Chapter III. Static Information Transmission

This chapter is concerned with the development of three stimulus sets and the measurement of information transfer per presentation with these stimuli. These stimulus sets are referred to as the 500-*msec*, 250-*msec*, and 125-*msec* stimulus sets reflecting the differences in their signal durations. Emphasis was placed on the 500-*msec* stimulus set in terms of its construction and subject training. In Sec. III-1, we give the background for this work. In Sec. III-2, we describe the 500-*msec* stimulus set and its corresponding response set. The probe experiments used for the construction of this stimulus set are summarized in Appendix B. In Sec. III-3 – III-5, the information transfer measurements with the three stimulus sets are presented and compared.

III-1 Background

This section provides an overview of the absolute identification (AI) paradigm, the computation of information transfer (IT), issues concerning IT estimation from experimental data, and principles for maximizing IT with human observers.

III-1.1 The Absolute Identification (AI) Paradigm

The AI paradigm of interest to us in this set of experiments involves a set of k stimuli S_i, $1 \le i \le k$, a set of k responses, R_j, $1 \le j \le k$, and a one-to-one mapping between the stimuli and responses. The stimuli are presented one at a time in random order with equal *a priori* probabilities and the subject is instructed to respond to each stimulus presentation with the response defined by the one-to-one mapping, i.e., to identify which of the k stimuli was presented. Without loss of generality, we can assume that the stimuli and responses are labeled such that the response corresponding to the stimulus S_i under the one-to-one mapping is R_i. In other words, R_i is the correct response to S_i.

In some experiments, the subject is provided with trial-by-trial correct-answer feedback. That is, on each trial the subject is informed of the correct response after making his or her own response. In general, identification performance may depend not only on the characteristics of the stimulus set, but also on the extent to which the response set and the mapping between stimuli and responses are "natural" (i.e., on the degree of "stimulus-response compatibility").

The stimulus set is said to be one-dimensional if only one attribute of the stimulus (e.g., intensity) is varied. In this case, it is only the value of this variable that needs to be identified by the subject. The stimulus set is said to be multi-dimensional if more than one attribute of the stimulus (e.g., intensity and frequency) is varied. In some experiments with multi-dimensional stimulus sets, the subject is required to identify the values of all the stimulus attributes that are varied in the stimulus set (e.g., both the intensity and frequency of the stimulus). In others, the subject is required to identify the values of only a subset of the attributes that are varied and to ignore the other attributes. In this case, the stimulus attributes to which the subject must attend and respond are referred to as "target" attributes; the attributes that are varied but are ignored in specifying the response set are said to be "roved." (Stimulus attributes that are not target attributes are referred to as "background" attributes, and background attributes that are not roved are referred to as "fixed," even though their values may be changed in proceeding from one experiment to the next.) Thus, for example, if both intensity and frequency are varied within the stimulus set, but only intensity is to be identified, the experiment would be referred to as an intensity identification experiment with roving frequency. (Note that in AI experiments with roving parameters, the entities S_i referred to above are not actually individual stimuli but rather classes of stimuli.) The extent to which identification performance for a given attribute is degraded by roving another attribute provides a measure of the perceptual interaction between the two attributes.

Independent of the particular type of AI experiment under consideration, we will summarize the results in terms of the first-order stimulus-response matrix, i.e., the $k \times k$ matrix in which the entry in row *i* and column *j* specifies the number of times stimulus S_i led to response R_j. In other words, we will assume that the trials are statistically independent and ignore all possible sequential effects.

Further discussion of such notions as stimulus-set dimensionality, perceptual interaction, and stimulus-response compatibility, is included in later sections when we discuss the issue of maximizing information transfer.

III-1.2 Information Concepts and Computation

Information is something we get when we learn something we didn't know before. Any communication act provides information only insofar as it reduces a condition of ignorance or uncertainty about the state of things under consideration. The amount of information in a stimulus set (IS) is defined by the weighted sum of $\log_2 P(S_i)$:

$$IS = -\sum_{i=1}^{k} P(S_i)\log_2 P(S_i) \quad , \qquad \text{(Eqn. 1)}$$

where $P(S_i)$ is the *a priori* probability of stimulus S_i, and k is the number of alternatives in the stimulus set. Information transfer measures the increase in information about the signal transmitted resulting from knowledge of the received signal. For a particular stimulus-response pair (S_i,R_j), it is given by $\log_2[(P(S_i/R_j))/(P(S_i))]$, where $P(S_i/R_j)$ is the probability of S_i given R_j, and, as above, $P(S_i)$ is the *a priori* probability of S_i. The average information transfer IT is thus given by

$$IT = \sum_{j=1}^{k}\sum_{i=1}^{k} P(S_i, R_j)\log_2\left(\frac{P(S_i/R_j)}{P(S_i)}\right) = \sum_{j=1}^{k}\sum_{i=1}^{k} P(S_i, R_j)\log_2\left(\frac{P(S_i, R_j)}{P(S_i)P(R_j)}\right) \quad , \qquad \text{(Eqn. 2)}$$

where $P(S_i,R_j)$ is the joint probability of stimulus S_i and R_j, and $P(R_j)$ is the probability of R_j. Note that the direction of communication is not important in computing IT because of the symmetry of S_i and R_j in the above equation. A related quantity, 2^{IT}, is interpreted as the number of stimulus categories that can be correctly identified. It is an abstraction, since 2^{IT} is not necessarily an integer. The values of IT and 2^{IT} are used interchangeably to characterize the outcome of an AI experiment.

III-1.3 Issues in IT Estimation

The first issue in obtaining a reliable estimate of IT for a given stimulus set concerns the total number of trials to be collected. The maximum likelihood estimate of IT from a confusion matrix is computed by approximating underlying probabilities with frequencies of occurrence:

$$IT_{est} = \sum_{j=1}^{k}\sum_{i=1}^{k}\left(\frac{n_{ij}}{n}\right)\log_2\left(\frac{n_{ij}\cdot n}{n_i\cdot n_j}\right), \qquad \text{(Eqn. 3)}$$

where n is the total number of trials in the experiment, n_{ij} is the number of times the joint event (S_i,R_j) occurs, and $n_i = \sum_{j=1}^{k} n_{ij}$ and $n_j = \sum_{i=1}^{k} n_{ij}$ are the row and column sums. These quantities can all be derived directly from the confusion matrix obtained in the AI experiment.

Unfortunately, IT_{est} is not only subject to statistical fluctuations, but it is also a biased estimate: it tends, for a limited number of trials, to overestimate IT. Further, the magnitude of the bias tends to greatly exceed the magnitude of the fluctuations (Rogers & Green, 1954; Rabinowitz, Houtsma, Durlach, & Delhorne, 1987).

According to Miller (1954), a useful first-order correction for the bias provided $n>5k^2$ is to subtract $\Delta IT = \frac{1}{2n\cdot\ln 2}(k-1)^2$ from IT_{est}. Miller (1954) also pointed out, however, that when $n < 5k^2$ and many of the n_{ij} $(i \neq j)$ values are near zero (i.e., transmission is good), ΔIT often results in too large a correction. Houtsma (1983) used computer simulations to estimate the asymptotic value of IT_{est} from limited experimental data, but the method does not work well when there are large differences among the amount of information each stimulus attribute contributes to the overall IT in a stimulus set involving many attributes (Tan, 1988). Thus, the best way of obtaining a reliable estimate of IT is still, if at all possible, to collect sufficient data to satisfy the constraint $n > 5k^2$.

In our experiments, when the number of alternatives in the stimulus set was small (i.e., $k \leq 10$), at least $n=5k^2$ trials were conducted for each subject. When the number of alternatives in the stimulus set was large (i.e., $30 \leq k \leq 120$) and collecting $5k^2$ trials appeared too time-consuming (i.e., $4{,}500 \leq n = 5k^2 \leq 72{,}000$), a different strategy was used. According to Rabinowitz (1995, personal communication and unpublished data), IT_{est} tends to reach an asymptote faster if the performance level is high. In the extreme case, when identification performance is perfect, very few trials are

needed to determine that IT = IS = $\log_2 k$ (assuming that the k alternatives are equally likely). As the percent-correct score decreases, IT decreases but the exact relationship depends upon the distribution of errors. For relatively large k (>10) and low error rate e (i.e., $0 \le e \le 0.05$, 95% ≤ percent-correct score ≤ 100%), the ratio of IT over IS deviates from 100% by less than twice the error rate in almost all cases (and frequently deviates by less than the error rate). Therefore, a conservative estimate of IT from percent-correct scores, denoted IT_{pc}, is given by

$$IT_{pc} = IS \times (1 - 2e) .$$

III-1.4 Principles for Increasing Information Transfer with Human Observers

Given a one-dimensional stimulus set, the information transfer for human observers is limited to roughly 2.3 to 3.2 *bits* corresponding to roughly 7±2 perfectly identifiable stimuli (Miller, 1956). This limit can be overcome by employing multi-dimensional stimulus sets. Therefore, the most important thing to do in increasing information transfer is to use stimuli with as many dimensions as possible. In this section, we discuss the principles for increasing IT for each dimension, the principles for recruiting additional dimensions, and the related issues of redundant coding, stimulus-uncertainty selection, and stimulus-response (SR) compatibility.

The principles for increasing IT for any dimension are (1) to use the entire variable range, and (2) to space stimuli in equal perceptual units. According to Braida & Durlach (1972, Fig. 4d and Fig. 8), IT increases monotonically with range of intensity for auditory intensity perception. Durlach, Delhorne, Wong, Ko, Rabinowitz, & Hollerbach (1989) also showed that IT increases with range of length for manual length identification by the finger-span method. Given the maximum stimulus range, the values of stimuli should ideally be equally spaced in terms of JNDs. If Weber's law holds for the stimulus variable under consideration (e.g., tonal intensity, force), equal perceptual distances can be accomplished by spacing stimuli logarithmically across the entire range.

For a multi-dimensional stimulus set, it is generally true that the greater the number of stimulus attributes, and the smaller the perceptual interaction among these attributes, the higher the value of IT. First, the stimulus set should include as many dimensions as possible. Although the concept of dimensionality is still not well defined (see Chap. V for discussion of this issue), it is

generally true that higher dimensionality is associated with a larger number of stimulus attributes. In this sense, human faces constitute a good example of a rich display; i.e., a display with a large number of dimensions. One can easily recognize hundreds of faces because there are many facial features that contribute to the overall "look" of a face. Pollack & Ficks (1954) obtained an IT of 7 *bits* (>> 3.2 *bits*) by employing an eight-dimensional auditory display. Thus it is generally fruitful to vary as many attributes as possible in a stimulus set. In addition, the perceptual interaction between stimulus attributes should be minimized. The effectiveness of an additional stimulus attribute can be judged by the additional IT it brings compared with the additional IS it contributes to the stimulus set. Each of the eight auditory dimensions employed by Pollack & Ficks (1954) contributed 1 *bit* to the overall IS, and roughly 0.77–0.97 *bits* to the overall IT. It has also been well established that a higher IT can be obtained by using a few steps along many stimulus attributes than using many steps along a few dimensions. Pollack & Ficks (1954) showed, with six-dimensional auditory displays, that overall IT increased by 25% when the coding went from binary to trinary, but only by an additional 5% when it went from trinary to quinary. In general, extreme subdivision of a stimulus dimension does not appear warranted.

Redundant coding can be used to increase IT without necessarily increasing IS at the same time. Eriksen & Hake (1955) showed that whereas IT for unidimensional identification of size, hue or brightness of visual stimuli is between 2.3 to 3.1 *bits*, IT for identifying visual stimuli with these three attributes varying in concert is 4.1 *bits*. (In all cases, the number of stimuli was either 17 or 20.) One is thus led to the following question: given a set of stimulus attributes, is it better to vary the attributes independently or in a totally correlated way in order to increase IT? Lockhead (1966) provided some data on this issue for visual identification of line length and position. When stimulus duration was 200 *msec* and the display was well lit, IT was 1.1 *bits* for length identification, 1.0 *bit* for position identification, 1.2 *bits* when the two were perfectly correlated, and ≥ 1.7 *bits* (a lowerbound estimate) when the two were varied independently. It thus seems that higher IT can be achieved by varying stimulus attributes independently. In this thesis, emphasis was placed on discovering effective stimulus attributes and determining the number of steps along each attribute. We did not explore the option of varying stimulus attributes redundantly.

Selection of stimulus uncertainty also affects the IT that can be achieved. Because the IT *vs.* IS curve is usually thought of as being monotonic and having an asymptotic value, in addition to the fact that it is always below the straight line defined by IT = IS, the general rule of thumb is to select an IS value that is higher than the expected IT. In our first set of probe experiments on amplitude identification with fixed or roving frequencies, an IS value that was at least 1 *bit* higher than the expected IT was selected. However, we then discovered (based on a limited amount of data) that the relationship between IT and IS may not be monotonic. In particular, it appeared that IT decreased as IS increased above our estimate of the maximum IT achievable. It may therefore be more efficient to select several IS values around the expected IT in order to reveal the maximum IT. This issue is discussed further in Sec. B-2 and Chapter V.

Finally, stimulus-response compatibility also affects IT. The term "stimulus-response (S-R) compatibility" was popularized by the research of Fitts and his colleagues, in which assignments of stimuli to responses were manipulated. In one of their studies, Deininger & Fitts (1955) studied three stimulus sets and three corresponding response sets in a perceptual-motor task where the subject was instructed to move a stylus along a certain path when a particular stimulus appeared. They found that performance is best (in terms of reaction time and errors) when the response set is spatially congruent with the stimulus set *and* the matching of the points in the stimulus space to those in the response space is spatially consistent. The authors demonstrated that when the mapping between stimuli and responses was more compatible, subjects performed faster with less errors. The phenomena of S-R compatibility, however, are not restricted to situations involving physical correspondence between the stimulus and response sets, as is evident in a recent book of reviews on this topic (Proctor & Reeve, 1990). It is generally accepted that compatibility effects reflect basic cognitive processes (i.e., mental representations and translations between them) that influence human performance in a wide variety of situations. Although the relative compatibility between two groups of stimulus-response sets can be determined by subject's performance in terms of speed and error rate, there are no universal rules for the design of the optimal response set for a given stimulus set. In our experiments, many attributes (e.g., frequency, amplitude, and site of stimulation, etc.) were associated with a stimulus. A response set that reflects the salient features of the stimulus set would hopefully lead to higher performance faster. Although it is not obvious that a higher S-R compatibility necessarily leads to a higher IT plateau after extensive

training, it does seem obvious that it reduces the training time required to approach such a plateau.

Further discussion of the characteristics of stimulus sets that are likely to lead to high IT is presented in Sec. V-2.

III-2 The Stimulus and Response Sets for a Stimulus Duration of 500-*msec*

A series of probe experiments were conducted to determine the effective stimulus attributes that can be used in constructing a relatively large stimulus set with easily identifiable stimuli. It was found that subjects could naturally categorize motions over a frequency range of near DC to 300 *Hz* into three perceptually distinctive categories: slow motion (up to about 6 *Hz*), fluttering motion (about 10 *Hz* to 70 *Hz*), and smooth vibration (above about 150 *Hz*). Therefore, multi-component stimuli were formed by simultaneously stimulating multiple fingers with waveforms containing sinusoids (varying in both frequency and amplitude) from the three frequency regions. The number of values to be used with each stimulus attribute was determined by employing the absolute identification paradigm with fixed and roving backgrounds. It was found that subjects could reliably identify two frequencies within each of the three frequency regions, two amplitudes with the low-frequency component, and one amplitude (i.e., fixed amplitude) with the mid- and high-frequency components, provided that masking is minimized by carefully balancing the signal strengths of components from different frequency regions. Based on the results obtained from the probe experiments and the intuitions gained from running these experiments, the 500-*msec* stimulus set was constructed. A corresponding response code was designed that reflected the underlying structure of this stimulus set. This section provides an overview of the stimulus and response sets with a duration of 500 *msec*. A more detailed description of the probe experiments can be found in Appendix B.

III-2.1 The 120 Stimuli in the 500-*msec* Stimulus Set

The structure of the 500-*msec* stimulus set is illustrated in Fig. III-1. Each stimulus (S) is defined by which finger(s) (L_i, i=1,2,3 for thumb, index finger, and middle finger, respectively) are stimulated

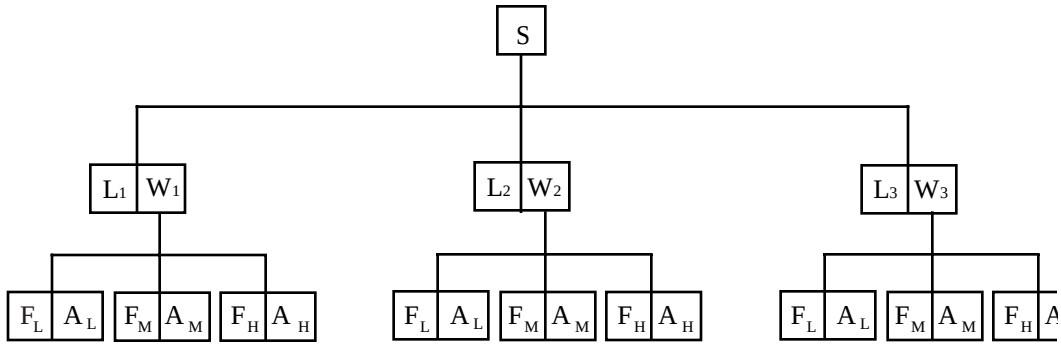

Figure III-1. The structure of a stimulus.

with which waveform (W_i, i=1,2,3). The value of L_i was 1 if the corresponding finger was stimulated and 0 otherwise. Four (4) stimulation sites were employed: either one of the three fingers was stimulated, or all three of them were stimulated with the same waveform. In other words, the possible combinations of (L_1, L_2, L_3) were (1,0,0), (0,1,0), (0,0,1) or (1,1,1), and in the case of (1,1,1), W_1=W_2=W_3. Each waveform, of which there were 30, was a broadband signal containing three sinusoidal components with frequencies and amplitudes denoted by (F_L, A_L), (F_M, A_M), and (F_H, A_H). The same set of waveforms was used to stimulate any of the fingers or all fingers. The combinations of 4 finger locations and 30 waveforms resulted in a total of 120 alternatives in the 500-*msec* stimulus set.

Among the 30 waveforms, each of which had a 10 *msec* rise-fall time, eight (8) used a single frequency (i.e., the amplitudes for two of the three components were zero), sixteen (16) used two frequencies (i.e., the amplitude for one of the three components was zero), and six (6) used three frequencies. Among the 8 single-frequency waveforms (Fig. III-2), the value of F_L was 2 or 4 *Hz*, the value of F_M was 10 or 30 *Hz*, and the value of F_H was 150 or 300 *Hz*. The amplitude for each F_L was 35 dB SL or 44 dB SL. The amplitude for each F_M or F_H was fixed because the perceptual qualities of the middle- and high-frequency components were not independent of amplitude. Double- and triple-frequency waveforms were constructed by combining single-frequency elements in different frequency regions (see Table III-1 for a complete listing). A 4-*Hz* signal was never combined with a 10-*Hz* signal because the former was found to interfere with the perception of the latter. Whenever middle and high-frequency components were combined, only the 300-*Hz*

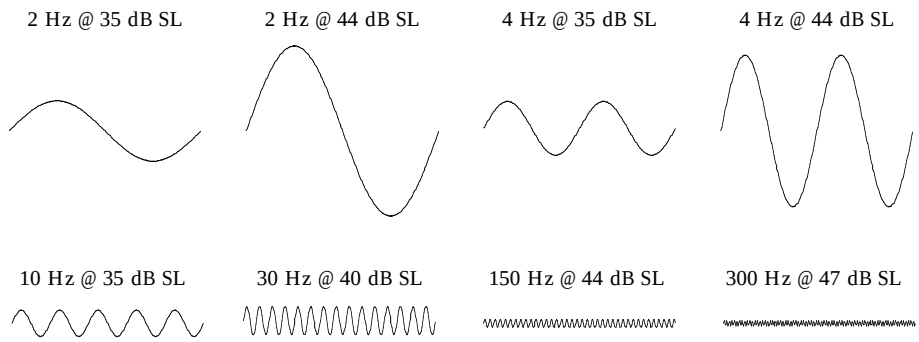

Figure III-2. Single-frequency waveforms.

TABLE III-1. The 30 waveforms for the 500-*msec* stimulus set. Units are (*Hz*, dB SL).

single-frequency waveforms	(2, 35) (2, 44) (4, 35) (4, 44)	Group1
	(10, 35) (30, 40)	Group2
	(150, 44) (300, 47)	Group3
double-frequency waveforms	(2, 35)+(10, 35) (2, 35)+(30, 40) (2, 35)+(150, 44) (2, 35)+(300, 44)	Group 4
	(2, 44)+(10, 40) (2, 44)+(30, 40) (2, 44)+(150, 44) (2, 44)+(300, 44)	Group 5
	(4, 35)+(30, 40) (4, 35)+(150, 44) (4, 35)+(300, 44)	Group 6
	(4, 44)+(30, 44) (4, 44)+(150, 44) (4, 44)+(300, 47)	Group 7
	(10, 35)+(300,44) (30, 40)+(300, 44)	Group 8
triple-frequency waveforms	(2, 35)+(10, 35)+(300, 44) (2, 35)+(30, 40)+(300, 47)	Group 9
	(2, 44)+(10, 40)+(300, 44) (2, 44)+(30, 40)+(300, 47)	Group 10
	(4, 35)+(30, 40)+(300, 47)	Group 11
	(4, 44)+(30, 40)+(300, 47)	Group 12

signal was used because the middle-frequency components were found to interfere with the identification of F_H. Finally, some amplitudes were adjusted in order to balance the relative strengths of different signal components and to minimize fatigue due to excessively strong signals.

The 30 waveforms have distinctive perceptual qualities. The 2-*Hz* and 4-*Hz* signals are perceived as slow motions with 1 or 2 cycles at small or large amplitudes. The 30-*Hz* signal is very rough

and seems to be beating on the fingertip. The 10-*Hz* signal is relatively mild, and gives rise to a wobbling sensation when combined with a 2-*Hz* signal. The 150-*Hz* vibration is relatively diffused and of lower pitch. The 300-*Hz* vibration is more focused and of higher pitch. When two or three frequencies are combined, the sensations associated with single-frequency components can still be discerned.

III-2.2 The Response Code

It was a challenge to design a response set and a stimulus-response mapping that was compatible with the 120 stimuli. Intuitively, it seemed that the response set should reflect the underlying structure of the stimulus set; e.g., each response should consist of two parts, one corresponding to stimulation site, and one to stimulating waveform. After preliminary experimentation, it seemed that a graphical response code might work better than text or numerical labels. Accordingly, graphic icons corresponding to the 30 waveforms were laid out as circular buttons on a digitizing tablet along with four icons “M”, “I’, “T” and “ALL” corresponding to the middle finger, index finger, thumb, and all fingers, respectively (Fig. III-3). A “DEL” icon was available for deleting responses if the subject felt that the wrong icon was accidentally pressed. An “ENTER” icon was used to terminate a trial. In general, the component with the lowest frequency was the same across a row of waveform icons, and the component with the highest frequency was the same across a column of waveform icons. Some exceptions were made in order to contain the waveform icons to a relatively small area (for ease of visual search). Subjects used a stylus to pick the appropriate response icons by pressing on them. Since the 150 *Hz* and 300 *Hz* waveforms did not reproduce very well at that scale, they were represented by blue and red dots in the actual response tablet.

III-3 Information Transfer Measurements with the 500-*msec* Stimulus Set

III-3.1 General Methods

Three subjects (S_1, S_2 and S_3) were trained and tested. S_1 (the author) is a 30 year old female graduate student at MIT; S_2 is a 42 year old male who also participated in our earlier study on

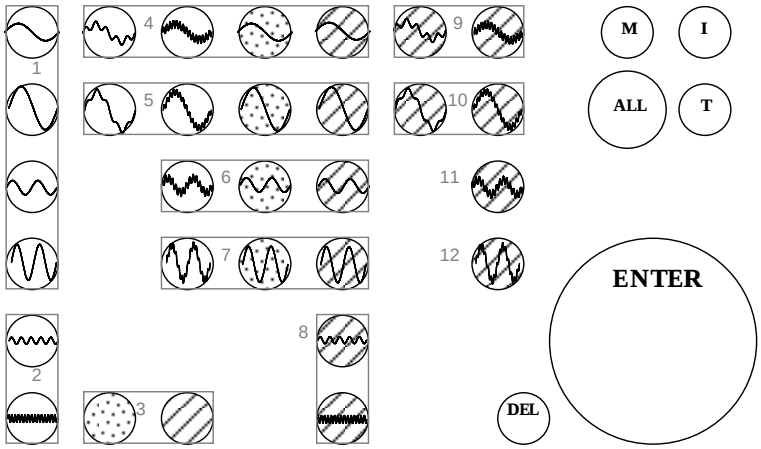

Figure III-3. Layout of responses for experiments using the 500-*msec* stimulus set. The numbers correspond to the grouping listed in the rightmost column of Table III-1. (The dot and line patterns represent the blue and red colors used in actual icons for the 150 *Hz* and 300 *Hz* waveforms, respectively.)

tactual reception of Morse code; and S_3 is a 20 year old undergraduate student at MIT. All subjects are right-handed with no known tactual impairments of their hands.

During all experiments, the TACTUATOR was visually blocked from the subjects. Subjects wore earplugs and earphones with pink noise to eliminate auditory cues. (The TACTUATOR produces no audible noise except at 300 *Hz*.)

For both training and testing, the standard AI paradigm with trial-by-trial correct-answer feedback was employed. There were two differences, however, between the paradigms used for training and testing. During training, each stimulus alternative was presented an equal number of times per run (i.e., randomization without replacement). This ensured that subjects had an equal opportunity to learn all the signals in the stimulus set. The side effect was that stimulus uncertainty decreased as a function of number of trials. During testing, stimuli were presented with equal *a priori* probabilities on each trial (randomization with replacement). Thus, stimulus uncertainty remained the same throughout an experimental run. The other difference was that during training, subjects were allowed to skip trials. Skipped trials were repeated later on. During testing, subjects were required to respond to all trials.

Training was conducted for all stimulus durations before testing was done. There was a two months gap between the end of training and the beginning of testing for subjects S_1 and S_2. Subject S_3 was tested immediately after he completed training. Results for training are presented in terms of percent-correct scores and total number of hours; results for testing are summarized in terms of IT. The discussion of the AI paradigm and IT computation presented earlier is applicable to the testing procedures. Because of the relatively large number of alternatives in the stimulus set (i.e., up to 120), IT was computed as $IT_{pc} = IS_k \times (1 - 2 \times \bar{e})$, where $IS_k = \log_2 k$, k is the number of equally-likely alternatives in the stimulus set),[1] and $\bar{e}$ is the average error rate. IT_{est} was also computed. However, because the amount of experimental data is relatively limited, IT_{est} should be treated as an upperbound.

III-3.2 Training Results

Subjects learned to identify the 120 alternatives in the 500-*msec* stimulus set in a number of steps. They were first trained to identify the 30 waveforms on the index finger, then to identify the same 30 waveforms when applied randomly to any one of the three fingers, and finally to identify both the finger locations and the waveforms of all 120 alternatives in the 500-*msec* stimulus set. For each stimulus set, training was terminated when a subject reached the performance criterion of either one run of 100% correct or three runs with percent-correct scores of 95% or higher (not necessarily consecutively).

Waveform Identification on the Index Finger

The waveforms in the 500-*msec* stimulus set were divided into 12 groups (see Table III-1 and Fig. III-3): the waveforms in the first group contained only F_L components; those in the second group contained only F_M components; and those in the third group contained only F_H components. Groups 4 to 8 contained double-frequency waveforms. Groups 9 to 12 contained triple-frequency waveforms. Subjects first practiced with and identified the 4 waveforms in group 1, then the 6 waveforms in groups 1 and 2, then the 8 waveforms in groups 1 to 3, and so on until the stimulus set contained all 30 waveforms.

1. $IS_k = \log_2 k$ is derived by substituting $P(S_i) = 1/k$ in Eqn. 1.

During practice, subjects could choose to feel any waveform on the index finger by selecting the corresponding waveform icon on the response tablet. Practice was self-terminated when they a subject felt ready to run the identification experiments. Each stimulus alternative was applied exactly 5 times to the index finger for each experimental run. In other words, the number of valid trials per run, upon which percent-correct scores were computed, was 20 for group 1, 30 for groups 1 and 2, etc. Subjects had to reach the performance criterion before new waveforms were added to the stimulus set. Results were summarized either as 100% or by averaging the last three percent-correct scores (Table III-2).[1] Also, the total number of hours of training was recorded for S_2 and S_3.[2] The three subjects were able to reach the performance criterion with the 30 waveforms in the stimulus set with an average accuracy of 100% (S_1), 100% (S_2, 9 *hours*), and 96% (S_3, 15 *hours*).

TABLE III-2. Average percent-correct scores from waveform identification on the index finger.

No. of Stimuli	S_1	S_2	S_3
4 (Group 1)	-	100%	100%
6 (Groups 1-2)	-	100%	100%
8 (Groups 1-3)	-	100%	96%
12 (Groups 1-4)	-	96%	100%
16 (Groups 1-5)	-	98%	97%
19 (Groups 1-6)	-	100%	98%
22 (Groups 1-7)	-	96%	97%
24 (Groups 1-8)	-	96%	96%
26 (Groups 1-9)	-	97%	96%
28 (Groups 1-10)	-	95%	95%
29 (Groups 1-11)	-	97%	93%
30 (Groups 1-12)	100%	100%	96%

Waveform Identification with Roving Fingers

To quickly check whether the 30 waveforms were readily identifiable when applied to the thumb or the middle finger, the subjects practiced with and identified these waveforms again when they

1. Because S_1 was highly experienced with these stimuli, she was only tested with all 30 waveforms in the stimulus set.
2. Total number of hours of training could not be accurately estimated for S_1 because she was involved in the development of all stimulus sets and, therefore, was over-exposed to the stimuli.

were applied randomly to any one of the three fingers. There were 90 alternatives (30 waveforms × 3 finger locations) in the stimulus set. The response set still consisted of the 30 waveform icons. During practice, the subject could select any combination of stimulation site and waveform by picking the "M", "I" or "T" icon followed by a waveform icon. During a training run, however, only the waveform response was required. Each of the 30 waveforms was applied exactly twice to each of the three fingers during each run, resulting in a a total of 180 non-skipped trials per run.

It was noticed that because of the difference in range of motion of the three fingers, adjustment in signal amplitude was needed to equalize loudness perception of the low-frequency signals. Informal testing was done in which S_1 was presented with one of the 4 waveforms in group 1 on the index finger, and a signal of the same frequency on the thumb or the middle finger. S_1 could adjust the amplitude of the signal on the thumb or the middle finger until it felt equally "loud" to the one on the index finger. It was found that equal tactual loudness could be achieved by increasing the amplitude of the 2 *Hz* and 4 *Hz* signal components by 2 dB for the middle finger and decreasing it by 2 dB for the thumb. With this modification, all subjects were able to reach the performance criterion with an average accuracy of 99% (S_1), 98% (S_2, 1 *hour*), and 94% (S_3, 1 *hour*). These results indicated that the thirty waveforms could be well identified using any one of the three fingers.

Identification of All 120 Alternatives

Subjects were now ready to be trained with all 120 alternatives in the 500-*msec* stimulus set. Each stimulus alternative was applied twice during each training run, resulting in a total of 240 non-skipped trials per run. Subjects were instructed to first select the "M", "I", "T" or "ALL" icon for site of stimulation, then the waveform icon corresponding to the stimulus as a response. All three subjects were able to reach the performance criterion with an average accuracy (over the last three runs) of 98% (S_1), 96% (S_2, 3.5 *hour*) and 96% (S_3, 3.5 *hour*).

Learning curves for the three subjects are presented in Fig. III-4. All subjects were able to reach the performance criterion within 10 training runs.

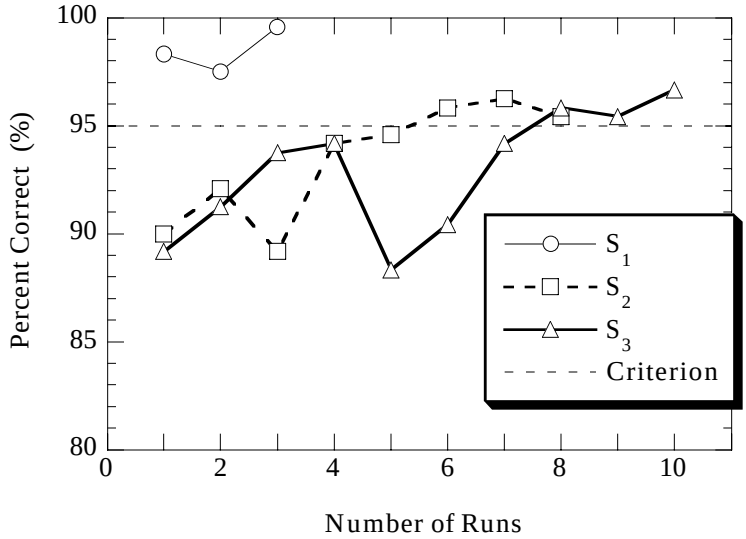

Figure III-4. Learning curves for each subject with all 120 stimuli in the 500-*msec* stimulus set.

The number of skipped trials for the last three runs were 12/26/22 for S_1, 3/0/3 for S_2, and 48/51/37 for S_3. S_2 rarely skipped a trial; when he did, either too few or too many response icons were in the response, indicating that efforts were made to respond to those trials. S_1 and S_3, however, skipped many trials intentionally (i.e., their responses were blank).

III-3.3 Test Results

During final testing, three runs (each containing exactly 240 trials) were collected for each subject. Results are summarized in Table III-3. The IT_{pc} averaged over all three subjects is 6.5 *bits*, corresponding to 90 perfectly identified items. The values of IT_{est} are not much higher than IT_{pc}, mainly because they cannot exceed the IS value of 6.9 *bits*.

TABLE III-3. Information Transfer with the 500-*msec* stimulus set.

Subject	***Percent-Correct Scores***	IT_{est}	***Average***	IT_{pc}	***Average*** IT_{pc}
S_1	99%, 99%, 98%	6.8 *bits*	99%	6.8 *bits*	**6.5 *bits***
S_2	94%, 95%, 96%	6.6 *bits*	95%	6.2 *bits*	
S_3	96%, 97%, 97%	6.6 *bits*	97%	6.5 *bits*	

III-4 Information Transfer Measurements with the 250-*msec* Stimulus Set

III-4.1 Stimuli

The 250-*msec* waveform set contained 30 waveforms that were very similar to those in the 500-*msec* waveform set, except that the frequency of the 4-*Hz* components was raised to 6 *Hz*, and that of the 10-*Hz* components was raised to 15 *Hz*. With these changes, the two F_L values of 2 and 6 *Hz* could be easily discriminated, as were the higher F_L value of 6 *Hz* and the lower F_M value of 15 *Hz*. A complete listing of the 30 waveforms is shown in Table III-4. The same four finger locations were used: all fingers, thumb alone, index finger alone, and middle finger alone. Therefore, there were again a total of 120 alternatives in the stimulus set. The response code shown in Fig. III-3 was modified to take account of the waveform changes.

TABLE III-4. The 30 waveforms for the 250-*msec* stimulus set. Units are (*Hz*, dB SL). Signals that are different from those in the 500-*msec* stimulus set are underlined.

single frequency	(2, 35) (2, 44) <u>(6, 35)</u> <u>(6, 44)</u> <u>(15, 35)</u> (30, 40) (150, 44) (300, 47)
double frequency	<u>(2, 35)+(15, 35)</u> (2, 35)+(30, 40) (2, 35)+(150, 44) (2, 35)+(300, 44) <u>(2, 44)+(15, 40)</u> (2, 44)+(30, 40) (2, 44)+(150, 44) (2, 44)+(300, 44) <u>(6, 35)+(30, 40)</u> <u>(6, 35)+(150, 44)</u> <u>(6, 35)+(300, 44)</u> <u>(6, 44)+(30, 44)</u> <u>(6, 44)+(150, 44)</u> <u>(6, 44)+(300, 47)</u> <u>(15, 35)+(300,44)</u> (30, 40)+(300, 44)
triple frequency	<u>(2, 35)+(15, 35)+(300, 44)</u> (2, 35)+(30, 40)+(300, 47) <u>(2, 44)+(15, 40)+(300, 44)</u> (2, 44)+(30, 40)+(300, 47) <u>(6, 35)+(30, 40)+(300, 47)</u> <u>(6, 44)+(30, 40)+(300, 47)</u>

III-4.2 Training Results

The training procedure was the same as that used with the 500-*msec* stimulus set. Subjects practiced with and identified all 120 stimuli in the 250-*msec* stimulus set. Each stimulus was presented twice during each training run. All subjects were able to reach the performance criterion with an average accuracy (over the last three runs) of 100% (S_1), 96% (S_2, 3.5 *hours*), and 93% (S_3, 6 *hours*).

Learning curves for the three subjects are presented in Fig. III-5. All subjects were able to reach the performance criterion within 20 training runs.

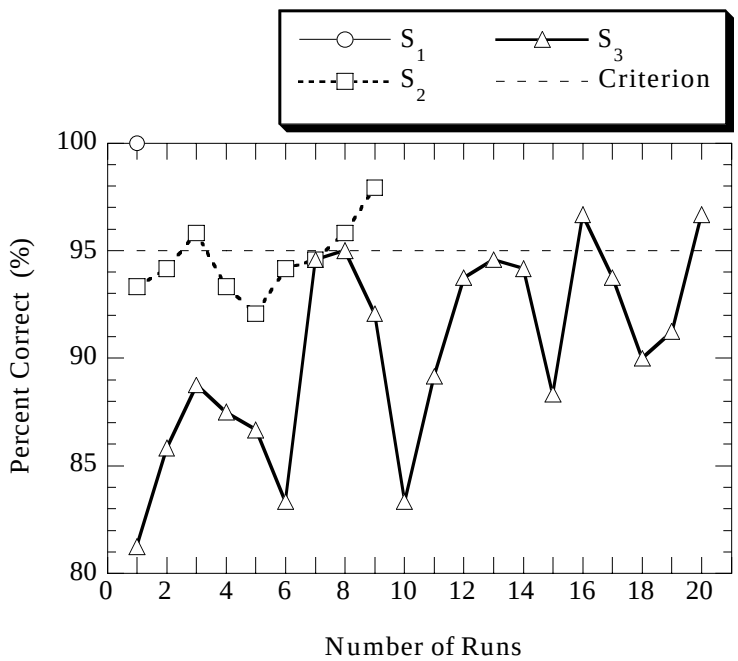

Figure III-5. Learning curves for each subject with all 120 stimuli in the 250-*msec* stimulus set.

The number of skipped trials for the last three runs were 22 for S_1 (one run of 100%), 4 / 0 / 1 for S_2, and 38 / 25 / 63 for S_3. Again, S_2 rarely skipped a trial; but S_1 and S_3 skipped many trials.

III-4.3 Test Results

During final testing, there were exactly 240 trials per experimental run. After the initial three runs, it appeared that S_2 and S_3 had not yet reached a performance plateau (Fig. III-6). Therefore, a total of 10 runs were collected with these two subjects. Scores for the last three runs are averaged and summarized in Table III-5. The IT_{pc} averaged over all three subjects is 6.4 *bits*, corresponding to 84 perfectly identified items. Thus, there is very little loss (i.e., 0.1 *bit*) in IT by shortening the stimulus duration by a factor of 2. The values of IT_{est} are, again, not much higher than IT_{pc}, mainly because they cannot exceed the IS value of 6.9 *bits*.

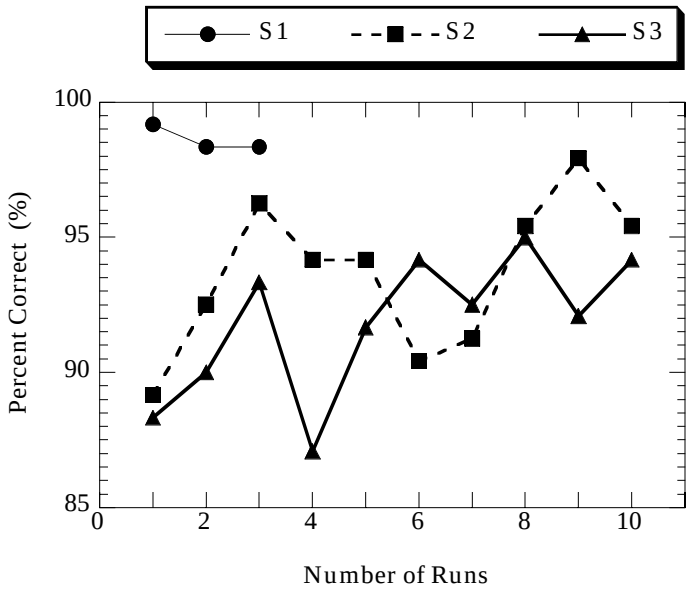

Figure III-6. Test results in terms of percent-correct scores for each subject with the 250-*msec* stimulus set.

TABLE III-5. Information Transfer estimated from the last three test runs conducted with the 250-*msec* stimulus set.

Subject	***Percent-Correct Scores***	***IT_{est}***	***Average***	***IT_{pc}***	***Average IT_{pc}***
S_1	99%, 98%, 98%	6.8 *bits*	99%	6.8 *bits*	***6.4 bits***
S_2	95%, 98%, 95%	6.7 *bits*	96%	6.3 *bits*	
S_3	92%, 95%, 94%	6.6 *bits*	94%	6.1 *bits*	

III-5 Information Transfer Measurements with the 125-*msec* Stimulus Set

III-5.1 Stimuli

With a signal duration of 125 *msec*, subjects could no longer reach the performance criterion of ≥ 95% with 30 waveforms that were similar to the ones in the 500-*msec* or 250-*msec* waveform sets. In order to keep performance at a high level, only one frequency value was used in each of the three frequency ranges. Direction of motion was also introduced as an additional signal attribute.[1] The resultant 125-*msec* waveform set contained 19 waveforms, as shown in Table III-6. A negative sign indicates that movements started in a direction that corresponds to finger flexion. The default used in this and all previous experiments was to start movements in the finger-extension direction. The direction attribute was only effective with the F_L components.

TABLE III-6. The 19 waveforms for the 125-*msec* stimulus set. Units are (*Hz*, dB SL).

single frequency	(4, 35) −(4, 35) (4, 44) −(4, 44) (30, 40) (300, 47)
double frequency	(4, 35)+(30, 40) −(4, 35)+(30, 40) (4, 44)+(30, 40) −(4, 44)+(30, 40) (4, 35)+(300, 47) −(4, 35)+(300, 47) (4, 44)+(300, 47) −(4, 44)+(300, 47) (30, 40)+(300, 47)
triple frequency	(4, 35)+(30, 40)+(300, 47) −(4, 35)+(30, 40)+(300, 47) (4, 44)+(30, 40)+(300, 47) −(4, 44)+(30, 40)+(300, 47)

Again, the same four finger locations were used: all fingers, thumb alone, index finger alone, and middle finger alone. Therefore, there were a total of 76 alternatives in the stimulus set. The response code shown in Fig. III-3 was modified accordingly.

III-5.2 Training Results

The training procedure was the same as that used before with the 500-*msec* and 250-*msec* stimulus sets. Each stimulus was presented twice during each training run. All subjects were able to reach the performance criterion with an average accuracy (over the last three runs) of 100% (S_1), 96% (S_2, 2.5 *hours*), and 96% (S_3, 1.5 *hours*).

1. Direction of movement was not an effective signal attribute for longer signal durations.

Learning curves for the three subjects are presented in Fig. III-7. All subjects were able to reach the performance criterion within 9 training runs.

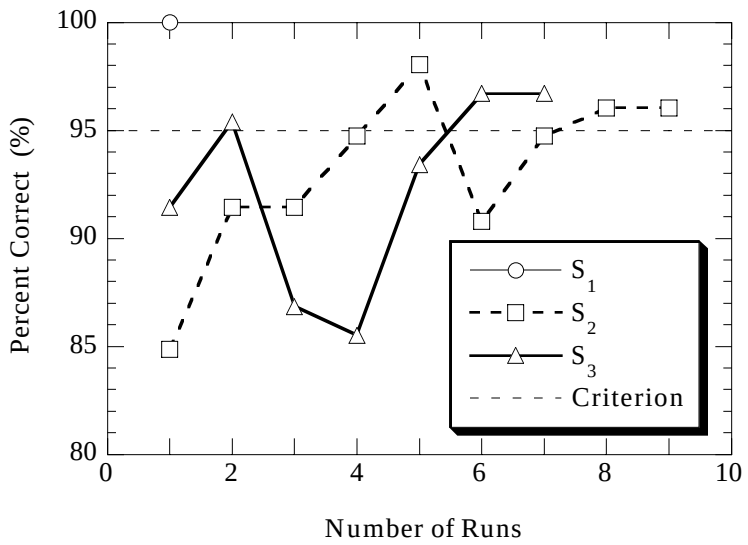

Figure III-7. Learning curves for each subject with all 76 stimuli in the 125-*msec* stimulus set.

The total number of trials skipped were 43 for S_1 (one run of 100%), 1/0/0 for S_2, and 59, 47, 41 for S_3. Although S_2 continued to skip very few trials, the percentage of skipped trials increased dramatically for S_1 and S_3.

III-5.3 Test Results

During final testing, there were exactly 152 trials per experimental run. After the initial three runs, it was not clear if S_2 and S_3 had reached a performance plateau (Fig. III-8). Therefore, one more run was collected with these two subjects. Scores for the last three runs are averaged and summarized in Table III-7. The IT_{pc} averaged over all three subjects is 5.6 *bits*, corresponding to approximately 50 perfectly identified items. Overall, there is an approximately 1 *bit* decrease in IT when the signal duration was reduced from 500 (or 250 *msec*) to 125 *msec*. The values of IT_{est} are, again, not much higher than those of IT_{pc} and limited by the IS value of 6.2 *bits*.

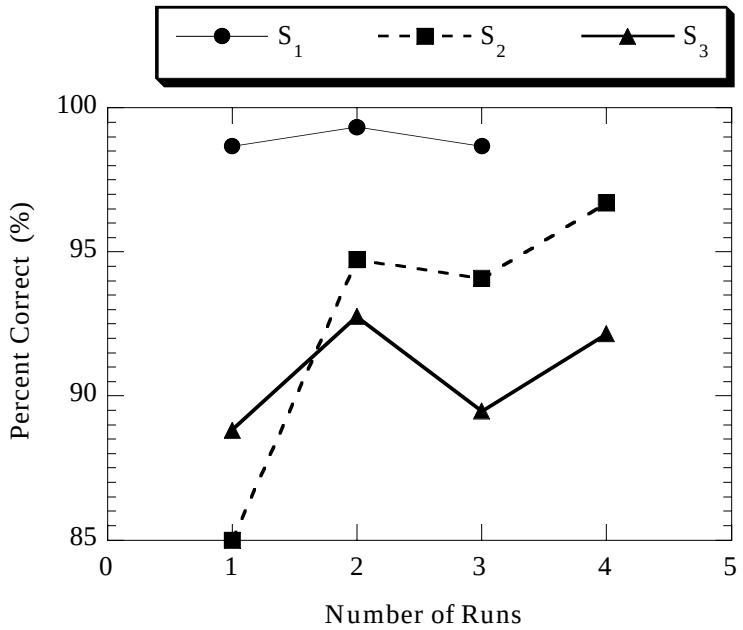

Figure III-8. Test results in terms of percent-correct scores for each subject with the 125-*msec* stimulus set.

TABLE III-7. Information Transfer with the 125-*msec* stimulus set.

Subject	***Percent-Correct Scores***	IT_{est}	***Average***	IT_{pc}	***Average*** IT_{pc}
S_1	99%, 99%, 99%	6.1 *bits*	99%	6.1 *bits*	**5.6 *bits***
S_2	95%, 94%, 97%	6.0 *bits*	95%	5.6 *bits*	
S_3	93%, 89%, 92%	5.9 *bits*	91%	5.1 *bits*	