Chapter V. Unresolved Issues

Three issues are discussed in this chapter: (1) selection of stimulus uncertainty for estimation of maximum information transfer, (2) definition of stimulus-set dimensionality in the context of increasing information transfer, and (3) relationship between motor output and reception of motional input sequences.

V-1 Selection of Stimulus Uncertainty

Consider first the case of a univariate stimulus set. It is well established that for such a set, IT is limited by the magic number 7±2 (Miller, 1956). Moreover, as illustrated in Fig. V-1 (left panel), it is generally accepted that IT increases monotonically with IS when IS is small and then plateaus when IS is large (see, for example, Garner, 1962, p. 75). With this picture in mind, attempts to determine the maximum IT (i.e., the plateau level) usually involve selecting an IS that is large relative to the expected value of IS at the knee of the IT *vs.* IS curve. However, two results suggest that this strategy may not be optimal.

First, we have observed in our probe experiments (see Appen. B) that IT tends to decrease slightly as IS is increased beyond the point at which the subject begins to make a significant number of identification errors (Fig. V-1, right panel). This implies that IT is not a monotonically increasing function of IS, and that to determine the maximum value of IT, it is necessary to judiciously select IS values around the expected value of maximum IT and to iteratively refine these choices. These IS values are likely to be much lower than those chosen by the usual approach of selecting IS>>IT.

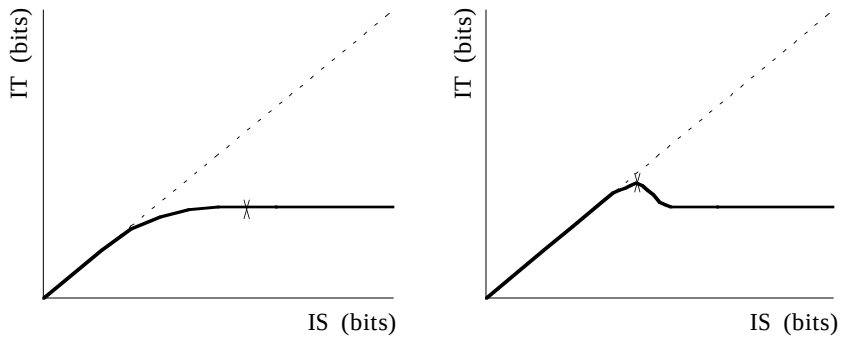

Figure V-1. Schematic illustrations of IT *vs.* IS relationships. The dashed lines indicate where IT=IS.

Second, using a smaller IS can greatly reduce the number of trials required to estimate IT, thereby increasing test efficiency. The saving in the total number of trials needed to obtain a relatively unbiased estimate of IT can come from two sources. The first source is the smaller IS. As discussed in Sec. III-1.3, Miller (1954) argued that at least $5k^2$ ($k = 2^{IS}$) trials are needed in order to effectively eliminate the bias in the estimate of IT. Based on this criterion, an IS of 5 *bits* would require 5120 trials (!), a number that is difficult to achieve in a practical experiment (Fig. V-2). Assuming, for example, that the maximum IT for a certain variable is around 3 *bits*, measuring IT with IS values of 1, 2, 3 and 4 *bits* (which would reveal the peak in IT) requires a total of 1,700 trials, which is much less than the number of trials needed if we selected an IS of 5 *bits*.

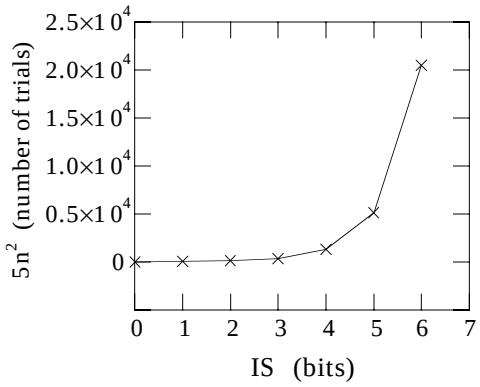

Figure V-2. Total number of trials needed for an unbiased estimate of information transfer.

Another result of using IS near, and not much greater than, IT is that overall identification accuracy is very high. This produces a further saving in the testing time that is required. According to Rabinowitz (1995, personal communication and unpublished data), estimated IT converges faster as a function of total number of trials when the overall performance level is high. In the extreme case, if performance were perfect, then the stimulus-response confusion matrix is purely diagonal and it converges after few trials.[1] When information transfer is estimated with a small IS (see the "cross" in the right panel of Fig. V-1), performance is close to being perfect (i.e., the "cross" is close to the dashed line). This means additional savings in total number of trials in order to obtain a good estimate of information transfer, a saving which can be very large. For example, our estimate of IT ≈ 6.5 *bits* with IS = 6.9 *bits* (with 500-*msec* stimuli) were obtained with 720 trials; Miller's criterion would have required at least 72,000 trials.

The existence of a maximum IT near where the IT *vs* IS curve deviates from a straight line of unit slope can be found in some earlier studies (e.g., Garner, 1953; Braida, 1969), but not in others (e.g., Pollack, 1952). We suspect that when this peak does exist, it is relatively small (i.e., probably less than 0.5 *bits* above the asymptotic value). Thus, in terms of estimating maximum IT, it does not lead to serious error if one used the asymptotic value. However, as noted above, the selection of stimulus uncertainty has a substantial impact on the number of trials needed to obtain an unbiased estimate of maximum IT. The effect of stimulus uncertainty on number of trials is even more evident when one considers the case of multivariate stimuli and larger maximum IT values.

To investigate this issue further, we need to first verify that a maximum IT exists with an intermediate value of IS for different sensory modalities under a variety of experimental conditions. We can then explore the possible sources of this local maximum.

V-2 Dimensionality of a Stimulus Set

As mentioned previously in Sec. III-1.4, it is generally well accepted that the magnitude of information transfer is related to the "richness" of a display. The word "dimension" has been used

1. Conversely, when IT/IS is small, the confusion matrix has many more entries (approaching k^2), that must be accurately determined.

extensively in classical literature to describe the richness of a display. For example, Miller (1956) showed with empirical results that information transfer for unidimensional stimulus sets is limited to 2–3 *bits*. Pollack & Ficks (1954) showed that a much higher IT could be obtained with an eight-dimensional auditory display. A closer examination of the way the word "dimension" is used in these two classical papers, however, reveals that dimension, and unidimensionality, are not adequately defined. According to Pollack & Ficks (1954), "when a stimulus aspect can be manipulated independently, it is often called a dimension of an elementary auditory display" (p. 155). This notion that dimensionality is defined by the number of independently manipulated variables in the display appears to be accepted implicitly by many investigators in the field. However, as will be seen below, in order for the notion of dimensionality to be truly useful in the domain of psychophysics, it is essential that it be defined in terms of perceptual properties of the stimulus set, not the physical properties.

In the following paragraphs, we consider three factors related to the "richness" of stimulus sets and to the amount of information transfer that is likely to be obtained with these sets. All of these factors need to be considered in the search for a definition of stimulus-set dimensionality that is both rigorous and relevant to the amount of information transfer that can be achieved when human observers are required to identify stimuli in the stimulus set.

One important factor in considering the "richness" of a display concerns the perceptual segmentation of a stimulus parameter. Consider, for example, identifying the site of a single-point stimulus (e.g., a pin prick) applied to some position along a line running up the back of one's body from ankle to neck. Although the subject might confuse points within a single body segment (e.g., along the calf of the legs), it is reasonable to think that very few mistakes would occur across the boundaries of a segment (e.g., between a point on the calf and one on the thigh). In such a case, it seems likely that Miller's 7±2 constraint would be applicable only to the locations within a given segment and that the total number of distinguishable locations would be roughly 7 times the number of segments. Moreover, in this case, the boundaries appear to be defined by our ability to label the body surface verbally. Boundaries can also be defined by salient changes in the sensation that is elicited as the physical parameter is varied. Consider the case in which one is listening to a broadband click presented binaurally with an interaural time delay τ, and the subject's task is to

identify τ. In the region $0 \leq \tau \leq 1$ *msec*, the subject will hear a single fused image within the head whose location will move from the center of the head to the leading ear as τ varies from 0 *msec* to 1 *msec*, and the identification response will be based on the lateralization of this image. In the region 1 *msec* $\leq \tau \leq$ 10 *msec*, the image will remain at the leading ear, but its character will change: as τ increases from 1 *msec* to 10 *msec*, the image will become increasingly rough and complex, eventually breaking up into two clicks that are temporally as well as spatially separated (i.e., a click to the leading ear followed by a click to the lagging ear). Before the click breaks up into two distinct clicks, the identification response will be based on the roughness or complexity of the image. Finally in the region $\tau \geq$10 *msec*, where two distinct temporally separated clicks are heard, and the identification response will be based on subjective estimates of the time duration between the clicks. In this example, most errors will occur within one of the three τ segments, but not across their boundaries, because of the distinctive perceptual qualities that are associated with these ranges of τ. In general, in considering the potential IT that can be obtained from these two examples, it is the number of perceptually distinct segments, not the number of physical variables, that is important. If the number of perfectly identifiable items that can be achieved per segment is l and there are m segments, then the total number of perfectly identifiable items will be roughly $m \cdot l$. In other words, segmentation leads to an additive increase in the number of perfectly identifiable items.

A second important factor in considering the "richness" of a display is the number of physical parameters that can be varied within a single-component stimulus. Traditionally, researchers have tried to increase IT by independently varying the physical parameters of a single-component stimulus. For example, greater IT can be obtained by varying both the frequency and the amplitude of a vibration than by transmitting either variable alone. In this case, the overall IT will be determined not only by the IT per parameter and the number of parameters, but also by the perceptual interaction between the parameters. For instance, "tactual loudness" is dependent on both the frequency and the amplitude of a vibration (Verrillo *et al.*, 1969). A more extreme case occurs in binaural lateralization. Not only is lateralization influenced by interaural amplitude difference as well as interaural time delay, but it is exceedingly difficult to distinguish between the two: these two physical variables appear to give rise to one perception, i.e., lateralization. In general, however, when the perceptual interaction among parameters is small, as in the case of the

multidimensional auditory displays employed by Pollack & Ficks (1954), the potential IT will be roughly equal to the product of the IT per parameter and the number of parameters. In other words, independently varying multiple parameters of a single-component stimulus can potentially lead to a multiplicative increase in the number of perfectly identifiable items.

The third important factor in considering the "richness" of a display concerns the use of multi-component stimuli. For example, one could apply two simultaneous single-point stimulations to different locations on the surface of the body. Or, as we have done in our main experiments, one can make use of both multiple stimulation sites (the different fingers) and multiple frequency components (the different stimulating waveforms).

In this thesis research, we have attempted to make use of all three means for optimizing information transfer. We identified three perceptually distinctive frequency segments (denoted F_L, F_M, and F_H, respectively) within the single physical variable of stimulation frequency. In the 500-*msec* stimulus set, single-frequency waveforms were used with multiple number of frequencies in each of the three segments. Within the F_L segment, frequency and amplitude of the waveform were varied independently. Double- and triple-frequency waveforms were used by combining single-frequency elements from different frequency segments. Finally, these waveforms were applied simultaneously to the thumb, index, or middle fingers, or to all three fingers. However, we do not yet know how to characterize our stimulus sets in terms of "dimensionality".

V-3 Relationship between Motor Output and Motional Inputs

One issue we had looked into during the study of tactual reception of Morse code was the extent to which one's ability to receive motional input sequences is dependent upon one's experience in outputting similar motions. Ideally, we want to look at a variety of daily tasks that require skilled motor output (e.g., typing), construct an apparatus that can deliver similar motions to the hand (e.g., a reverse typewriter), recruit subjects who are skilled with the output task as well as those who are inexperienced, and study how well they can learn to receive the motional inputs. The difference in performance between the experienced and inexperienced subjects (as revealed, for example, in the learning curves) may be an indication of the influence motor output skills have on reception of motional stimulation.

In the study described in Appendix A, we studied how well subjects could receive motional stimulation delivered to the fingertip as up-down displacements. The movement patterns were designed according to the way a "straight keyer" is used to send Morse code. The signal was on whenever the fingertip was down. The timing and duration of the signal was determined by the Morse code for the letter being transmitted. Experienced Morse code operators and inexperienced subjects were trained to receive single letters, random letter sequences, words, and sentences in Morse code through such up-down finger motions. We found that experienced subjects generally performed better than inexperienced subjects. The biggest difference between the two subject groups was their ability to process words and sentences. The experienced subjects were very good at "chunking", a skill they developed via auditory reception of Morse code over many years. The inexperienced subjects, however, could barely receive words at rates above 10 *wpm* and could not receive sentences at all. It seems clear that the main differences between the experienced and inexperienced subjects were related to central processing abilities. Note that most Morse code operators are highly trained in both sending (manually) and receiving (auditorily) the code. Their ability to process continuous *input* streams of Morse code with audio tones obviously contributed to their ability to do the same with up-down finger motions. Thus, in this case, it was not possible to attribute the performance difference in the two subject groups directly to the experienced subjects' ability to *output* Morse code manually.

In future studies of this issue, we will need to account for a number of factors when considering the difference between subjects who are highly trained with certain motor output tasks and those who are not. The first factor is the nature of the code involved in such tasks. For example, radiotelegraphy and stenography require knowledge of special codes with which the average person is not familiar. Typing, however, uses a code with which many people are familiar. We should avoid using tasks that require special knowledge of codes, because it will be difficult to determine if a difference in performance between experienced and naive subjects is due to the difference in their ability to output motions, their knowledge of the code, or both.

The second factor is the readiness of an experienced subject's ability to transfer skills acquired through one modality to the other. For example, Morse code operators are highly skilled at

receiving the code auditorily. To the extent that they can learn to “hear” Morse code through the tactual channel, they may process tactual Morse code very differently than a naive subject would.

Finally, the issue of training must be considered. A person who is good at a motor output task has probably been trained extensively (for months and years). A laboratory subject cannot be expected to receive the same kind of training with an experimental device. Thus, we need a strategy to be able to compare performances among different subjects despite the different amount of training they have received.