as a function of the stimulus uncertainty. Observation time 1000 msec, frequency groups size 0.8 imp/sec. Filled circles: intervals read in normal order. Open circles: Intervals read in shuffled order. Line through origin: maximum transinformation.

extrapolation. With increasing group size the transinformation decreases asymptotically towards zero. However, in the group size range 0.4 imp/sec–3.5 imp/sec the relationship is approximately linear. Similar results were found with other values of n and d. The number of muscle lengths n was varied from 4 to 25, and the reading time d was varied from 0 msec to 1000 msec. In this range of the parameters, the transinformation decreases approximately linearly with the frequency group size from 0.4 imp/sec to 3.5 imp/sec. The slopes of these lines were also similar to the lines in Fig. 8. Thus, accepting the extrapolation, the difference in transinformation between a discrete output with frequency group size of 0.8 imp/sec and a continuous output is approximately 0.05 bits. The frequency groups size was kept constant at 0.8 imp/sec in all the rest of the calculations.

(b) The number of muscle lengths n was varied while the reading time d and the frequency group size were kept constant. The results from one set of calculations with reading time d = 1000 msec are shown in Fig. 9. The abscissa in the figure is the "stimulus uncertainty," i.e., $\log_2 n$. The figure reveals that the transinformation is only a little less than the possible maximum up to a stimulus uncertainty near 3 bits. Then the transinformation levels off to a constant value as the stimulus uncertainty increases. Similar results were found with other values of the reading time *d* between 30 msec and 1000 msec. Again, it is tempting to extrapolate the results to an infinite number of groups, that is to a continuous rectangular input distribution.

(*c*) The number of muscle lengths *n* and the frequency group size were kept constant while the reading time *d* was varied. Figs. 6 and 7 show the result of one set of calculations with $n = 7$. The abscissa is the actual mean observation time $\bar{d}_k$. The figures reveal that the transinformation increases as a function of the observation time, at least above 50 msec. The transinformation calculated from the original sequence of intervals increases rapidly and reaches a plateau corresponding to observation times near 500 msec. A further increase in observation time gives only a small additional increase in the transinformation. When the intervals are read in shuffled order, the transinformation for a given observation time is much smaller. The transinformation increases more slowly with the observation time and reaches the maximum value later. The rate of transinformation increases with decreasing observation time and reaches a plateau value of 25 bits/sec near observation times of 50 msec in the normal sequences.

These results were similar in all three cells for different values of *n*.

With observation time less than 50 msec the results are somewhat anomalous. The transinformation increases somewhat as the observation time decreases. The most probable explanation of these results is that the method is useless for small reading times. Some other findings gave similar indications. The actual mean observation time $\bar{d}_k$ was always greater than the preset parameter *d*. The difference was about 15 msec for all reading times from 80 msec and upwards. Below 80 msec the difference increases and was 23 msec for $d = 0$ msec. The intervals were added by the computer until the accumulated interimpulse time d_k was greater than the preset reading time *d*. Thus the computer always read a whole number of intervals. When *d* was small, the transinformation was probably overestimated by this procedure. It never happened that no impulse was fired in the observation time.

DISCUSSION

The main problem in this paper has been to estimate the amount of information about the muscle length which is transmitted from the primary endings in the muscle spindles to the cerebellum through one DSCT neuron. Assuming a frequency code, the amount of information which has been transmitted from the muscle to one DSCT neuron has been calculated. Whether this information actually is conveyed to the next cell in the pathway, is dependent upon the decoding mechanism in this cell.

Some support can be given to the frequency assumption and thus to the validity of the estimates of transinformation.

(*a*) The DSCT neurons seem to decode the signals from the primary afferent fibers according to a frequency code. This is apparent in the model which has been proposed for the synaptic mechanisms in the DSCT cells (Walløe et al., 1969).

(*b*) The DSCT fibers terminate as mossy fibers on the granule cells in the cerebellar cortex. With reasonable assumptions about the synaptic mechanisms in the granule cells (similar to the mechanisms in the DSCT cells), it has been shown that the firing rate of the granule cells will increase approximately linearly with the input frequency (Walløe, 1968).

The transinformations have been calculated by two different methods.

Figs. 6 and 7 show the similarity and the difference between the results obtained. The only common assumption of the two methods is that of frequency coding. Each method is also based on a number of additional assumptions. Considering the number of different assumptions involved, the results show a high degree of similarity. The transinformations are probably overestimated for short observation times by the stimulus-response matrix method. The reason is that the computer neither starts nor stops the observation period in the middle of an interspike interval, but always reads a whole number of intervals. On the other hand, the transinformations are underestimated for short observation times by the frequency-convergence method. The transinformations calculated by this method are apparently zero for short observation times. But even very short intervals will carry some information (Stein, 1967).

The frequency-convergence method overestimates the transinformations for long observation times. The reason is that the linear relationship found between $1/\Delta f$ and the observation time (Fig. 4) is not valid for long observation times.

From Fig. 6 it is reasonable to conclude that the transinformation is very small for short observation times, up to 30 msec or more. With longer observation times the transinformation increases steeply to about 2 bits. The increase is much slower for observation times greater than 200 msec. The rate of transinformation has a peak value of the order 20 bits/sec with observation times near 100 msec (with 50 msec and 150 msec as reasonable limits). This value shows that the firing pattern conveys enough information to distinguish four different muscle lengths if the signal is read in a frequency code with observation times of 100 msec. With observation times of 1000 msec the rate of transinformation is less than 4 bits/sec, corresponding to 16 different muscle lengths. Thus the decoder will have maximum efficiency if it uses the information which arrives every 100 msec.

The transinformations are also dependent upon the stimulus uncertainties. With only a few possible stimulus intensities, equivalent to a small input uncertainty, all information was transmitted. The transinformation is near the maximum value up to stimulus uncertainties about 3 bits, and it levels off at a value about 2.6 bits as presented in Fig. 9.

These results are in close agreement with Mountcastle and Werner's results from first-order touch fibers. Beyond a stimulus uncertainty of about 2 bits, information transmission levelled off at a maximum value in their investigation.

There is also a surprising similarity between these findings and the information transmitting capacities of humans determined by the method of absolute judgements. In a variety of sense continua the human capacity as regards stimulus intensity is between 2 and 3 bits of information (Garner, 1953; Miller, 1956; Garner, 1962). The present results show that with a constant observation time not more than a maximum value of information is transmitted, no matter how great the number of stimulus intensity levels is. With observation times less than 1000 msec, not more than 6–7 different intensity levels could be distinguished in the neuronal response.

The results also show that it can make sense to calculate transinformations even if the number of possible categories in the response has to be chosen arbitrarily. The transinformations are only to a small degree dependent on one particular choice of response categories (Fig. 8). This result has in fact been an unstated assumption in most transinformation calculations from psychophysical experiments and also in Werner and Mountcastle's calculations.

It is possible to apply Stein's formulas (1967) for the information capacity on the DSCT data.

Stein has derived approximate equations for the transinformation in a nerve fiber from theoretical considerations. His starting point is a renewal process (Cox and Miller, 1965). He makes assumptions about the probability distributions of the input stimuli, and calculates the transinformation. The *information capacity* is defined as the maximum value of the transinformation considering all possible stimulus distributions. Stein shows that

$$I = \log_2 \left\{ [t/(2\pi e)]^{\frac{1}{2}} \int_{\mu_o}^{\mu_s} \frac{d\mu}{\sqrt{\mu a \sigma}} \right\} \qquad (20)$$

is an approximate expression for the information capacity. It becomes increasingly accurate for large values of the observation time t. μ is the mean interspike interval, and σ^2 the variance. μ_o is the lower value and μ_s the upper value of the interval. a is defined by equation 21:

$$a = 1 + 2 \sum_{k=2}^{\infty} R_{1k}, \qquad (21)$$

where R_{1k} is the serial correlation coefficient between interval No. 1 and k. R_{1k}, and therefore also a, may be a function of μ.

The information capacity is defined as the maximum transinformation, considering all possible distributions of the input. However, Stein has shown that these optimum distributions are approximately rectangular, and his information capacities can thus be compared with the present transinformations.

Stein's formulas are applied on the DSCT data in the following manner. The curve in Fig. 7 is plotted according to equation 20 with $a = 1$. The experimentally determined relation $\sigma = 0.4\,\mu$ is used in the formula. Equation 22 was the resulting equation.

$$I = \tfrac{1}{2} \log_2 t + \log_2 1.21 (\mu_o^{-1/2} - \mu_s^{-1/2}). \qquad (22)$$

μ_o and μ_s are chosen from one of the three cells. μ_o is put equal to 12 msec (corresponding to an upper frequency limit of 83 imp/sec), and μ_s equal to 50 msec (corresponding to a lower frequency limit of 20 imp/sec). I is more sensitive to changes in μ_o than in μ_s. If the maximum frequency is doubled, corresponding to a μ_o of 6 msec, I will increase by about 0.9 bits for all values of t. The curve in Fig. 6 is plotted according to equation 20 with the same assumptions and the same choice of parameters as above. In addition a was estimated by equation 21. The estimates for R_{12} and R_{13} of the present material are acceptably accurate (Walløe et al., 1969). R_{12} is to some degree dependent upon μ with low numerical values corresponding to low activity in the nerve (large μ). For higher levels of activity R_{12} is approximately constant for most values of μ. As a first approximation $R_{12} = -0.6$ is chosen for all μ. In an autoregressive Markov chain with $R_{12} = -0.6$, R_{13} should be $+0.36$, R_{14} should be -0.216, etc. In this case

$$\sum_{k=2}^{\infty} R_{1k} = -0.375,$$

and consequently $a = 0.25$.

However, Walløe et al. (1969) have shown that an auto-regressive Markov chain is a bad description of the experimental results. R_{13} in the train of impulses is probably about 0.25, not 0.36. The higher order serial correlation coefficients can only be guessed from the 10 experimental determinations of R_{1k} with k from 4 to 20. The results from one train of impulses is shown in Fig. 10. The train contained about 1600 intervals, that is 78 parts with 20 intervals each. All R_{kl} with k from 1 to 19, l from 2 to 20, and $k < l$, are calculated, and the accumulated sum is plotted as a function of n. The expected values for an autoregressive scheme with the same R_{12} is also plotted. The first point on the curve is calculated from R_{12}, R_{23}, etc. up to $R_{19\,20}$, i.e., 19 different values of R. The next point is in the same manner calculated from 18 additional values, and so on. It is difficult to make guesses about $\sum_{k=2}^{\infty} R_{1k}$ from this figure. However, all the 10 long stretch levels that have been investigated, give rise to curves similar to Fig. 10. The sum fluctuates around -0.45 for all k from 4 and upwards. One possible guess of $\sum_{k=2}^{\infty} R_{1k}$ therefore is -0.45, which gives $a = 0.10$. The transinformations in

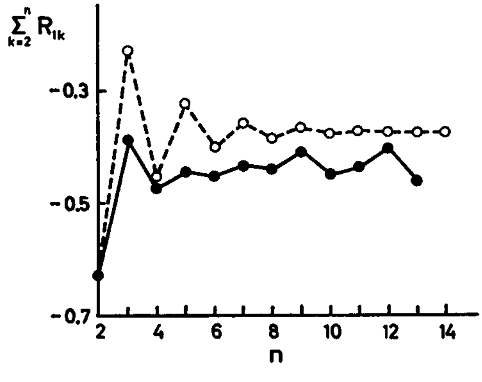

FIGURE 10 $\sum_{k=2}^{n} R_{1k}$ as a function of n. Solid line: experimental data. Broken line: expected values for an autoregressive process with the same value of R_{12}. Explanation of symbols: see text.

Fig. 6 are calculated from equation 20 with $a = 0.10$. Since a is supposed to be independent of μ, the equation is similar to equation 22 with addition of $A = -\frac{1}{2} \log_2 a$. With $a = 0.10$, $A = 1.66$ bits, and with $a = 0.25$, $A = 1.00$ bit. Small changes in $\sum_{k=2}^{\infty} R_{1k}$ from 0.45 towards greater negative values have great influence on A. When the sum approaches -0.50, A increases towards ∞. The curve in Fig. 6 is therefore only a rough estimate where the shape of the curve is more important than the numerical values of the transinformations.

Figures 6 and 7 show that the transinformations obtained by Stein's formulas are in reasonable agreement with the results obtained by the two other methods.

Fig. 6 reveals a striking similarity between Mountcastle and Werner's results from first-order touch fibers and the results from a similar stimulus-response matrix method applied on the DSCT cells. Mountcastle and Werner's results are plotted in Fig. 6 only. These results are replotted from Fig. 22 in their report (1965). The four points represent transinformations calculated with 14 different stimulus levels (equivalent to a stimulus incertainty of 3.8 bits). Other possible choices from their figures would give similar curves.

Considering the number of different assumptions involved in the four sets of calculations presented in Figs. 6 and 7, the results show a high degree of similarity.

Figs. 6 and 7 show that the transinformations and the rates of transinformation are smaller and the maximum rate occurs later in the shuffled series than in the normal series of intervals. The difference in transinformation is about 1.5 bits independent of the observation time. This extra information is only available to the granule cells if the signals from one DSCT fiber are read at a time. The corresponding histological requirement appears to be provided for by the rather unique synaptic arrangements of the cerebellar glomeruli. Thus, the particular time structure of the signal in the DSCT neurons may have physiological significance.

An interesting question is, "How much information is lost in the synapse between a first-order fiber and a second-order fiber, again assuming a frequency code and a certain observation time?" It is not possible to give a final answer to this question from the available experimental data. However, some rough estimates can be made. A numerical example based on Stein's formulas may illustrate the situation.

The coefficient of variation in the primary afferents from deefferented muscle spindles is about 0.04 and approximately independent of the mean frequency. With an observation time of 200 msec, each primary afferent transmits about 4.6 bits of information ($\mu_o = 12$ msec, $\mu_s = 50$ msec, $a = 1$, $\sigma = 0.04\,\mu$).

About 15 first-order neurons converge on one second-order neuron (Walløe et al., 1969; Eide et al., 1969). If these 15 neurons were operating as 15 independent channels, the information received by one second-order neuron would be the sum of the information conveyed by the 15 first-order neurons, that is about 70 bits. If the 15 spindle afferents transmit the same kind of information (i.e. muscle length), but the firing patterns still are completely independent, the second-order neuron will receive an amount of information equivalent to a record 15 times as long from one first-order cell. In other words, the reading time necessary to obtain a certain amount of information is reduced by a factor of 15. 15 spindle afferents from the same muscle transmit about 6.6 bits in 200 msec. Stein's formulas reveal that the effective coefficient of variation in the second-order neuron is one-third of the numerical value, or about 0.13. With an observation time of 200 msec, each DSCT neuron thus transmits about 3.0 bits of information (Fig. 6, broken line). Apparently, the loss in the synapse is about 3.6 bits. i.e., more than half of the information is lost. However, the physiological loss is probably less than this.

Firstly, each primary afferent fiber transmits its information to a number of DSCT cells. The degree of divergence in the system is not known, and quantitative estimates cannot be made at present. However, the firing pattern in these DSCT neurons are neither independent nor completely dependent of each other. Consequently, the information transmitted through M DSCT cells, partly activated from the same primary afferents, will be less than equivalent to a record M times as long as that from one afferent, but greater than the information transmitted through only one DSCT cell.

Secondly, the firing pattern in the primary afferent fibers is much more irregular in preparations with intact efferent innervation to the spindles, than in preparations with deefferented spindles.

Some information is, of course, bound to be lost in the transmission from first-order neurons to second-order neurons, if not otherwise because some impulses will happen to arrive in the refractory periods of the second-order neurons.

Other kinds of information may be lost in the synapses. The signal in each primary afferent may contain precise information about the local conditions near that particular spindles. One of the main functions of this synapse may indeed be to "average" the signals from different spindles.

In psychophysical studies the relation between a just-noticeable stimulus increment (ΔS) and stimulus intensity (S) has been determined experimentally for many stimulus continua in man. The function relation $\Delta S/S = f(S)$ is called the Weber function of the system. The information from the muscle spindles is not brought to consciousness, and it is consequently not possible to determine a Weber function for the system. However, in the physiological range of muscle lengths, the frequency in the second-order neuron increases linearly with the length. With a constant Δf in equation 12, the observation times were equal for all muscle lengths. Other possible reading mechanisms suggested did not result in equal observation times. It is a fair assumption that this sensory system has a constant sensitivity for all muscle lengths. Thus, these results suggest that the central nervous system is capable of discriminating between states separated by an equal increment in spike frequency. Further, these results indirectly give support to the assumption that information about muscle lengths is coded in a frequency code. These results are also in close agreement with Werner and Mountcastle's (1965) findings from first order touch fibers in the cat.

I would like to thank Jan K. S. Jansen, Peter B. C. Matthews, and Richard B. Stein for stimulating discussions and Olav Tjugen for valuable assistance.

The work described in this paper was partly done as fulfillment of a doctoral degree Walløe, 1968.

The work has been supported by grants from Nansenfondet, Oslo, Norway.

Received for publication 24 June 1969 and in revised form 2 December 1969.

REFERENCES

Cox, D. R., and H. D. Miller. 1965. The Theory of Stochastic Processes. Methuen & Co. Ltd., London.

Darian-Smith, I., M. J. Rowe, and B. J. Sessle. 1968. *Science (Washington)*. **160**:791.

Eide, E., L. Fedina, J. Jansen, A. Lundberg, and L. Vyklický. 1969. *Acta Physiol. Scand.* **77**:145.

Garner, W. R. 1953. *J. Exp. Psychol.* **46**:373.

Garner, W. R. 1962. Uncertainty and Structure as Psychological Concepts. John Wiley and Sons, Inc., New York.

Grüsser, O.-J. 1962. *Kybernetik.* **1**:209.

Grüsser, O.-J., K. A. Hellner, and U. Grüsser-Cornehls. 1962. *Kybernetik.* **1**:175.

Jansen, J. K. S., K. Nicolaysen, and T. Rudjord. 1966. *J. Neurophysiol.* **29**:1061.

MacKay, D. M., and W. S. McCulloch. 1952. *Bull. Math. Biophys.* **14**:127.

Matthews, P. B. C. 1964. *Physiol. Rev.* **44**:219.

Miller, G. A. 1956. *Psychol. Rev.* **63**:81.

Mountcastle, V. B., G. F. Poggio, and G. Werner. 1962. *In* Information Processing in the Nervous System. Proceedings of the International Union of Physiological Science, 22nd Congress. Leiden, Netherlands. 3:196.

Rapoport, A., and W. J. Horvath. 1960. *Inform. Control.* **3**:335.

Reza, F. M. 1961. An Introduction to Information Theory. McGraw-Hill Book Company, New York.

Stein, R. B. 1967. *Biophys. J.* **7**:797.

Walløe, L. 1968. Transfer of Signals through a Second Order Sensory Neuron. Ph.D. Thesis. University of Oslo, Oslo, Norway.

Walløe, L., J. K. S. Jansen, and K. Nygaard. 1969. *Kybernetik.* **6**:130.

Werner, G., and V. B. Mountcastle. 1963. *J. Neurophysiol.* **26**:958.

Werner, G., and V. B. Mountcastle. 1965. *J. Neurophysiol.* **28**:359.