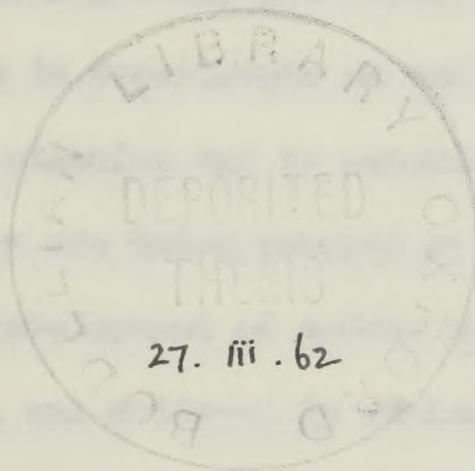


**A STUDY OF THE SENSORY THRESHOLD**

**Michel Treisman**



**ABSTRACT of a Thesis  
submitted to the  
University of Oxford  
for the degree of  
Doctor of Philosophy,  
1962.**

(Gobard, 1955) that sensitivity to an electric discharge is affected

ABSTRACT

This thesis is concerned with the problem of defining the processes by which a subject selects a response to a stimulus. Initially certain topics in two areas of research, sensory scaling and threshold measurement, are reviewed. In relation to the first the conflict between Fechner's law, which is derived from threshold measurements, and the more recent 'power law' given by 'direct methods' of scaling, is examined. It is shown that this conflict is due not to differences in the data from which the laws are derived, but to differences in the assumptions underlying the treatment of the data, and it is suggested that the logarithmic function may prove more useful for explaining a number of the findings in experiments on scaling, and that the same central effect of the stimulus may be considered to determine the response whether scales are being derived or thresholds measured. A short account of the development of models of the sensory discrimination process is then given, and evidence is reviewed which appears sufficient to justify applying the theory of signal-detection to the sensory threshold, and rejecting the neural quantum theory.

To investigate the threshold mechanism further it was decided to examine the effect of an accessory stimulus given in controlled temporal relation to a critical stimulus on the threshold for the latter. The first experiment was an attempt to confirm Motokawa's finding (Gebhard, 1953) that sensitivity to an electric phosphene is affected

by a preceding flash of light, reaching a maximum 1-3 seconds after the flash, the exact time depending on the colour of the light. The absolute threshold for the electric phosphene was measured by the method of limits at intervals of 1-9 seconds after a flash of blue or white light. Motokawa found that sensitivity was maximal when the interval between the stimuli was 2 seconds for the white light, or 3 seconds for blue light. Neither particular could be confirmed. There was no peak of sensitivity in the range studied, nor was there any marked difference in the curves relating threshold level to inter-stimulus interval for the two colours of light. Instead a monotonic relation between threshold and inter-stimulus interval was found, for this range, the threshold falling as the inter-stimulus interval decreased, slowly at long intervals and more rapidly at the shorter ones. This raised two questions: why could Motokawa's findings not be confirmed? and what was the nature of the effect which had been found?

The experimental technique of Motokawa and his colleagues is reviewed, and it is suggested that their findings may be an artefact of certain features of their procedure, in particular the use of large steps where a threshold change is not expected, and small steps where a threshold is expected to appear.

Two possibilities are considered for the threshold-lowering effect: it might depend on the use of two visual stimuli and reflect a peripheral interaction at the retina, or it might be a central effect,

not specifically visual. To investigate this the light flash was replaced, in the second experiment, by the ring of a bell of the same duration and given at the same inter-stimulus intervals. The same threshold-lowering effect of the accessory stimulus was found as in the first experiment, showing that the effect could be produced by an accessory stimulus in another modality. Further support for this conclusion was provided by Experiment 3, in which both the light-flash and the bell were used, the former at the same intervals as in Experiment 1, but the latter always preceding the critical stimulus by 1 second. Here the threshold did not vary significantly: there was little or no residual effect of the light when the bell was given at a shorter inter-stimulus interval. It appeared that it was the interval of time since the most recent accessory stimulus, whatever its modality, that mainly determined the degree to which the threshold fell.

In Experiment 4 the generality of the threshold-lowering effect was examined further. Using a visual accessory and auditory critical stimulus the possibility that the effect would occur only with visual critical stimuli was excluded.

These findings raised two questions: What processes underlie the change in threshold? What determines the relation between threshold level and inter-stimulus interval? Possible answers to the first question were suggested in terms of the analysis offered by signal detection theory. In this theory it is assumed that the decision whether

or not a stimulus has been administered is made by a process equivalent to comparing the afferent input with a criterion, and the criterion is computed in a way which may take account of the parameters, the variance ( $\sigma^2$ ) and mean ( $M$ ), of the 'noise' and 'signal + noise' distributions. The possibilities considered were that a change was produced in the functioning of the afferent paths at the time that the message from the critical stimulus was travelling along them, equivalent to a reduction in  $\sigma_N$  or  $M_N$ , with a consequent change in the value of the criterion,  $c$ , computed by the subject (a 'distribution effect'), or that the computation of  $c$  was affected directly (a 'criterion effect').

Two hypotheses were proposed in answer to the second question (the relation between threshold level and inter-stimulus interval):

(a) The accessory stimulus might have an 'arousing' or 'alerting' effect, causing an immediate central change with a small latency, which then decayed with time. (b) The subject might make use of the accessory stimulus as a 'temporal reference point' or 'warning'. By virtue of temporal information which he might possess he could use the accessory stimulus to determine when he could expect the critical stimulus to arrive. Here two subsidiary hypotheses might be suggested: (i) The subject lowers his threshold for the whole of the 'waiting time'. (ii) He may lower his threshold only when he expects the critical stimulus. Since his time-keeping ability is of limited precision, and the range of error will be greater for long times than for short, the

'range of expectation', if he allows for this, will be smaller the shorter the inter-stimulus interval. In either case, if the reduction in threshold is inversely related to the period for which it is reduced, a relation between threshold level and inter-stimulus interval of the sort found would be produced.

In Experiment 5 an attempt was made to test the 'warning' hypothesis directly by using a small range of randomly varied inter-stimulus intervals instead of a single fixed interval as previously, in order to see whether a fall in threshold during the 'range of expectation' could be demonstrated. Ranges of 0.5-1.5 seconds were used but no consistent effect was found. In Experiment 6 the same procedure was used but the inter-stimulus intervals were varied at random over a large range - 5 seconds - to prevent the accessory stimulus providing temporal information. Though it was found that the fall in threshold previously shown over this range disappeared, the possibility that this was due to 'habituation' of an 'arousal response' could not be excluded.

The exact relation between the fall in threshold and the inter-stimulus interval was next considered. It was suggested that a close and simple relation between the threshold and the 'just noticeable difference' for time discrimination would be difficult to reconcile with the arousal hypothesis, and would constitute evidence for the hypothesis of a 'range of expectation'. Studies of temporal discrimination were

briefly reviewed; it appeared that few results were in agreement, save that Weber's law had usually been found not to hold. Since no reliable measures of the temporal differential threshold were available two experiments were performed to measure the j.n.d. in a situation as similar to Experiments 1, 2, and 4 as possible. In each case Weber's law was found to apply for a range of intervals from 0.5 to 9 seconds. These experiments also provided new evidence on the causation of the classical 'indifference interval', and this was examined further in Experiment 9.

It was found that the fall in threshold could be described as an inverse function of either the j.n.d. or the inter-stimulus interval itself, with the latter giving a possibly somewhat better fit. This suggested that the 'range of expectation' might be a constant proportion of the inter-stimulus interval, though the inverse function found was not necessarily incompatible with an arousal explanation.

A threshold reduction consequent on a decrease in 'noise' (and 'signal + noise') variance should show a corresponding fall in threshold variance, whereas the variance might be expected to be constant if the threshold reduction were secondary to a change in  $M_N$  or directly due to a change in  $c$ . In Experiment 10 this was examined using the method of constant stimuli. No significant decrease in the threshold variance was found; it appeared to be constant and there was some evidence that rises might occur as the threshold fell. Explanations

were offered for this finding. It was also shown that the fall in threshold could not be attributed to an increase in 'guessing'.

The 'distribution' and 'criterion' hypotheses were tested further in Experiments 11, 12, and 14. On the 'distribution hypothesis' the threshold curve would be expected to show a minimum, with a rise in threshold at shorter intervals, due to the latency of any reflex effect on the afferent paths evoked by the accessory stimulus. In Experiment 11 inter-stimulus intervals ranging from 0.0-1.5 seconds were used, but the threshold fell continuously to the zero interval. Experiment 12 tested the possibility that this was an artefact due to temporal information or arousal provided or caused by the pre-warning; but despite the use of pre-warnings varied randomly over a range of 2-7 seconds the fall in threshold at the zero interval was confirmed. It remained possible that the conduction of part at least of the afferent message from the critical stimulus might have been sufficiently slow to be affected by a rapid response to the simultaneous accessory stimulus. This was excluded by Experiment 14, which showed that an accessory stimulus administered as much as 500 msec. after the critical stimulus could lower the threshold for the latter. Thus it appeared that a direct effect was exerted on the computation of the criterion, and that stimulus traces could be kept in a short-term peripheral store for 0.5-1.0 seconds, and could then benefit by a lowered decision criterion. Certain experiments in the literature bearing on the short-

term store were discussed.

The 'warning' and 'arousal' hypotheses were tested by an experiment for which the 'arousal' and 'waiting time' hypotheses on the one hand and the 'range of expectation' hypothesis on the other made different predictions. In Experiment 13 thresholds determined with fixed inter-stimulus intervals were compared with the thresholds given when the inter-stimulus intervals were randomly varied over a small range, 1 second. On the assumption that subjects lowered their thresholds for a range of time about the expected time of arrival of the stimulus, it was predicted that the threshold would fall as the longer intervals decreased, but would be relatively constant at the shorter intervals. This prediction, which was not made by the 'arousal' or 'waiting time' hypotheses, was confirmed.

Both arousal and warning hypotheses predict a reduced threshold-lowering effect of accessory stimuli which are themselves within the absolute threshold range. A review of some relevant experiments showed that such an effect had not been satisfactorily established. In Experiment 15 its occurrence was demonstrated and a model was proposed to explain certain features of these and other results. Changes in the intensity of supraliminal accessory stimuli would not affect their value as temporal reference points, but would seem likely to alter the degree of arousal caused by them. In Experiment 16 the threshold changes induced by a barely supraliminal and a moderately intense

accessory stimulus were compared. No significant differences in the falls in threshold produced by them could be shown, though significant differences in reaction time were found.

Finally it was assumed that the temporal reference hypothesis applied, and that the subject computed a criterion which would tend to keep the probability of his making false positive responses when the stimulus was expected at or below a minimal rate; this allowed the extent of the changes in frequency of response to a stimulus as the inter-stimulus interval was varied to be predicted. When these predictions were compared with the results of earlier experiments it was found that the agreement between them was in most cases satisfactorily close. The agreement appeared slightly better when it was assumed that the 'range of expectation' was a constant proportion of the inter-stimulus interval than when it was related to the differential threshold for duration.

It is concluded that the evidence appears to be sufficient to accept that the subject reduces his threshold about the expected time of arrival of the stimulus, for a period which is related to the inter-stimulus interval, and that the threshold reduction is produced by lowering the criterion applied to threshold judgments for that period. It also appears possible that the range of expectation may be a constant proportion of the inter-stimulus interval, and that the subject may reduce his criterion to an extent calculated to maintain the probability

of his making a false positive response to an input during the range of expectation at or below a limiting value.

A STUDY OF THE SENSORY RESPONSES

Michel Treisman

Department of Psychology  
University of Cambridge  
Downing Street, Cambridge  
England

CONTENTS

		Page
<u>Abstract</u> .....		1
<u>Chapter 1. Stimuli and responses</u> .....		5
1.1 Introduction.....		5
1.2 Sensory coding.....		7
1.3 The two psychophysical laws.....		8
<b>A STUDY OF THE SENSORY THRESHOLD</b> .....		13
1.5 Psychophysical judgment scales.....		19
1.6 The sensory threshold.....		22
<b>Michel Treisman</b>		
1.7 The theory of the neural quantum.....		25
1.8 Further consideration of the statistical decision hypothesis.....		34
<u>Chapter 2. Sensory stimulation and the phosphene threshold</u> .....		39
2.1 Introduction.....		39
2.2 The effect of a flash of light on the phosphene threshold.....		43
2.3 The effect of a ball on the phosphene threshold.....		48
2.4 Light and ball combined.....		
2.5 The status of the subject.....		
2.6 Conclusion.....		56
<u>Chapter 3. The history of the muscular threshold curves</u> .....		57
3.1 Introduction.....		57
3.2 The effect of visual sensory stimulation on an auditory threshold.....		62

Thesis submitted to the University of Oxford for the degree of Doctor of Philosophy, 1962.

CONTENTS

Page

Abstract..... 1

Chapter 1. Stimuli and responses ..... 1

    1.1 Introduction..... 1

    1.2 Sensory scaling..... 1

    1.3 The two psychophysical laws..... 5

    1.4 Two arguments for the power law..... 13

    1.5 Psychophysical judgment models..... 19

    1.6 The sensory threshold..... 32

    1.7 The theory of the neural quantum..... 45

    1.8 Further consideration of the  
        statistical decision hypothesis..... 54

Chapter 2. Accessory stimulation and the phosphene threshold 59

    2.1 Introduction..... 59

    2.2 The effect of a flash of light on the  
        phosphene threshold..... 63

    2.3. The effect of a bell on the  
        phosphene threshold..... 68

    2.4 Light and bell combined..... 72

    2.5 The status of the colour-sensitive effect 76

    2.6 Conclusion..... 86

Chapter 3. The Nature of the monotonic threshold change . 87

    3.1 Introduction..... 87

    3.2 The effect of visual accessory stimulation  
        on an auditory threshold..... 93

	Page
<u>Chapter 3 (contd)</u>	
3.3 The use of ranges of time intervals.....	96
3.4 Randomized time intervals.....	105
3.5 Conclusion.....	113
<u>Chapter 4. The discrimination of temporal duration.....</u>	115
4.1 Introduction.....	115
4.2 Reproduction of time intervals.....	117
4.3 Production of time intervals.....	127
4.4 A further experiment on temporal duration.....	136
4.5 The indifference interval.....	151
4.6 Temporal discrimination and threshold level.....	158
4.7 Conclusion.....	166
<u>Chapter 5. Threshold variance.....</u>	168
5.1 Introduction.....	168
5.2 Further consideration of the earlier experiments.....	180
5.3 Changes in threshold variance.....	182
<u>Chapter 6. Short Inter-stimulus intervals.....</u>	194
6.1 Introduction.....	194
6.2 Short inter-stimulus intervals.....	198
6.3 Randomized pre-warnings.....	203
<u>Chapter 7. Randomized Inter-stimulus intervals.....</u>	210
7.1 Introduction.....	210
7.2 Randomized inter-stimulus intervals.....	211

	Page
<u>Chapter 8. The effect of an accessory stimulus</u> <u>subsequent to the critical stimulus.....</u>	222
8.1 Introduction.....	222
8.2 Negative inter-stimulus intervals.....	226
8.3 The retroactive effect of accessory stimulation.....	241
8.4 Short-term information storage.....	257
<u>Chapter 9. Variation in the intensity of the accessory</u> <u>stimulus.....</u>	267
9.1 Introduction.....	267
9.2 'Intraliminal' accessory stimuli.....	271
9.3 Supraliminal accessory stimuli.....	287
<u>Chapter 10. Conclusion.....</u>	297
10.1 A further test of the warning hypothesis.....	297
10.2 Conclusion.....	308
<u>References.....</u>	313
<u>Acknowledgements.....</u>	330

CHAPTER 1

STIMULI AND RESPONSES

1.1. Introduction.

One of the most often used concepts in Psychology is that of a stimulus evoking a response, yet we know relatively little about the processes which intervene when this occurs. The work that follows is an attempt to investigate these processes in human subjects. It seemed that the best strategy might be to frame very general hypotheses about the intervening processes in simple situations, set subjects simple tasks in relation to stimuli, and see whether their performance confirmed or refuted the predictions of alternative hypotheses. Tasks of this sort have been used in sensory scaling and in the investigation of the threshold. This made it of interest to examine these fields briefly, in order to see both what hypotheses are commonly held, and which procedures can be considered more or less complex.

1.2. Sensory scaling.

Fechner wished to determine the relation of 'sensation' to the physical stimulus. Since it could not be measured directly he believed a measure of sensation must be inferred, and that this could be done by postulating that just noticeable increments in the physical stimulus corresponded to equal increments on a scale of the corresponding sensation. He applied this assumption to Weber's law and found that the magnitude of a sensation was then a logarithmic function of the physical stimulus. However, in the last thirty years this view has been strongly opposed.

Procedures for measuring sensation directly have been developed (Stevens, 1955, 1957), and their application has produced considerable evidence for a different psychophysical law: with these methods the magnitude of the sensation is given by a power function of stimulus intensity. These two opposing and apparently irreconcilable laws both have strong supporters. Stevens (1957, 1958, 1959a, 1959d, 1960a, 1960b, 1960c) believes it can be demonstrated that the 'just noticeable difference' is not constant and that this shows Fechner to be wrong. He argues that "in terms of their psychological magnitude, as measured by the scale we first set up by fractionation and magnitude estimation, the j.n.d.'s get larger as we go up the scale. Q.E.D., the j.n.d. is not constant in subjective size". This depends on the claim that the direct methods are valid measures of sensation, but Fechner's 'discriminability' scale is merely "the transforming of unreliability, inconstancy, or confusion into units of measure" (Stevens, 1957). Garner (1958, 1959) has defended the discriminability criterion. He argues that both scales have face validity, so that the conflict cannot be resolved by showing that they disagree. But the discriminability methods give reliable results while fractionation does not, therefore it is "the discriminability criterion . . . which is directly meaningful to normal observers".

So long as this major conflict in scaling theory is unresolved, it will be difficult to construct a model of the psychophysical judgment process that could appear valid for the whole field of sensory scaling.

Thus it is of some importance to attempt to clarify the nature of the dispute and see whether it can be resolved.

Both protagonists appear to agree that scaling is a process of describing the variation in an underlying 'sensory experience' or 'psychological magnitude', that the two types of procedure lead to two incompatible scales, and that they are disputing which is, in some sense, the 'correct' one. A possible source of confusion here may lie in their retention of the term 'sensation', or its synonyms, with the implication that subjects' reports must be interpreted as evidence about a unique, non-physical dimension. One alternative is to restrict the account to observables - stimuli and responses. The psychophysical function could then be regarded as defining an intervening variable, in the terminology of MacCorquodale & Meehl (1948). That is, the response would be a function of the psychophysical scale, which would itself be a function of the stimulus, in a mathematical description of the relation between stimulus and response; it would be no more than a device for facilitating this description, with no reference beyond it. This account could be amplified by introducing inferred intervening neurophysiological processes. It is then almost tautologous to suppose that there is some locus in the central nervous system at which the responses are selected, that the stimuli produce a central effect at this locus which varies along some dimension relevant to response selection, and that this variation depends mainly on the intensity of the stimulus. This

neurophysiological response-determining central effect is a hypothetical construct (MacCorquodale & Meehl, 1948), and the psychophysical scale becomes an inferred measure of it. In terms of such a model sensory scaling would be an attempt not to measure 'sensation' but to describe the transformations and operations on the stimulus input occurring in the nervous system when the stimulus is determining the selection of a response. The 'psychophysical law' forms part of such a description.

It follows that at least two such 'laws' would be required: the familiar function relating the magnitude of the central response-determining effect to the stimulus magnitude, and a second one relating it to the choice of response. This has not usually been made explicit; the latter has been treated by implication as simple and **direct** - a transition from 'Yes' to 'No' at some value of the central effect, or a linear transformation into 'magnitude estimates', etc. - and the former has been supposed to be complex - a power or logarithmic function. There is a mathematical argument (Luce, 1959a) which would limit the possible psychophysical laws to these two. Assuming certain plausible principles for the construction of mathematical theories, Luce shows that if the derived scales are to have interval or ratio properties (and are not log interval scales, which are not found in practice), and are functions of ratio scales of **physical** stimuli, then they must be themselves ratio scales, related by a power function, or interval scales, related by log or power functions.

One possible reason why both these functions occur might be that the 'direct' and 'indirect' procedures which produce them draw on different central effects of the stimulus to determine the response (Luce & Edwards, 1958). Another is that there might be a systematic difference in the mathematical assumptions made when scales are derived from the data provided by the different procedures.

### 1.3. The two psychophysical laws.

Fechner's law can be written as  $R = c \log (S + a)$ , and the power law as  $R = c (S + a)^n$ , where 'S' is the stimulus value, 'R' is the psychophysical scale (measuring the hypothetical neural effect), 'c' is a constant related to the unit chosen, and 'a' is an empirical constant which has been found necessary to ensure a good fit to data for both functions (Stevens, 1957; Gregory, 1956; Ekman, 1958, 1959; Stevens & Stevens, 1960). Since these laws do not specify explicitly how values of 'R' are to determine responses, there is an interesting equivalence between them: if we choose the appropriate 'response function' (i.e. the law relating the values of R to some feature of response) to supplement the 'stimulus function' (the psychophysical law), then both laws will fit the same relation between stimuli and responses. If a feature of response (e.g. meeting a criterion for the j.n.d., such as giving 75 per cent 'greater' responses) can be interpreted as indicating addition of a fixed interval on an assumed log function R scale, it can equally well be regarded as indicating

multiplication by an equivalent fixed ratio on a power function scale of the same variable. The two scales would afford alternative and equally valid descriptions of the same situation, just as a set of lights whose successive brightness differences are constant in log millilamberts could equally well be considered to be separated by equal ratios in millilamberts. This suggests that the production of different psychophysical laws by the 'direct' and 'discriminability' procedures might be due to their implicitly or explicitly assuming different 'response functions' in deriving scales from the data.

a) Discriminability scales.

In the case of Fechner's law the assumption is, of course, perfectly explicit. Fechner proposed that the just noticeable increment in the intensity of a sensation was constant. This means that when the feature of response which is taken as the criterion for discrimination is present, the difference,  $\Delta R$ , between the  $R$  values corresponding to the two just discriminable stimuli,  $S$  and  $(S + \Delta S)$ , has a certain constant value. This requirement (that a discriminative response will occur when  $\Delta R = k$ ) in effect defines the  $R$  scale for a given set of stimuli and responses. The  $R$  value corresponding to a stimulus is the number of j.n.d.s separating it from the absolute threshold, and  $R$  is given by the cumulated j.n.d. function. When Weber's law holds this response criterion makes  $R$  a log function of  $S$ . (Luce & Edwards, 1958). Had Fechner chosen to say that the response was determined by the

occurrence of a given ratio between the R values, i.e. when  $\Delta R/R = k$  (where k is a constant), this response function would have defined a different and equally valid R scale. When Weber's law holds R would have been a power function of S.

b) The direct methods: ratio production

In fractionation the subject is presented with a standard stimulus S, and is asked to adjust a variable stimulus to a value  $S_p$  which is to be subjectively a fraction p (e.g.  $\frac{1}{2}$ ) of the standard, S. The assumption made in deriving a power function from the data is that the subject performs this task in terms of the R scale; as Stevens & Davis (1938) put it, "the response of an observer who says 'this is half as great as that' is one which, for the purpose of erecting a subjective scale, can probably be accepted at its face value." These authors are assuming that the response is determined by the occurrence of a constant ratio,  $R_p/R = p$ , on the R scale, and they regard this not as an arbitrary choice of definition on their part, but as justified by the subjects' beliefs about what they are doing. This 'response function' may appear natural and necessary if subjects are believed to be producing valid introspective reports about their sensations. But if R is a variable in a mathematical description, or an inferred measure of a hypothetical neural process, the choice of this function is arbitrary.

It was known to Fechner (Stevens, 1957) that if the constant ratio requirement ( $\Delta R/R = k$ ) is applied to difference threshold data

which accord with Weber's law this will give a power function. Another interesting comparison between fractionation and difference threshold data follows from the algebraic method of deriving a power function from fractionation data proposed by Ekman (1958).

If a power function holds we can write  $R = c (S + a)^n \dots\dots\dots(1)$

and  $R_p = c (S_p + a)^n \dots\dots\dots(2)$

Dividing (2) by (1) (and applying the constant ratio requirement) we get

$\frac{(S_p + a)^n}{(S + a)^n} = \frac{R_p}{R} = p \dots\dots\dots(3)$

This gives  $\frac{(S_p + a)}{(S + a)} = k \dots\dots\dots(4)$

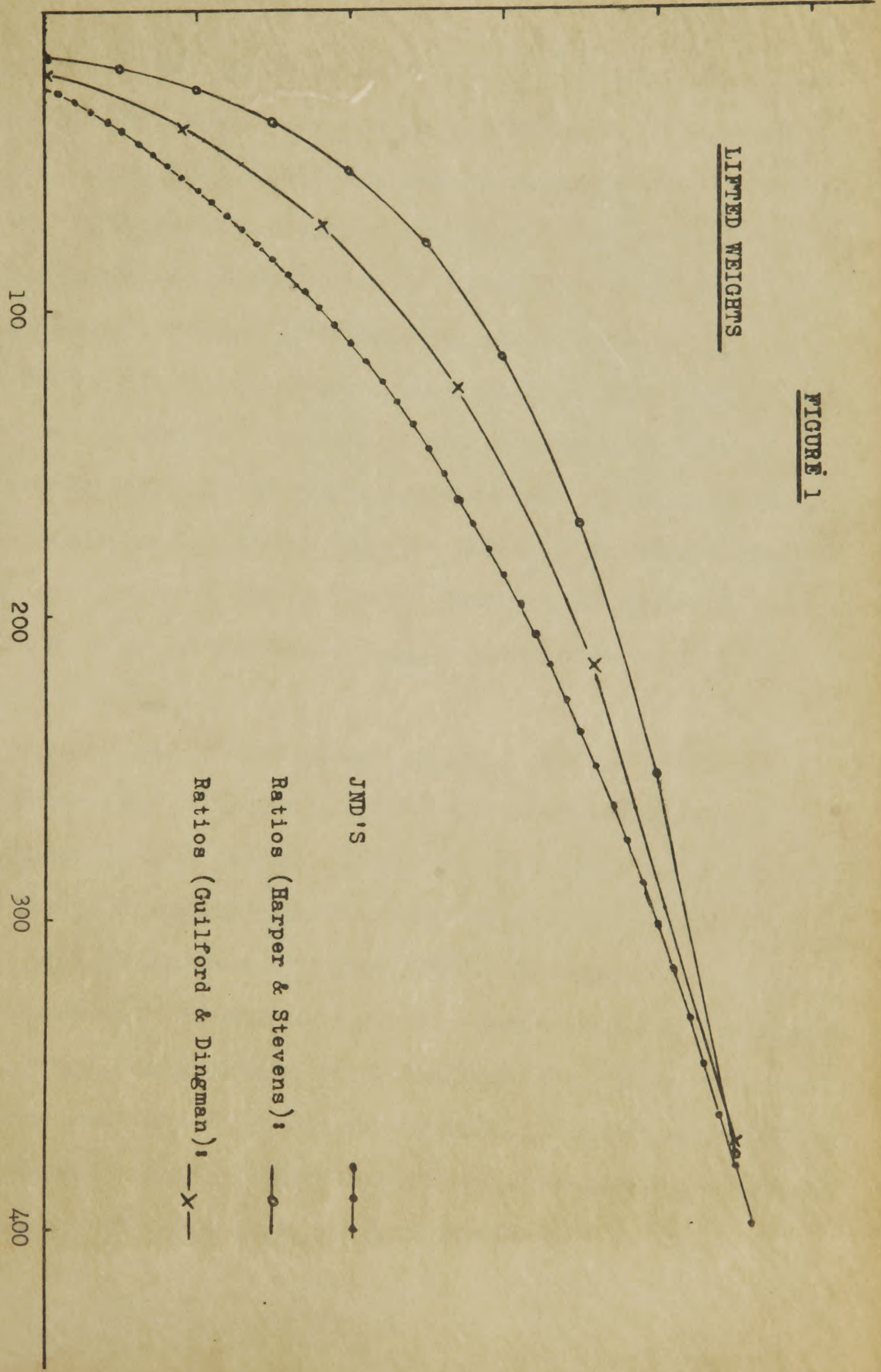
where  $k = p^{1/n} \dots\dots\dots(5)$

or  $S_p = kS + (k-1)a \dots\dots\dots(6)$

Therefore, as Ekman points out,  $S_p$  must be a linear function of  $S$  if we are to get a power function. If equation (4) or (6) holds for the data and we apply the constant ratio assumption this makes the  $R$  scale a power function. It is of interest that these equations have the same form as Weber's law. This suggests a very considerable similarity in the performance of subjects required to indicate stimuli 'just noticeably greater than' or 'twice as great as' a standard. Thus there is no evidence that the difference between the psychophysical laws reflects the utilisation of different central effects of the stimulus in the

LIFTED WEIGHTS

FIGURE 1



JND'S

Ratios (Harper & Stevens):

Ratios (Gullford & Dingman):

—●—

—○—

—X—

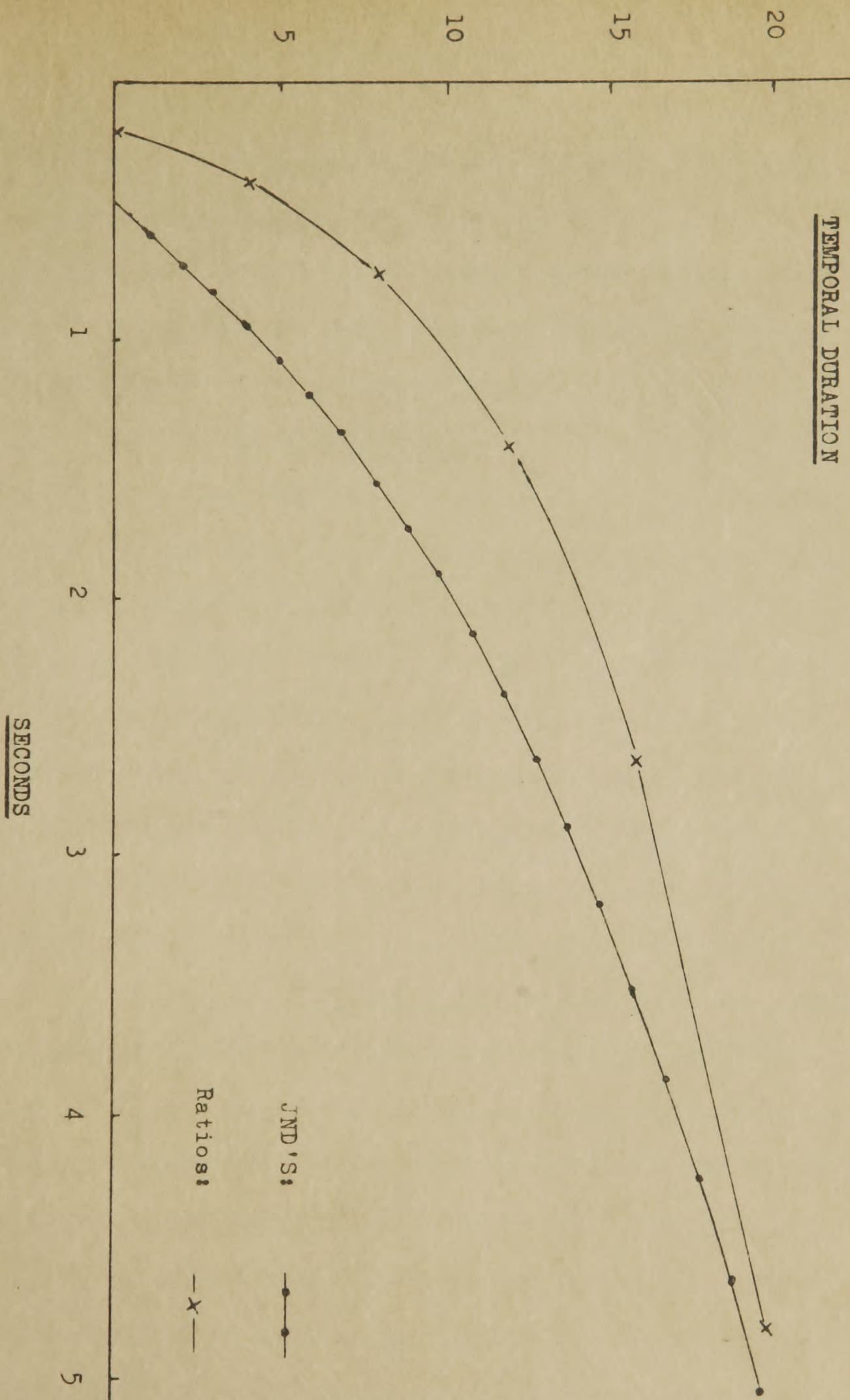
GRAMS

different procedures. On the contrary the disparity can be attributed to the failure to assume the same response criteria in deriving scales from the different data, so that the comparison usually made is invalid.

If fractionation data give a linear relation between  $S_p$  and  $S$  (as they should if a power law can be derived from them), then applying Fechner's constant difference requirement ( $R_p - R = \Delta R_f = k$ ) to them should give a log function (Luce & Edwards, 1958). This could then be compared with the log function that discrimination data for the same dimension would give if they fit Weber's law. If both procedures give the same value for the constant 'a', and  $\Delta R_f$  is given a suitable value in terms of j.n.d. units, then the functions should be the same. To illustrate this point scales of lifted weight and temporal duration are shown in Figures 1 and 2.

In Figure 1 three curves are given. One was prepared by cumulating the j.n.d.'s (probable errors) for the discrimination of lifted weights given by Oberlin (1936). The other two are cumulated ' $\Delta R_f$ ' curves derived from fractionation experiments by Harper & Stevens (1948) and Guilford & Dingman (1954). To prepare these curves the relation between the standards and the weights judged half as heavy (which was linear in each case) was plotted, the smallest half judgment was taken as zero, and the successive stimulus differences corresponding to an apparent doubling in weight were given equal value on the ordinate and cumulated. The value given to  $\Delta R_f$  for Stevens' and Harper's

FIGURE 2



results was 5 j.n.d.'s, and for Guilford & Dingman 9 j.n.d.'s. These values were chosen by eye; they do not necessarily give the best comparison. In Figure 2 the same procedure was followed, using half judgments of temporal duration given by Gregg (1951), and j.n.d.'s from Experiment 7. Here  $\Delta R_f = 4$  j.n.d.'s.

It is interesting to note how similar these scales are, despite their various origins. Similar agreement would be shown with loudness: in his review of loudness measurement Stevens (1955) concludes that loudness doubles for every 10 db, as scaled by direct procedures; this implies that if the constant difference assumption were used the loudness scale resulting would be linearly related to the decibel scale. Over most of the range the integrated j.n.d. scale (Riesz, 1928; Wever, 1949), and other scales of loudness based on the discriminability criterion (Garner, 1958), are approximately linear on a decibel scale.

If fractionation data are scaled using the constant difference assumption, the parameter  $n$  will still be present, though not as an exponent, and it will serve the same purpose as in the power function. Equation (4) shows that when the subject is asked to select stimuli representing a fraction  $p$  of the standard, he in fact selects a fraction  $k$ . We can use  $n$  to express the divergence between  $k$  and  $p$  ( $k = p^{1/n}$ ), or we can raise both sides of the equation to the power  $n$ , (giving equation (3)), and so produce a function of the stimulus scale

$(R = c(S + a)^n)$  in terms of which the variable stimulus is related to the standard by the fraction,  $p$ , asked for in the instructions.

Similarly, taking logs, equation (4) becomes

$$\log(S_p + a) = \log(S + a) + \log k \dots\dots\dots (7)$$

If we were to measure the stimulus on a log scale, the equivalent version of the fractionation instruction would be to ask the subject to add  $\log p$  to the value of the standard stimulus. In this case the divergence between  $\log p$  and the quantity actually added -  $\log k$  - would be expressed by  $n$  ( $\log k = \frac{\log p}{n}$ ). If we prefer to multiply through by  $n$  we get

$$n \log (S_p + a) = n \log (S + a) + \log p \dots\dots\dots (8)$$

i.e. we again have a function in terms of which the subject has followed the instructions.

It now seems that there is no essential conflict between the data produced by the two approaches; either scale could be derived from the data of either procedure. However, of the direct methods, only ratio production has so far been considered.

### c) Other direct methods

The procedures of absolute judgment (Garner & Hake, 1951) or category scaling (Stevens & Galanter, 1957) and equisection are grouped among the direct methods by Garner (1958) but are placed in a special group of 'partition scales' by Stevens (1960a). Category scaling might seem in some ways very similar to magnitude estimation:

in each case the subject applies verbal labels (usually numbers) to the presented stimuli. One difference is that in the former the number of categories to be used is prescribed by the experimenter, but in the latter the subject chooses his own responses. There is also considerable similarity between equisection and fractionation; the main difference here is that in the former the sensory range presented is bounded at both ends. However, on most perceptual continua both 'partition' methods, as conventionally scaled, give log functions or functions of intermediate form but not the power functions which are given by magnitude estimation or fractionation (Torgerson, 1960; Stevens & Galanter, 1957).

It is of interest that in both category scaling and magnitude estimation the same assumptions are made in deriving a scale from the data, namely that the instructions to the subjects determine the operations on the R scale which should be assumed, and that the numbers used as responses are a linear function of the magnitude of R. Again this would be plausible if subjects were giving valid reports on operations on their 'sensory experiences'. But if R is regarded as an inferred measure of a hypothetical construct, we could equally well vary the 'response functions' so as to get the same psychophysical scale from either type of procedure. Thus if in category scaling the number applied to a stimulus by a subject was supposed to approximate the log of its R value, the R scale would be a power function, while magnitude estimation would

give a log function R scale if the estimates were regarded as exponentially related to the R values.

Although it appears that the original conflict resolves into a trivial difference between alternative but equally valid mathematical transformations of the data, three problems remain: Why should category scaling and the 'direct' methods give different R scales for the same response criteria (or require different response transformations to give the same R scale)? This is found with most dimensions; Stevens (1957) calls these 'prothetic' and he labels the few dimensions on which magnitude scales and category scales both agree with the cumulated j.n.d. function 'metathetic'. Why does this difference occur? Finally, though neither scale is 'right' or 'wrong' are there any reasons to prefer to use one rather than the other?

Before trying to answer these there are two arguments for the 'validity' of the power function that must be considered. An argument which would seem to imply a log function (Gregory & Cane, 1955; Gregory, 1956) will be discussed later.

#### 1.4. Two arguments for the power law

##### a) Cross-modality matching

Subjects can be asked to equate the apparent strengths of stimuli on two different dimensions, such as brightness and loudness, to produce an equal-sensation function. Stevens (1959a, 1960a, 1960b,

1960c) has laid great weight on the results of a number of such experiments which, he believes, 'testify to the basic validity' of the power function scales. His argument is: if, given an appropriate choice of units, the laws for the two dimensions are  $R_1 = S_1^m$ , and  $R_2 = S_2^n$ , and the subject equates subjective values at various levels (i.e. does the equivalent of writing  $R_1 = R_2$ ), then the resulting function will have the form  $S_1^m = S_2^n$ . Taking logs we get

$$\log S_1 = \frac{n}{m} \log S_2 \quad \dots\dots\dots (9)$$

Therefore, in log-log coordinates the resulting function should be a straight line with a slope given by the ratio of the two exponents.

This prediction was tested by matching loudness and vibration (Stevens, 1959b), loudness and electric shock (Stevens, 1959b), loudness and brightness (Stevens, 1957), and a number of continua with force of handgrip (Stevens, Mack & Stevens, 1960). In all cases linear plots with slopes near those predicted were found. Stevens (1959a) concludes: "By predicting successfully from numerical estimations to cross-modality matchings, and from one kind of matching to another, we have in effect closed an important circle in a process of validation", though he also notes (Stevens, 1959b) that this success is a sufficient rather than a necessary condition for validating the subjective scales since it verifies only their relative forms. However, he considers that the similar procedure of ratio matching is more compelling and does verify that the subjective scales are power functions. This

procedure requires the subject to match apparent ratios on two continua; if a ratio  $r$  on one is always matched by a ratio  $r^n$  on the other he considers this shows that both psychological magnitudes are power functions of their respective stimuli. Data in agreement with this were obtained when ratios of loudness were matched to ratios of brightness (Stevens, 1957, 1958).

The direct matching results are evidence for the reliability of the methods used, but they do not validate the power function as compared with the log function, since the same prediction is given by the latter. In this case the laws for the two dimensions would be  $R_1 = m \log S_1$ , and  $R_2 = n \log S_2$ . If the subject matches these so that  $R_1 = R_2 + k$ , the resulting function will have the form  $\log S_1 = \frac{n}{m} \log S_2 + k/m$ , which gives the same prediction as equation (9). Stevens believes that his scales are ratio scales of subjective intensity, and it is true that if the power function is used  $R_1$  and  $R_2$  must be treated as ratio scales (i.e. we cannot write  $R_1 = R_2 + k$  where  $k$  is an undefined constant) if the matching prediction is to be given. This contrasts with the situation when the log function is used. In this case the prediction is given despite the presence of 'k', showing that here only interval properties are required of  $R_1$  and  $R_2$ .

The ratio matching data are also compatible with the use of the log function since they can be equally well interpreted as implying that a separation 'log r' on one scale is always matched with 'n log r'

on the other.

b) Luce's power law.

Luce (1959a, 1959b) has suggested that the problem of the psychophysical law might be solved by an axiomatic approach. He accepts Stevens' arguments for the 'empirical generality' of the power function, and finds that applying his axiom to Weber's law "leads to what appears to be the correct form" - a power law. It is necessary to examine his implication that by applying a novel approach to discriminability data he has produced independent evidence of the validity of the power function as a scale of "subjective sensation".

His argument deals with the probabilities ( $P$ ) of selecting an element ( $x$ ) from a finite set ( $T$ ) or subset ( $S$ ) of alternatives (stimuli or responses). He proposes as an axiom that if the probability of selecting  $x$  from  $x$  and  $y$  (where these are any elements of the set  $T$ ),  $P(x,y) \neq 0,1$ , then  $P_T(R)$  - the probability that an element selected from  $T$  will be in the subset  $R$  - is given by  $P_T(R) = P_S(R)P_T(S)$ , and derives consequences from this. He finds that if certain conditions hold the elements of a set can then be given values on a ratio scale, the  $v$  scale, defined by  $v(x) = k P_T(x)$ , where  $k$  is any positive constant. If his axiom applies one advantage of this scale is that  $P(x,y)$  is then given by  $P(x,y) = \frac{v(x)}{v(x) + v(y)}$ . To apply this to the problem of the form of the psychophysical scale he supposes that his axiom could apply to psychophysical data, that  $v$  is a continuous function of physical

intensity, and that Weber's law is true, and shows that then  $v(x) = A\sqrt{x} + C$ , where  $A$  is a constant,  $x$  is the S scale value of the stimulus, and  $C$  is equivalent to 'a' in equation (1).

There are two immediate difficulties to identifying this scale with Stevens' power function. Firstly, there is the variability of the exponent,  $B$ . Its value depends on the criterion used to define the j.n.d. Let us consider threshold data collected by the method of constant stimulus differences, with the subject responding "greater" or "less", and let us suppose that a normal ogive fits the resulting psychometric function. The threshold is commonly taken as the stimulus giving 75%, or 84% (one standard deviation), of "greater" responses. If the Weber fraction, (or rather Luce's  $c(\pi)$ ), is either 0.01, or 0.1, for the 84% criterion, it is easy to show that the value of the exponent  $B$  given by his equations will then vary from 158 to 227, in the first case, and from 16 to 26 in the second, as we vary the threshold criterion from 51% to 99.9%. Thus Luce's equations would generate not one but a family of functions from the same data, depending on the arbitrary choice of threshold criterion.

The second difficulty is also indicated by these figures: as Luce points out, his exponent is from one to two orders of magnitude bigger than those obtained by the direct methods, which are mostly less than 2 (Stevens, 1957, 1960a, 1960b). His comment is: "The exact meaning of this discrepancy is at the moment uncertain, and further

It is apparent that Luce's argument and justification lead to the same type of function because in each the same requirement is applied work, much of it empirical, will be needed to understand it" (Luce, 1959b).

However, the discrepancy appears to be a necessary consequence of the way the scale is derived. Luce's  $v$  is best considered as an intervening variable: it is a function of the stimulus, and is intended to facilitate the calculation of probabilities of selecting stimuli from stimulus sets when Weber's law holds and Luce's axiom applies. It is defined by Luce's requirement that  $P(x,y) = \frac{v(x)}{v(x) + v(y)}$ . When  $x$  and  $y$  are two stimuli separated by one j.n.d., then  $P(x,y)$ , the threshold criterion, is a constant, and consequently  $v(x)/v(y)$  must be constant. Since  $x/y$  is also constant, by Weber's law, the two scales must be related by a power function. This is strictly parallel to applying  $R_p/R = p$  to equation (4) to get a similar function. It is now apparent why the exponents are so different. If we suppose  $P(x,y) = 0.75$ , a common threshold criterion, then  $v(x)/v(y) = 3$ . If  $p$  in equation (5) is given this value, then since  $k$  will be a little over 1, by Weber's law,  $n$  must obviously be large. In contrast, a ratio production scale is designed to ensure that when the subject doubles the apparent strength of a stimulus its  $R$  scale value should increase by a factor of 2. Since the subject in such an experiment will make an adjustment considerably greater than one j.n.d., the exponent of the resulting power function could not possibly be as great as when the scale value must be approximately trebled for each increase of one j.n.d.

It is apparent that Luce's argument and fractionation lead to the same type of function because in each the same requirement is applied to data as similar as Weber's law and equation (4). The difference lies in the reasons for choosing this requirement; it is here that Luce makes his individual contribution. He considers a log function,  $u$ , but concludes that " $v$  is a much more basic scale than Fechner's. For example,  $v$  relates to the probabilities in a particularly simple way, . . . and it is a ratio scale, whereas  $u$  is only an interval scale."

### 1.5 Psychophysical judgment models

It is true that for the calculations Luce wishes to perform the power function is more convenient, and ratio scales have an advantage over interval scales in that more statistical procedures can validly be performed on them, though decibel scales have all the same been sometimes preferred to their ratio equivalents, in other fields. If these were the only considerations a choice between the two valid alternative descriptions might not seem to have much importance.

However, it is possible to use the complex input-output relations provided by performance in psychophysical judgment tasks as evidence about the sorts of operations applied to the input information. The subject acts as a computer of sorts, since a measure of the input reliably determines a related output. If we were investigating a computer of familiar type we would expect more errors and failures to occur when long or complex calculations are required than when they are short and simple. Some tasks might require a complex operation if a log function

measure is used but a simple transformation of a power function, or vice versa. If our computer performed well when one function, but not the other, required simple operations, and badly when complex transformations of the function were necessary, this would be good reason for describing it as employing that function. Applying this analogy to the human subject, we can examine the details of his performance to see whether the consequences of the underlying neural transformations best match the implications of taking the log function or the power function as the basis of our model of them, i.e. whether performance is in fact variable, prone to error, and subject to failure where they suppose operations that are difficult and involved, and more reliable and ready where simple operations are implied, whether either model needs fewer and simpler operations to provide descriptions of the performance (addition rather than multiplication, symmetrical rather than skew distributions, etc.), and whether the explanations the models suggest or require are general, or arbitrary and ad hoc.

The first problem that must be considered is why the power law exponents have the values found. It is an old view that sensory judgments reflect learning based on experience with stimulus intensities. This is made more precise in Warren's (1958) formulation: "The physical-correlate theory of sensory intensity considers that judgments of sensory

magnitudes are based upon experience with the manner in which sensory excitation is correlated with the amount of some physical attribute associated with the stimulus," e.g. experience of the relations between distance from the source on the one hand, and the luminance of a surface, or the loudness of a noise on the other. This does not perhaps go quite far enough since physical attributes are known to us through the sensory excitations they cause, so that we have a closed circle. However, we could suppose learning to have an important effect with dimensions such as visual length, and area, which in our early experience are frequently subject to additive operations. If the developing individual needs to be able to predict the extent and nature of changes in external stimuli, he will be most successful with such dimensions when he can convert the basic neural measure of the stimulus into a form on which he can perform analogous additive operations. Thus, to have rewarding interactions with the environment, he must come to think and act in terms of linear functions of visual length and similar dimensions. Other sensory excitations could be correlated with these.

This view is supported by the values of many of the exponents found. Thus visual length, visual area, temporal duration (Stevens, 1957) and, when density is constant (Warren & Warren, 1956), or cues to the size of the weights are eliminated (Joy, see Baker & Dudek, 1955), heaviness, all have exponents of approximately 1. The inverse square laws for light and sound predict exponents of 0.5 for loudness and

brightness, and values close to this are obtained (Warren & Warren, 1958; Warren, Sersen & Pores, 1958; Warren & Poulton, 1960). Experience of the effects of dilution on sapid substances suggests a linear function for taste: exponents found range from 0.8 to 1.3 (Stevens, 1960a; Warren, 1958). The concentration in air of an odorous substance diffusing freely should obey an inverse square law: the exponent for olfaction approximates 0.5 (Stevens, 1960a, 1960b; Reese & Stevens, 1960).

Though this hypothesis is compatible with both models, and is offered by Warren (1958) as a rationale for the power functions, Stevens (1960b) does not accept it, and prefers to suppose that these functions reveal the "dynamic operating characteristic, defined by the input-output relation" of the "sensory transducer". He argues that sensory learning is implausible when applied to an unfamiliar dimension such as electric shock to the fingers, and does not account for the difference between two similar dimensions such as warmth (exponent 1.6) and cold (exponent 1.0). He also objects that the exponent of the sone scale is 0.3, not 0.5, as found by Warren, Sersen, & Pores (1958). This is not conclusive, however, as confidence limits are not usually given with published exponents, there is considerable variation between the results of different experimenters (Stevens, 1955), and the direct methods have been shown to be subject to many biases (Garner, 1958; Warren & Warren, 1956; Warren, 1958; Warren, Sersen & Pores, 1958; Warren & Poulton, 1960). This sensitivity to details of the experimental

situation is clearly shown by two magnitude estimation brightness scales obtained by Stevens & Galanter (1957). With a  $5^\circ$  target the exponent was 0.36, but when a point source was used it was 0.47.

Unfortunately there is no systematic evidence on the effect of experience on psychophysical judgments (Hanes, 1949; Stevens, 1957; Warren & Warren, 1958) that might decide this issue.

a) The magnitude scales

To apply the log function model to the magnitude estimation procedure we must describe  $R$  as a log function of  $S$ , and suppose the subject uses an exponential transformation of  $R$  to determine the response to be made as the magnitude estimate. The alternative model is simpler:  $R$  is a power function of  $S$ , and the response is linearly related to  $R$ . The criterion of simplicity favours the latter description, though it could be argued against this that a description of this procedure ought to be relatively complex, since subjects seem to find it surprisingly difficult, giving variable results, and some subjects are unable to learn to do it (Stevens, 1956). One may note too that distributions of estimates are skewed toward the higher numbers, and Stevens prefers geometric to arithmetic means, as the former tend to coincide with the medians (Stevens, 1956; Reese & Stevens, 1960). This suggests that distributions might be symmetrical, and arithmetic means could be used, if a log function scale were employed.

Ratio production would be described as the addition of  $\log p$ ,

where  $p$  is the ratio, to the log function  $R$  scale value of the standard, or as multiplication of the power function value by  $p$ , the variable stimulus corresponding to the new  $R$  value then being selected.

b) The prothetic-metathetic distinction

Stevens (1957) distinguishes two types of continua. On 'prothetic' continua, such as loudness, the magnitude scale given by the direct methods is a power function, non-linearly related to the cumulated j.n.d. scale, so that the size of the j.n.d., in magnitude scale units, increases as one goes up the scale. On 'metathetic' continua, such as pitch, both scales are the same, so that the size of the j.n.d. is constant on the magnitude scale. The existence of these two types of continua is a problem for any theory of scaling.

So far only prothetic continua have been discussed. If the log function model is applied to these, the sensory learning hypothesis might explain why the subject comes to apply an exponential transformation to the  $R$  scale to produce magnitude estimates, or adds  $\frac{\log p}{n}$  to fractionate. If there were no environmental reward for some particular transformation of  $R$  (as when the dimension is not associated with any important independent physical attribute, nor subject to additive operations), we might assume that none would be learnt; asked to perform a task such as magnitude estimation or fractionation the subject would operate directly in the simplest way on the  $R$  scale (i.e., the function given by the cumulated j.n.d. scale, which would be approximately

logarithmic if Weber's law held), giving a linear function, or the requested proportion, of it. He would behave, in effect, as though j.n.d.'s were the units of his scale. This would then be a metathetic continuum, with all procedures producing linearly related scales. The same finding might be expected in scaling dimensions, such as attitudes or complex judgments, to which no physical attributes correspond simply. This prediction receives some confirmation from work by Ekman and Kuennapas (1960). They scaled the aesthetic value of handwriting specimens by two methods, and found a linear relation between a 'direct' scale calculated from ratio estimations, and a discriminability scale derived from paired comparisons.

If the power function model is employed the description of the prothetic continua is simpler: with the constant ratio response requirement both fractionation and j.n.d. data will give a power function, which can be made to have the same index. The discrepancy will now appear in the metathetic continua: to the extent that Weber's law holds the application of the constant ratio assumption to j.n.d.'s will give a power function, but when this assumption is applied to fractionation data (or when category scaling is employed) they will produce not this function but the cumulated j.n.d. curve (Stevens & Volkman, 1940; Stevens, 1959a). A power function could be recovered from these data but we would need to assume a relatively complex response requirement: that when the subject is asked to make a half-judgment he selects the

square root of the value of the standard on the power function R scale.

Expressed in terms of this model the sensory learning explanation would state that the problem is not the discrepancy on the metathetic continua but the occurrence of the single prothetic scale; it would have to describe the metathetic square root transformation as basic, and the prothetic linear response transformation as more complex and a result of learning. If we had only the power function model this account would seem arbitrary and implausible; it would be most unlikely to be suggested by the description this model gives.

Stevens (1960b) has treated the difference between the two groups of scales as indicating that "in at least some instances it seems evident that the existence of two basic kinds of physiological mechanisms is reflected in the behaviour of the psychological scales", one being additive, the other substitutive, at the neural level. There is little physiological evidence for this; and neural coding is probably considerably more complex than it suggests (Wever, 1949; Licklider, 1959). But even if it were the case, it would not explain why the two scale types should employ different response transformations.

### c) Partition scales

Stevens places category scaling and equisection in a different class from the other direct methods, as involving partition of the psychophysical continuum (Stevens, 1960a), although procedurally category scaling might seem in some ways more similar to magnitude estimation, and

equisection to fractionation, than they are to one another. However they both produce scales nonlinearly related to the magnitude scales (on prothetic continua), being either approximately a log function of  $S$ , or intermediate between the log and power functions (Stevens & Galanter, 1957; Torgerson, 1960). Category scaling is very sensitive to stimulus spacing, but though they varied this widely Stevens & Galanter could not produce the power function scale given by the direct methods. This is a difficulty for their assumption that the subject is attempting to partition a power function  $R$  scale into equal intervals. It leads Stevens (1960a) to conclude that "observers are so constituted that they are unable to partition a prothetic continuum without a systematic bias". He attributes this bias to variation in the subject's sensitivity to differences: "Near the lower end of the scale where discrimination is good the categories tend to be narrow . . . Near the upper end, where a given stimulus difference is less easy to detect, the categories broaden". He gives two related reasons for this effect: the subject fails to reproduce the power function "for the basic reason that his ability to tell one magnitude from another varies over the scale and affects the width of his categories. Since he can easily tell 0.5 sec. from 1.0 sec. he tends to put them in different categories; since he can only with difficulty tell 3.5 sec. from 4.0 sec. he tends to put them in the same category" (Stevens, 1957) or because "a given difference at the low end of the scale is more noticeable or impressive

than the same difference at the high end of the scale" (Stevens, 1959a).

One difficulty with this explanation is that no reason is given why the direct methods should not be affected in the same way by the greater impressiveness of the lower end of the scale. If it leads to the category labels attached to smaller stimuli being too large, why does it not bias estimates of their magnitude? If this impressiveness leads the subject to bisect the distance from 0.1 to 6 sones at 0.8 sones, it should prevent him selecting 3 sones as being half as loud as 6 sones.

The argument about the confusion of neighbouring stimuli depends on their being linearly spaced in the example, so that the greater stimuli are more closely spaced, on a discrimination scale, than the lesser. This makes the discrepancy a function of stimulus spacing; consequently the category scale should depart most from the power function when the smaller stimuli are well spaced and the greater ones bunched, and the disparity between the two scales should disappear when the stimuli are equally spaced in terms of a discrimination scale (i.e. when they are a constant number of j.n.d.s apart). But Stevens & Galanter (1957) note that the disparity is most marked (a log function is given) when the stimuli are logarithmically spaced, and it is reduced by bunching the stimuli at the upper end of the range.

An experiment by Torgerson (1960) throws further doubt on this explanation. He asked subjects to scale a series of grays in four ways: by magnitude estimation, and by category scaling, for 'lightness' and

for 'darkness'. Using the conventional methods of treating the data he obtained, for both 'lightness' and 'darkness', a power function from magnitude estimation, and an approximately logarithmic function from category scaling. In each case the category scale has the same relation to the magnitude scale. As he points out, to explain these results Stevens would have to suppose that the lower end of the scale is more impressive and shows better discrimination when lightness is being scaled, and the upper end when the subject is judging darkness.

The log function model suggests a simpler explanation. To get the power function of the stimulus favoured by sensory experience, on a prothetic continuum, magnitude estimation requires an exponential transformation of  $R$ , and fractionation the addition of a constant ( $\log p$ ). Though these may be difficult (Stevens, 1956), partitioning appears to present an even more complex requirement: the subject must select the series of decreasing intervals on  $R$  which when exponentially transformed will give equal intervals on the power function of the stimulus scale. If this task is too complex the subject might instead select equal intervals on  $R$  - this would give a log function, as customarily scaled - or he could fail to make the intervals decrease sufficiently, which would produce the intermediate functions found. On metathetic continua there would be no reason to depart from equal intervals on  $R$ , so that the category scale would be linearly related to the magnitude and j.n.d. scales, as is found.

This explanation is an application of the hypothesis of sensory

learning, and the assumption that operations appearing complex in terms of the model will be more difficult for the subject than those which seem simple. If stated in power function terms it would require us to accept that it is easier and simpler for the subject to select equal ratios on the power function and assign category labels as logs of R values, than to select equal intervals and use a linear transformation. Presented in this fashion it would appear implausible.

d) Directional scaling

Torgerson (1960) found that category scale (log function) values for 'lightness' were the reverse of those for the 'darkness' of the same stimuli (i.e., 'lightness value' =  $k - \text{'darkness value'}$ ) but values on the two magnitude estimation (power function) scales were related to one another as reciprocals. On the log function model this change is simply accounted for by supposing that when the subject changes from lightness to darkness he multiplies the R values by -1 before taking the exponential function or assigning category labels. The category labels are linearly related to equal intervals on the R scale. The alternative description is that the change is accomplished by taking reciprocals of all values on the 'lightness' power scale. Category labels would be the logs of intervals separated by constant ratios on this scale.

e) The discriminability of stimuli

The discriminability of stimuli (Crossman, 1955; Murdock, 1960), and the determination of adaptation levels (Helson, 1959), appear to be

simply expressed in terms of logarithmic functions.

It now seems plausible that a curve of the form given by cumulating j.n.d.'s may be the more useful basis for an account of sensory scaling. The differences between 'discriminability' and 'direct' scales appear to be mainly due to differences in the mathematical treatment of the data; no evidence of any major differences in the central effects of the stimuli in the different procedures has been found. On the contrary, the success of attempts to fit power functions by the direct methods suggests that ratio production data may fit equation (4) as well as threshold data usually fit Weber's law; the central effects involved are at least similar. It also seems that if j.n.d.'s are treated as equal units, or an equivalent procedure is applied to 'direct response' data, and the approximately logarithmic function resulting is used as a basis for describing the processes involved in making psychophysical judgments, the 'prothetic-metathetic' distinction can be plausibly interpreted as reflecting differences in the learning situations with which different dimensions are associated in the subjects' experience, and the incongruity between 'partition' and 'magnitude' scales can be explained by differences in the complexity of the operations they require. Though this account is formally equivalent to the power function model, the applications of the two which have been considered suggest that the log function model may prove the more useful, both in affording explanations of known findings, and suggesting new predictions.

This brief review of scaling theory has suggested that a similar central effect of the stimulus may determine response whether scales are being derived or thresholds measured. Since the demands on the subject made by the latter procedure appear the simpler, the mechanisms underlying the determination of the sensory threshold are next considered.

#### 1.6. The sensory threshold.

The traditional view of the threshold judgment was that it was a report by a subject about a sensation. Such reports might vary but this was not attributed to variability of the sensation. The early experiments had shown that as the intensity of a stimulus or the size of a stimulus difference was increased there was no sudden transition from failure to report it to reliable perception of it. Instead the probability of positive report increased in a way often approximated by the normal cumulative distribution function. This variability in response was interpreted as erratic introspection. Thus in Külpe's (1901) account, sensory discrimination is described as being "the introspection of different contents [of consciousness] and the report of their difference". The reports vary because "the many unnoticed fluctuations of the attention will lead to accidental errors of observation" of the sensation, and also because of expectation and habituation, described as "predispositions of consciousness", practice and fatigue. A further source of variability lay in the processes intervening between the stimulus and "the nervous excitations running parallel to sensation" since "the same stimulus may

produce different central excitations, different stimuli the same central effect".

A more useful account of response at the threshold had its genesis in the work of Thurstone (1927a, 1927b, 1927c, 1927d, 1928, 1948a). In his model of the psychophysical judgment the presentation of a stimulus to the subject resulted in the occurrence of a process, called by him the 'discriminal process', which was used by the organism to identify, distinguish or discriminate stimuli. These processes could be assumed to be psychic or physiological without prejudice to his argument. They could be ordered on a psychological continuum, i.e. on a dimension relevant to the determination of the response. Thus if a stimulus  $S_a$  produced a discriminational process  $R_a$ , and a stimulus  $S_b$  produced  $R_b$ , and these were compared, the subject would report that  $S_a$  was greater than  $S_b$  if the discriminational difference on the psychological continuum ( $R_a - R_b$ ) was positive, and the reverse if it was negative. Furthermore, if  $S_a$  was repeatedly presented to the subject  $R_a$  would not necessarily be evoked each time: there would be variation in the processes evoked, giving a 'discriminal distribution' corresponding to each stimulus. Thurstone further assumed that this distribution was normal on the psychological continuum. Thus in a series of comparative judgments of two stimuli, each stimulus would determine a normal discriminational distribution, centred on its modal discriminational process, and the result of each comparison would be determined by which of the two processes evoked on that occasion stood higher on the psychological continuum.

The normality assumption has some empirical support: Thurstone (1927a) found that scaling by the method of comparative judgments led, in most of his experiments, to consistent assignments of psychological distances when stimuli were scaled in alternative ways and normality was assumed, but not if other distribution assumptions were made; it is also very frequently found that when thresholds are determined the normal distribution function gives a good fit for the psychometric function (Guilford, 1954). Though Thurstone (1928) showed that if the threshold increases with stimulus strength, as in Weber's law, the psychometric function should be slightly positively skewed, rather than the symmetrical normal ogive, the difference is so slight for the small stimulus ranges covered by the psychometric function that it does not appear possible to demonstrate it reliably experimentally (Blackwell, 1952).

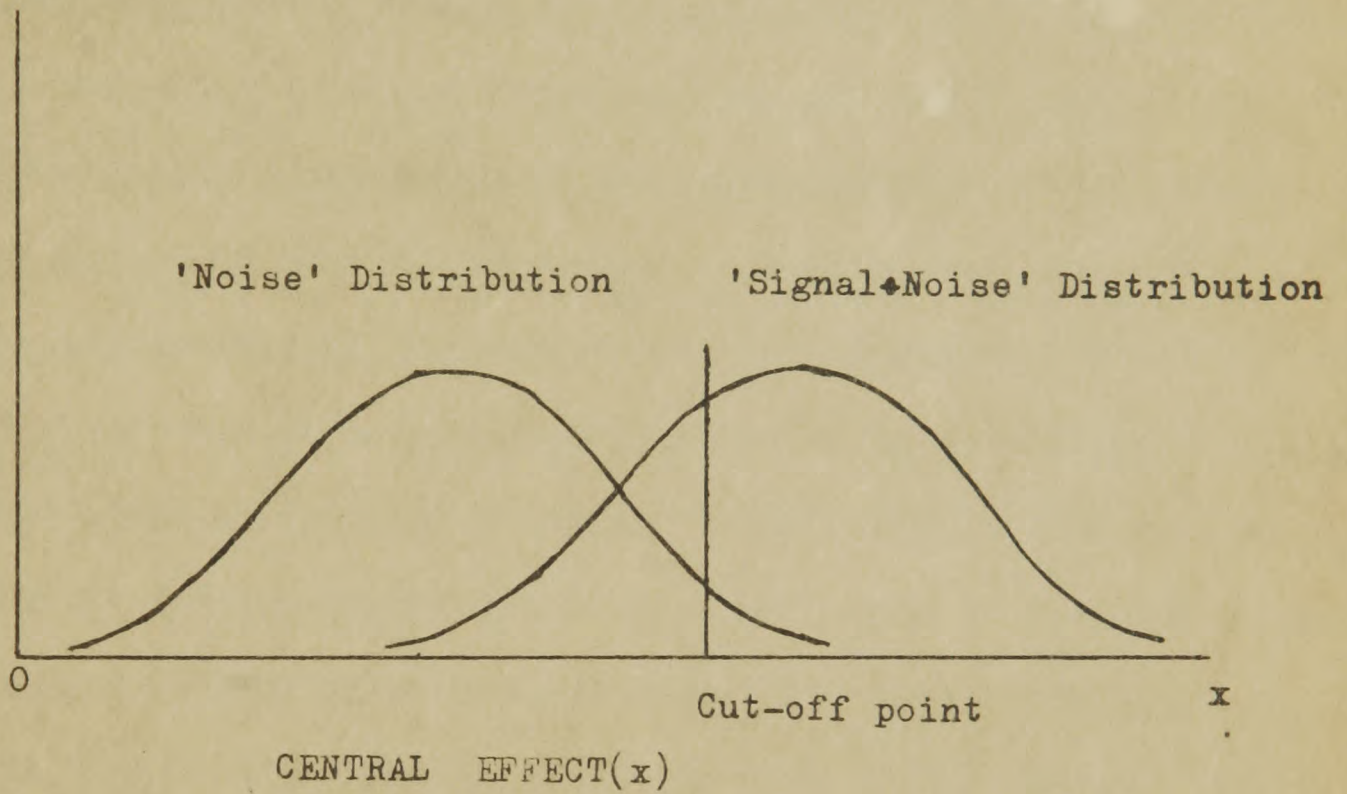
The virtue of Thurstone's approach was that he postulated a minimal number of intervening elements, and defined their relations to stimuli and responses, rather than merely naming them, so that the model could in principle be tested. The features of response followed directly from the postulated relations rather than requiring additional assumptions. He had provided the basis for a model in which the variability of response is attributed to variation in a response-determining central neural effect consequent on randomness in the afferent system (and lack of exact control of the stimulus).

Thurstone's interest was in scaling procedures, and his model

was developed for the method of pair comparisons; however it has been extended by a number of authors to apply to the determination of thresholds (Mueller, 1950; Smith & Wilson, 1953; Tanner & Swets, 1954; Cane, 1956). It applies almost without modification to the differential threshold. Here the subject is presented with two stimuli and asked to report whether one is greater or less than the other. Each is assumed to produce a central effect which may be taken to be normally distributed, and each judgment is determined by the sign of the discriminial difference on that occasion. If the central effects of the two stimuli are uncorrelated, the distribution of discriminial differences is normal with variance equal to the sum of the variances of the two stimuli, and mean equal to the difference between their mean effects. To measure an incremental threshold the subject is presented with a standard stimulus, to which an increment in the quality to be judged may have been added, and he is required to say whether an increment is present or not. The absolute threshold is measured in a similar fashion save that there is no standard stimulus. In these two cases the subject may give 'incorrect' responses, such as reporting the presence of an increment when none had been added. (The traditional explanation for such 'false positives' was of course that they were due to 'guessing': the subject's attention had wandered, he had not observed his sensation correctly, and had guessed instead.) Two further assumptions allow the model to be applied to these two cases. The first relates to the way in which the response is determined. If the

FIGURE 3

Probability  
Density



subject is presented with a stimulus which may be  $S$ , or  $(S + \Delta S)$ , and he has to decide which, the result may be considered to be determined not by a discriminial difference, as when a direct comparison is made between two stimuli, but by the position on the psychophysical continuum of the discriminial process evoked on each occasion. It is supposed that the subject selects a 'critical value' or 'cut-off point' (see Figure 3),  $c$ , on the continuum; when values of the central effect greater than  $c$  are evoked he reports that the increment is present, when the values are less he reports it absent. The rate of positive response, when  $(S + \Delta S)$  is presented, is then given by that proportion of the distribution centred on  $(R + \Delta R)$  which corresponds to values of the central effect greater than  $c$ , and the false positive rate is similarly given by the corresponding proportion of the distribution centred on  $R$ . The second assumption is that in the absence of a stimulus, random activity in the afferent system produces a variable central effect, or 'noise distribution', which can be considered to play the same role in discrimination at the absolute threshold as does the variable effect of the standard stimulus in the determination of the incremental threshold.

These ideas have been most fully developed in the 'decision-making theory' of the threshold (Tanner & Swets, 1954; Swets, Tanner & Birdsall, 1955; Tanner, Swets & Green, 1956; Licklider, 1959; Green, 1960) which is based on the 'theory of signal detectability' (Peterson & Birdsall, 1953; Birdsall, 1955). The latter applies the theory of

statistical decision to the problem of distinguishing signals from noise in communication channels. Its application to psychophysics is based on the assumption that a threshold discrimination can be considered as a problem of this sort. It supposes that the result of each observation made by the subject is a particular value of a variable,  $x$ , which can be considered as a measure of neural activity, and which is regarded as continuous. At each observation either 'noise alone' (N) or 'signal plus noise' (SN) is presented. Each produces a distribution of values of  $x$  which is described by a conditional probability density function  $f_N(x)$  or  $f_{SN}(x)$ . The greatest amount of information is then transmitted when the a posteriori probability of the signal consequent on each observation is reported. This is a function of the likelihood ratio,  $l(x)$ , i.e. the ratio of the probability density of  $x_i$  when SN is presented, to the probability density when N alone is presented. If the subject must report in two categories: 'signal present' (symbolised by 'A') or 'signal absent' ('CA') the theory shows that optimum performance is achieved by selecting a critical value of likelihood ratio and reporting the signal present for those observations which give a value of  $l(x)$  greater than this, and absent when the value is less. If  $x$  is monotonically related to  $l(x)$  the same performance is achieved by taking the value of  $x$  corresponding to the critical likelihood ratio as the critical value, and applying the corresponding decision rule. The probability of correct detections,  $P_{SN}(A)$ , is then given by the integral of  $f_{SN}(x)$  for values of  $x$  greater than the

critical value, and the probability of false positive responses,  $P_N(A)$ , by the corresponding integral for  $f_N(x)$ . For a signal of fixed intensity the relation of  $P_{SN}(A)$  to  $P_N(A)$  is determined by the value of  $l(x)$  chosen as criterion. The form of this relation can be exhibited by plotting  $P_{SN}(A)$  against  $P_N(A)$  for different values of the criterion; the resulting curve is known as the 'receiver operating characteristic' (R.O.C.).

Thus the theory shows that an optimal decision rule will depend on selection of a critical value of  $l(x)$ ; how an appropriate value for the critical likelihood ratio should be determined depends on the definition of 'optimum performance' used. A number of alternative definitions are considered and are shown to be similar in effect. The one most often applied in psychophysics requires that the total expected value of the performance is to be maximised, where the individual values are  $V_{SN.A}$  (the value of a detection),  $V_{N.CA}$  (the value of a correct rejection),  $K_{SN.CA}$  (the cost of a miss), and  $K_{N.A}$  (the cost of a false positive). For this definition the critical value of  $l(x)$ ,  $\beta$ , is given by

$$\beta = \frac{P(N)}{P(SN)} \frac{(V_{N.CA} + K_{N.A})}{(V_{SN.A} + K_{SN.CA})} \dots \dots \dots (10)$$

where  $P(SN)$  and  $P(N)$  are the a priori probabilities of SN and N.  $\beta$  will be the slope of the corresponding point on the R.O.C.

Of the different signal-in-noise cases which have been considered by the theory (Peterson & Birdsall, 1953), the one most often applied is the case of 'the signal known exactly' in Gaussian noise. Here it is assumed that both N and SN distributions are normal with equal

variances, so that  $\sigma_N = \sigma_{SN}$ . The standardised difference between the means of the two distributions,  $d'$ , is given by  $d' = (M_{SN} - M_N) / \sigma_N$ . As the intensity of the stimulus increases  $d'$  increases, and the shape of the R.O.C. alters accordingly.

A number of studies have provided evidence of some weight for the view of threshold discrimination as a statistical decision process.

Smith & Wilson (1953) found that observers could vary their thresholds if instructed to be 'liberal' or 'conservative'. There was a correlated change in false positive rate, but the threshold change could not be made to disappear by correcting for the false positives on the assumption that they represented 'guesses'. Tanner & Swets (1954), and Swets, Tanner & Birdsall (1955), using a visual incremental threshold, attempted to vary the critical values employed by their subjects by altering monetary rewards and costs assigned to the four possible outcomes. They succeeded in this and from the resulting R.O.C. curves were able to estimate, for each subject, the values of  $d'$  which corresponded to the different intensities of  $\Delta S$ . These values were consistent with independent estimates of  $d'$  obtained from a four-alternative forced choice procedure on the assumption that the subject selects the interval in which the greatest value of  $x$  occurs. The changes in threshold could not be accounted for by the correlated change in false positive rates. Swets (1959) found similar consistency between estimates of  $d'$  given by forced-choice and 'yes-no' procedures for a 1,000 c.p.s. tone presented

in white noise. Tanner, Swets & Green (1956) varied the a priori probability,  $P(SN)$ , of the signal, a 1,000 c.p.s. tone in white noise, and obtained R.O.C. curves, which, again, could not be accounted for by assuming that false positives were guesses. Though  $\beta$  varied as  $P(SN)$  was altered the change in its value was somewhat less than is predicted by equation (10) (Green, 1960).

If an observer is allowed to make a number of observations of a signal of fixed intensity in white noise, and uses the mean value in determining his response, then statistical decision theory would require his performance to be limited by the variance of the corresponding sampling distribution. Swets, Shipley, McKey, & Green (1958) confirmed this prediction for a 1,000 c.p.s. tone in white noise. As the number of observations made increased from 1 to 5,  $d'$  increased as the square root of this number, as expected.

It was shown by Hecht, Shlaer, & Pirenne (1942) that a major portion of the variability in the response to a flash of light near the absolute threshold could be attributed to quantum fluctuation in a light stimulus of constant nominal intensity, and Rose (1948) suggested that the differential sensitivity of the eye was limited by random fluctuation in the number of quanta absorbed from the stimulus. Barlow (1956) measured the absolute threshold for visual intensity by the method of constant stimuli. The subject used two criteria: the latter gave a considerably lower threshold, and the fall could not be accounted for by

the increase in the false positive rate (from 0 to 1%). Barlow showed that this result, and the results of Hecht, Shlaer, & Pirenne (1942) and Bouman & van der Velden (1947), were compatible with the hypothesis that the absolute threshold was limited by a low level of intrinsic retinal noise occurring even in complete darkness.

On the assumption that visual thresholds were "efficient statistical judgments of constant fallibility" (Barlow, 1957a), and that the incremental threshold was limited by statistical fluctuations in the number of quanta absorbed, he was able to predict the form of the curve relating incremental threshold to background intensity. Since quantal fluctuation follows the Poisson distribution, the assumption that the threshold increment corresponds to a constant multiple of the standard deviation of the variation in effect of the background illumination gives the prediction that the relation between threshold increment and background intensity will be given by  $\Delta I \propto I^{\frac{1}{2}}$ . This was confirmed, for stimuli of short duration and small area, with low background intensities, both for rods (Barlow, 1957a) and cones (Barlow, 1958).

When standard and comparison stimuli are presented simultaneously and the subject must report the presence or absence of an increment in the comparison stimulus, the response can be considered to be determined both by the variance of the 'distribution of discriminial differences' and by a choice of critical value. Gregory (1956) investigated differential sensitivity for light intensity in this fashion. A white screen was

divided into two areas,  $A_1$  and  $A_2$ , which were viewed simultaneously and for the same duration on each trial.  $A_2$  was illuminated with light of intensity  $I$  and  $A_1$  with  $(I + \Delta I)$ . He made the assumption that the background and test field could each be regarded as providing a set of independent observations approximately proportional in number to the corresponding areas. If the variance of each observation is taken as  $V$ , then the variance of the mean for test field or background is proportional to  $V/A_1$  or  $V/A_2$ , and the distribution of differences has a variance proportional to  $V(1/A_1 + 1/A_2)$ . This gave the prediction that  $\Delta I$ , the threshold increment, would be linearly related to  $\sqrt{(1/A_1 + 1/A_2)}$ , which Gregory was able to confirm.

Gregory & Cane (1955) and Gregory (1956) present a theoretical equation for this discrimination. They assume that the judgment is based on central effects of the stimulus intensities  $I$  and  $(I + \Delta I)$ , which they suggest may be mean neural impulse rates and which they represent by  $r$  and  $(r + \Delta r)$ , and that  $C$ , the 'critical value', is a constant, its constancy implying that "some fixed difference between impulse rates is required before discrimination is established, this difference being independent of the intensity  $I$ ". Then, when  $\Delta I$  meets the chosen threshold criterion (such as 80% positive responses),

$$\frac{\Delta(r + \Delta r) - \bar{r} - C}{\sqrt{\left(\frac{1}{A_1} + \frac{1}{A_2}\right) V}} = \text{const.} \quad \dots\dots\dots (11)$$

Marriott (1958) has criticised this equation on the grounds that it leads to implausible consequences: 1) Although, for a threshold criterion such as 80% of positive responses,  $\Delta I$  will decrease as the area of the stimulus increases, this would not be true for a 50% criterion. For this case the threshold would be independent of changes in area. 2) For criteria of less than 50%  $\Delta I$  would increase as the area increased. It might be added that though the threshold would be independent of changes in area for the 50% criterion, the false positive rate would decrease as the area increased.

These difficulties arise from the assumption that  $C$  is a constant and so independent of the variance of the distribution of differences. If it were supposed instead that the critical value was selected to give maximum detection consistent with a limiting rate of false positives, then  $C$  would not be constant but would be a function of the term  $\sqrt{(1/A_1 + 1/A_2)}$ , (which will be referred to as  $\sigma$ ).  $C$  can then be replaced by  $c\sigma$ , in which case the value of  $c$  determines the proportion of false positives that will be obtained. Writing the 'constant' in equation (11) as 'b', the equation for the threshold becomes:

$$(r + \Delta r) - r = (b + c)\sigma \quad \dots\dots\dots (12)$$

This form meets both Marriott's objections. As the distribution of  $(r + \Delta r) - r$  is assumed to be approximately normal,  $b$  is the normal deviate corresponding to the frequency of positive response required

by the threshold criterion. When this is 50%,  $b = 0$ , and when it is less than this  $b$  is negative; but as long as the threshold criterion exceeds the false positive rate ( $b + c$ ) will be positive. Thus for all criteria the threshold increment will be directly proportional to  $\sigma$ , and will be affected by change in area in the same way.

Gregory & Cane develop a further argument of some interest. On the assumption that  $r$  is logarithmically related to  $(I + k)$ , where  $k$  is supposed to represent sensory noise, they derive Ricco's and Piper's laws for areal summation at the absolute threshold from equation (11). Their argument follows equally well from equation (12), so that Marriott's (1958) objection, that it depended on an implausible basis, falls away. It is of importance, since if correct it would afford both a parsimonious derivation of the laws of areal summation and a cogent demonstration of the advantage of employing a logarithmic rather than a power function to define the psychological continuum. The argument is that if we take exponents in equation (12), ignoring  $A_2$ , which is the total visual field, and letting  $I = 0$  at the absolute threshold, we get:

$$1 + \Delta I/k = e^{(b + c)V^{\frac{1}{2}}A^{-\frac{1}{2}}} \dots\dots\dots (13)$$

Defining the unit of  $I$  as  $k = 1$ , and writing  $(b + c)V^{\frac{1}{2}} = d$ , we get:

$$\Delta I = d/A^{\frac{1}{2}} + d^2/2A + d^3/6A^{3/2} + \dots\dots\dots (14)$$

Gregory & Cane obtain Ricco's law by multiplying throughout by  $A$ , and Piper's law by multiplying by  $A^{\frac{1}{2}}$ . In each case a constant appears among

the terms on the right, and it is assumed that the other terms are relatively unimportant. However in the first case it is the second term which provides the constant, and in the second case the first term. This causes the following difficulty: the values of the terms on the right depend on the values of both  $d$  and  $A$ . While  $d$  and  $A$  can have values such that the first term is considerably larger than any other term, and decrease in the value of  $A$  will increase the size of the second term relative to the first term, it will cause the values of other terms to increase as well, and at no time will the second term be more than a fraction of the total value of  $\Delta I$  (for example, if  $d = 1$ , and  $A = 10$ , the first terms are 0.3, 0.05, 0.005 . . . . If  $A$  is made 0.1, the terms are 3.2, 5.0, 5.3 . . .). Thus Ricco's law does not follow simply from these assumptions; taken alone, they appear to exclude it, so that they hold no implications for the form of the psychophysical law.

Although these studies provide considerable evidence of the usefulness of applying the statistical theory an alternative view of the threshold must also be considered.

#### 1.7. The theory of the neural quantum.

In 1926 Boring suggested that if the physiological correlate of auditory intensity was assumed to be the number of fibres stimulated, as he preferred to believe, or, alternatively, the correlate of pitch was taken to be the specific fibre stimulated, this would imply a quantum theory of intensity, in the first case, or of pitch, in the second. He

believed that if the neural correlate varied in discrete jumps, each dependent on the stimulation of an additional fibre, so must the corresponding sensation and he proposed that this might show as 'critical points' in the response function, if the correct technique were used. The first evidence for this hypothesis was provided in 1930 by von Békésy (1960). He appears to have believed Fechner to have shown "that sensory magnitude can change only in quantum fashion", and he suggested that each such just noticeable difference for loudness might correspond to the stimulation of an extra cell on the basilar membrane. If no extra cell was excited there would be no change in sensory magnitude, if the response of the ear was increased by a whole cell, a just noticeable sensory quantum would be added. He supposed that a stimulus,  $S$ , would fully excite a certain number of cells, and might partially excite a further cell. The partially excited cell would not be able to produce an increase in sensory magnitude until fully excited. An increment,  $\Delta S$ , added to the stimulus would produce no perceptible change if it were less than the deficit in excitation of the partially excited cell; if however it was as big or exceeded this the extra cell would be fully excited and the subject would report a j.n.d. On this theory the psychometric function would be a step function, the change from 0 to 100% positive response taking place at a value of  $\Delta S$  equal to the deficit in the stimulation of the partially excited cell. In order to ensure that the effect of a given increment was constant on each

presentation, he gave the two tones,  $S$ , and  $(S + \Delta S)$ , in immediate succession and for only 0.3 seconds each. However, even with these precautions he did not obtain a step function, so he assumed that the excitatory effect of  $S$  varied widely from presentation to presentation, causing  $q$ , the degree of excitation of the partially excited cell, to take all possible values equally often. Then the probability that the subject would report the occurrence of a j.n.d. of loudness when  $\Delta S$  was added to the stimulus would be the probability that  $\Delta S \geq (Q - q)$ , where  $Q$  is the total excitation required to excite a single cell, and this would increase linearly from 0 to 1 as  $S$  increased from 0 to  $Q$ . If the subject were to adopt a two-quantum criterion and perceive no change until at least two additional cells had been excited then the probability of positive response would remain at 0 as  $\Delta S$  increased from 0 to  $Q$ , and would then increase linearly to 1 as the stimulus increment increased to  $2Q$ . As evidence that this occurred von Békésy produced linear functions of each sort, obtained from well practised subjects.

Stevens, Morgan & Volkmann (1941) produced more extensive evidence for the neural quantum theory. They amended it by making the quantum central rather than peripheral, and functional rather than anatomic. They felt the quantum could not be identified with the stimulation of an additional cell at the basilar membrane since there were more fibres in the auditory nerve than there were quanta in a sensory

attribute, the size of the quantum might vary from session to session, and results with binaural listening were inconsistent with the hypothesis of a peripheral quantum. If using two ears merely summated the probabilities of stimulating a fibre in one ear or the other, then with binaural presentation of the stimulus the size of the quantum should not alter, but the psychometric function would become curvilinear. What they found was that the size of the quantum decreased to about two-thirds of its monaural value, and the psychometric function remained rectilinear. Otherwise their theory is little different from that of von Békésy. They state that in order to demonstrate the neural quantum special experimental methods are necessary since any difficulty in the task, distraction, change of mood, or other disturbing factor may introduce extraneous variability which converts the psychometric function into an ogive. Their 'quantal method' consisted in presenting to the subject blocks of 25 increments, either of pitch or of loudness, all of them known by the subject to be of the same size. These were added to a continuous tone at intervals of 3 seconds, the increment being a 0.15 second increase in intensity, or a 0.3 second change in frequency. The subject responded by pressing a key whenever he noted a change in the steady tone. They assumed there would be large fluctuations in the subject's sensitivity so that the number of quanta stimulated by the steady tone would vary from moment to moment, and all values of  $q$  would be equally probable. To avoid the fluctuation in sensitivity causing

apparent changes whenever it increased or decreased the effect of the continuous tone by a quantum they supposed a well practised subject would use a two-quantum criterion. This led to the predictions that there would be a value of the stimulus increment,  $S_0$ , below which no positive responses would be given, and a value,  $S_1$ , at and above which only positive responses would be given. Between  $S_0$  and  $S_1$  the psychometric function would be rectilinear.  $S_0$  would correspond to one quantum, and  $S_1$  to two, so that the 'quantal index' (Corso, 1956),  $S_1/(S_1 - S_0)$ , would be integral and equal to the number of quanta constituting the subject's criterion.

Some features of the theory are worthy of comment. By supposing the quantum to be central and suggesting that several fibres might be involved in producing a quantal discrimination the hypothesis is made more general, since it can now be applied to any sensory discrimination, but loses some of the initial plausibility which was given to it by its original derivation from a discontinuous feature of nervous functioning, the all-or-none response of the individual cell (Boring, 1926). There is no reason to suppose that the central nervous system might not employ continuous variables, such as impulse rate, to determine responses. Some of the assumptions made appear implausible. It is supposed that the number of quanta excited by the steady tone varies greatly from moment to moment because of large fluctuations in the subject's sensitivity. As this is not due to variation in the size of

the quantum, which remains constant (though it may vary between sessions), it must presumably be attributed to variation in the central excitatory effect of the continuous stimulus. But it is also necessary to suppose that the central excitatory effect of an increment,  $\Delta S$ , of constant nominal intensity, is constant, undergoing no variation whatever (since otherwise the psychometric function would be an ogive). This is implausible when compared with the evidence that in the eye irreducible variability in the stimulus plays a large part in producing the response variation (Hecht, Schlaer, & Pirenne, 1942), or the evidence for random and spontaneous activity in the auditory and other afferent pathways (Galambos & Davis, 1944; Kuffler, FitzHugh & Barlow, 1957; Granit, 1955a). In effect the theory requires the central effects of  $S$  or  $(S + \Delta S)$  to be located on a scale on which there are a number of equidistant 'quantal points'. If two successively evoked values on this scale are separated by two or more 'quantal points' a positive response is given. Considered as a signal detection system this would appear likely to give considerably less than optimal performance, and to be unable to account for the results of many of the experiments suggested by statistical decision theory (Tanner & Swets, 1954; Swets, Tanner & Birdsall, 1955; Tanner, Swets & Green, 1956).

The first experimental evidence for the theory was provided by von Békésy (1960), who presented rectilinear functions for intensity discrimination but did not consider how good a fit the normal function

would have given for the same data. Stevens, Morgan, & Volkman (1941) present rectilinear functions for auditory frequency and intensity discrimination. They compared the goodness of fit of the rectilinear function and the normal ogive for 15 psychometric functions, using the chi-square test; combining the results for all the curves they obtained strong evidence for rectilinearity: the probability for the rectilinear functions was 0.931, and for the normal ogive 0.008. The results also support their other prediction: the intercept ratio ( $S_1/S_0$ ) was close to 2, ranging from 1.89 to 2.34 for the 15 sets of data. An interesting finding was that the subjective magnitude of the increments varied. Increments heard 80% of the time appeared larger than those heard 20% of the time. Thus the assumption of Boring and von Békésy that a 'sensory quantum' corresponded to the neural quantum could not be maintained. Using the 'quantal method' for the loudness incremental threshold Miller & Garner (1944) obtained rectilinear functions from two subjects (the experimenters) with intercept ratios close to 2. When they randomized the presentation of  $\Delta S$  they obtained a more variable psychometric function, to which they fitted three straight lines on the assumption that the variability was due to fluctuation in the subject's criterion between 1 and 3 quanta. Later experimenters do not appear to have produced such strong evidence for the quantum theory. Koester & Schoenfeld (1947) examined pitch sensitivity by the quantal method using the experimenters as well practised subjects. They found

neither the rectilinear functions nor the integral intercept ratios required by the theory. Corso (1951) obtained 70 psychometric functions from 20 subjects selected for their good initial pitch and loudness discrimination. Despite the use of the quantal method only 9 curves did not depart significantly from linearity on the chi-square test, and of these 9 only one had an intercept ratio with a value close to 2. Blackwell (1953a) measured differential brightness thresholds in four well practised subjects. Each value of the luminance increment was presented in a block of 20 repetitions. He found that the results, considered as a whole, did not confirm the predictions of the theory. Other studies reviewed by Corso (1956) do not appear to offer much further support. Meisser (1957) obtained 52 psychometric functions from 4 subjects. He measured the loudness threshold by the quantal method and concluded that the results were 'predominantly favorable' to the quantal hypothesis since the mean intercept ratio was 2.15. This was based on lines fitted by eye to curves of which he says "most of them are jumbles of points with no clear character".

Since the theory makes its predictions for optimal conditions only, positive evidence must be given much more weight than negative. This makes it important to assess carefully the validity of the positive evidence produced. Lewis & Burke (1949) discuss the use of the chi-square test made by Stevens, Morgan & Volkmann (1941). They consider 6 of the functions in detail and show that for these chi-square for the

phi-gamma hypothesis is seriously over-estimated because of two errors: theoretical proportions that were very large or very small were not combined with the neighbouring proportions before applying the test, and extreme proportions were omitted when calculating the theoretical curve but included when computing chi-square, an incorrect procedure. Because of these errors the normal function was rejected at the 3% level of confidence whereas if chi-square is correctly calculated it is acceptable at the 30% level of confidence. Blackwell (1952, 1953b) measured differential brightness thresholds by presenting blocks of 20 increments of a given magnitude and found an effect which he calls 'positive channelization'. This is a tendency for the frequency of 'Yes' responses to increase towards the end of a block in which the stimuli elicit more than 50% of positive responses. As he points out, positive and negative channelization would tend to produce a spurious increase in probabilities near unity and decrease in probabilities near zero, which would make an ogive tend to appear rectilinear, so that straight line functions might be expected as an artifact of the 'quantal method'. Neisser (1957) has demonstrated that response dependencies are obtained when the quantal method is applied to auditory intensity discrimination.

Though the neural quantum theory of the threshold cannot be disproved by negative experimental results, in view of the lack of any very compelling positive support for it, the possibly artifactual

nature of the rectilinear curves obtained, the implausibility of some of the assumptions required by it, and its inability to account for the results of the experiments suggested by statistical decision theory, it will not be considered further.

#### 1.8. Further consideration of the statistical decision hypothesis.

The statistical decision theory of the threshold can be considered as a conceptual model not of the physiological processes occurring when a stimulus determines a response but of the set of logical and mathematical operations - determination of the sign of a discriminial difference, application of a decision rule, etc. - which these processes must approximate for the subject to achieve the best possible performance. Confirmation of predictions made by the theory serves to show only that the functioning of the discrimination mechanism is such as to make its input-output relations tend towards the theoretical optimum. The explanation for this might be that operations whose effect is equivalent to the computations prescribed by the theory most often prove advantageous to the organism and come to be retained.

If it is accepted that the pattern of performance shown in sensory discrimination tasks is what would be expected to follow from the application of the rules of statistical decision theory, this raises two main problems. The first is the definition of the operations which we may suppose subjects actually to employ to approximate those contained in the theory. This requires us to discover what measures, both of his

own performance and of the external situation, the subject takes, and how he applies them to the determination of his responses. For example, the theory prescribes the choice of a value of likelihood ratio as the critical value. To apply a criterion of this sort the distributions of both  $N$  and  $SN$  must be known. It might be simpler for the subject to select instead a value of  $x$ , the measure of the central effect of the stimulus, as the criterion,  $c$ , and apply the decision rule to this, choosing  $c$  to give a satisfactorily low rate of false positives over a period of time. This would be equivalent to choosing a value of likelihood ratio, since  $x$  is monotonically related to  $l(x)$ , and it could be applied in the many instances in which features of the stimulus must be known to the subject only very inexactly. It would correspond to the adoption of the Neyman-Pearson criterion by the subject (Neyman & Pearson, 1933; Birdsall, 1955) in place of the 'expected value' criterion usually assumed (Licklider, 1959). Equation (10) specifies an ideal relation between  $\beta$  and  $P(SN)$ . To apply this relation the subject must know the a priori probability of the signal, compute  $\beta$  as required by the equation, and employ  $l(x) = \beta$  as the critical value, either by calculating  $l(x)$  for each value of  $x$ , or by finding the value of  $x$ ,  $c$ , for which  $l(c) = \beta$ , and applying the appropriate decision rule. Here again he could operate more simply by taking as his measure the response on one trial only, and varying the value of  $c$  after each trial, reducing  $c$  if the response was positive, increasing it after

a negative response. This would have the effect of altering the critical value in a direction determined by  $P(SN)$  and so approximating the effect of computing equation (10). The closeness of the approximation would depend on the function used to relate the previous response to the size of the adjustment applied to  $c$ . Thus the occurrence of response dependencies could be understood as the expression of an operation by which a requirement for optimal decision making is approximated (Howarth & Bulmer, 1956).

Even when the operations the subject may be actually supposed to perform have been specified, we do not know how these 'computations' have been carried out, by what 'digital' or 'analogue' methods. This is related to the second main problem, that of giving physiological and anatomical identity to the states and processes postulated in abstract form by the model. One element for which suggestions of some plausibility have been offered is the 'noise' which obscures the signal. Some variability must always be present in the production of the physical stimulus; however variability in the response of the afferent pathways may also play an important role in limiting sensitivity.

Fechner believed that the 'Augenschwarz' constituted the limit to visual sensitivity at the absolute threshold (Guilford, 1954; Barlow, 1957<sup>a</sup>), and the concept of variability of nervous response to a constant stimulus is included in the account of the threshold given by Külpe (1901). An implication of the all-or-none nature of the nerve

impulse is that randomness in the performance of the afferent pathways would manifest as a variable spontaneous discharge in the absence of a stimulus, and variability of the response to a stimulus. The occurrence of an apparently spontaneous discharge has been described in afferent pathways from the retina, muscle spindles, temperature receptors, olfactory bulb, vestibular organ, vibrissae (Granit, 1955a), and ear (Galambos & Davis, 1944). In an analysis of the spontaneous discharge recorded from fibres in the optic nerve of the decerebrate cat in the dark, Kuffler, FitzHugh & Barlow (1957) showed that the intervals between successive impulses varied in an apparently random fashion over a wide range. A gamma distribution fitted the distribution of intervals. Buller, Nicholls & Ström (1953) examined the low frequency discharge found in the afferent fibre from the frog's muscle spindle when it is held at a low tension. They found that successive intervals appeared to constitute a random sequence, giving an approximately exponential distribution. FitzHugh (1957) has described the variability in response in the optic nerve fibre in the cat to repeated flashes of the same intensity, though this should be attributed at least partly to quantum fluctuation in the stimulus (Pirenne, 1956).

Possible sources of the sensory noise are thermal agitation acting at the receptors or at synaptic junctions in the afferent pathways, and the arrival at afferent synapses of impulses originating outside the afferent system. Barlow (1956) has suggested that most of

the noise occurring in the visual pathways at the absolute threshold may be attributable to thermal decomposition of molecules of visual pigment in the receptors, and Buller, Nicholls & Ström (1953) showed that the random discharge in the afferent fibre from the frog's muscle spindle could be attributed to fluctuations of potential in the terminal nerve membrane resulting from thermal agitation of ions.

To investigate the threshold mechanism further it appeared that it might be useful to have a reliable technique for varying the subject's performance in sensory discrimination tasks. One procedure which appeared potentially useful for this was the employment of accessory stimulation to modify sensitivity to a stimulus at threshold level. Though considerable work has been done on the interaction of stimuli, it has been mainly concerned with demonstrating that stimulation of one modality may affect performance in another, and the experiments have not usually been in a form allowing the results to be related directly to theories of the threshold mechanisms. Explanations have usually been couched in terms of hypothetical processes such as 'irradiation' or 'energy flow' which are difficult to relate to the known facts of physiology and vague in their implications (Ryan, 1940; Gilbert, 1941; Harris, 1950; London, 1954; Symons, 1954). It seemed that a useful starting-point might be to attempt to analyse the relation between some feature of the accessory stimulation, such as its temporal relation to the threshold stimulus, and the changes induced in the subject's sensory discrimination.

CHAPTER 2

## ACCESSORY STIMULATION AND THE PHOSPHENE THRESHOLD

2.1. Introduction.

Although there has been considerable work on interactions between stimuli (Ryan, 1940; Gilbert, 1941; Harris, 1950; London, 1954) the consequences of variation in the time interval between the accessory and critical stimuli has not often been studied systematically. In an investigation of attention, Newhall (1923) included two experiments of interest in this regard. In the first he attempted to vary the degree of 'expectant attention' by using situations "which made the observer aware of the time at which the stimulus would occur with three different degrees of completeness". The critical stimulus was an approximately liminal pressure to the pulp of the finger of constant magnitude, randomly alternated with an equal number of blanks, and the three conditions were: 1) A constant 4 second interval separated the stimuli; 2) A lever was visible through a slot in a screen 1.7 seconds before each stimulus was administered; 3) The lever could be observed moving through a path which took it behind a screen at the moment the stimulus was administered. The percentages of positive response for these three conditions (combining the results of his two subjects), were respectively 52, 76, and 86 per cent, and the corresponding false positive rates were 8, 7, and 6 per cent. In a similar experiment on a visual incremental threshold the stimulus was a 0.1 second increase in the brightness

of an area 8 mm. square, and the three conditions were: 1) A click preceded the stimulus by 9 seconds; 2) A train of 5 clicks at one second intervals ended 4 seconds before the stimulus came; 3) The train ended one second before the stimulus. The corresponding percentages of positive response (combining the results for his four subjects) were 48, 51, and 63 per cent, with false positive rates of 5, 5, and 4 per cent. These results are of interest for two reasons: an apparent threshold change is shown without concurrent change in the false positive rate, and the differences between the conditions associated with the apparent variations in threshold suggest that the temporal factor may be of importance.

Child & Wendt (1938) studied the effect of a 0.1 second flash of light on the absolute threshold for auditory intensity. The auditory stimulus was a 1,000 c.p.s. tone presented for 165 msec. at one of five intensities approximately at threshold level. The flash preceded the tone by 2, 1, 0.5, or 0 seconds, or followed it by 0.5 seconds (the '-0.5 seconds' condition), or was omitted. The order of intervals and intensity of stimuli were randomized. The combined results for their 11 subjects show the lowest percentage of positive response for the 'tone alone' condition and the highest when the flash precedes the tone by 0.5 seconds. There was a clearly significant rise in response for the 0, 0.5, and 1 second intervals, a "questionable" rise at -0.5 seconds, and little effect at 2 seconds. With the light the

average false positive rate was 3.3 percent, with the tone alone 1.8 per cent. Thus they appeared to show a facilitating effect of the flash for which 0.5 seconds was the optimal interval. They thought this to be most probably due to central summation of excitatory processes, but also considered other possibilities, such as the relaxation of a hypothetical tonic contraction of the intratympanic muscles in response to the visual stimulus, or "complex phenomena of attention and judgment", such as an increase in alertness.

In her review of this and other studies on intersensory facilitation<sup>t</sup> Gilbert (1941) concludes that a short heteromodal stimulus momentarily reduces sensitivity in another modality, and then increases it after an optimal interval of about 0.5 seconds. Perhaps the most extensive investigation of the effect of the interval between two stimuli on the threshold for the second, and one which to some extent conforms with Gilbert's conclusion for heteromodal stimulation, has been that carried out on the threshold for the electric phosphene by Motokawa and his colleagues (Motokawa, 1949a, 1949b, 1949c, 1949d, 1950, 1951a, 1951b; Motokawa & Iwama, 1949a; Gebhard, 1953), though here the extra stimulus is also visual. In these experiments the absolute threshold for the phosphene consequent on the passage of a direct current through the eye is measured when administration of the current is regularly preceded by a flash of light. They found that the pre-illumination caused a brief depression of sensitivity, which then increased to a maximum before

returning to the resting level. The interval after the end of the pre-illumination at which the sensitivity was maximal (the "crest time") was independent of the intensity of the pre-illuminating flash, and was a function of its wave-length. The duration of the flash had no effect for durations between 0.5 and 10 seconds; increase in its intensity increased the enhancement of sensitivity shown at the crest time. When the pre-illumination field was red the crest time was 1 second, for yellow it was 1.5 seconds, white or green gave 2 seconds and blue 3 seconds.

As well as being of importance in relation to the problems of colour vision, these findings apparently demonstrate an interesting temporal relation between an accessory stimulus and a threshold. It seemed that it might be useful to attempt to clarify how the accessory stimulus caused the threshold to vary. A tentative hypothesis might be that the extra stimulus reduced sensory noise arising at the retina, and so allowed the critical value to be lowered without an increase in false positive responses. It might do this by a direct effect of the light flash on retinal elements, or by a centrally mediated effect. Motokawa believes the flash acts peripherally, since if one eye is pre-illuminated no facilitation is shown by the other (Motokawa, 1949d). However this would not exclude a centrally mediated effect elicited by the flash and acting on the same retina, perhaps through centrifugal fibres (Granit, 1955b; Dodt, 1956).

A necessary preliminary to testing this hypothesis is the repetition of Motokawa's observations.

## 2.2. The effect of a flash of light on the phosphene threshold.

In the technique used by Motokawa (Gebhard, 1953; Motokawa, 1949a, 1949b, 1949c, 1949d, 1951b) the subject is usually dark-adapted, though weak light adaptation was sometimes used (Motokawa, 1949c), and the stimulus is a constant current, 100 msec. square pulse administered by means of two electrodes, which are usually placed on glabella and occiput. To determine the threshold a modified form of the 'method of limits' is used. The applied voltage is decreased step by step (descending series, only, are used) until discrimination becomes difficult, when the subject is allowed to compare the stimulus with a blank. To investigate the effect of a previous light-flash on this threshold a patch of opaque glass, usually about  $2^{\circ}$  in size, is illuminated for 0.5 to 10 seconds, at an intensity of up to 1470 lux, before each administration of the stimulus. For each threshold determination the interval between the end of the pre-illumination and the stimulus is kept constant. Well-trained subjects were used.

### Experiment 1

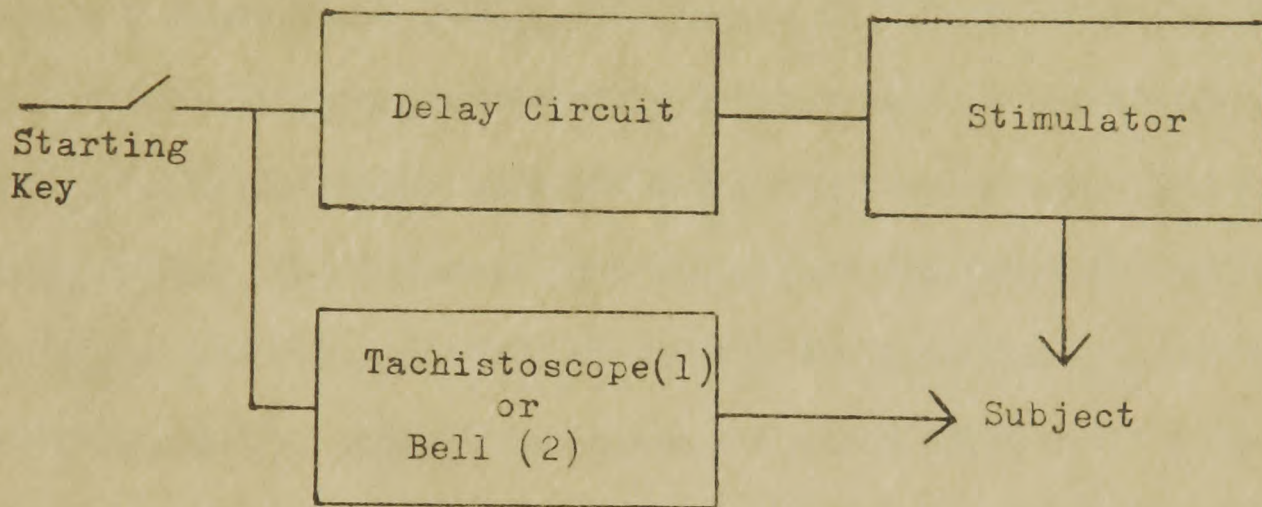
This experiment was an attempt to reproduce the main features of Motokawa's experiments.

### Apparatus and Procedure

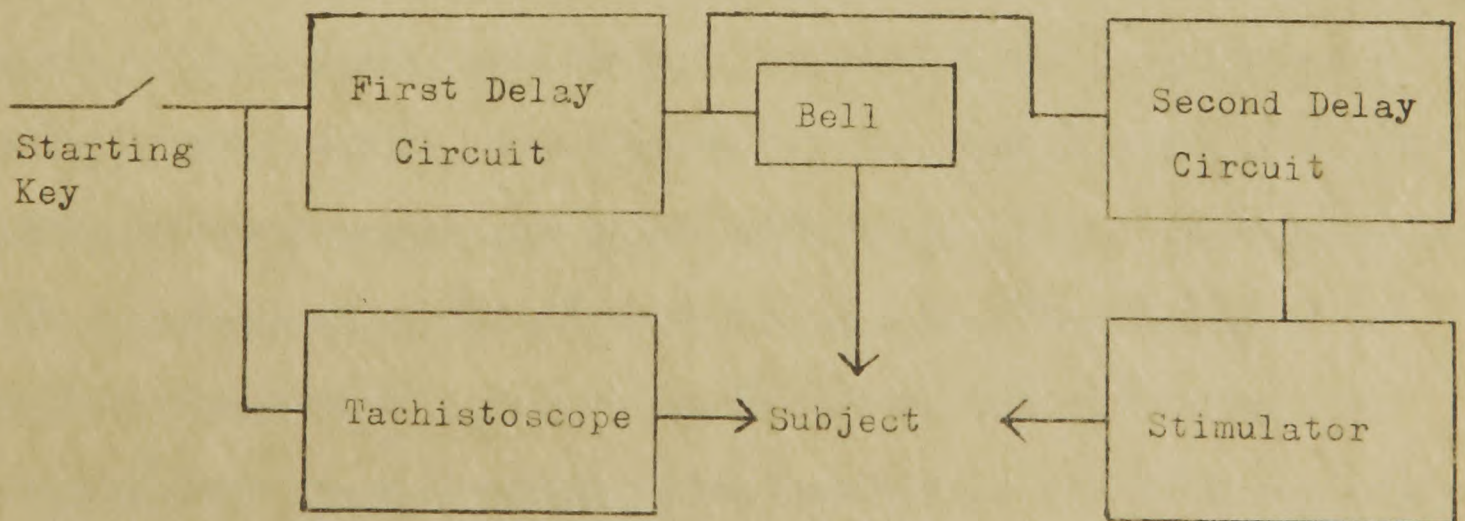
A constant current generator, which has been described

FIGURE 4

Experiments 1 and 2



Experiment 3



elsewhere (Howarth, 1954), was used to give single 100 msec. square pulses. These were administered through a large aluminium electrode, padded with saline-soaked cottonwool and applied to the subject's forehead, and a saline-soaked cottonwool electrode held in the hand. An electronic tachistoscope (Humphrey, Dawe, & Mandell, 1955) was used to provide a light flash lasting 750 msec. The spectrum of the light from the short-decay phosphor fluorescent tubes used is continuous and, compared with daylight, somewhat deficient in red, appearing a faintly bluish white. The interval from the end of the pre-illumination to the closing of the relay which administered the stimulus was determined by a standard timing circuit employing a Mullard Z803U cold cathode trigger tube. A metronome beating once a second was present and audible to the subject throughout. The arrangement of the apparatus is shown in Figure 4.

The experiment was conducted in a darkened but not completely dark room, as a low level of illumination gives more stable and reliable thresholds (Howarth, 1954). The walls of the room, where they were visible to the subject, were dimly illuminated by the experimenter's viewing light; however at no point did their brightness exceed 0.0125 millilamberts as measured with an SEI photometer. The tachistoscope box, which filled approximately  $90^\circ$  horizontally and  $45^\circ$  vertically of the subject's visual field, was considerably darker than the walls; its luminance could not be measured on the photometer. The flash was produced by illuminating a screen 6 inches square 3 feet from the subject, so that

it subtended about  $9^{\circ}30'$  horizontally and vertically. Its brightness was 75 millilamberts. The pre-illumination could be made blue by interposing an Ilford Tricolour blue filter; this reduced the brightness to about 5 millilamberts and the area to 3 inches square (approximately  $4^{\circ}45'$  across). A similar reduction in brightness and area could be achieved by interposing a neutral density filter.

The subject sat facing the screen. Each trial was begun by the experimenter saying 'coming'; this was followed about one second later by the flash. After a fixed interval the stimulus was administered, and the subject reported 'Yes' or 'No'. Trials were ordered as required by the 'method of limits': starting at a randomly varied stimulus strength well above threshold a series of trials would be given with the current reduced each time by an equal linear step until the subject failed to see three successive stimuli. The weakest stimulus strength to which a positive response was given was taken as the threshold. If reversals of response occurred the threshold was determined as described by Guilford (1954). Only descending runs were used.

Thresholds were determined in this way for the intervals 1, 2, 3, 4, and 9 seconds, in that order. Such a block of five threshold determinations was repeated four times in a session. Each session began with a short practice period to make the subject familiar with the phosphene and procedure and give a preliminary estimate of the threshold. When this was found to be greater than 100 microamps, the step size used

in decreasing the current was 20 microamps., when it was less than 100 a step size of 10 microamps. was used. A session lasted about an hour.

The subjects included the experimenters, who were well practised, and members of the department and undergraduates, many of whom had no previous experience of this type of experiment. Four subjects did four sessions with no filter, two subjects did two sessions with the neutral density filter, and five subjects did six sessions with the blue filter.

### Results.

As the results for the two subjects with the neutral density filter fell within the range of the results given without a filter they were combined in Condition (a), the sessions with the blue filter constituting Condition (b). In combining the results from different subjects the logarithms of the threshold values were used to avoid a bias arising from individual differences in absolute threshold level, and the standard deviation was taken as  $S.D. = \sqrt{\sum(x_s - \bar{x}_s)^2 / N(n-1)}$ , where  $x_s$  represents the individual threshold readings,  $\bar{x}_s$  the mean of the  $n$  threshold determinations in each session, and  $N$  the number of sessions. The mean threshold (in log. microamps.), standard deviation and standard error for each time interval in each condition are given in Table 1, and the thresholds for both conditions combined, with standard errors, are shown as curve B in Figure 5.

FIGURE 5

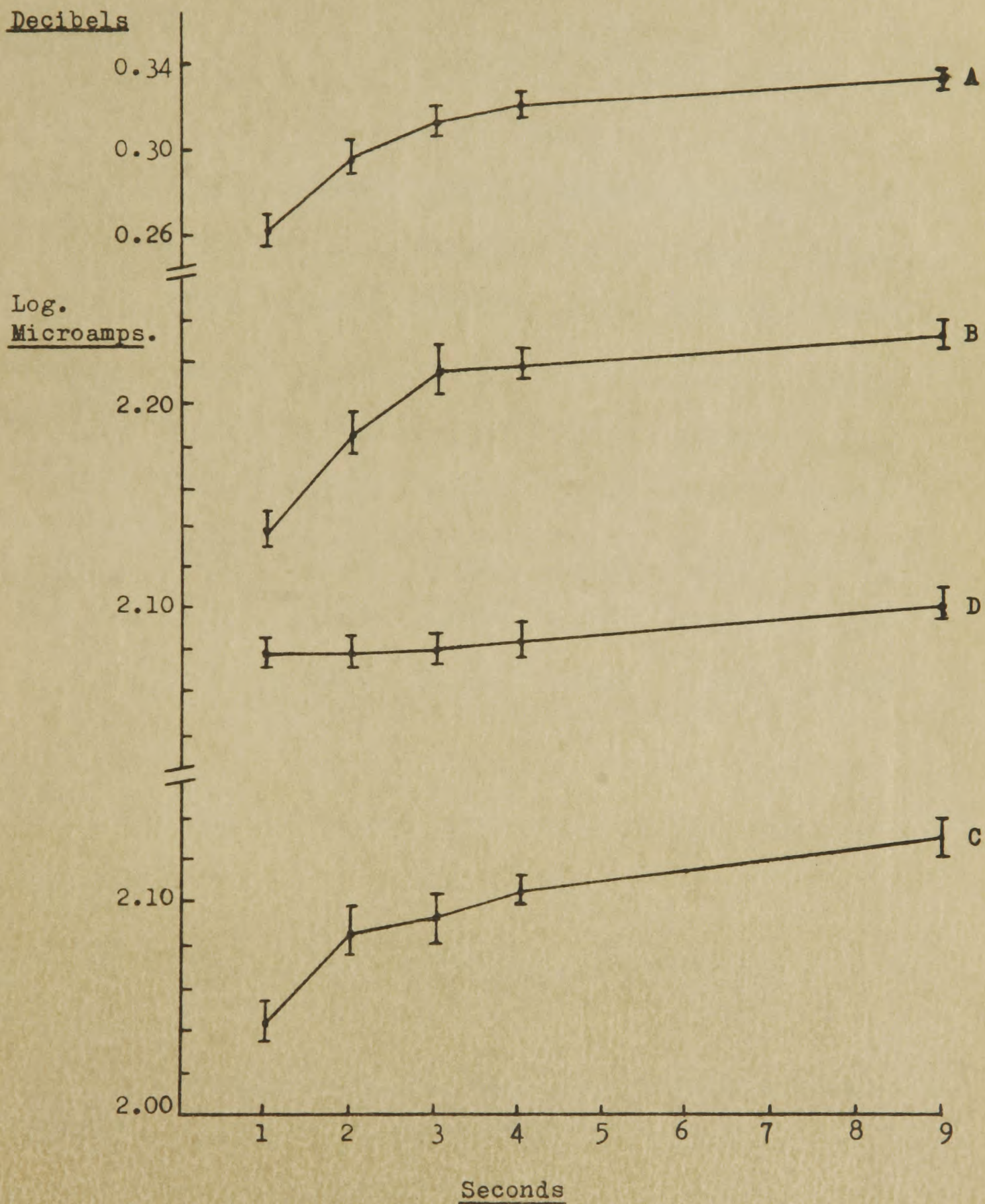


TABLE 1

<u>Condition</u>		<u>Time Interval (seconds)</u>					<u>Number of sessions</u>
		<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>2</u>	
a: White light	Mean	2.1000	2.1562	2.1653	2.1733	2.1903	6
	S.D.	0.0606	0.0366	0.0688	0.0491	0.0528	
	S.E.	0.0124	0.0075	0.0140	0.0100	0.0108	
b: Blue light	Mean	2.1796	2.2167	2.2679	2.2646	2.2800	6
	S.D.	0.0670	0.0840	0.0730	0.0671	0.0518	
	S.E.	0.0137	0.0172	0.0149	0.0137	0.0106	
a + b: White or Blue light	Mean	2.1398	2.1864	2.2167	2.2190	2.2352	12
	S.D.	0.0639	0.0648	0.0709	0.0588	0.0523	
	S.E.	0.0092	0.0094	0.0102	0.0085	0.0076	

## Discussion.

The threshold is lowest at the shortest interval, one second, and rises, rapidly at first and then more slowly, as the interval lengthens. This result does not agree with those reported by Motokawa, who finds a peak of sensitivity at two seconds, for white light, or three seconds, for blue light, with higher thresholds at times shorter or longer than these. A possible explanation for the divergence with the white light might have been that the subjects used in this experiment had shorter crest times, falling below 2 seconds, but in this case a peak should still have appeared with the blue light. As it is, the relation between threshold and interval is much the same for both conditions, except that the thresholds are all higher with blue light than with white. This is most probably a consequence of the use of different subjects in the two conditions.

The failure to obtain Motokawa's curves will be discussed later. A problem of immediate interest is whether the relation found is due to peripheral interaction at the retina or represents a centrally mediated effect.

### 2.3. The effect of a bell on the phosphene threshold.

#### Experiment 2.

The aim of this experiment was to see whether the effect found in Experiment 1 would be shown when the accessory stimulation was given in a different modality.

### Apparatus and procedure

In this experiment the tachistoscope was not used to provide a flash; instead a loud ring from an electric bell was employed. The duration of the ring was controlled by a relay with a condenser and resistance in the coil circuit so that the relay opened automatically after 0.75 seconds. In Condition (a) the metronome was present, beating once a second, in Condition (b) it was not used. In all other respects the apparatus and procedure were the same as in Experiment 1.

Five subjects did five sessions in Condition (a), and five subjects did six sessions in Condition (b). There were eight subjects altogether, of whom two served in both conditions.

### Results

The mean thresholds for each condition and for the experiment as a whole are given (in log microamps.) in Table 2 and as curve C in Figure 5. Time intervals are measured from the end of the ring.

### Discussion

These results are very similar to those of Experiment 1. Thus it appears that the effect shown there can be produced by an accessory stimulus in another modality. This makes it very unlikely to be the same, but with a much earlier peak, as that described by Motokawa, since the latter is considered to represent an action of the light flash on colour-sensitive elements of the retina.

The analysis of variance shows significant effects of the

TABLE 2

<u>Condition</u>		<u>Time interval (seconds)</u>					<u>Number of sessions</u>
		<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	
a: Bell + metronome	Mean	2.0211	2.0781	2.0744	2.0764	2.0990	5
	S.D.	0.0613	0.0625	0.0651	0.0512	0.0562	
	S.E.	0.0137	0.0140	0.0146	0.0114	0.0126	
b: Bell alone	Mean	2.0594	2.0908	2.1050	2.1269	2.1365	6
	S.D.	0.0750	0.0652	0.0758	0.0470	0.0577	
	S.E.	0.0153	0.0133	0.0155	0.0096	0.0118	
a + b:	Mean	2.0420	2.0850	2.0911	2.1040	2.1194	11
	S.D.	0.0716	0.0640	0.0711	0.0489	0.0571	
	S.E.	0.0108	0.0096	0.0107	0.0074	0.0086	

Analysis of variance of the thresholds

<u>Source</u>	<u>d.f.</u>	<u>Mean square</u>	<u>F</u>	<u>Significance</u>
Intervals	4	0.00928	18.20	<u>p &lt; 0.001</u>
Conditions	1	0.01560	30.59	<u>p &lt; 0.001</u>
Between subjects (Condition (a))	4	0.14040	275.29	<u>p &lt; 0.001</u>
Between subjects (Condition (b))	5	0.03560	69.80	<u>p &lt; 0.001</u>
Intervals x Conditions	4	0.00058	1.14	N.S.
Residual	36	0.00051		

time intervals, the subjects, and the two conditions. The higher thresholds in Condition (b) can probably be attributed to differences between the two groups of subjects. However, the lack of any significant interaction between intervals and conditions indicates that the presence or absence of the metronome has had little effect on the relation between threshold and time interval.

A possible reason for the failure to find Motokawa's effect might be that the interaction between the light-flash and the colour-sensitive elements is obscured by the monotonic relationship between threshold and delay interval shown in this and the previous experiment. If the latter effect could be neutralised, the former might be revealed. Another possible explanation depends on a feature of Motokawa's use of the method of limits. He does not use a constant step size, but decreases the applied voltage in large steps initially, when the phosphenes are readily perceptible, and by small steps when he is near the threshold. Kohata, Komatsu, & Motokawa (1956) suggest that this saves the retina from a depressing effect produced by an excessive number of strong stimuli. They found that if they used equal steps, or administered a number of strong stimuli before the threshold determination, the variability of the threshold response increased and the colour-sensitive effect disappeared. However the effect could be demonstrated when a constant step size was used provided thresholds were determined with the use of ascending runs only, so that no strong stimuli were given.

These possibilities are examined in the next experiment.

#### 2.4. Light and bell combined.

##### Experiment 3.

In this experiment the light-flash was given as in Condition (a) of Experiment 1. If the threshold change shown in Experiments 1 and 2 was a function only of the fixed time interval between a recent accessory stimulus and the critical stimulus then the addition of an auditory accessory stimulus coming after the flash but a fixed interval before the electric stimulus should make the threshold constant, and this might allow an effect of the light as such to be shown more clearly.

##### Apparatus and procedure.

A second delay circuit of the same type as the first, and the bell, were added to the apparatus of Experiment 1, as shown in Figure 4. The sequence of events on each trial began with the experimenter saying 'coming' and about a second later pressing a switch which simultaneously started the first delay circuit and the 0.75 second flash. After an interval the delay circuit closed a relay which started the bell ringing for 0.75 seconds and also closed the second delay circuit. After 1.75 seconds the second circuit administered the stimulus. The first circuit could be adjusted so that the interval between the end of the flash and the stimulus was equal to 1, 2, 3, 4, or 9 seconds. In each case the ringing ended one second before the stimulus. No filter was used with the flash.

In Condition (a) thresholds were determined by the method of limits, using descending series only, as before. In Condition (b), however, only ascending series were used. Each series started well below threshold and the stimulus strength was increased in 10 microamp. steps until three successive positive responses had been given. Eight subjects did ten sessions in Condition (a), and six of these subjects did six sessions in Condition (b).

### Results.

The thresholds, in log. microamps., are given in Table 3, and the results for both conditions combined are shown as curve D in Figure 5.

### Discussion.

In both conditions the rapid rise in threshold shown in the earlier experiments has been reduced to negligible proportions by the introduction of the bell. The threshold at 9 seconds is slightly higher than at the shorter intervals but this effect is not significant on analysis of variance. It appears that it is the interval of time since the most recent accessory stimulus that mainly determines the degree to which the threshold is lowered; any residual effect of the earlier stimulus, if present, is very slight. Thus this experiment confirms the effectiveness of auditory accessory stimulation shown in Experiment 2. The analysis of variance shows significant effects of subjects and conditions. The higher thresholds in Condition (b) might be due to the

TABLE 3

<u>Condition</u>		<u>Time interval (seconds)</u>					<u>Number of Sessions</u>
		<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>2</u>	
a: Descending Series	Mean	2.0535	2.0425	2.0629	2.0633	2.0728	10
	S.D.	0.0628	0.0688	0.0641	0.0547	0.0737	
	S.E.	0.0099	0.0109	0.0101	0.0086	0.0117	
b: Ascending Series	Mean	2.1223	2.1422	2.1111	2.1255	2.1564	6
	S.D.	0.0449	0.0299	0.0473	0.0682	0.0407	
	S.E.	0.0092	0.0061	0.0097	0.0139	0.0083	
a + b:	Mean	2.0793	2.0799	2.0810	2.0866	2.1041	16
	S.D.	0.0567	0.0574	0.0584	0.0601	0.0634	
	S.E.	0.0071	0.0072	0.0073	0.0075	0.0079	

Analysis of variance of the thresholds.

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Intervals	4	0.00155	1.58	NS ( $p < 0.2$ )
Conditions	1	0.09920	101.22	$p < 0.001$
Between subjects (Condition (a))	9	0.15401	157.15	$p < 0.001$
Between subjects (Condition (b))	5	0.03162	32.27	$p < 0.001$
Intervals x Conditions	4	0.00148	1.51	NS
Residual	56	0.00098	-	

use of ascending rather than descending series, or to the absence of some of the subjects used in Condition (a).

This experiment provides a control for two features of the design of the earlier experiments which might have caused part at least of the rise in threshold. Thresholds were always determined for the intervals 1, 2, 3, 4, and 9 seconds in that order; if there had been a tendency for the threshold to rise gradually during the course of a session, this would have resulted in an apparent rise of threshold from 1 to 9 seconds. Furthermore, the intervals between successive presentations of the stimulus were not controlled, so that it was given more frequently at the shorter intervals. If this more rapid presentation had in some way lowered the threshold, an apparent fall at the shorter intervals would have been produced. However both these factors should have been operating at full strength in the present experiment; the lack of any significant effect shows they cannot have been important.

Despite the effective suppression of the rapid rise in threshold found earlier, the colour-sensitive effect is not shown. In Condition (a) the threshold at 2 seconds is slightly lower than at 1 or 3 seconds, but not significantly so. In Condition (b), in which, since strong stimuli have been avoided by the use of ascending runs only, Motokawa's effect should be most clearly expected, the threshold is higher at 2 seconds than at either 1 or 3 seconds.

## 2.5. The status of the colour-sensitive effect.

The failure to find the effect described by Motokawa in either Experiment 1 or Experiment 3 requires explanation. One reason for this could be that some essential feature of his method has not been accurately reproduced. In designing the experiments care was taken to choose the intensity and duration of the flash, and the area illuminated, to be within the ranges used by Motokawa. Although in most of his experiments the subject is in the dark, he has also obtained his effect with weak background illumination (Motokawa, 1949c); this condition was preferred in the present experiments as it gives a more stable threshold. Motokawa used a small group of well-practised subjects; both trained and naive subjects were used in these experiments but no systematic difference between their results was found. Another feature of his procedure (Motokawa, 1951b) is the administration of a click, 0.5 seconds before the stimulus, in order to reduce the subject's variability. The metronome in Experiment 1 was intended to serve a similar function. However its beat was not coordinated with the initiation of each trial, so that its relation to the time of arrival of the stimulus was not fixed. It is perhaps for this reason that it had so little effect in Experiment 2. However, the bell in Experiment 3 should have had the same effect as this click. The main divergence in method appears to lie in the manner in which the method of limits has been applied. Motokawa uses a large step size at the beginning of a descending series and makes the stimulus

changes much smaller when he is near the threshold. When the subject finds the judgments difficult he is allowed to change to a procedure in which each stimulus is compared with a blank. The technique used here has employed equal steps throughout. Kohata, Komatsu, & Motokawa (1956) claim that this method results in the administration of too many strong stimuli, which produce a state of retinal fatigue in which the eye does not respond to weak stimuli, and consequently the effect cannot be shown. They conclude that descending series with large initial steps, or ascending series only, must be used. However Condition (b) of Experiment 3 used ascending series only, and provided no indication of the colour-sensitive effect.

A possible explanation for the dependence of the effect on the variability of step size in Motokawa's method derives from the form of the psychometric function. Since the probability of response changes gradually over a transition zone the threshold found by the method of limits will depend to some extent on the step size used; where steps are small the number of responses will be large and the probability of finding that sequence of responses which is the experimenter's criterion for terminating the series, and which defines the threshold, will be increased. Therefore the use of small steps where the threshold is expected, and larger steps before that stage, will in itself create a probability of finding the threshold where it is expected by the experimenter. This probability may have been increased by some other

features of Motokawa's technique. The use of well-practised subjects only, very often the experimenters, who may have had some knowledge of the results to be expected, may have increased the danger, often present in psychophysical experiments, of influencing the subject's responses by unintended suggestions. The concept of 'multiple thresholds' is also important in this respect. Motokawa (1951b) believes that each light-sensitive process stimulated will determine a different threshold; if, for example, a descending series of judgments is made, each stimulus coming 0.5 seconds after a blue flash, a point will be reached at which the positive responses become negative. He would take this to be the threshold for the blue process. If the descending series is then continued he would expect positive responses to recur; these would be attributed to the red process, which has also been stimulated and is more sensitive at this time interval, and if the series is continued further the threshold for this process would then be reached. Motokawa attributes some failures to repeat his results to this phenomenon: the experimenter, finding an apparent threshold, is satisfied, and does not continue the series until he reaches the true threshold at its lower level. Motokawa avoids this in some cases by taking large steps to skip over the apparent thresholds, where these are to be expected. This account of the variability of response which is found in a continued series of judgments is unattractive. It supposes responses will become negative when the blue process, for instance, ceases to respond, although

the red process is so sensitive that at still weaker stimulus strengths it will produce positive responses on its own. A zone of variability of response is found in all threshold determinations, even in cases where it would seem less plausible to postulate a number of independent sensitive processes; this variability may be better accounted for by statistical variation in the central effect of a stimulus of constant intensity. The 'multiple threshold' hypothesis implies that the psychometric function given when the method of constant stimuli is employed should not be the usual smooth curve, but should show a series of maxima, corresponding to the thresholds. However, when this experiment is performed, a function without notable discontinuities is produced (Riggs, Cornsweet & Lewis, 1957).

An important consequence of this account is that it seems almost to require a conscientious experimenter who reaches the criterion for the threshold, but considers it may be no more than an apparent threshold, to continue the series further in search of the true threshold at a lower level. As a check that this does not systematically alter, for whatever reason, the relation of threshold to time interval, some of the data given above were re-scored, the weakest stimulus eliciting a positive response in a series being taken as the threshold, without regard for the number of negative responses given to stronger stimuli in the series. However this produced no notable change in the relation of threshold to interval. For example, the resulting thresholds for

Experiment 3, Condition (b), for the intervals of 1 to 9 seconds, in that order, are 2.109, 2.123, 2.085, 2.108 and 2.124 log. microamps. The rise for the 2 second interval is still in evidence.

A feature of the subjects' training is that they are encouraged to attempt to see stimuli which are as weak as possible. These are described as "sometimes very similar in appearance to intrinsic light of the retina, but have a characteristic that they appear in definite relation to the electric stimulus, while intrinsic light appears at random" (Motokawa, 1951b). In view of this the precision achieved by the subjects is remarkable; their variability is considerably less than is usually found, being of the order of 1 per cent of the stimulating voltage at the absolute threshold. Kohata, Komatsu & Motokawa (1956) reported that if they used equal steps on a logarithmic scale in their descending series variability increased to about 10 per cent; by comparison the standard deviation of ten values determined by their routine method was 0.27 per cent of the mean. They conclude that using equal steps induces variability and thus obscures their effect. They found that some subjects could not be trained to give this low variability; in such cases the results were not used. Another interesting finding is reported by Motokawa, Ebe, Arakawa, & Oikawa (1951). They used a test patch of 2' of arc and an equal energy spectrum of an intensity such that only the middle of the range was visible: "the other parts of the spectrum evoked no light sensation. Never the less, the responses to

these subthreshold lights could be determined by our method with the same degree of accuracy as those to the superthreshold lights", the usual curves being given. This is surprising in view of the small number of quanta required at the absolute threshold, and the corresponding importance of the quantum variability of the stimulus (Hecht, Schlaer & Pirenne, 1942).

A number of attempts to confirm Motokawa's findings have been made. Michaels (1957a, 1960) has investigated the phenomenon of retinal induction reported by Motokawa (1949c; 1950a) and confirmed its occurrence. He used two subjects in each experiment, and a rather complex training process was necessary to get positive results. Initially the subjects "tended to discriminate a diffuse brightening as a phosphene", but with this criterion their results showed a variability of 10-20 per cent, and he concluded that "these shifts and inconsistencies were due to an unstable frame of reference". This was dealt with by a four month training period to "motivate the Ss to search for different cues and criteria for judgment" which reduced the threshold to a low level at which the subjects "used as a criterion more of a recognition of a change in state rather than an actual perception of light". He found the method of constant stimuli, the method of limits, and Motokawa's method of allowing comparisons with blanks were all unsatisfactory, and developed a new psychophysical method, the 'method of successive approximations', with which he was able to obtain positive results. In this procedure he started with a stimulus

"well above threshold (10-30  $\mu$  amp.)" and reduced its intensity in steps of 4 microamps. until the range in which the threshold lay was determined. The threshold was then "bracketed" by presenting stimuli alternatively above and below it, reducing their separation each time, until they were 2-4 microamps. apart and three successive, or four out of five, positive responses were given to the larger, and a corresponding number of negative responses to the smaller stimulus. In this way it was possible to estimate the threshold to within  $\pm 3$  per cent.

This technique might be criticised. In Experiments 1 and 2 the variation in threshold between the 1 and 9 second intervals was of the order of 20-40 microamps.; to determine the threshold it was thought advisable for the starting point of the descending series to be 100-150 microamps. above the expected threshold. A starting-point only 10-30 microamps. above the expected threshold may in itself have tended to determine where that threshold would be found. In the experiments reported above standard deviations are of the order of 10-20 per cent of the mean. The finding that over a range of 2-4 microamps. the probability of response of Michaels' subjects regularly changed from 0-20 per cent to 80-100 per cent is surprising. Again, it seems possible that the operation of response dependencies (Verplanck, Collier & Cotton, 1952) could have resulted in the experimenter's criterion being met. Thus if the stimuli are being alternated and the subject has a tendency to alternate responses this might result in an apparent "bracketing" of the threshold. Some of

Michaels' observations suggest that response dependencies may well have been present. Thus he notes that if a run of three or four stimuli just above threshold give positive responses, the third or fourth will tend to give a negative response, and this 'reversal tendency' is so strong that a stimulus coming in this position may have to be increased considerably in strength to give a positive response. Like Motokawa's, this procedure is dependent on the subject's responses. Michaels (1957b) notes "I feel uneasy about an experiment controlled by the subject rather than the experimenter. The nature of the discrimination is such, however, that I think it is the only way I was able to get stable thresholds".

Korzun (1957) has also repeated Motokawa's work. He used four subjects, and presents results for two. He confirmed the colour-sensitive effect and investigated the consequences of using very intense flashes. These caused the crest times to be 0.5-1.0 seconds later than usual, though the author refrains from claiming greater precision because of the "fatigue of the subject and ... fluctuations of the threshold". He gives little information about his psychophysical procedure. The experimenter's criterion for the threshold was that, of two stimuli, the subject did not notice one and "barely registered" the other, but he does not report in what order, or at what step sizes, the stimuli were presented. Lacking this information we must assume that the technique was similar to that of Motokawa, and subject therefore to the same criticisms.

Riggs, Cornsweet & Lewis (1957) have performed an extensive and carefully controlled series of experiments to test Motokawa's results. They used a very similar technique, the only important differences being that equal steps on a logarithmic scale were used, a rigid protocol was followed in giving the stimuli and deciding the point of termination of the series, and no data were discarded on the grounds of variability. The criterion for the stopping point, which involved comparisons between the weak stimulus and a blank, was similar to that of Motokawa. Using red, blue, green and white flashes with dark-adapted subjects they were able to confirm the reduction in sensitivity immediately after the flash. Within one second excitability returned to normal in some subjects and became supernormal in others. This enhancement often persisted for several seconds, but it showed no clear peak nor the specific wave-length effects reported by Motokawa, nor did they find the very small variabilities he reports. They believe that his wave-length effects must be attributed to his use of different step sizes and a flexible protocol modified in accordance with the responses of the subject.

An impressive though indirect degree of support is given to Motokawa's findings by the results of a number of physiological experiments in which the response of the optic nerve or retina is examined directly (Motokawa, 1949a; Motokawa, Iwama & Ebe, 1952; Motokawa & Iwama, 1949b, 1949c, 1951; Motokawa, Iwama & Tukahara, 1951). In these experiments an electric pulse was passed through the eye at a fixed

interval after a flash of light and the response observed with a cathode ray oscillograph and loudspeaker. With the excised frog's eye Motokawa and his colleagues found a peak in the optic nerve response at 1.5 minutes after a red light, 2.5 minutes after a green light and 4 minutes after blue. Recording directly from the retina of the decerebrate cat they found red, yellow, green and blue crest times of 1, 1.5, 2.25, and 3 seconds respectively. The responses were very dependent on the physiological state of the animal and could readily disappear. In these experiments much longer electric pulses were used to stimulate the eye than in the human: 1.0 seconds for the frog and 0.5 seconds for the cat.

It is interesting though somewhat surprising that the responses in the frog should be so very much slower than in the cat. An important point in evaluating these experiments is the method used to determine the response curves. This is by observing what stimulus strength is necessary to give a threshold response, i.e. one just perceptible to the observers. Motokawa, Iwama & Tukahara (1951) describe their technique as "it was tried to adjust the intensity of each electrical stimulus to just the threshold value at each moment in order to obtain a just perceptible response ... For a trained experimenter it was not so difficult, as it might appear, to adjust the intensity of each stimulus to the threshold strength". This technique may not have been very dissimilar from setting the experimenters a difficult threshold discrimination with a flexible protocol and full knowledge of the expected result. Thus,

until these results are confirmed with the use of an objective technique for distinguishing faint signals from neural noise and avoiding observer bias, such as the superimposition of traces, they must remain in doubt.

Thus consideration of both the psychophysical and physiological experiments of Motokawa and his colleagues suggests that their results could have been produced as artefacts of faulty observational methods. Taken with the fact that it has not proved possible to obtain the same effects when reliable psychophysical procedures have been used, their findings must be treated with grave reserve.

## 2.6. Conclusion.

In the experiments described in this chapter an attempt was made to demonstrate the colour-sensitive effects of a flash of light on the phosphene threshold first described by Motokawa. This was not achieved. Instead an effect on the threshold was found which was monotonic over the range investigated, the threshold being lower the shorter the interval between the accessory stimulus and the electrical pulse. This was shown not to be due to a specific after-effect of light on sensitive mechanisms at the retina, since it was equally well given by auditory accessory stimulation. Consideration of the technique of the experiments which have produced the colour-sensitive effects suggested that their results may well be an artefact of the use of faulty procedures. This, taken together with the present inability to demonstrate Motokawa's effect, suggests that it would be better to direct further experimentation to the problem of the proper explanation of the monotonic effect found.

### CHAPTER 3

#### THE NATURE OF THE MONOTONIC THRESHOLD CHANGE

##### 3.1 Introduction

In the last chapter a relation between the accessory stimulus and the absolute phosphene threshold was found in which the threshold fell as the interval of time between the accessory and critical stimuli was made shorter, over the range of intervals examined, so that a monotonic, negatively accelerated curve was obtained when the threshold level was plotted against the time interval. It was also shown that this effect was not dependent on the accessory stimulation being in the same modality as the critical stimulus. Thus we now have the problem of framing hypotheses to explain this 'threshold-lowering effect' of accessory stimulation.

In signal-detection theory terms the fall in threshold would correspond to a downward shift of the 'critical value'. This could be due to a direct effect on the 'criterion computer', or it could be secondary to a change in  $f_N(x)$  or  $f_{SN}(x)$ . A change in  $\sigma_N$  or  $M_N$  would alter  $f_N(x)$  and, since it is assumed that  $\sigma_N \doteq \sigma_{SN}$ , might affect  $f_{SN}(x)$  too. A plausible explanation for the threshold shift is offered by the hypothesis that the accessory stimulus produces an immediate, reflex-like change in the subject, analogous to, or part of, the 'alerting' or 'arousal' response, causing a temporary general increase in the responsiveness of all the sensory systems. This effect would be assumed to be a 'ballistic' physiological change, building up, with appropriate latency, to a maximum, and then gradually decaying, so that the extent to which

the threshold is lowered, at a given interval after the accessory stimulus, would be a measure of the strength of the effect remaining after that interval, and the form of the threshold curve would represent the time-course of the physiological change. The effect might be central, a change in the neural mechanisms which deal with the afferent input, or peripheral, an effect on the receptors, mediated perhaps by centrifugal fibres (Granit, 1955b; Dodt, 1956; Galambos, 1956; Kerr & Hagbarth, 1955; Loewenstein, 1956). It might correspond to a temporary suppression or reduction of physiological noise in the afferent pathways, or some other change resulting in an enhanced signal/noise ratio.

The 'arousal hypothesis' postulates a general effect acting on all sensory systems. An alternative 'immediate effect' hypothesis is that the effect found is specific to the threshold and modality which have been examined, either because it reflects the workings of a sensory mechanism peculiar to the eye, or because it is an artifact of the technique used, such as a side-effect of some essentially non-sensory physiological change. For example, the accessory stimulus might have caused changes in the vascular bed about the eyes. Since a constant current generator was used gross changes in skin and tissue resistance should have had no effect on the results, but a redistribution of blood in the vessels, due to differential vasoconstriction or dilatation, might have so altered the paths of current flow in the eye and surrounding tissues that the proportion of the current actually passing through the sensitive elements of the

retina was increased, and an apparent lowering of threshold produced.

A second, quite different hypothesis would require the accessory stimulus to function as a 'warning', i.e. as a source of information which can be used by the subject to modify his strategy for dealing with the stimulus. Thus, if he knows the length of the time interval, the accessory stimulus will mark its inception, and he can use this information to lower his threshold at the time when the stimulus can be expected. This hypothesis covers two possibilities. The first is that the threshold-lowering is due to a 'preparatory' effect: the afferent paths must be got ready in time for the arrival of the stimulus, perhaps by reducing physiological noise, or in some other fashion, so as to facilitate its discrimination. As with the 'arousal' hypothesis we can think of the change produced as an alteration in the 'noise' or 'signal' distributions, such as a reduction of the mean value of noise, or of the scatter of the distributions, which will improve discrimination at the moment that the critical stimulus is administered.

The second, 'judgmental' possibility is that the change in threshold would not be produced by an anticipatory alteration in the sensory mechanisms but by a change in the critical value used to determine the response, this change being a consequence of information which the accessory stimulus has the effect of giving the subject. Thus, if the

critical value is in part determined by the a priori probability of the stimulus,  $P(SN)$ , which we may take to correspond to the subject's assessment of the probability of his receiving the stimulus, this latter will be greater at the moment that the stimulus is expected to arrive than at other times, so that we may suppose that the critical value with which the input occurring at that time is compared will be correspondingly lower.

Both possibilities require an extra assumption to account for the form of the threshold curve shown in Experiments 1 and 2. This is supplied by the known imprecision of the subjective estimation of time intervals (Woodrow, 1951). If we can suppose that the measure of time used by the sensory discrimination mechanism in determining the moment when the stimulus can be expected to arrive has similar characteristics to that illustrated by subjects' judgments of temporal duration, then the subject will not be able to determine the moment of arrival of the stimulus with absolute precision, the best he can do will be to determine a range of time over which he may consider the arrival of a stimulus highly likely. Since differential sensitivity to time is approximately described by Weber's law, this range of time would increase with the length of the time interval between the accessory and critical stimuli. For the 'preparatory' hypothesis this would imply that the threshold-lowering effect would have to be exerted for longer as the inter-stimulus interval increases, if it is begun at the beginning and continued to the

end of the 'range of expectation'. It is plausible to suppose that it is 'costly' to the organism to produce this effect, since otherwise it could be done continuously, so we may suppose that the longer the effect must be maintained the less intense it will be, and accordingly, threshold reduction will be less at longer intervals.

A similar argument applies to the 'judgmental' hypothesis. If we consider each 'range of expectation' as made up of equal 'moments' during which the stimulus might be presented, then for short ranges the subjective probability of the stimulus arriving at any one 'moment' would be greater than for longer ranges constituted of a greater number of 'moments'. Correspondingly, the critical value could be lowered more for inputs occurring during a short range than for those from a long range, so that some inverse relation would be expected between the extent of the threshold fall and the length of the time interval between accessory and critical stimulus. As a first approximation, we might suppose that simple inverse proportionality would obtain between them.

The different implications of the 'immediate effect' and 'warning' hypotheses can be illustrated by supposing that we could observe the threshold continuously during the period following the administration of the accessory stimulus. On the first hypothesis we should find an immediate fall in threshold to a minimum at some time less than 1 second, followed by a gradual return to an asymptotic level along a course similar to that illustrated by curves A, B, and C in Figure 5. On the second

hypothesis the threshold would remain constant at the asymptotic level from the moment the accessory stimulus is administered until some time before the expected time of arrival of the critical stimulus. Here we should find the threshold fall to a minimum, remain at this level for a greater or lesser time, and then return to the asymptotic level. This would reflect either the time course of the inception of 'preparation' at the beginning of the 'range of expectation', and its decay at the end, or the corresponding 'judgmental' change as the subjective probability of the stimulus varies.

A second important difference in their implications is that in the first case the threshold fall should be found whether or not the subject is told in advance what the length of the time interval will be, but in the second case withholding this information should disrupt the effect.

These implications will be tested in two experiments later in this chapter, and further implications will be pointed out and tested in later chapters, in an endeavour to obtain evidence allowing us to choose between the various possibilities that have been discussed. Before considering these experiments, however, it is necessary to examine the question of whether the effect is peculiar to vision, or, indeed, the phosphene threshold, or can be obtained when the critical stimulus is in another modality and the mode of stimulation is different.

### 3.2. The effect of visual accessory stimulation on an auditory threshold.

#### Experiment 4.

The object of this experiment was to see whether the monotonic threshold would be obtained, if a different modality, hearing, was used for the critical stimulus, and an incremental, rather than an absolute threshold was determined.

#### Apparatus and procedure.

The apparatus was designed to administer to the subject a constant tone at 60-70 db.SL, to which a short increment of intensity could be added. A 500 c.p.s. tone from a Muirhead audio-oscillator was passed through a one stage amplifier using a Mazda 6F33 pentode. The amplification was governed by the suppressor grid voltage which could be switched between two positions on a potentiometer to give an increase or decrease in the intensity of the tone. The accessory stimulation was provided by a neon bulb lighting up; it stayed on until the response had been made. The time interval between the neon bulb coming on and the stimulus being switched in was determined by the Z803U timing circuit used in the earlier experiments. The switching was done by a relay closing a circuit; after 300 msec. a second Z803U timing circuit broke the circuit, restoring the constant tone to its resting intensity. The rise and decay of the 300 msec. intensity increment had a time constant of about 100 msec. so that there were no audible clicks. The form of the pulse was monitored on a Cossor oscilloscope. The tone was administered to the subject through

Brown Type K moving coil earphones.

The subject sat in a darkened room facing a screen 3 feet square, in the centre of which was the neon bulb. The procedure was similar to that of Experiment 1: thresholds were determined by the method of limits, using descending series only, at intervals of 1, 2, 3, 4, and 9 seconds after the beginning of the accessory stimulation. The stimulus was decreased by equal intervals on a decibel scale. Four such blocks of threshold determinations constituted a session. No pre-warning was given. Nine sessions were obtained from 7 subjects.

#### Results.

The mean thresholds in decibels are given in Table 4 and as curve A in Figure 5.

#### Discussion.

It will be seen that these results show much the same fall in threshold level as the time interval shortens that was found in Experiments 1 and 2. As the critical stimulus used here is in a different modality it appears the result is not peculiar to a particular modality, such as vision, nor is it an artefact of the use of an electric stimulus. The generality of the effect is further supported if the results of Newhall (1923) with a cutaneous pressure stimulus, described in Section 2.1, are taken to represent an effect of the same sort. The experiment also confirms the effectiveness of heteromodal accessory stimulation shown in Experiments 2 and 3. It is interesting that the accessory stimulus was

effective response being left as continuously varying with level. This is in accord with the hypothesis that is investigated by setting the beginning of a fixed interval time for the subject. Another hypothesis already stated is TABLE 4 which shows the mean response times for right-handed subjects as a function of the Time Interval (seconds).

	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>9</u>
<u>Means</u>	0.262	0.296	0.312	0.320	0.331
<u>S.D.</u>	0.0456	0.0450	0.0424	0.0410	0.0313
<u>S.E.</u>	0.0076	0.0075	0.0071	0.0068	0.0052

The two hypotheses imply that if the subject receives a fixed interval of the threshold units about the expected time of arrival of the stimulus in detail it would be possible to show that the hypothesis was correct. The threshold level would rise gradually over the course of the time interval increased; as the learning hypothesis was valid in a dog, with the threshold level at about the expected time of arrival of the stimulus, and rising, at intervals longer or shorter than those, to the asymptotic level. In the following experiment an attempt is made to see whether any report can be obtained for one or other of these possibilities.

#### Experiment 5

In this experiment it is assumed that the dog will learn the time of a small number of the threshold units by using information obtained from the stimulus and a small number of units. The threshold level would rise gradually over the course of the time interval increased; as the learning hypothesis was valid in a dog, with the threshold level at about the expected time of arrival of the stimulus, and rising, at intervals longer or shorter than those, to the asymptotic level. In the following experiment an attempt is made to see whether any report can be obtained for one or other of these possibilities.

effective despite being kept on continuously during each trial. This is in accord with the hypothesis that it functions by marking the beginning of a fixed interval known to the subject. However it is not strongly against the 'arousal' hypothesis since sensory adaptation might prevent a continuous stimulus being continuously arousing.

Since a modality-specific effect has thus been excluded, we have now to discriminate between the 'general sensory arousal' and 'warning' hypotheses.

### 3.3. The use of ranges of time intervals.

The two hypotheses imply that if one could examine a small section of the threshold curve about the expected time of arrival of the stimulus in detail it would be different in each case. If the 'arousal' hypothesis were correct the threshold would rise gradually over the range as the time interval increased; on the 'warning' hypothesis there would be a dip, with the threshold lowest at about the expected time of arrival of the stimulus, and rising, at intervals longer or shorter than these, to the asymptotic level. In the following experiment an attempt is made to see whether any support can be obtained for one or other of these predictions.

#### Experiment 5.

In this experiment an attempt was made to map the form of a small section of the threshold curve by using inter-stimulus intervals randomised over a small range of times. One difficulty, in principle,

with this procedure, is that if the performance previously found depends on the use of a known fixed interval by the subject, then presenting stimuli at discriminably different times may alter his expectations, and consequently the function which is desired to observe. Thus, although a clear positive result, if obtained, would be of great interest, a negative result will not require us to reject the 'warning' hypothesis.

#### Apparatus.

In this experiment the constant current generator used in the first three experiments was again employed to produce an electric phosphene as the critical stimulus. The subject wore earphones, and the accessory stimulus, a 50 c.p.s. tone at about 30 db.SL, was administered through them, continuing until the subject had responded. Randomization of the intervals between the inception of the tone and the administration of the phosphene was achieved by an arrangement of two Z803U timing circuits and four relays. When the experimenter pressed the operating button the first relay, which was self-holding, closed, simultaneously starting the administration of the accessory stimulus, a bank of Dekatron counters, and the first timing circuit. After a pre-set 'minimum' interval this timing circuit closed a second self-holding relay. No stimulus could be administered before this relay had closed, so that this timer determined the shortest possible inter-stimulus interval. A second timer operated continuously throughout each session, momentarily closing a third relay, which re-set the timer, at a regular pre-selected interval, the 'range'.

When all three relays were simultaneously closed, a fourth relay, which was self-holding, then closed, triggering the electronic stimulator and stopping the bank of Dekatrons, thus recording the time interval to the nearest millisecond. With this arrangement the shortest possible interval between the beginning of the tone and the administration of the electric stimulus was determined by the 'minimum' setting on the first timer, and the longest possible interval was the sum of 'minimum' and 'range'. The actual interval on any trial depended on the moment in the second timing circuit's cycle at which the experimenter pressed the operating button. To ensure that the experimenter's operations were independent of the cycling, he wore earphones giving a continuous 90 c.p.s. masking tone which prevented his hearing the relays clicking, so that a random distribution of intervals, of rectangular form, was achieved.

#### Procedure.

The experimental situation was as in Experiment 1, save that the tachistoscope was replaced by a white screen 3 feet square, which the subject faced. Each trial was preceded by the experimenter saying "Coming". Since the time intervals were randomised the method of limits could not be used, so instead the changing probability of response to a stimulus of constant strength was used as an index of change of threshold. Each session began with a practice period during which a number of fairly strong stimuli were given to accustom the subject to the range of time intervals to be used, and an approximate threshold reading was taken by the method

of limits. This was used to select two stimulus strengths separated by 20-40 microamps. which would, it was hoped, give frequencies of positive response just above and just below 50 per cent. In the course of each session about 240 stimuli were administered; they varied in intensity between three values, 'strong'(S), 'weak'(W), and zero, randomly ordered by reference to random number tables (Fisher & Yates, 1948). There were about 40 per cent, each of 'strong' and 'weak' stimuli, and about 20 per cent. at zero intensity.

The midpoint of the range of time intervals given was always 3 seconds, but three different ranges were used, in the expectation that some ranges might show the effect, if it existed, more clearly than others. Only one or two well-practised subjects who had shown the monotonic change clearly in the earlier experiments were used, in the hope of thus getting more reliable and stable performance. In Condition (a) a range of 1.5 seconds was used, the intervals being rectangularly distributed between 2.25 and 3.75 seconds, the stimulus was a 100 msec. square pulse, and 3 sessions were obtained from two subjects. In Condition (b) the range was 1 second, from 2.5 to 3.5 seconds, the stimulus was a 30 msec. square pulse, and 4 sessions were spent on one subject. In Condition (c) the range was 0.5 seconds (2.75 to 3.25 seconds), the stimulus was a 30 msec. pulse, and one subject did 7 sessions.

### Results.

In treating the results each range was divided into five equal

subdivisions and for each sub-range the proportion of the W and S stimuli given which had been seen was calculated. Since the time intervals were not under the experimenter's control there was some unavoidable variation in the number of W and S stimuli falling in each sub-range. In combining the results from the individual sessions in each condition they were treated as follows: for each sub-range the total number of W (or S) stimuli given in all the sessions was recorded, and the percentage of this total seen,  $P_W$ , (or  $P_S$ ) was calculated. The standard error was taken as  $\sqrt{P_W Q_W / N_W}$  or  $\sqrt{P_S Q_S / N_S}$ . In this way the contribution of each session to a sub-range is weighted in accordance with its reliability. Since the total number of W stimuli,  $N_W$ , for each sub-range, did not always equal  $N_S$ , the arithmetical mean of the two percentages was taken to give a mean value for the percentage seen ( $P_M = \frac{1}{2}(P_W + P_S)$ ) for both stimulus strengths which would not be affected by the random variation in the distribution of the W and S stimuli. The standard error of  $P_M$  was estimated as  $S.E. = \frac{1}{2} \sqrt{P_W Q_W / N_W + P_S Q_S / N_S}$ , which is a slight under-estimation. The proportions of the W and S stimuli seen in the sub-ranges in the individual sessions were normalised by an angular transformation (Fisher & Yates, 1948), and an analysis of variance performed on them. For this purpose the random variation in the number of stimuli falling in each range and intensity subdivision in each session was disregarded.

The values of  $P_M$  for the three conditions are plotted as curves A, B, and C in Figure 6, and the results are given in Table 5.

Percentage Response

65

55

60

50

60

50

65

55

FIGURE 6

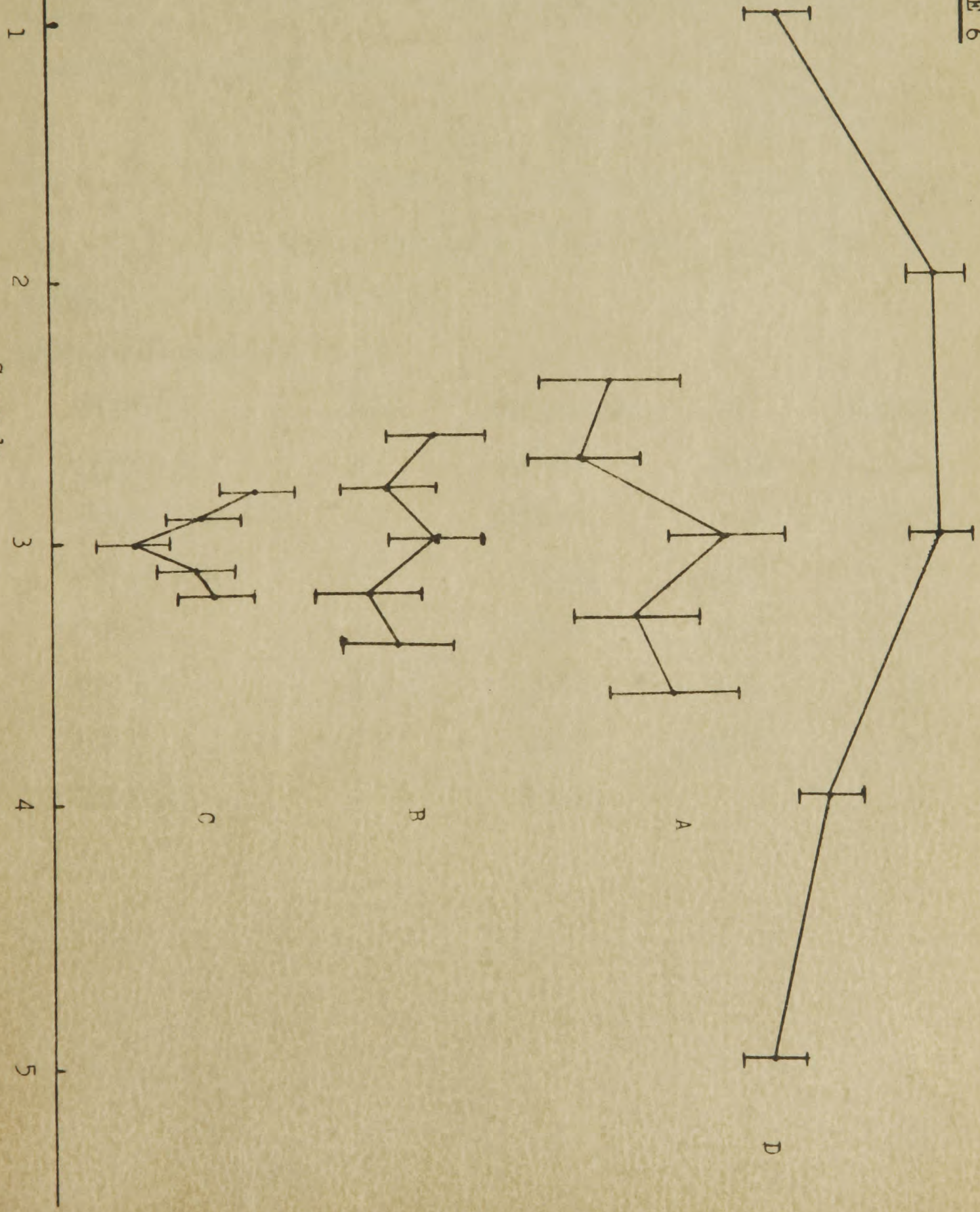


TABLE 5Condition (a)

Sub-range: 1(2.25-2.54) 2(2.55-2.84) 3(2.85-3.14) 4(3.15-3.44) 5(3.45-3.75)  
(seconds)

Percentage Seen (and S.E.)

$P_W$ :	31.3 $\pm$ 8.2	23.9 $\pm$ 6.3	40.0 $\pm$ 6.9	36.2 $\pm$ 7.0	43.8 $\pm$ 7.2
$P_S$ :	69.1 $\pm$ 7.1	72.6 $\pm$ 6.3	78.0 $\pm$ 5.9	67.9 $\pm$ 6.4	66.0 $\pm$ 6.7
$P_M$ :	50.2 $\pm$ 5.4	48.2 $\pm$ 4.4	59.0 $\pm$ 4.5	52.1 $\pm$ 4.8	54.9 $\pm$ 4.9

False Positives: 3.7 per cent.

Analysis of Variance

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Times (T)	4	42.71	0.64	N.S.
Intensities (I)	1	3497.04	52.33	<u><math>p &lt; 0.001</math></u>
Sessions (S)	2	32.46	0.49	N.S.
T x I	4	51.75	0.77	N.S.
T x S	8	16.23	0.24	N.S.
S x I	2	6.88	0.10	N.S.
Residual	8	66.83		

TABLE 5 (contd)Condition (b)

Sub-range: 1(2.50-2.69) 2(2.70-2.89) 3(2.90-3.09) 4(3.10-3.29) 5(3.30-3.50)  
(seconds)

Percentage seen (and S.E.)

$P_W$ :	35.3 $\pm$ 5.8	25.4 $\pm$ 5.5	34.7 $\pm$ 5.5	34.3 $\pm$ 5.6	35.8 $\pm$ 5.9
$P_S$ :	77.9 $\pm$ 5.0	80.6 $\pm$ 4.8	78.5 $\pm$ 4.6	68.8 $\pm$ 5.2	71.4 $\pm$ 6.0
$P_M$ :	56.6 $\pm$ 3.8	53.0 $\pm$ 3.7	56.6 $\pm$ 3.6	51.5 $\pm$ 3.8	53.6 $\pm$ 4.2

False Positives: 2.7 per cent.

Analysis of Variance

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Intensities (I)	1	7606.6	98.40	<u><math>p &lt; 0.001</math></u>
Times (T)	4	32.5	0.42	N.S.
Sessions (S)	3	233.8	3.02	N.S. ( $p < 0.1$ )
T x I	4	67.2	0.87	N.S.
T x S	12	109.4	1.42	N.S.
I x S	3	43.8	0.57	N.S.
Residual	12	77.3		

TABLE 5 (contd)

Condition (c)

Sub-ranges: 1(2.75-2.84) 2(2.85-2.94) 3(2.95-3.04) 4(3.05-3.14) 5(3.15-3.25)  
(seconds)

Percentage seen (and S.E.)

$P_W$ :	58.3 $\pm$ 4.4	45.5 $\pm$ 4.5	41.6 $\pm$ 4.0	46.7 $\pm$ 4.5	51.2 $\pm$ 4.4
$P_S$ :	72.9 $\pm$ 3.9	77.4 $\pm$ 3.8	71.0 $\pm$ 3.9	75.5 $\pm$ 3.7	73.7 $\pm$ 4.1
$P_M$ :	65.6 $\pm$ 2.9	61.5 $\pm$ 2.9	56.3 $\pm$ 2.8	61.1 $\pm$ 2.9	62.5 $\pm$ 3.0

False Positives: 1.9 per cent.

Analysis of Variance

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Intensities (I)	1	4742.95	111.73	<u>p &lt; 0.001</u>
Times (T)	4	73.18	1.72	N.S. (p < 0.2)
Sessions (S)	6	393.73	9.28	<u>p &lt; 0.001</u>
T x I	4	85.20	2.01	N.S. (p < 0.2)
T x S	24	67.09	1.58	N.S. (p < 0.2)
I x S	6	28.32	0.67	N.S.
Residual	24	42.45		

### Discussion

As in each condition there was no significant interaction between Times and Intensities,  $P_M$ , the mean frequency of response for both stimulus strengths, provides the best indication of the threshold changes in each case.

In Condition (a) the results show a rise in the probability of seeing the stimulus in the middle of the range as compared with lesser or greater time intervals. However this is not significant, only the two stimulus strengths producing a significant effect on analysis of variance. These preliminary results suggested the possibility that a 'warning' effect was present which might be more clearly shown if a smaller range of time intervals were used; but Condition (b) does not support this. Though the range here is approximately equivalent to sub-ranges 2-4 in Condition (a), the suggestion of a 'warning effect' curve is not repeated, and there is again no significant effect of time interval. In the even smaller range used in Condition (c) the absence of a significant effect of time interval is again found, while the shape of the curve given when the frequency with which the stimulus is seen is plotted against time interval (Curve C in Figure 6) is just the opposite of what would be expected on the 'warning effect' hypothesis.

The failure to find, in these 14 sessions, any significant effect of warning interval suggested that it would be unprofitable to

continue the experiment further. It may be that no drop in threshold in the vicinity of the mean time of arrival of the stimulus could be shown because no such drop occurs, or because the experiment was not well designed to reveal it, the variation in time interval necessarily required by the technique used so altering the form of the threshold-depression that a detectable fall did not occur over the range of times used. Thus it appears that the direct approach tried in this experiment is unlikely to prove useful in investigating the hypothetical threshold drop; and more indirect procedures must be sought.

On the 'arousal response' hypothesis this experiment should have shown a fall in threshold at the shorter end of each range. However, as this would not have been very great with these short ranges, the failure to find it may reflect no more than the insensitivity of the technique used.

#### 3.4 Randomized time intervals.

Though the use of a randomized range of time intervals has not proved successful in testing one difference between the implications of the two hypotheses, it provides a technique for testing the second difference mentioned in Section 3.1, namely that a 'warning effect' should be disrupted, but an 'arousal response' unaffected if the subject were not informed what interval would intervene between the stimuli on each trial. In the experiments in which a monotonic change of threshold has been found, information about the length of the inter-stimulus interval has been

available to the subject, since he was informed of the interval before each threshold determination by the method of limits, which also afforded him instances of the time interval in the presentation of the initial suprathreshold stimuli which introduced the descending series used.

Since the major part of the fall in threshold shown in curves A, B, and C in Figure 5 takes place between 1 and 4 seconds, randomizing the distribution of time intervals over a range equal to or greater than this should produce a curve showing much the same monotonic changing relation of threshold inter-stimulus interval, if the threshold curve does no more than record the decay of a change in the state of the afferent systems resulting immediately from the onset of the accessory stimulus. However if the effect found is due to the judicious employment by the subject of information about when to expect the stimulus, depriving him of that information should result in a constant threshold level at all intervals, i.e. the accessory stimulation should become irrelevant to his performance.

#### Experiment 6.

The object of this experiment was to see whether the threshold curve obtained in Experiments 1, 2, and 4, would be shown if the inter-stimulus intervals were randomized over a sufficient range instead of being known to the subject. In view of the relation between alertness and reaction time (Woodworth & Schlosberg, 1955) it seemed plausible that if the threshold change were due to an 'arousal response' a parallel effect on

reaction time might be found, so this was recorded as well.

### Apparatus and Procedure.

The apparatus used was the same as in Experiment 5, with the addition of a second bank of Dekatrons, which was started by the closing of the fourth relay, and stopped by the subject pressing a reaction key, thus measuring his reaction time. The procedure was also similar. The stimulus was a 30 msec. square pulse, the 'minimum' was 0.5 seconds, and the 'range' 5 seconds, so that the intervals were rectangularly distributed between 0.5 and 5.5 seconds. Five subjects did 10 sessions, one doing the experiment four times, another three times, and the others once each.

### Results.

The results were treated as in Experiment 5. They are given in Table 6. Since the analysis of variance showed no significant interaction between Times and Intensities, the mean percentage of positive response ( $P_M$ ) and its standard error were calculated, and are plotted as curve D in Figure 6.

The mean reaction times, which are given for positive responses only, were obtained by averaging the mean values for the sub-ranges in the different sessions without weighting; the standard error calculated excludes intersubject variability. Occasional reaction times were of the order of several seconds in duration; presumably they represented second thoughts or false positive responses. Any reaction time longer than 1 second was arbitrarily omitted when calculating the means. These are based on 1,075 responses, 11 having been omitted.

TABLE 6

Sub-range: 1(0.50-1.49) 2(1.50-2.49) 3(2.50-3.49) 4(3.50-4.49) 5(4.50-5.50)  
(seconds)

Percentage seen (and S.E.)

$P_W$ :	37.3 $\pm$ 3.5	47.4 $\pm$ 3.4	53.0 $\pm$ 3.5	39.4 $\pm$ 3.6	34.0 $\pm$ 3.3
$P_S$ :	64.8 $\pm$ 3.4	78.3 $\pm$ 3.0	72.9 $\pm$ 3.2	69.1 $\pm$ 3.3	65.6 $\pm$ 3.5
$P_M$ :	51.0 $\pm$ 2.5	62.9 $\pm$ 2.3	63.0 $\pm$ 2.4	54.3 $\pm$ 2.5	49.8 $\pm$ 2.4

False Positives: 5.6 per cent.

Reaction Time (seconds), S.E., and number of positive responses.

W:	0.453 0.014	0.406 0.008	0.414 0.007	0.421 0.008	0.467 0.014
	70	100	104	69	71
S:	0.455 0.009	0.415 0.009	0.400 0.006	0.412 0.008	0.428 0.007
	123	148	140	133	117

Analysis of variance of an angular transformation of the percentage of response.

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Times (T)	4	284.40	4.18	<u>p &lt; 0.01</u>
Intensities (I)	1	7857.05	115.58	<u>p &lt; 0.001</u>
Sessions (S)	9	755.48	11.11	<u>p &lt; 0.001</u>
Residual	85	67.98		

TABLE 6 (contd)

Analysis of variance of a logarithmic transformation of the reaction times

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Times (T)	4	0.00874	11.68	<u>p &lt; 0.001</u>
Intensities (I)	1	0.00292	3.90	N.S. (p < 0.1)
Sessions (S)	9	0.02774	37.07	<u>p &lt; 0.001</u>
T x I	4	0.00147	1.96	N.S.
T x S	36	0.00123	1.64	N.S. (p < 0.1)
S x I	9	0.00105	1.41	N.S.
Residual	36	0.00075		

In the analysis of variance of an angular transformation of the percentages of response, the first order interaction terms showed no suggestion of significance. Their sums of squares and degrees of freedom were therefore combined with the residual to give an improved estimate of the remainder variance for testing the significance of the main factors.

### Discussion

If the 'arousal' hypothesis were correct the threshold should rise monotonically from 0.5 to 5.5 seconds after the inception of the accessory stimulus, and the reaction time become, perhaps, progressively longer across the same range. This, however, is not found, so that the result of this experiment provides evidence against this hypothesis and in favour of the assumption that information about the length of the inter-stimulus interval is necessary for the monotonic change in threshold to occur. This evidence is not conclusive, however. Firstly, the range used goes down to 0.5 seconds, while in the earlier experiments no interval of less than 1 second was employed. It is conceivable that had smaller intervals been used previously the threshold might have been found to rise sharply below 1 second. If that were the case, the curve found here might be considered compatible with it, at least for the strong stimuli. To exclude this possibility the form of the threshold curve for intervals of less than 1 second must be determined. Secondly, it might be argued that informing the subject that randomization would occur might have induced a 'set' having the effect of preventing an arousal reaction that would otherwise have been

made. To make this more precise, it might be supposed that the initial 'arousal response' to the accessory stimulus would normally tend to habituate and disappear (Humphrey, 1933; Pavlov, 1927), but that this is prevented, when known inter-stimulus intervals are used, by a 'reinforcing' effect of having the stimulus arrive at about the expected time. When this reinforcement is absent, as in the present experiment, the habituation would not be opposed, the response would disappear, and the threshold-lowering effect of short intervals would, therefore, not be found. This sequence could be compared with the habituation of the cochlear nucleus response to a repeated auditory stimulus described by Hernández-Peón, Jouvét, & Scherrer (1957), and the reversal of this habituation when an event of significance to the animal, a shock to the forepaw, was made to follow the auditory stimulus regularly. If this were happening it might be expected that the smaller the range of intervals used, the less the habituation would be, and the more the result would approximate the appropriate segment of the threshold curve. No effect of this sort is shown by the results of Experiment 5, but this could be due to limitations in the sensitivity of the method. Thus, though the experiment favours the 'warning' hypothesis, it needs further support.

Though the monotonic effect has not been obtained, it is of interest that another effect is shown, both by thresholds and reaction times; the percentage of positive response is higher, and the mean reaction time lower, toward the centre of the range, than at the two extremes. In

both cases this effect is significant on analysis of variance. Variation between sessions affects both probability of response and mean reaction time, and variation of stimulus intensity affects the former but does not have a clearly significant effect on the latter.

Thus it appears that in this experiment occurrence of the accessory stimulus has favoured stimuli towards the middle of the range, tending to give both lower thresholds and reduced reaction times there. An explanation for this effect on the threshold can be suggested which is an extension of the presuppositions of the 'warning' hypothesis. We need suppose no more here than that on each trial the subject's expectation of when the stimulus will occur is based on his experience with the immediately preceding trials, and that after each trial on which he has detected a stimulus he alters his expectation in the direction of his last error; thus, if he expected the stimulus too soon, he will now expect it later; if too late, sooner. Since the intervals used were rectangularly distributed, if his expectation on any trial is not at the midpoint of the range (3 seconds here), and is thought of as dividing the range into two parts, the stimulus is more likely to come in the larger subdivision of the range which includes the midpoint than in the lesser subdivision, and the more extreme the subject's expectation the more likely this is. Thus he will more often adjust his expectation towards the midpoint than away from it, so that it should tend to vary about that region. If threshold shifts reflect changes in the critical value used in a decision procedure, and these changes can be the

consequence of changes in the subjective probability of the stimulus at different times, the threshold will be lower when the stimulus is expected. Or if a 'preparatory adjustment' is made when the stimulus is expected, this will have the same effect. Thus the threshold curve found in this experiment would reflect the distribution of the subjects' expectancies or subjective probabilities during the experiment. It has been suggested that reaction time is affected by the subject's expectancy (Mowrer, 1940, 1941; Mowrer, Rayman & Bliss, 1940; Hick, 1952) so that the same variation in subjective probability that is assumed to account for the threshold changes could be related to the very similar changes in reaction time found. In instructing the subjects no stress was laid on the reaction time; they were not even informed that it was being recorded. This was to avoid the possibility of reducing the accuracy of their threshold judgments. This may be related to the failure to find significant evidence of the customary effect of stimulus strength on the length of the reaction time (Woodworth & Schlosberg, 1955), and adds to the interest of the significant effect of the time interval shown.

### 3.5 Conclusion

In this chapter possible explanations for the monotonic change in threshold found earlier have been considered. An effect peculiar to the phosphene or to vision was excluded by using an auditory critical stimulus. This left two main alternatives: a 'ballistic' physiological effect or 'arousal response', or a 'warning effect', dependent on the subject making

use of the information about the length of the time interval given him. Two differences between the implications of these hypotheses were pointed out: Experiment 5 attempted to discover whether a 'threshold depression' about the time of arrival of the stimulus or a monotonic change, could be shown, but this effort was unsuccessful; possibly because the dispersion of the intervals in this experiment altered the phenomenon it was desired to observe; Experiment 6 tested the effect of not giving the subject exact information about the length of the inter-stimulus interval. As expected on the second hypothesis this resulted in the disappearance of the curve previously found; however this result is not conclusive and more evidence is required if one of the hypotheses is to be rejected.

which the subject can estimate the point at which the interval will end. If the data which have been obtained would be shown to exhibit a definite relation of this sort, this would tend to support the 'threshold' hypothesis, whereas a failure to find this would tend to support the 'monotonic' hypothesis. It is interesting to observe that in some of the experiments in which the subject is given information about the length of the inter-stimulus interval, the results will not be very different from those obtained when no such information is given. This suggests that the 'threshold' hypothesis may be correct, and that the 'monotonic' hypothesis is incorrect. The 'threshold' hypothesis can be tested by the 'step' method, which is described in the appendix.

The difficulty in doing this is that the subject must be given

## CHAPTER 4

## THE DISCRIMINATION OF TEMPORAL DURATION

4.1 Introduction.

The attempt made, in Experiment 5, to record directly changes in threshold level expected to occur about the time of arrival of the stimulus proved unsuccessful, and while Experiment 6 provided some evidence for the 'warning' hypothesis, it was not conclusive, so that it is necessary to test further consequences of the hypotheses. The 'warning' hypothesis would account for the threshold curve by referring to the inexactness of time judgments on the assumption that the fall in threshold at a given inter-stimulus interval may be a function of the accuracy with which the subject can estimate the moment at which the interval will end. If the data which have been obtained could be shown to exhibit a simple relation of this sort, this would tend to support the 'warning' hypothesis, whereas a failure to find this would cast some doubt on it; it is therefore of interest to examine this possibility. It seems initially plausible to assume that the variation in the accuracy of time estimation which may underlie the effect of the inter-stimulus interval on threshold judgments will not be very different from the variation which is shown in experiments on temporal discrimination. Thus the problem becomes one of seeing whether the threshold curves can be related to the 'just noticeable difference' for temporal duration.

One difficulty in doing this is that the exact relation between

the differential threshold and the length of the time interval does not seem clearly established, various experimenters having obtained conflicting results. The early investigators wished to establish whether Weber's law applied to differential sensitivity to temporal duration. In some cases, however, they found a minimum value for the Weber fraction at a short time interval. Thus Mach obtained a minimum of 5 per cent. at 0.375 seconds, and Vierordt a minimum of 3 per cent. between 1 and 1.5 seconds (Boring, 1942). Glass, on the other hand, found that Weber's law held for intervals between 0.7 and 15.4 seconds (Woodrow, 1930). Fraisse (1957) criticises the early investigators for using unstandardised methods and too few subjects. However a minimum value for the Weber fraction appears again in an experiment by Woodrow (1930), who used the method of reproduction for durations between 0.2 and 30 seconds. He found a Weber fraction of 10-11 per cent, for 0.2 seconds; this fell to a minimum of 7-8 per cent. for 0.6 seconds and then increased for durations up to 4 seconds, remaining constant at a value of 16-17 per cent. for time intervals greater than this. Although there is some support for this finding (Woodrow, 1951), Gilliland & Humphreys (1943) found that their subjects' variability declined as duration increased over a range of 9-180 seconds, while Fraisse (1948) reports that the Weber fraction passes through a maximum towards the middle of the range of times presented and decreases towards the extremes. He used two ranges, 0.2-1.5 seconds and 0.3-12.0 seconds; in the first the variation was between 12 and 14 per cent, in the second between 12 and 20 per cent.

The various divergences may possibly be due to procedural differences, which readily affect temporal judgments. It is not clear, therefore, what function should be taken to hold in considering the experiments in Chapters 2 and 3. Thus it seems necessary to determine the relation between the just noticeable difference and time interval in an experimental situation as similar as possible to those in which the effect of the inter-stimulus interval on threshold level was examined.

#### 4.2 Reproduction of time intervals.

##### Experiment 7

The experiment was designed to be similar in procedure to Experiments 1, 2, and 4, save that instead of being required to report whether or not a stimulus had been presented, the subject had to terminate a comparison time interval after a standard had been presented. This was based on the assumption that the use of descending series of the method of limits had had the effect of providing the subject with information about the time interval on those trials in which the accessory stimulus was followed by a suprathreshold critical stimulus, and that he could later make use of this information to estimate the end of the interval when the critical stimulus was at threshold level. The same set of times was used, with 0.5 seconds added in order to include an interval in the region of Woodrow's minimum, and because it was intended at a later stage to examine the threshold curve for times below 1 second. In order to avoid any artefact resulting from the particular stimuli chosen to bound

an empty interval (Woodrow, 1951; Kowalski, 1943), a continuous stimulus was used.

#### Apparatus and procedure.

Each trial was begun by the experimenter pressing the operating button. This closed a self-holding relay which simultaneously caused a neon bulb to light up, a bank of Dekatrons to start counting 1,000 c.p.s. pulses from an oscillator, and triggered a 2803U timing circuit. At the conclusion of the standard interval this timing circuit extinguished the light, stopped the Dekatrons, and started a second timing circuit. After an interval of 2 seconds this switched the bulb on again, and simultaneously started a second bank of Dekatron counters. The light could be switched off and this bank of Dekatrons stopped by the subject depressing a reaction key. The Dekatrons recorded the standard interval and the reproduction on each trial to the nearest millisecond.

The subject sat 18 inches in front of a white screen, 18 inches by 20 inches in size, with the neon bulb at its centre. The experimenter announced each trial by saying 'coming'. The neon lamp then went on for the standard interval, off for two seconds, and on again, and the subject terminated it by tapping the reaction key. Six intervals were used, 0.5, 1, 2, 3, 4, and 9 seconds, in that order. A series of eight successive presentations of each interval was given before passing on to the next interval. A block of 6 intervals given 8 times each was repeated 3 times in a session of about an hour, so that the subject reproduced each interval

24 times in the course of the session. Eight sessions were obtained, from eight of the subjects who had taken part in the earlier experiments.

### Results

The length of the standard presented on each trial was recorded, and as a check on the accuracy of the apparatus the mean standard, and its standard deviation, for each series of 8 trials was calculated, and these were averaged for the 8 sessions. The average means and standard deviations for the experiment as a whole, in seconds, were:

<u>Means:</u>	0.505	1.007	2.000	3.004	4.010	9.001
<u>S.D.:</u>	0.0003	0.0025	0.0053	0.0037	0.0067	0.0142

As the standard deviations are less than 1-2 per cent. of those given by the subjects' reproductions they are ignored in considering the latter.

The mean and standard deviation of each series of 8 reproductions were calculated and the average values for the eight subjects are given in Table 7. They are given separately for the first, second, and third blocks of judgments and for the sessions as a whole. The mean values of the standard deviations for the whole session were taken as the just noticeable differences ( $\Delta T$ ), and Weber fractions were calculated as the ratio of the j.n.d. to  $T_m$ , where  $T_1$ ,  $T_2$ ,  $T_3$ , and  $T_m$  represent the mean reproductions for the first, second, and third blocks, and for the session as a whole. The percentage increase of the means between the first and third blocks was calculated as  $100(T_3 - T_1)/T_1$ . The best fitting regression

line for the 'Weber function'  $\Delta T/(T+a) = k$ , was fitted to the mean standard deviations and reproductions for the whole experiment, and its goodness of fit tested by analysis of variance. The constant errors of the subjects' mean reproductions for each block were expressed as percentages of the standard and their variance analysed. An analysis of variance was also performed on the standard deviations, after they had been expressed as proportions of the corresponding time intervals. This was done since if Weber's law in its classic form fitted the results exactly, no significant effect of time interval would then be found.

#### Discussion.

The Weber fractions do not fall to a minimal value at the short intervals, as was found by Mach, Vierordt, Woodrow, Blakely, and Stott (Boring, 1942; Woodrow, 1930, 1951). They are smaller for the longer intervals and larger for the shorter intervals than the results given by Woodrow (1930), and they show no tendency to decline below 4 seconds. Instead they tend to increase as the time interval decreases, producing a significant effect of Times in the analysis of the variance of the standard deviations. The Weber function relating the standard deviations to the mean reproductions is highly significant on analysis of variance while the deviations from linear regression are not significant. These results are very similar to those found with many sensory continua: an increase of the Weber fraction for small values of the stimulus to which a Weber function which includes the constant 'a' can be fitted. The

TABLE 7

Reproductions of time intervalsStandard Intervals (seconds):

	<u>0.5</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>9</u>
<u>First Block</u>						
Mean	0.555	0.940	1.802	2.836	3.914	8.545
S.D.	0.090	0.114	0.204	0.363	0.406	0.748
<u>Second Block</u>						
Mean	0.593	1.137	1.971	3.008	3.974	8.837
S.D.	0.102	0.113	0.259	0.316	0.378	0.614
<u>Third Block</u>						
Mean	0.625	1.105	2.037	3.086	4.130	9.055
S.D.	0.122	0.149	0.241	0.287	0.343	0.619
<u>Whole Session</u>						
Mean	0.591	1.061	1.937	2.977	4.006	8.812
S.D.	0.104	0.125	0.235	0.322	0.376	0.661

Weber fractions (per cent.):

	17.6	11.8	12.1	10.8	9.4	7.5
--	------	------	------	------	-----	-----

Increase of Means between First and Third Blocks:

Seconds:	0.070	0.165	0.235	0.250	0.216	0.510
----------	-------	-------	-------	-------	-------	-------

Per cent.:	12.6	17.6	13.0	8.8	5.5	6.0
------------	------	------	------	-----	-----	-----

Weber function:  $\Delta T = 0.068 (T + 1.268)$  ..... (15)

Analysis of Variance (Regression):

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Linear regression:	1	1.6460	63.31	<u>p &lt; 0.001</u>
Deviations from linear regression:	4	0.0080	0.31	NS
Residual:	42	0.0260	-	

TABLE 7 (contd)Analysis of Variance (Percentage Constant Errors)

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Times (T)	5	1369	24.89	<u>p &lt; 0.001</u>
Subjects (S)	7	1890	34.36	<u>p &lt; 0.001</u>
Blocks (B)	2	1420	25.82	<u>p &lt; 0.001</u>
T x S	35	217	3.95	<u>p &lt; 0.001</u>
T x B	10	110	2.00	<u>p &lt; 0.05</u>
S x B	14	154	2.80	<u>p &lt; 0.01</u>
Residual	70	55	-	

Mean Constant Error: + 2.7 per cent.

Analysis of Variance (Standard Deviations):

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Times (T)	5	0.052522	34.35	<u>p &lt; 0.001</u>
Subjects (S)	7	0.035744	23.38	<u>p &lt; 0.001</u>
Blocks (B)	2	0.001497	0.98	NS
T x S	35	0.004471	2.92	<u>p &lt; 0.001</u>
T x B	10	0.002894	1.89	NS (p < 0.1)
S x B	14	0.003684	2.41	<u>p &lt; 0.01</u>
Residual	70	0.001529	-	

analysis of the variance of the standard deviations also shows significant effects of Subjects, and of the interactions which include Subjects, but no effect of the Blocks.

The appearance of a minimal value of the Weber fraction in some experiments but not in others suggests that this might depend on some feature of the experimental procedure. A factor which might possibly be suggested is the mode of response required. In the method of reproduction, as used by Woodrow and Vierordt, the subject had to delimit an interval by making two taps. Possibly a fairly rapidly repeated movement may be organised by the motor system as a 'rhythmic unit' with less variability than is contributed to the reproduction by two separately organised movements separated by a longer interval. In the present experiment the comparison interval began independently of the subject, and he was only required to terminate it, so that the occurrence of favoured rhythms of response could not affect the result. However a minimum has been found using the method of comparison (Woodrow, 1951), and is not always shown when the subject is required to make two taps (Fraisse, 1948).

The means of all the reproductions show two classic related features of temporal estimation (Boring, 1942; Woodrow, 1951): there is a transition point, the 'indifference interval', between intervals which are over-estimated and intervals which are under-estimated, and it is the shorter intervals which are over-estimated and the longer which are under-estimated. If we disregard the small over-estimation at 4 seconds, this

point would appear to lie between 1 and 2 seconds here.

Although early studies provided widely differing values for the indifference interval (Woodrow, 1934), most weight was given to studies suggesting values between 0.6 and 0.8 seconds (Fraisse, 1957). Wundt believed that this interval represented a specific physiological unit of time, corresponding to the time required for the process of association or for the movement of a leg in rapid walking. Hollingworth proposed that it was a consequence of the 'central tendency of judgment', the shorter intervals in a presented range being over-estimated and the longer under-estimated, with the indifference interval lying approximately at the median (Fraisse, 1957). However Woodrow (1934) tested this explanation by using a separate group of subjects for each time interval in a range. He obtained a value of approximately 0.6 seconds, with over-estimation of shorter and under-estimation of longer intervals. As no subject was presented with more than one interval he considered this excluded Hollingsworth's explanation. Fraisse (1948) presented a range of intervals in random order for reproduction by the subject. He used ranges of 0.2 - 1.5 seconds or of 0.3 - 12.0 seconds; with the first he obtained an indifference interval of 1.138 seconds, with the second 3.65 seconds, with the familiar pattern of under- and over-estimation in each case. He believes that this finding can be explained in terms of the central tendency of judgment, and that this does not conflict with Woodrow's finding, provided it is assumed that in the absence of a range of presented

intervals the indifference interval is determined by the average effect of the durations "perçus dans toutes les occasions de la vie". This is consistent with the value of 0.6 seconds obtained by Woodrow since Fraisse believes that only durations of about 0.4 - 2 seconds are "perceived".

Some support is given to this explanation by a number of experiments which employ longer durations and find values of the indifference interval which greatly exceed that found by Woodrow. Thus Gilliland & Humphreys (1943), using a range of 9 - 180 seconds, found over-estimation of short and under-estimation of long intervals, Clausen (1950) found that his (schizophrenic) subjects over-estimated 5 seconds and under-estimated 10 and 15 seconds, and Hirsh, Bilger, and Deatherage (1956) obtained an indifference interval of 8 seconds with a range of 1 - 16 seconds. To explain this last finding Woodrow reverts to the hypothesis that the indifference interval may tend to be in the middle of the range of intervals used (Hirsh, Bilger, and Deatherage, 1956). Weber (1933) notes that Swift & McGeoch, using time intervals from 30 seconds to 10 minutes, and the method of estimation, found a tendency in all individuals to over-estimate the times between 30 seconds and 5 minutes.

The present experiment affords an opportunity to test Hollingworth's explanation, since each session consists of three similar blocks. If the range of durations presented was exerting a systematic effect on the judgments in the manner suggested by the 'central tendency'

hypothesis there should be an increased tendency to over-estimate the short, and under-estimate the long intervals in successive blocks. However, when the mean reproductions for each block are considered, the hypothesis is not supported. The mean reproductions of the three shortest intervals do increase in successive blocks, as required by the hypothesis, but so do the reproductions of the longer intervals. The only difference is that the rate of increase is apparently greater for the shorter intervals. The effect of this tendency for reproductions to increase at different rates is that whereas the indifference interval in the first block lies between 0.5 and 1 second, in the second block there is such a point between 1 and 2 seconds (with an over-estimation at 3 seconds), and in the third block all the intervals are over-estimated. That this tendency for the reproductions to lengthen is highly significant is shown by the effect of Blocks in the analysis of the variance of the constant errors. The greater rate of increase for the shorter intervals is confirmed by the significant interaction between Times and Blocks. An increase in response time during the course of the session, resulting perhaps from fatigue, could cause a proportionately greater rate of lengthening for shorter intervals. However this does not seem likely to be the explanation for the results found here, since it would tend to add a constant quantity to each reproduction, whereas the absolute increases in the mean reproductions are not of the same order for different time intervals, but increase with the time interval.

If this pattern of change should prove to be a general feature of repeated time judgments it might provide a simple account of a number of the features of temporal discrimination experiments. Assuming only that there is an initial tendency to under-estimate all intervals, the lengthening of reproductions as the session proceeds, at a greater rate for the shorter intervals, would fairly soon result in the shortest intervals being over-estimated, and as the session continued the indifference interval would shift towards the longer intervals. Experiments in which long time intervals are used will tend to take longer, and therefore to produce larger values of the indifference interval. Thus the ubiquity of the finding that it is the shorter intervals that are over-estimated and the longer that are under-estimated, and the variable values of the indifference interval shown in different experiments would both be results of this same pattern of change. Individual subjects might vary in the rate at which their reproductions lengthen, or in the degree of their initial under-estimation, which would be consistent with the frequent finding that some subjects under-estimate and others over-estimate all the intervals (Woodrow, 1934; Fraisse, 1948).

#### 4.3 Production of time intervals.

A possible objection to relating the results of Experiments 1, 2, and 4, to the Weber function obtained in Experiment 7 is that the conditions of the latter experiments may have favoured reproductions of

greater accuracy than was possible in the earlier experiments. In Experiment 7 the subject was asked to reproduce time intervals, which was not the ostensible task in the threshold experiments; the interval between standard and reproduction was 2 seconds, whereas the intervals between successive trials were usually greater than this in the threshold estimation experiments; and in the latter case, if the subject missed or was uncertain of the critical stimulus this in effect deprived the next trial of an immediately preceding standard. This makes it necessary to consider whether using less favourable conditions would produce any important deterioration in the subjects' performance.

If the standards were omitted altogether, the subject being told the length of the interval only, we would have the method of production. Gilliland & Humphreys (1943) compared the methods of production, reproduction and estimation using the same subjects and time intervals. They did not find a clearly significant effect of the method used and conclude that reproduction "is perhaps easier" than the other two methods, but they do not give the exact differences found. Their results showed much more variability than was found here: the "per cent error" for 9 seconds was 31.2 per cent for adults and 47.5 per cent for children. It is possible that a significant effect could have been obscured by the variability. Clausen (1950) compared these three methods, using schizophrenic subjects. He found that reproduction gave less consistent results than the other two methods, but he considered the means only and

did not give the standard deviations.

Since these experiments offer no clear indication of the probable effect of omitting the standard time intervals in the present situation, it seemed necessary to investigate this experimentally.

### Experiment 8

This experiment was intended to be as similar to Experiment 7 as possible, save that the method of production was used instead of the method of reproduction.

#### Apparatus and procedure

The apparatus used in Experiment 7 was modified so that the first timing circuit was by-passed. Each trial began with the experimenter saying 'coming'. Two seconds later the neon light was switched on and a bank of Dekatrons simultaneously triggered. The light was switched off and the counters stopped when the subject depressed the reaction key.

Each session consisted, as in the previous experiment, of three blocks of 8 productions of each of the six intervals (0.5, 1, 2, 3, 4, and 9 seconds, in that order). Before each series of 8 time productions the subject was told what time interval he was to attempt to produce. Three sessions were obtained from 3 subjects, two of whom had served in the previous experiment.

### Results

The results of this experiment have been treated in the same way

TABLE 8

Production of Time IntervalsStandard Intervals (seconds):

	<u>0.5</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>9</u>
<u>First Block</u>						
<u>Mean</u>	0.616	1.210	2.017	3.446	4.486	10.420
<u>S.D.</u>	0.087	0.168	0.211	0.315	0.322	0.821
<u>Second Block</u>						
<u>Mean</u>	0.709	1.211	2.397	3.399	4.526	10.715
<u>S.D.</u>	0.133	0.233	0.229	0.353	0.323	0.806
<u>Third Block</u>						
<u>Mean</u>	0.725	1.454	2.370	3.611	4.625	10.554
<u>S.D.</u>	0.140	0.221	0.250	0.303	0.328	0.513
<u>Whole Session</u>						
<u>Mean</u>	0.683	1.292	2.261	3.485	4.546	10.563
<u>S.D.</u>	0.120	0.207	0.230	0.324	0.324	0.715

Weber fractions (per cent):

	17.6	16.0	10.2	9.3	7.1	6.8
<u>Increase of Means between first and third blocks:</u>						
Seconds:	0.109	0.244	0.353	0.165	0.139	0.134
Per cent:	17.7	20.2	17.5	4.8	3.1	1.3

Weber functions  $\Delta T = 0.057 (T + 1.770)$  ..... (16)

TABLE 8 (contd)

Analysis of variance (Regression):

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Linear Regression:	1	0.6399	68.08	<u>p &lt; 0.001</u>
Deviations from linear regression:	4	0.0027	0.29	NS
Residual:	12	0.0094	-	

Analysis of variance (Percentage Constant Errors):

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Times (T)	5	827	8.89	<u>p &lt; 0.001</u>
Subjects (S)	2	166	1.78	NS (p < 0.2)
Blocks (B)	2	704	7.57	<u>p &lt; 0.01</u>
T x S	10	365	3.92	<u>p &lt; 0.01</u>
T x B	10	138	1.48	NS
S x B	4	801	8.61	<u>p &lt; 0.001</u>
Residual	20	93	-	

Mean Constant Error: + 20.9 per cent.

Analysis of variance (Standard Deviations):

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Times (T)	5	0.04178	15.42	<u>p &lt; 0.001</u>
Subjects (S)	2	0.01675	6.18	<u>p &lt; 0.01</u>
Blocks (B)	2	0.00438	1.62	NS
T x S	10	0.00721	2.66	<u>p &lt; 0.05</u>
T x B	10	0.00217	0.80	NS
S x B	4	0.00046	0.17	NS
Residual	20	0.00271	-	

TABLE 8 (contd.)

Combined Data of Experiments 7 and 8

Common Weber Function:  $\Delta T = 0.061(T + 1.571)$  ..... (17)

Analysis of variance (Regression for combined data):

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Common linear regression	1	2.0675	102.76	<u>p &lt; 0.001</u>
Disparity of the common regression from the two individual regressions:	2	0.0144	0.72	NS
Residual:	62	0.0201	-	

as in Experiment 7. The Weber function,  $\Delta T/(T+a)=k$ , has been calculated, taking the mean standard deviations for the whole experiment as ' $\Delta T$ ', and relating them to the mean productions,  $T_m$ . The significance of this regression was tested by analysing the variance of the standard deviations. The 'constant errors' between the means of the series of productions of a time interval and the required intervals were expressed as proportions of the latter, and subjected to analysis of variance, as were also the standard deviations. Finally a Weber function was calculated for the data of Experiments 7 and 8 taken together, and the variance of the standard deviations from the two experiments was analysed to see whether there was a significant disparity between the common regression function and the individual functions calculated for the two experiments separately. The results are given in Table 8.

### Discussion

These results are similar in many respects to those obtained in Experiment 7. The Weber fractions are very much like those previously found. There is again a constant increase as the time interval shortens, rather than a decline to a minimal value, as found by Woodrow (1930). The Weber function relating the standard deviations to the time intervals is similar to that found in the earlier experiment, differing mainly in having a larger value for 'a', and there is again no evidence of departure from linearity of regression. Analysis of variance shows no evidence that the Weber functions for the two experiments differ

significantly from the combined function calculated for both sets of data taken together. Though it is possible that a difference might have been shown had more subjects been examined, it does seem that the subjects' variability is not very different whether they are reproducing standard intervals or producing time intervals without a standard. This result serves to demonstrate that the temporal information available to the subjects in Experiments 1 - 4 probably provided a sufficient basis for temporal discrimination by them.

The mean productions show the same tendency to increase in successive blocks that was found when the method of reproduction was used, and again the rate of increase is greater for short intervals than for long. Analysis of the variance of the constant errors shows a significant effect of Blocks, which is due to the lengthening of the productions. There are also significant interactions of Subjects with Times and with Blocks, but the greater rate of increase for the shorter intervals was not sufficient to produce a significant interaction of Times and Blocks, perhaps because too few subjects were used. There is again a significant effect of Times, due to the proportionately greater constant errors at the shorter intervals.

All the intervals were over-estimated in this experiment, the difference between the mean constant errors here and in Experiment 7 being highly significant ( $p < 0.001$ ). This suggests that the most important effect of presenting the standard may be not to increase accuracy, in the

sense of reducing variability, but to reduce constant errors. This can be no more than a suggestion, however, since the same group of subjects was not used for the two experiments. If the results for the two subjects who served in both experiments are compared they show greater productions than reproductions in every case. However the possibility of a practice effect as the cause of this cannot be excluded, so that this finding requires confirmation.

These two experiments have provided the information about the subjects' temporal discrimination which was needed for the further consideration of the results of the threshold estimation experiments. However they have also revealed an unanticipated pattern of change in the lengths of reproductions or productions during the course of a session which could, if it were found to be general, provide an account of some of the features of the classic problem of the 'indifference interval'. The design of these experiments was not well adapted to investigate this effect, however, since the order of presentation of the time intervals was not counter-balanced. If the tendency for reproductions or productions to lengthen were greater at the beginning of a session and decreased during its course, the intervals presented first would appear to lengthen most rapidly, so that the greater rate of increase which has been found for the shorter intervals could be an artefact of the design. This suspicion is strengthened by finding that whereas the mean increase between the First and Second Blocks in Experiment 7 is 8.0 per cent., and in

Experiment 8 is 6.1 per cent., the corresponding increases between the Second and Third Blocks are 2.5 and 4.7 per cent.

The experiment was interrupted, for non-experimental reasons, before a larger number of subjects could be examined, and when it was resumed it seemed that, in view of the interest of the 'lengthening effect', it might be more useful to re-design it to exclude the possibility of an artefact than to continue it as it was.

#### 4.4 A further experiment on temporal discrimination.

Besides the use of a counterbalanced order of presentation for the time intervals a number of other changes were made in the design of the next experiment. It was decided to use a different range of times. On Hollingworth's hypothesis the indifference interval is determined by the range of times. This was not supported by the increase at all time intervals shown in the two experiments, but another test would be provided by using a range from which the longer times are omitted, and seeing whether this results in lower values being obtained for the indifference interval than have been found with the longer range.

It is of interest to consider whether the increase in the means is like 'fatigue', presumably showing a similar course in repeated sessions, or a 'practice' or 'adaptation' effect which might show changes in successive sessions. In the hope of throwing some light on this it was decided to study a small number of subjects in a series of sessions.

In Experiments 7 and 8 a visual stimulus was used to fill the

time interval. In the threshold experiments both visual and auditory accessory stimulation was used, which raises the question whether this could have an effect on time estimation. Goodfellow (1934) investigated the difference limen for a duration of 1 second presented as an empty interval bounded by two clicks or by two dim flashes. Using the methods of limits, of reproduction, and of constant stimuli he found in each case that the differential threshold was smaller for auditory than for visual stimuli. The method of reproduction gave a small positive constant error when the stimulation was auditory, and a small negative error when it was visual. Clausen (1950) presented intervals of 5, 10, or 15 seconds filled with a buzzer or a dim lamp. He employed the methods of production, reproduction, and estimation, but in no case found a significant effect of the stimulation used on the means. Hirsh, Bilger, & Deatherage (1956) used intervals of 1-16 seconds filled with auditory or visual stimulation and the method of reproduction. They found that the mean reproductions did not indicate a difference in the subjects' responses for the two types of stimulation. However Goldstone, Boardman, & Lhamon (1959), investigating estimates of a second obtained by using the method of limits without a standard, found that a significantly smaller estimate was obtained when a continuous auditory stimulus was used than was given with continuous visual stimulation; and this has been confirmed by Goldstone, Jernigan, Lhamon, & Boardman (1959) using a whole wall as the visual stimulus. These results suggest that the differential threshold may be

affected by the stimulus used but that the constant error will not be much different, if a standard is used. But in the absence of a standard, it is possible that auditory stimulation may result in a considerably smaller estimate of the time interval. This made it seem advisable to use an auditory stimulus in the next experiment, to see whether any major feature of the judgments might depend on the stimulus used, and especially whether the increase in the means would occur in the absence of visual stimulation.

Finally, it seemed desirable to see whether the large constant error found in Experiment 8 was a feature of the method of production which would be shown if this method was compared with the method of reproduction used by the same subjects in the same sessions.

### Experiment 9

In this experiment the methods of production and of reproduction were both used, with a range of times from 0.25 to 3 seconds.

#### Apparatus and procedure

Four Z803U timing circuits were used to provide the standard time intervals for the 'Reproduction Series' (RS) condition. When the experimenter pressed the operating button a self-holding relay closed, which started the first timing circuit, and switched a neon light on. The light served as a warning to the subject, and it stayed on until he had responded. Two seconds later the timing circuit switched on a 500 c.p.s. tone at 60-70 db SL, provided by a Muirhead audio-oscillator, which was

heard by the subject through Brown Type K moving coil earphones. The tone lasted for a pre-selected interval of 0.25, 0.5, 1 or 3 seconds, and was succeeded by a silent interval which varied randomly between 0.5 and 2.5 seconds. The random interval was produced by the two timing circuits described in Experiment 5. It was followed by the tone commencing again, and a bank of Dekatron counters was simultaneously triggered. The subject terminated the tone and stopped the counters by pressing a reaction key. In the 'Production Series' (PS) condition the first two timers were by-passed and the standard omitted. In each trial the warning light was followed, after a random interval of 0.5-2.5 seconds, by the commencement of the tone which was to be terminated by the subject.

The subject sat in a darkened, sound-shielded room, wearing earphones and facing a white screen, 3 feet square, with the neon bulb at its centre. Two conditions were used, the 'Reproduction Series' in which a standard time interval was presented, and the 'Production Series' in which it was omitted. Each series consisted of 10 trials; the subject was informed of the time interval he was to produce or reproduce at the beginning of each series.

Two subjects were used, each for four sessions. After a preliminary practice session the four sessions were given on successive days, each lasting about an hour. The order of the time intervals was counter-balanced both within and between the sessions. For each time

interval a series of trials with the standard would be followed by a series without, or the reverse, and this order was constant throughout a given session. Thus one subject had the following series of trials in the first session: 0.25 sec. RS; 0.25 sec. PS; 0.5 sec. RS; 0.5 sec. PS; 1 sec. RS; 1 sec. PS; 3 sec. RS; 3 sec. PS; 3 sec. RS; 3 sec. PS; 1 sec. RS; 1 sec. PS; 0.5 sec. RS; 0.5 sec. PS; 0.25 sec. RS; 0.25 sec. PS. The second session began with 3 sec. PS, the third with 3 sec. RS, and the fourth with 0.25 sec. PS. The second subject had the same four sessions but in the reverse order.

### Results

The mean and standard deviation of each series of judgments were calculated, and the average values for the two blocks under the two conditions at the different standard intervals are given in Table 9. Weber Functions were calculated for the two conditions, and also for the combined data, and analyses of the variance of the standard deviations and constant errors are also given. In these analyses of variance all interactions which had a probability greater than 0.2 on the first analysis were combined to give an improved estimate of the residual variance. In half the sessions each 'Reproduction Series' at a given standard interval was followed by a 'Production Series', and in the other half the order was the reverse. Product-moment correlations between the means of such successive series for each standard interval are given in the table.

TABLE 9

Productions and Reproductions of Standard Intervals (Means and Standard Deviations):Standard Intervals (seconds):

	<u>0.25</u>		<u>0.5</u>		<u>1</u>		<u>3</u>	
	<u>Mean</u>	<u>S.D.</u>	<u>Mean</u>	<u>S.D.</u>	<u>Mean</u>	<u>S.D.</u>	<u>Mean</u>	<u>S.D.</u>
<u>Reproduction Series</u>								
<u>First Block:</u>	0.302	0.057	0.534	0.066	1.029	0.110	2.954	0.223
<u>Second Block:</u>	0.333	0.058	0.578	0.078	1.055	0.113	2.950	0.228
<u>Whole Session:</u>	0.317	0.057	0.556	0.072	1.042	0.112	2.952	0.226
<u>Production Series</u>								
<u>First Block:</u>	0.295	0.065	0.501	0.079	1.034	0.144	3.031	0.236
<u>Second Block:</u>	0.324	0.055	0.545	0.095	1.101	0.148	3.178	0.246
<u>Whole Session:</u>	0.309	0.060	0.523	0.087	1.067	0.146	3.105	0.241

Weber Fractions (per cent.):First Session:

<u>RS:</u>	24.3	13.0	13.9	9.1
<u>PS:</u>	19.5	19.0	18.4	8.8

All Sessions:

<u>RS:</u>	18.1	12.9	10.7	7.6
<u>PS:</u>	19.3	16.7	13.7	7.8

Increase of Means Between First and Second Blocks:

<u>RS: Seconds:</u>	0.031	0.044	0.026	-0.004
<u>Per cent:</u>	10.3	8.2	2.5	-0.1
<u>PS: Seconds:</u>	0.029	0.044	0.067	0.147
<u>Per cent:</u>	9.8	8.8	6.5	4.8

Sessions:

	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>
<u>Per cent:</u>	15.0	-3.6	-1.0	19.7

TABLE 9 (contd)

Mean Constant Errors (per cent.):

<u>Sessions:</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>All</u>
<u>Reproduction Series</u>					
<u>First Block:</u>	-0.1	8.5	12.6	7.4	7.1
<u>Second Block:</u>	7.8	9.4	16.4	18.9	13.1
<u>Whole Session:</u>	3.9	8.9	14.5	13.2	10.1
<u>Production Series</u>					
<u>First Block:</u>	-5.4	10.4	16.3	1.0	5.6
<u>Second Block:</u>	14.7	1.2	9.8	28.5	13.6
<u>Whole Session:</u>	4.6	5.8	13.0	14.7	9.5

Mean Standard Deviations:

<u>Sessions:</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>
<u>RS:</u>	0.135	0.130	0.110	0.092
<u>PS:</u>	0.153	0.126	0.122	0.133

Weber Function (RS):  $\Delta T = 0.064 (T + 0.614)$  ..... (18)

Analysis of Variance (RS Regression):

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Linear regression	1	0.27924	105.77	<u>p &lt; 0.001</u>
Deviations from linear regression	2	0.00063	0.24	NS
Residual	60	0.00264	-	

Weber Function (PS):  $\Delta T = 0.061 (T + 0.923)$  ..... (19)

Analysis of Variance (PS Regression):

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Linear regression	1	0.29487	95.43	<u>p &lt; 0.001</u>
Deviations from linear regression	2	0.00679	2.20	NS (p < 0.1)
Residual	60	0.00309	-	

TABLE 9 (contd)

Weber Function (Combined data):  $\Delta T = 0.063 (T + 0.758)$  ..... (20)

Analysis of Variance (Common regression):

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Common linear regression	1	0.58288	201.69	<u>p &lt; 0.001</u>
Deviations of the two individual regressions from the common regression	2	0.00204	0.71	NS
Residual	124	0.00289	-	

Analysis of Variance (standard deviations):

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Times (T)	3	0.136046	68.02	<u>p &lt; 0.001</u>
Conditions (C)	1	0.012680	6.34	<u>p &lt; 0.05</u>
Sessions (Se)	3	0.005226	2.61	NS (p < 0.1)
Subjects (Su)	1	0.003741	1.87	NS (p < 0.2)
C x Se	3	0.010164	5.08	<u>p &lt; 0.01</u>
C x Su	1	0.007907	3.95	NS (p < 0.1)
T x C x Se	9	0.004762	2.38	<u>p &lt; 0.05</u>
T x Se x Su	9	0.003214	1.61	NS (p < 0.2)
C x Se x Su	3	0.004074	2.04	NS (p < 0.2)
Residual	94	0.002000	-	

Analysis of Variance (Percentage Constant Error):

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Times (T)	3	0.36060	57.24	<u>p &lt; 0.001</u>
Conditions (C)	1	0.00103	0.16	NS
Blocks (B)	1	0.15645	24.83	<u>p &lt; 0.001</u>
Sessions (Se)	3	0.07448	11.82	<u>p &lt; 0.001</u>
Subjects (Su)	1	0.18720	29.71	<u>p &lt; 0.001</u>
T x C	3	0.02280	3.62	<u>p &lt; 0.05</u>
T x B	3	0.01497	2.38	NS (p < 0.1)
T x Se	9	0.01808	2.87	<u>p &lt; 0.01</u>
T x Su	3	0.02696	4.28	<u>p &lt; 0.01</u>

TABLE 9 (contd)Analysis of Variance (Percentage Constant Error): (contd)

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
C x Su	1	0.04340	6.89	<u>p &lt; 0.05</u>
B x Se	3	0.10668	16.93	<u>p &lt; 0.001</u>
T x C x Su	3	0.03657	5.80	<u>p &lt; 0.01</u>
T x B x Se	9	0.02095	3.33	<u>p &lt; 0.01</u>
T x B x Su	3	0.02419	3.83	<u>p &lt; 0.05</u>
T x Se x Su	9	0.02096	3.33	<u>p &lt; 0.01</u>
C x B x Se	3	0.03975	6.31	<u>p &lt; 0.01</u>
B x Su x Se	3	0.03387	5.38	<u>p &lt; 0.01</u>
T x C x Su x Se	9	0.01129	1.79	NS (p < 0.1)
Residual	57	0.00630	-	

Product-Moment Correlations Between the Means of Successive Series at the same Standard Interval:

<u>Standard Intervals:</u>	<u>0.25</u>	<u>0.5</u>	<u>1</u>	<u>3</u>
RS preceding PS:	0.75 <u>p &lt; 0.02</u>	0.82 <u>p &lt; 0.02</u>	0.57 NS	0.12 NS
PS preceding RS:	0.78 <u>p &lt; 0.01</u>	0.63 NS (p < 0.1)	-0.19 NS	0.12 NS
All:	0.76 <u>p &lt; 0.001</u>	0.74 <u>p &lt; 0.01</u>	0.22 NS	0.12 NS

## Discussion

As in Experiments 7 and 8 the Weber fractions increase as the time interval shortens, so that no minimal value is found, even though the shortest standard interval is now 0.25 seconds. The Weber fractions, and the corresponding standard deviations, are somewhat smaller than the values obtained for the same time intervals in the two earlier experiments. However there is a decrease in the standard deviations in successive sessions, giving a significant interaction between Sessions and Conditions. Since the subjects in the earlier experiments did only one session each, this suggests that only the results for the first sessions here should be used in making the comparison. The Weber fractions calculated for the first session alone are all larger than the corresponding figures for the four sessions, and are of approximately the same order as the values found in Experiments 7 and 8 (being, on the average, 0.15 per cent. less), so that there is no evidence that any important change in the accuracy of time estimation was brought about by using auditory rather than visual stimulation. The decrease in the standard deviations in the successive sessions is greater in the 'Reproduction Series'. Though the decrease does not give a significant effect of Sessions in the analysis of variance, there is a significant interaction of Conditions and Sessions. This implies that practice has an effect on at least one of the conditions. Thus it seems that the effect of repeated sessions is to decrease the subject's variability in reproducing time intervals, though any decrease in

variability of the productions is less. The Weber fractions are slightly larger for the productions, giving a significant effect of Conditions. Weber functions were calculated for the standard deviations and the means obtained from the experiment as a whole. The regressions are again significant with no evidence of departure from linearity. The slopes of the two functions are very similar, the main difference again being that the value of 'a' is greater for productions than for reproductions. A Weber function calculated from all the data gives a satisfactory fit.

Both the mean productions and the mean reproductions show an increase in the course of the session, producing a highly significant effect of Blocks in the analysis of variance. The rate of increase appears to be greater at the shorter intervals, though the Times by Blocks interaction does not quite achieve significance. However the interactions of Times and Blocks with Sessions, and with Subjects, are both significant. The indifference interval lies between 1 and 3 seconds with the Reproduction Series, and is greater than 3 seconds when the method of production is used. In each case the value that would be given by interpolation increases from the first to the second block. This is in substantial accord with the results of the earlier experiments, and does not show the marked decrease in the indifference interval that might have been expected on the 'central tendency' hypothesis. A significant Times effect, due to the different constant errors at different intervals,

is present. The analysis of the variance of the constant errors shows no significant effect of Conditions. Thus in this experiment there has been no tendency for the use of one or other method to affect the average size of the constant error; the difference between the constant errors found in Experiments 7 and 8 is not shown here. However the constant errors show a number of significant interactions. The table of mean constant errors illustrates the significant effects of Blocks, Sessions, the interaction of Blocks and Sessions, the interaction of Blocks, Sessions, and Conditions, and the lack of effect of Conditions. There is an increase in the means, in the first three sessions, which is shown, in a similar way, both by productions and by reproductions, and occurs in both blocks. The interaction of Blocks and Sessions reveals an important effect. In the first and fourth sessions, time intervals in the first block were presented in ascending order, from 0.25 seconds to 3 seconds, and in the second block in descending order, from 3 seconds to 0.25 seconds. In the second and third sessions this order was reversed: in the first block the time intervals were presented in descending order, and in the second block in ascending order. This appears to have modified the tendency for means in the second block to be longer than those in the first block. In the first and fourth sessions the rate of increase of the means is greater than in the second and third sessions. This can be understood in terms of an effect similar to the 'assimilation' effect that is produced by an extra stimulus interpolated between standard and

comparison stimuli (Woodworth & Schlosberg, 1955). If the interpolated stimulus is larger than the standard it tends to increase the point of subjective equality, if smaller, to decrease it. Here the previous series appear to exert a similar effect: when they are larger than the time intervals being presented or produced they increase the mean reproductions or productions, when smaller, they decrease them. If the series were randomly ordered such a 'series effect' would produce a 'central tendency of judgment', since the smaller series would more often be preceded by larger series, and the larger by smaller series. In the present experiment the counter-balancing allows this effect to be seen superimposed on the significant overall increase in means during the session. The use of a standard reduces the effect; it is more marked with the Production Series than with the Reproduction Series, and this difference is significant.

In the method of production we may assume that the subject employs a 'stored standard' or system of 'traces' derived from his previous experience with time intervals. When the method of reproduction is applied it is usually assumed that the subject is 'copying' or 'reproducing' each standard presented. However, with a very well learned skill it might be more economical for the subject to use the presented standard as an 'instruction' identifying the appropriate 'trace', and to use this 'stored standard' in making the reproduction. In the present experiment the use of a standard resulted in slightly smaller standard

deviations, a greater decrease in the standard deviation with practice, and greater resistance to the 'series effect'. But it did not produce any significant effect of Conditions, nor did it prevent the constant error becoming greater in successive sessions, at approximately the same rate in both conditions. This suggests that the 'stored standards' may have played a part in determining the reproductions. An alternative explanation for the similarity between productions and reproductions is that the presented standard might not only determine the subsequent reproduction, but might also affect later productions, making them essentially 'delayed reproductions'. If the latter is true there should be some similarity between the mean of a reproduction series and the mean of an immediately following production series, but not between the means of series presented in the reverse order. If, however, the 'stored standards' mainly determine performance the relation would be relatively unaffected by the order in which the two series are given. This can be tested in the present experiment, since the series were always given in pairs, one of each sort at each time interval. In half the sessions they were given in one order, and in the other half in the other order. The correlation coefficients calculated between the means of successive series in the two orders are similar for both orders for the two shortest intervals, but are not significant for the two long intervals. In no case did the two orders give significantly different correlations. This is against the assumption that the productions were really 'delayed

reproductions'; it favours the hypothesis that the 'stored standard' may play an important part even in determining reproductions, and suggests that it may undergo changes at a rate sufficiently slow to allow a high correlation between successive series of 10 reproductions or productions of a 0.25 second interval, but which greatly reduces the correlation for intervals of more than a second.

In this experiment the increase in the means during the course of the session found in Experiments 7 and 8 has been confirmed. It appears that the greater rate of increase for the shorter durations is not an artefact of giving the shorter intervals earlier in the session, since it is again found (though not significantly) despite the use of a counter-balanced order of presentation.

The marked decrease in the indifference interval expected on the 'central tendency' hypothesis was not found. However a 'series effect' was shown, of a sort which would produce a 'central tendency of judgment' in an appropriately designed experiment. Here its effect was subject to the counter-balancing.

Practice was found to reduce the standard deviations, when a standard was used, and to cause a similar increase in the means in both conditions. There did not seem to be much change in the rate of increase of the means in successive sessions with the same ordering of time intervals; the change from the first to the fourth session is from 15.0 per cent. to 19.7 per cent. This suggests that the effect is recurrent,

like fatigue, but no very certain conclusion can be drawn here, because a change may be obscured by the interaction of the 'series effect' and the counter-balancing.

Despite the use of auditory rather than visual stimulation to fill the time intervals there was little change in the differential threshold, when the first sessions are compared with the earlier experiments, so that Goodfellow's (1934) result was not confirmed. Similarly, the results do not support the finding of under-estimation with auditory stimulation in the absence of a standard reported by Goldstone, Boardman, & Lhamon (1959). The mean constant error for the first session reproduction series is slightly greater than that for Experiment 7, in which a visual stimulus was used. Neither in the production nor in the reproduction is there a general tendency to under-estimate, and over-estimation increases in successive sessions. This does not exclude the possibility that under-estimation would have been shown had the method of production only been used; but if intermittent presentation of the standard can prevent this under-estimation it does not seem necessary to consider it in interpreting the threshold experiments.

Taking this experiment with the two previous ones, the features of temporal discrimination now seem sufficiently well established to be applied to the interpretation of the threshold experiments.

#### 4.5 The indifference interval.

In the last three experiments it has been found each time that

the subjects' productions or reproductions lengthen during the course of the session, and that the rate of increase is greater for small intervals than for large. Though this last finding was significant only in Experiment 7, it appeared in each experiment and can therefore probably be accepted. It had the effect that when the indifference interval lay within the range of time intervals used, it got larger in successive blocks.

This suggested a possible explanation for some of the features of the indifference interval which have been found by previous experimenters. If we assume that there is an initial tendency to under-estimate all intervals, a tendency for the means to increase of the sort found here would result in short intervals being over-estimated sooner than long intervals, so that the indifference interval obtained would be a function of the intervals used, the length of the session, and the initial state of the subjects. This would be in accordance with the ubiquity of the finding that it is the shorter intervals which are over-estimated and the longer intervals which are under-estimated, and with the varied values of the indifference interval which have been reported.

These findings could also be explained by the 'central tendency of judgment', and a 'series effect' of this sort was found in Experiment 9. However when this effect was counter-balanced the means still increased to a significant extent in the second block, and the tendency for the shorter intervals to increase more rapidly than the long was still shown.

The 'lengthening effect' would provide the readier explanation for the finding of an indifference interval when no range of intervals is used, as in the experiment by Woodrow (1934) in which he employed a different group of subjects for each time interval. Probably both effects contribute to varying extents to any result, depending on the way the experiment is designed.

Why do the means of short intervals lengthen more rapidly? This could be accounted for, in Experiments 7 and 8, by the 'series effect'. In those experiments the short intervals at the beginning of the second and third blocks followed directly after the long intervals at the end of the first and second blocks, while the long intervals were always preceded by shorter series. Because of this lack of counter-balancing the 'series effect' would have tended to increase the short and decrease the long intervals. It might even be possible to extend this explanation to Experiment 9, despite the counter-balancing, if some extra arbitrary assumptions were made, such as that the 'series effect' acts more strongly on small than on large intervals, and produces larger increases than decreases. This is not very plausible, however, and still leaves the increase in all means in successive blocks to be accounted for.

Another explanation which could be offered for both the increase in the means and the more rapid rate of increase at the short intervals, is that a change takes place in the 'stored standards' during the course of a session. It has been suggested that the basis of the

appreciation of time might lie in some regular process in the nervous system, such as a repetitive discharge from pace-maker cells, which could constitute a 'psychological clock' and evidence has been produced tending to show that time judgments depend on the velocity of an underlying continuous chemical reaction (Piéron, 1952; Hoagland, 1932). Whatever the nature of the 'psychological clock', the production of a time interval by the subject must presumably require that, on one occasion, a measure of the clock's output during the period separating two events must be taken and stored, and that on another occasion this measure must be so applied that an equivalent output by the clock is made to determine the length of the subject's reproduction. On this model a change in the subject's performance could be due to a change in the process of measuring the standard, or in the 'stored standard', or in the way it is used to determine the duration of the subject's production. Since the means were found to increase when the method of production only was used, in Experiment 8, it appears that the change is not an alteration in the reception of the standard; thus it may be a change in the 'stored standard' or in the process of applying it. These alternatives cannot be distinguished on the data available. Since the Weber function,  $\Delta T/(T+a) = k$ , has been found to hold for the data in these experiments, the psychological scale for time could be taken as being  $R = \log(T+a)$ , and the subject's stored knowledge of time standards can be thought of as though it were in the form of marks or notches on a 'measuring-stick', marked with units of this

scale. The assumption we must then add is that during the course of a session a positive quantity, ' $\log q$ ', is added to the measuring-stick, so as to alter the origin of the scale, increasing each point on it by the same amount. This quantity, ' $\log q$ ', could perhaps be supposed to correspond to an increase in the level of relevant neural 'noise'.

If a given interval,  $T_1$ , is presented, stored as ' $\log(T_1+a)$ ', and then used to determine the duration of a reproduction, the latter process will be equivalent to taking the antilog of the measure, subtracting the constant ' $a$ ', and applying the result to the output of the 'clock'. We suppose that when the measure has altered to ' $\log q + \log(T_1+a)$ ' during the course of the session, ' $\log q$ ', not being a usual feature of the measure, is not allowed for, so that the time interval produced is equivalent to ' $q(T_1+a)-a$ '. This will represent a proportionate increase, as compared with  $T_1$ , that is larger the smaller  $T_1$  is.

One prediction that follows from this model is that if the increase during a session is calculated as a proportion of  $(T_1+a)$  instead of as a proportion of  $T_1$ , then it should be found to be constant for all intervals. To test this these figures have been calculated for Experiments 7 and 8, and for the two conditions of Experiment 9, using the values of ' $a$ ' given by the corresponding Weber functions in each case, and they are given in Table 10.

It can be seen that, in both Experiments 7 and 8, there is still a greater increase at the shorter intervals, but that in Experiment 9

TABLE 10Percentage Increases in the means between the First and Final Blocks

$$A = 100(T_3 - T_1)/T_1; B = 100(T_3 - T_1)/(T_1 + a).$$

Standard Intervals:    0.5      1      2      3      4      2

Experiment 7

A:	12.6	17.6	13.0	8.8	5.5	6.0
B:	3.8	7.5	7.7	6.1	4.2	5.2

Experiment 8

A:	17.7	20.2	17.5	4.8	3.1	1.3
B:	4.6	8.2	9.3	3.2	2.2	1.1

Experiment 9

Standard Intervals:    0.25      0.5      1      3

Reproduction Series

A:	10.3	8.2	2.5	-0.1
B:	3.4	3.8	1.6	-0.1

Production Series

A:	9.8	8.8	6.5	4.8
B:	2.4	3.1	3.4	3.7

the increase is greater at the shorter intervals in the Reproduction Series, but at the longer intervals in the Production Series. These differences might possibly be due to the 'series effect', which is not controlled by counter-balancing in Experiments 7 and 8, and the error involved in estimating 'a'. The residual differences are greater in Experiment 8 than in Experiment 7, which is in keeping with the finding that the 'series effect' modified the Production Series, in Experiment 9, more than the Reproduction Series. Thus the figures in this table are not incompatible with the model.

If we take 'log q' to correspond to neural 'noise' resulting from 'fatigue' during the course of the sessions, then we might suppose that the processes involved in making temporal estimations or reproductions are adjusted for the 'resting level' of 'noise' obtaining in the everyday circumstances in which the subject usually estimates durations. At the beginning of an experimental session the subject would be in an unusually favourable situation, rested, protected from distraction, concentrated on the task. Thus 'log q' might be smaller than usual, giving the initial tendency to under-estimate intervals previously assumed. In successive sessions, as the task becomes less novel, and the subject more bored, the initial concentration and interest may become less, giving a larger initial value of 'log q' in each session. This would account for the tendency for the mean constant error to increase on successive days which was noted in Experiment 9; and it is in keeping with the finding that the

increase appeared to be present in the first block, with little suggestion of an increased rate of lengthening of the means from day to day.

This explanation is speculative and requires considerable experiment to evaluate it. However, as this would constitute a divagation from the original object of the investigation, it was decided to leave the problem at this stage, and return to the consideration of the threshold experiments.

#### 4.6 Temporal discrimination and threshold level.

In Section 3.1 the hypothesis that the accessory stimulus functioned as a 'warning', marking the beginning of the time interval, was discussed, and it was suggested that, in this case, the fall in threshold would be a function of the precision with which the subject could predict the moment of arrival of the stimulus. Since he can at best determine a range of time over which the stimulus is probable, and not an instant at which it is certain, his adjustment of the 'critical value' determining the threshold would be related to the size of this range. At the larger intervals, where the range will be wider, the reduction of the critical value would be correspondingly less. The subject might, of course, vary the critical value continuously, as the subjective probability he attaches to the occurrence of the stimulus varies, or in some other fashion, but as a first approximation we will assume that the critical value is either at an 'asymptotic' resting level, corresponding to the threshold that would be given with a very long

inter-stimulus interval, or drops rapidly to a lower level which is held constant during the range of time that the stimulus is considered probable. On the assumption that the temporal discrimination function obtained in Experiments 7 and 8 gives a sufficiently close estimate of the degree of the hypothesized uncertainty in the use of the inter-stimulus interval to determine when the critical stimulus is to be expected, and that the changes in threshold in Experiments 1, 2, and 4, can be used as estimates of the corresponding changes in the critical value, we can now examine whether the falls in threshold, on the log microamp. and decibel scales used in those experiments, were inversely proportional to the 'range of expectation' of the stimulus.

To estimate this range of expectation we can take it to be equivalent to a constant number of j.n.d.s. As the j.n.d.s the values of  $\Delta T$  given by equation (17) will be used, these being derived from the two experiments most similar in design to the threshold experiments. One feature of this Weber function was the fairly large value, 1.57 seconds, of the additive constant 'a'. This constant is sometimes taken as an expression of the 'background noise' which, it has been suggested, is the limiting factor at the absolute threshold (Gregory, 1956; Cane & Gregory, 1957; Ekman, 1958, 1959). However 'noise' could arise not only in the process of 'receiving' the stimulus, but as variation in a 'trace', or in the processes involved in applying the trace to a task, such as reproduction. Thus the value obtained might reflect, at least partly, the

subject's response variability, and in the rather different use of the 'stored standards' which might be taking place in threshold discrimination, the value might be different, and possibly considerably less. In view of this possibility it seemed of interest to see what relation would be given if the range of expectation was estimated on the assumption that the value of 'a' is negligible, i.e. if the range were taken as a constant proportion of the time interval, T.

An inverse relation between the fall in threshold and the range of expectation can be expressed by the theoretical equations:

$$(A-L) \Delta T = k \quad \dots\dots\dots (21)$$

and  $(A-L)T = k' \quad \dots\dots\dots (22)$

where the range is derived from the temporal discrimination function found earlier, or taken as a constant proportion of the time interval, respectively, A is the 'asymptotic' threshold level when there is no temporal information, L is the level to which the threshold falls, and k and k' are constants. The results of fitting these equations to the thresholds determined in Experiment 1, Conditions (a) and (b), Experiment 2, Conditions (a) and (b), and Experiment 4, are given in Table 11. The equations were fitted by the method of least squares, and an analysis of variance performed to determine whether there was any significant departure from linearity of regression for either equation by comparing the variance due to departures from linearity with the residual variance after the effects of Times and Subjects had been extracted. In each case the

TABLE 11Experiment 1(a)Regression on  $(1/\Delta T)$ :  $L = 2.226 - 0.018(1/\Delta T)$ Regression on  $(1/T)$ :  $L = 2.200 - 0.099(1/T)$ Analysis of variance (thresholds)

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Linear regression $(1/\Delta T)$ :	1	0.025990	126.78	<u><math>p &lt; 0.001</math></u>
Deviations from linear regression $(1/\Delta T)$ :	3	0.001157	5.64	<u><math>p &lt; 0.01</math></u>
Linear regression $(1/T)$ :	1	0.027701	135.13	<u><math>p &lt; 0.001</math></u>
Deviations from linear regression $(1/T)$ :	3	0.000585	2.85	NS ( $p < 0.1$ )
Residual:	20	0.000205	-	

Experiment 1(b)Regression on  $(1/\Delta T)$ :  $L = 2.327 - 0.022(1/\Delta T)$ Regression on  $(1/T)$ :  $L = 2.294 - 0.118(1/T)$ Analysis of variance (thresholds)

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Linear regression $(1/\Delta T)$ :	1	0.039025	58.51	<u><math>p &lt; 0.001</math></u>
Deviations from linear regression $(1/\Delta T)$ :	3	0.002540	3.81	<u><math>p &lt; 0.05</math></u>
Linear regression $(1/T)$ :	1	0.039583	59.34	<u><math>p &lt; 0.001</math></u>
Deviations from linear regression $(1/T)$ :	3	0.002354	3.53	<u><math>p &lt; 0.05</math></u>
Residual:	20	0.000667	-	

TABLE 11 (contd)

Experiment 2(a)Regression on  $(1/\Delta T)$ :  $L = 2.125 - 0.015(1/\Delta T)$ Regression on  $(1/T)$ :  $L = 2.105 - 0.080(1/T)$ Analysis of variance (thresholds)

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Linear regression $(1/\Delta T)$ :	1	0.013940	56.44	<u><math>p &lt; 0.001</math></u>
Deviations from linear regression $(1/\Delta T)$ :	3	0.001088	4.40	<u><math>p &lt; 0.05</math></u>
Linear regression $(1/T)$ :	1	0.015321	62.03	<u><math>p &lt; 0.001</math></u>
Deviations from linear regression $(1/T)$ :	3	0.000627	2.54	NS ( $p < 0.1$ )
Residual:	16	0.000247	-	

Experiment 2(b)Regression on  $(1/\Delta T)$ :  $L = 2.167 - 0.017(1/\Delta T)$ Regression on  $(1/T)$ :  $L = 2.142 - 0.086(1/T)$ Analysis of variance (thresholds)

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Linear regression $(1/\Delta T)$ :	1	0.021625	31.94	<u><math>p &lt; 0.001</math></u>
Deviations from linear regression $(1/\Delta T)$ :	3	0.000421	0.62	NS
Linear regression $(1/T)$ :	1	0.021194	31.31	<u><math>p &lt; 0.001</math></u>
Deviations from linear regression $(1/T)$ :	3	0.000565	0.83	NS
Residual:	20	0.000677	-	

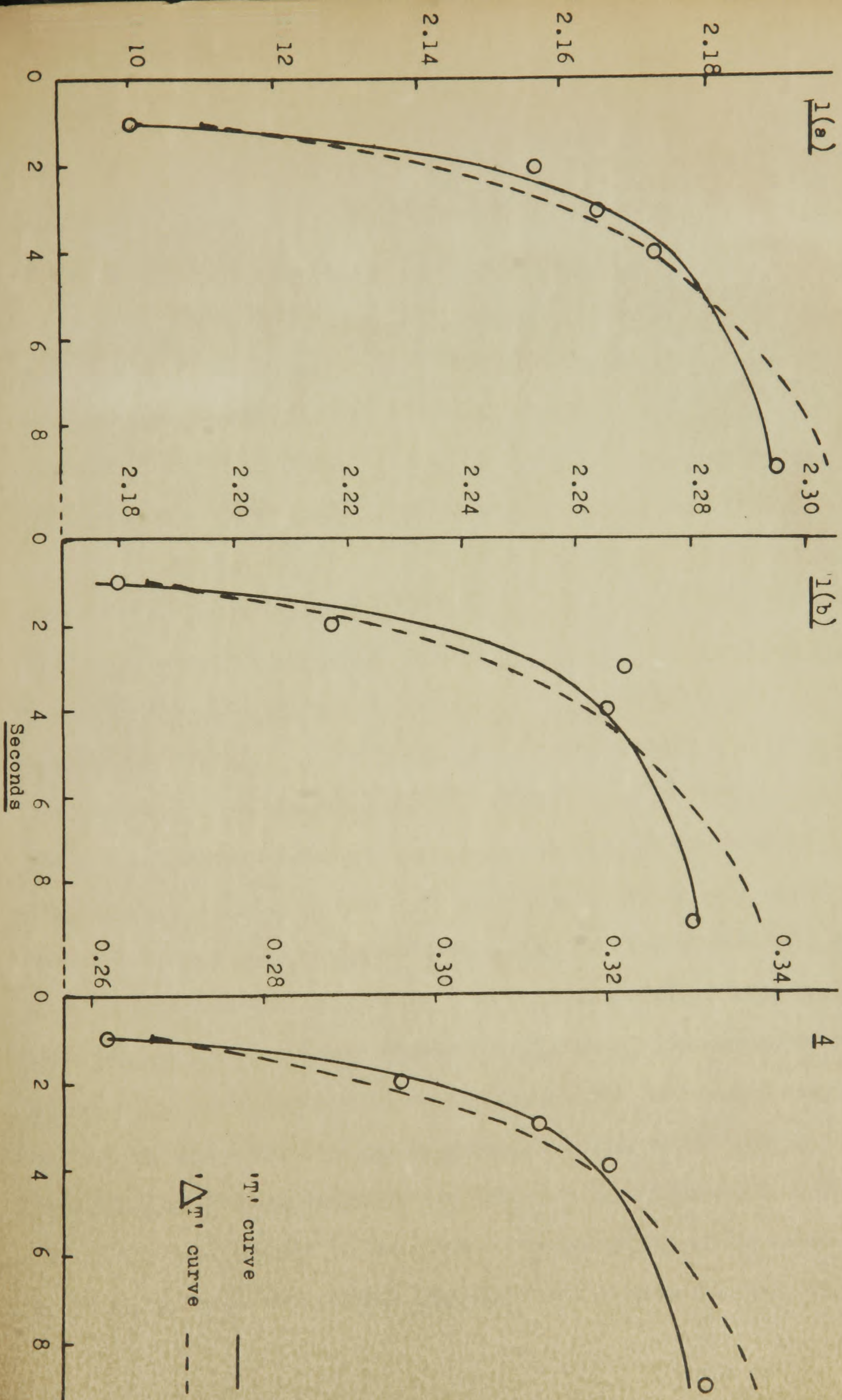
TABLE 11 (contd)Experiment 4Regression on  $(1/\Delta T)$ :  $L = 0.360 - 0.015(1/\Delta T)$ Regression on  $(1/T)$ :  $L = 0.337 - 0.075(1/T)$ Analysis of variance (thresholds)

<u>Source</u>	<u>d.f.</u>	<u>Mean Square</u>	<u>F</u>	<u>Significance</u>
Linear regression $(1/\Delta T)$ :	1	0.025092	55.51	<u><math>p &lt; 0.001</math></u>
Deviations from linear regression $(1/\Delta T)$ :	3	0.000509	1.13	NS
Linear regression $(1/T)$ :	1	0.024021	53.14	<u><math>p &lt; 0.001</math></u>
Deviations from linear regression $(1/T)$ :	3	0.000866	1.92	NS ( $p < 0.2$ )
Residual:	32	0.000452	-	

All experiments

<u>Regressions</u>	<u>Chi-square</u>	<u>d.f.</u>	<u>Significance</u>
$\Delta T$ :	11.280	1	<u><math>p &lt; 0.001</math></u>
T:	1.129	1	NS

Log.  $\mu$  A. **FIGURE 7**

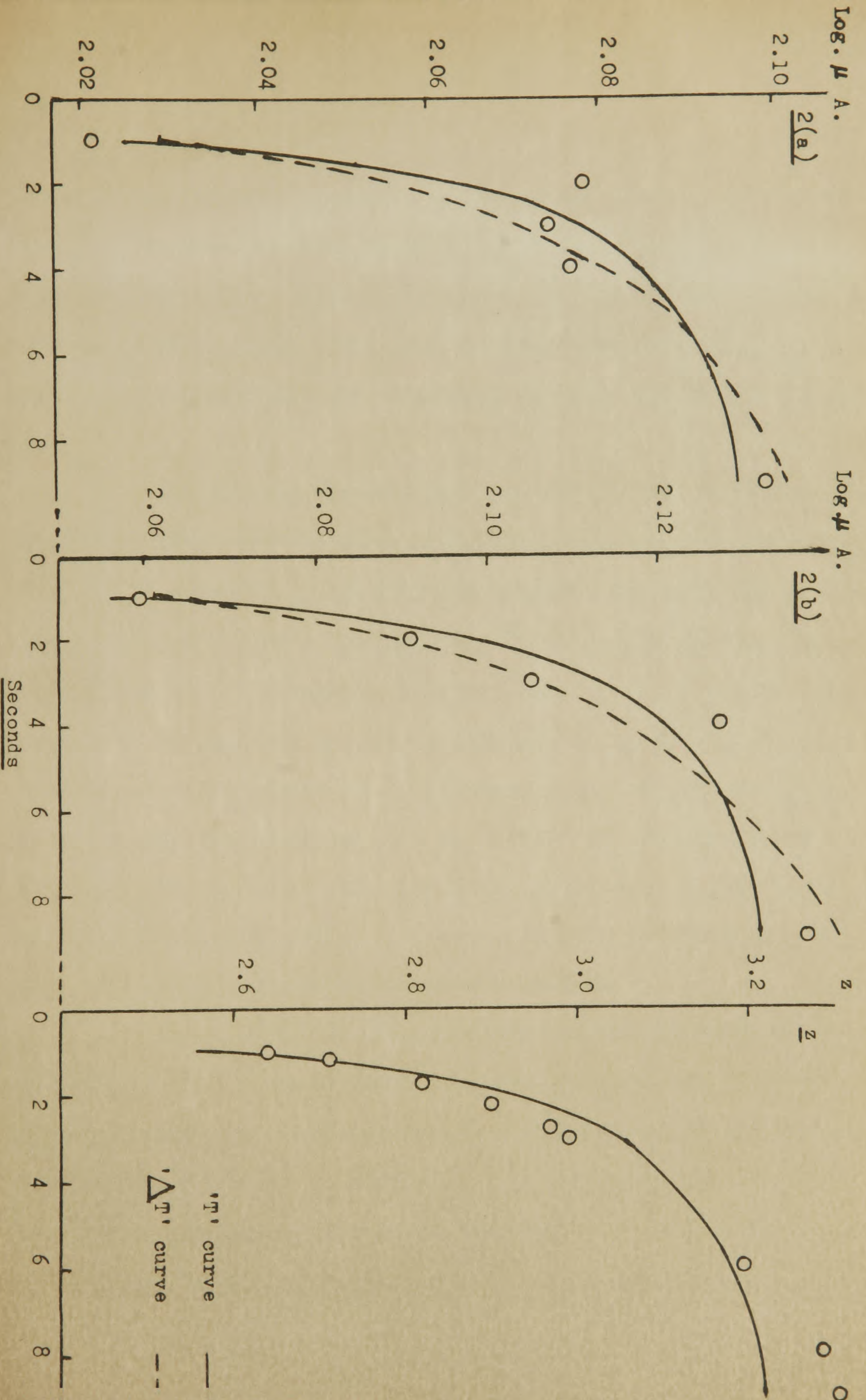


equations determined are given in the forms  $L = A - k(1/\Delta T)$  or  $L = A - k'(1/T)$ . The corresponding curves are plotted in Figures 7 and 8.

Both equations provide highly significant regressions for each experiment. Deviations from linearity of regression are significant in three out of five regressions on  $(1/\Delta T)$  but in only one of the regressions on  $(1/T)$ , though in some of the cases where the deviations are not significant their probability levels are nevertheless fairly low.

On inspection of the figures it can be noticed that for each set of data, the experimental points at the two extreme time intervals fall below the  $\Delta T$  regression, while they appear to be more randomly distributed in relation to the  $T$  curves. To test whether there was a significant tendency for the deviations from the regressions to be in the same direction for corresponding time intervals in the different sets of results, a chi-square test was performed, the 25 mean thresholds shown in Figures 7 and 8 being grouped into two classes, according as they fell above or below their respective regressions, for each time interval. The 1 and 9 second points, and the 2, 3, and 4 second points, were grouped together, to ensure adequate expected frequencies. The values of chi-square which were obtained are highly significant for the  $\Delta T$  regressions, but not for the  $T$  regressions, suggesting that the deviations of the data from the best-fitting curves in the different experiments are random in the latter case but not in the former. The propriety of this test can, however, be considered somewhat dubious, since it does not take note of the

FIGURE 8



extent of the divergences, small deviations being given as much weight as large, and because the grouping of the points is 'post hoc'. However, taken with the greater tendency for departures from the  $\Delta T$  regressions to be significant in the analyses of variance, these results do provide some slight evidence that the regressions on  $(1/T)$  tend to give the better fit.

This finding suggests that the value of the additive constant 'a', in the Weber function properly describing the temporal discrimination which may be occurring in the threshold judgment situation, may be somewhat smaller than the estimate obtained from the temporal production and reproduction tasks. Should the limiting factor in the threshold fall prove to be the accuracy of temporal discrimination, and should the lesser value of 'a' be confirmed, this would suggest that 'a' represents, at least partially, variance resulting from the functioning of the mechanisms involved in the tasks by which sensitivity to duration is determined, and that when the time interval is used as a 'warning', and treated centrally only, this extra source of variation is by-passed.

Had no simple or close relation been found between the fall in threshold and the discrimination of time intervals, the 'warning' hypothesis would have been more difficult to maintain, and the 'arousal response' hypothesis would accordingly have been supported. Had the relation to  $\Delta T$  been clearly closer than that to  $T$ , this would have seemed good evidence for an hypothesis based on temporal discrimination, since a

close relation of this sort would not appear a very likely consequence of 'arousal'. However, the best fit appears to be given when the falls in threshold are related directly to the time interval. This could mean that the temporal discrimination function which properly applies here is close to Weber's law, in its classic formulation; but it is not implausible that change in the degree of the state consequent on arousal might be inversely related to the passage of time. Thus the finding that has emerged does not decisively exclude nor markedly favour either hypothesis.

#### 4.7 Conclusion.

As the relationship which has been found between the fall in threshold and the time interval fails to indicate clearly whether the underlying processes are of an 'arousal' or of a 'warning' type, we are left with the problem of deciding between these hypotheses. In order to do this it is necessary to find predictions which are in conflict and which can therefore be used to make experimental tests of the hypotheses.

Some of their consequences which may prove useful for this purpose are:

- 1) Different mechanisms of threshold change may have different implications for threshold variability.
- 2) The effect of randomising the inter-stimulus intervals over a small range may have different effects on the threshold when this range is smaller and when it is larger than the 'range of expectation', if the 'warning' hypothesis holds.
- 3) If the fall in threshold is a consequence of some physiological change

which must have been effected at the time that the critical stimulus arrives, the latency of this process should produce a point of inflection in the curve relating threshold level to time interval, if this is plotted for short time intervals, and the effect of the accessory stimulation on the threshold should disappear when the inter-stimulus interval is zero.

- 4) A fortiori, in this case, if the accessory stimulus follows the critical stimulus it should have no effect on the threshold.
- 5) An 'arousal' effect may vary with the strength of the accessory stimulus, but if the latter serves only to convey temporal information its strength should be irrelevant to its effect, provided it is itself above threshold.

These predictions will be elaborated and experimentally evaluated in the following chapters.

## CHAPTER 5

### THRESHOLD VARIANCE

#### 5.1 Introduction

This chapter will consider the changes in the standard deviation and in the false positive rate that may occur as the threshold varies. If we suppose that a process of distinguishing signals from noise determines the level of the threshold, then there are a number of ways in which variations in this process could alter the threshold, and these would have different effects on the threshold variance. If, however, the older theory of the threshold is asserted then changes in threshold level must presumably be attributed to changes in the subject's tendency to 'guess', increased guessing lowering the threshold at the cost of increasing both its variance and the false positive rate.

In the account given by signal-detection theory the threshold at any time is determined by the value of  $c$ , the cut-off point or criterion, and this value depends both on the features of the  $N$  and  $SN$  distributions, and on the procedure for computing the optimal criterion. In statistical decision theory the criterion computed is a value of likelihood ratio,  $\beta$ , which corresponds to a criterion value,  $c$ , of  $x$ , that measure of the central effect of the stimulus which is used to determine the response. The distribution of  $l(x)$  on  $x$  depends on both the  $N$  and  $SN$  distributions, so that the value of  $x$  which corresponds to  $\beta$  is affected by any change in these. Thus the threshold could be made

to vary by a change in  $M_N$  or  $M_{SN}$ , the means of the two distributions, or in  $\sigma_N$  or  $\sigma_{SN}$ , their standard deviations.  $\beta$  will also, of course, be affected by changes in the quantities from which it is computed, such as values, costs, and a priori probabilities of signals and noise. It follows that any process which affects the thresholds, such as those postulated by the 'arousal' and 'warning' hypotheses, must, in terms of this model, correspond to a change in one or more of these factors. To see what changes to expect in the standard deviation as the threshold falls the implications of the different hypotheses must be worked out in terms of the signal-detection model.

Equation (10) gives the formula for calculating the critical likelihood ratio when the optimal criterion is taken to be that which maximises the total expected value of the subject's performance. If an equation of this sort were found to fit it would not imply that in making threshold decisions the subject makes unconscious mathematical inferences exactly equivalent to working out the equation, only that whatever processes underlie his performance are equivalent in their effects to the consequences of the equation. The expected value criterion is that which has most usually been assumed in the applications of signal-detection theory to the threshold (e.g. Tanner & Swets, 1954; Tanner, Swets & Green, 1956; Licklider, 1959; Green, 1960), but it is not the only optimal criterion which could be applied. An important alternative (Birdsall, 1955), based on the work of Neyman & Pearson (1933), requires

that  $P_{SN}(A)$  should be maximised, subject to  $P_N(A)$  not exceeding a constant limiting value,  $\epsilon$ . In this case the criterion is that value of  $l(x)$  for which

$$P_N(A) = \epsilon \quad \dots\dots\dots (23)$$

Neyman & Pearson did not wish to weight their criterion by the a priori probabilities, but Peterson & Birdsall (1953) present a modification of this criterion in which it is weighted by  $P(N)$ . They require that  $P(N.A) \leq P_0$ , where  $P_0$  limits the false positive rate, and that  $P(SN.A)$  should be maximal. The condition for the criterion can be expressed as

$$P_N(A) \leq P_0/P(N) = k \quad \dots\dots\dots (24)$$

and the criterion is then that value of  $l(x)$  for which  $P_N(A) = k$ .

However it is not necessary to suppose that the subject's 'criterion computer' (Swets, Tanner & Birdsall, 1955) computes one of these formulae or a close approximation to it; the subject might use some other system of weighting or values than they allow for, and this must be remembered in considering actual cases. Though Tanner & Swets (1954) found that variation of the monetary rewards and costs given to subjects for their performance, and Tanner, Swets & Green (1956) that variation in  $P(SN)$  led to variation in the value of  $\beta$  determining the subjects' responses, this does not allow us to conclude that the 'criterion computer' employs the expected value criterion; the explicit values, costs, and a priori probabilities prescribed by the experimenters could have been used to affect a criterion not usually calculated from equivalent

'subjective' terms. It might be noted that though Green (1960) found that the variation in  $\beta$  was in the expected direction, it was not of the expected extent.

However, whatever measures the subject may be supposed to compute the criterion from, they will be subjective in the sense that they are not necessarily the same as the values prescribed by the experimenter, the subject will be using his own approximations to or estimates of them. Thus though the objective value of  $P(SN)$  will change from 0 to 1 at the exact end of the inter-stimulus interval this will not be true of the subjective value, which will at best rise to a somewhat lower value over a range of time about the end of the interval. Similarly, even though in the method of limits  $P(SN) = 1$ , and this may be known to the subject, he is better considered as working with a subjective a priori probability of the stimulus which is less than this (on the 'expected value' criterion  $P(SN) = 1$  would give  $\beta = 0$ ). In considering how the threshold may change, the various ways in which the criterion might be computed must be borne in mind.

If the arousal hypothesis is maintained, the effect of the accessory stimulus is to cause an almost immediate reduction in  $c$ , which then gradually returns to its 'asymptotic' or resting level,  $c_a$ . This is not readily compatible with the 'expected value' criterion; the latter implies that, if the values and costs are taken to be constant,  $\beta$  will vary with changes in  $P(N)/P(SN)$ . Unless the subject is denied all ability

to estimate temporal duration, which seems unreasonable, the subjective value of  $P(SN)$  must be supposed to vary during the course of, say, a 5 second inter-stimulus interval, from a value close to zero at the beginning of the interval to a value near unity towards its end, with  $\beta$  correspondingly initially very high and later very low. It would seem to require very considerable changes in the  $N$  and  $SN$  distributions to outweigh this, so that it seems more plausible to combine the arousal hypothesis with a criterion such as that of Neyman & Pearson, which has the advantage of being independent of  $P(SN)$ , and of the form of the  $SN$  distribution.

When the method of constant stimuli is used the standard deviation of the subjects' responses can be taken to be directly related to the standard deviation of the 'signal + noise' distribution. The relation of the standard deviation of threshold estimates obtained by the method of limits is rather more complex; however, it can probably be taken to be approximately linearly related to  $\sigma_{SN}^2$  (Cane, 1956). This suggests that in any investigation of changes in variance it would be better to use the method of constant stimuli; this would also allow false positive rates to be measured, so it is of some interest to consider the features they might show.

Signal-detection theory assumes that a shift in thresholds must be accompanied by a change in the false positive rate, as shown by the R.O.C. curves which relate  $P_{SN}(A)$  to  $P_N(A)$ . However, the theory was

developed for situations where it could be assumed that each response made by the subject was based on one decision. In the present situation this expectation must be modified. During an inter-stimulus interval in which the subject is watching for the critical stimulus the 'noise' could exceed  $c$  at any time. While it remains below the criterion, no response is made until the end of the trial is reached, but if it exceeds  $c$  this will cause a positive response. The decision mechanism that relates the continuous afferent input to  $c$  could operate continuously, or it could be 'quantal', dividing the time that passes into 'units' or 'moments', for each of which a threshold decision is made. If it were assumed to be intermittent this would make it easier to calculate possible effects on the false positive rate. Can this assumption be supported?

It has been suggested that minimal sampling times may be used in a number of types of perceptual analyses. Thus Attneave & McKeynolds (1950) studied the maximum perceptible beat frequency between two flickering lights, and obtained a value for this of 5 c.p.s. Callaway & Alexander (1960) presented regular trains of flashes, or clicks, or both to their subjects, who were instructed to respond to the cessation of the train. When the reaction times were plotted against impulse frequency the curves showed perturbations, and frequently maxima, at about 10 c.p.s., which the authors interpret as evidence of a discontinuous sampling of sensory data at about this rate. Cheatham and White (1952) presented short trains of flashes to their subjects, at 10-30 c.p.s.,

and found that the subjective flash rate did not exceed 6-8 flashes per second. They concluded that the upper limit to the number of flashes that could be perceived was about one per 100 msec. In a similar experiment using sound pulses at the same rates, they again found a limit to the subjective rate for perception of the sound pulses, at about 9-11 pulses per second (Cheatham & White, 1954), and the finding has been extended to trains of touch stimuli. They conclude that there is approximately the same upper limit for all three senses, about 80 msec. per perceived stimulus (White & Cheatham, 1959). Broadbent (1958) considers that work on the intelligibility of interrupted speech suggests a sensory sampling time of 1/3rd second. Hirsh (1959) found that the order in which two sounds were initiated could be distinguished with an accuracy of 75 per cent when the interval separating them was about 20 msec.; if this were interpreted in terms of a central sampling time it would indicate a rather small value. Broadbent & Ladefoged (1959) suggest however, that practice and optimal experimental conditions may have allowed the subjects to use a feature of the stimulus analysed at the periphery in making the judgment; they quote evidence suggesting that in the absence of such conditions a time of 150 msec. is necessary to distinguish temporal order at the 75 per cent level. Stroud (1955) has presented data on the apparent brightness of sequences of flashes at different rates, and reviewed a number of different studies which lead him to suggest that the sensory sampling or integration time can vary between

0.05 and 0.20 seconds, with a value of about 0.10 seconds when there is no feature of the stimulus input tending to modify its length.

Though these studies suggest the possibility that a sampling time may play a part in a number of perceptual-motor tasks, they do not justify the conclusion that sampling necessarily occurs during the time that a subject may be watching for a stimulus. Most of these observations are difficult to interpret unambiguously and might quite well have other explanations; and if sampling occurs in some cases it does not necessarily occur in others. However the assumption that the inter-stimulus interval is divided into sampling times makes it easier to calculate possible changes in false positive rate, and in this respect probably supplies a reasonably close approximation to what might be expected from a continuous function. Therefore it will be tentatively assumed, as a matter of convenience, that the time intervals are divided up into 'moments' of about 0.1 seconds each, for each of which a threshold decision is made.

In each trial the inter-stimulus interval offers a number of opportunities to make a false positive response during a period in which the stimulus will occur only once. Since subjects must often, in the course of their daily activities, watch for stimuli whose appearance is uncertain, and whose time of arrival is indefinite, we may suppose that  $c$  is usually given a value sufficiently high to make the occurrence of an embarrassing number of false positives unlikely, even when watch must be maintained for some time. One consequence of this is that for any form

of the 'arousal' hypothesis the false positive rate should be less the shorter the inter-stimulus interval. Since  $c$  undergoes the same changes whatever the time at which the trial is terminated, the false positive rate for an interval  $T_2$  will always equal the false positive rate for the interval  $T_1$  plus the rate for the added duration  $(T_2 - T_1)$ , where  $T_2 > T_1$ . When the value of  $c$  is initially such that the false positive rate will be low for periods of observation of the order we have used, differences between the rates for different inter-stimulus intervals are unlikely to be detectable without large numbers of observations. Thus if we suppose that the value of  $c$  is such as to give a false positive rate of 1 per cent for an inter-stimulus interval of 10 seconds, and disregard the fact that subjects may wait somewhat longer than the length of the interval before ending a trial, then, if  $P_N(A)$  is kept constant, the false positive rate should be 0.5 per cent for an interval of 5 seconds, and 0.1 per cent for 1 second. Percentages as small as these would probably be indistinguishable in practice. However a small maintained increase in  $P_N(A)$  might produce a relatively large increase in the false positive rate, since each 'moment' would provide an increased opportunity for a false positive to occur. For example, if the value of  $c$  giving a 1 per cent false positive rate for an interval of 10 seconds were reduced by  $\frac{1}{2}\sigma_N$  for that period,  $P_N(A)$  would increase from 0.0001 to 0.0006, causing the false positive rate to increase from 1 per cent to 6 per cent.

The ways in which 'arousal' might affect the threshold include:

(a) The effect of the accessory stimulus might be to reduce variability in the afferent pathways, and other sources of 'noise' variance, with the effect of reducing  $\sigma_N$ , which would then gradually return to the resting level. If the value of  $c$  is altered as  $\sigma_N$  changes, so as to keep  $P_N(A)$  constant, as would be the case if the Neyman-Pearson criterion is being applied, there would be a fall in threshold for short inter-stimulus intervals. There should be no effect on the false positive rate, as compared with a hypothetical 'resting rate', i.e. the rate that would be given during a period of equal length in which the subject operated in the same fashion, but which did not begin with an initial stimulus. Such a 'resting rate' might be obtained from a vigilance task if the subject's performance could be considered to be constant throughout the period of watch. It may be noted that in this case, since  $P_N(A)$  is constant, false positives are equally likely to occur at any time during the inter-stimulus interval. Since the variable central effect of the stimulus is at least partly due to its being added to the 'noise' distribution, we can suppose that if  $\sigma_N$  is reduced  $\sigma_{SN}$  will also undergo a reduction. Therefore the threshold variance would decrease as the threshold fell. (b) The effect of the accessory stimulus might be to cause a reduction in  $M_N$ , the mean of the 'noise' distribution, with a corresponding downward shift of  $c$ . Here the false positive rate, and the variability of the threshold response would be unaffected. (c) That part of the decision process which computes or maintains the optimum

value of  $c$  could be affected, resulting in a downward shift of  $c$ , with the two distributions remaining unchanged. This would have the effect of lowering the threshold without causing any change in threshold variance, and of increasing the false positive rate as compared with the resting condition. It would also have the effect that the false positive rate would increase disproportionately rapidly for the short intervals, the increase from, say, 1 to 2 seconds being greater than that from 2 to 3 seconds. Thus the false positives would be more likely to occur in the earlier parts of an interval than in the later, despite the fact that the stimulus is not probable until the end of the interval. However, since false positive rates are usually fairly low it is difficult to confirm predictions of small differences between them.

If the first two alternatives are considered as ways in which a 'reflex' change might operate, the third could be described as an 'alarm' effect: there is no change in the efficiency of the afferent paths, instead the response is a change in the criterion for threshold decisions, with an increased tolerance of false positives.

These are not the only assumptions that could be made - for instance, one could suppose that  $\sigma_N$  and  $M_N$  are both reduced - but they seem the simplest and most plausible alternatives. As they stand, they imply that the threshold variance would either stay constant or fall as the threshold falls.

The possibilities suggested by the warning hypothesis are:

especially when procedures such as the method of limits are used.

(d) During the 'range of expectation' the preparatory effect might produce a reduction of  $\sigma_N$ . If  $c$  is varied to keep  $P_N(A)$  constant it will be correspondingly reduced. Here we need suppose no criterion more complicated than that prescribed by equation (23). The 'criterion computer' would take account only of  $P_N(A)$ ,  $M_N$ , and  $\sigma_N$ . The subject makes an adjustment for the increased  $P(SN)$  by lowering  $\sigma_N$ , but we need not assume that it also enters into the computation of the criterion. The effect of this would be to lower the threshold, and to reduce its variance correspondingly. Since  $P_N(A)$  remains constant the false positive rate should not be affected. (e)  $M_N$  is lowered by the preparatory effect, and  $c$  correspondingly reduced. This should not affect false positive rate nor threshold variance. (f) A judgmental effect would correspond not to an alteration of the state of the afferent pathways but to a change in the computation of the criterion, due to the change in the subjective probability of the stimulus,  $P(SN)$ , in the range of expectation. Here we might assume a criterion such as that in equation (10), or that prescribed by equation (24). However the first prescribes a likelihood ratio, which can only be used to specify a value of  $c$  if both  $f_N(x)$  and  $f_{SN}(x)$  are known, whereas the latter criterion can be used when the subject does not know the exact form of the SN distribution. This makes a criterion such as the latter more plausible for experiments such as these, where the subject is not given extensive preliminary practice with a limited range of stimuli, and

especially when procedures such as the method of limits are used.

During the range of expectation  $P(SN)$  will increase for each 'moment' in the range,  $P(N)$  will decrease correspondingly, and  $P_N(A)$  can be raised accordingly, giving a reduction in  $c$  which will be greater the greater the increase in  $P(SN)$ , i.e. the smaller the range. Since there will be an inverse relation between the length of the range and the increase in  $P_N(A)$  there will be little effect on the false positive rate, and the threshold variance will, of course, not change.

Thus hypotheses (a) and (d) appear to predict a fall in the variance, and (b), (c), (e), and (f) constancy, as the threshold falls.

## 5.2 Further consideration of the earlier experiments.

In Experiments 1, 2, and 4 the standard deviation appears to rise as the time interval decreases, but not in Experiment 3, in which the threshold did not fall. To test whether the increases were significant, the regressions of the standard deviation ( $S$ ) of the threshold on inter-stimulus interval ( $T$ ) were calculated and tested by analysis of variance in each case. The results, in Table 12, show that each regression coefficient is negative. In Experiments 1 and 2 neither equation is significant, though the probability is less than 0.1 when their data are combined. Experiment 4, however, is highly significant. Thus these three experiments suggest that there may be a slight increase in variance as the threshold falls. The method of limits, which gives a variance not very simply related to  $\sigma_{SN}^2$  (Cane, 1956), was used, so that

It is of interest to examine the threshold responses using the method of constant stimuli. This will also allow the false positive rate to be measured.

TABLE 12

3.3 Changes in threshold responses

<u>Experiment</u>	<u>Regression equation</u>	<u>Significance</u>
1	$S = 0.0671 - 0.00155T$	NS
2	$S = 0.0672 - 0.00144T$	NS ( $p < 0.2$ )
1 + 2	$S = 0.0672 - 0.00150T$	NS ( $p < 0.1$ )
4	$S = 0.0482 - 0.00191T$	<u><math>p &lt; 0.001</math></u>

not be attributable to the difference of procedure since the ANS methods were compared in the first part of the experiment.

Experiment 16(a)

Apparatus and procedure

The same apparatus was used as in Experiment 2 to provide the stimulus. This was a 30 msec. increase in the frequency of a constant tone of 500 c.p.s. at 60 dB SPL. The rise and decay of the pulse had a time constant of 25 msec. The subject sat in a comfortably illuminated sound-attenuated room 25 inches from a white screen, 70 inches by 24 inches in size, at the center of which were two rear holes. When the experimenter passed the operating button this screen in the upper light, providing a pre-reading, and started a timing circuit which switched on the second hole and started timer, to provide the necessary stimulus. The inter-stimulus interval, which was 1, 3, or 9 seconds, was

it is of interest to examine the threshold variance using the method of constant stimuli. This would also allow the false positive rate to be measured.

### 5.3 Changes in threshold variance.

It was thought that significant differences in the variance might be more likely to appear if a small number of experienced subjects were used. The experiment was designed in two parts. To make sure that any difference between the results that might be obtained with the method of constant stimuli, and those which were obtained earlier, should not be attributable to the difference of procedure alone, the two methods were compared in the first part of the experiment.

#### Experiment 10(a)

##### Apparatus and procedure

The same apparatus was used as in Experiment 4 to provide the stimulus. This was a 30 msec. increase in the intensity of a constant tone of 500 c.p.s. at 60-70 db. SL. The rise and decay of the pulse had a time constant of 25 msec. The subject sat in a moderately illuminated sound-shielded room 18 inches from a white screen, 20 inches by 24 inches in size, at the centre of which were two neon bulbs. When the experimenter pressed the operating button this switched on the upper light, providing a pre-warning, and started a timing circuit which switched on the second bulb one second later, to provide the accessory stimulus. The inter-stimulus interval, which was 1, 3, or 9 seconds, was

determined by a second timer; at the end of this time the amplifier was switched in to start the tone pulse, which was terminated by a third timer 30 msec. later. The form of this pulse was monitored on a Cossor Oscillograph. The two neon bulbs stayed on until the subject had responded 'Yes' or 'No'.

Each session began with a few practice runs. Then three descending series of stimuli at intervals of 0.1 db. were presented by the method of limits ('ML'), one for each of the three time intervals as in Experiment 4. These were followed by three series of stimuli, one for each inter-stimulus interval, presented by the method of constant stimuli ('MCS'). Each of these series consisted of 22 stimuli which were presented in a random order determined with the help of random number tables (Fisher & Yates, 1948). There were 10 'strong', 10 'weak' and 2 'blank' stimuli. The values for the 'strong' and 'weak' stimuli were chosen at the beginning of each session, on the basis of the ML series. The interval between them was 0.10 - 0.15 db., and values were selected that were judged likely to be just above and just below the subject's threshold. Each block of three ML series and three MCS series was repeated 4 times during the course of a session, which lasted 1½-2 hours. Two experienced subjects were used, the one doing 4 and the other 5 sessions. In 6 sessions the series were presented in the order 1, 3, 9 seconds, and in 3 sessions in the order 9, 3, 1 seconds.

## Results

In each session there were four ML threshold determinations for each of the three time intervals. The means and standard deviations of these were calculated for each session, and for the 9 sessions, as in the earlier experiments. The proportion of positive responses for the 'weak' and 'strong' stimuli at a given time interval in each session were converted into probits and a straight line fitted; the 50 per cent point was taken as the threshold and the reciprocal of the probit regression coefficient as the standard deviation (Finney, 1952), and average values were calculated for the 9 sessions. The false positive rate was calculated for the 72 'blank' stimuli at each time interval. Analyses of variance were performed on the standard deviations obtained with each method. The order in which the time intervals were presented had no apparent influence on the results, so that this has been disregarded in presenting the results, which are given in Table 13.

## Discussion

The fall in threshold as the inter-stimulus interval decreases is confirmed by these results for both psychophysical methods. The MCS thresholds are all lower than the corresponding ML thresholds. This difference might possibly not have been found had ascending series also been used for the latter. The standard deviations obtained by the method of limits are considerably smaller than those given by the method of constant stimuli. This is in keeping with the fact that the ML thresholds

TABLE 13

<u>Time Intervals (seconds):</u>	1	3	9
<u>Method of Limits</u>			
Mean Thresholds (db):	0.578	0.677	0.733
Mean S.D.:	0.081	0.082	0.088
<u>Method of Constant Stimuli</u>			
Mean Thresholds (db):	0.454	0.617	0.663
Mean S.D.:	0.262	0.213	0.131
False positives (per cent):	1.4	2.8	4.2

Analysis of variance (ML standard deviations)

<u>Source</u>	<u>d.f.</u>	<u>M.S.</u>	<u>F</u>	<u>Significance</u>
Sessions	8	0.001090	0.66	NS
Times	2	0.000301	0.18	NS
Residual	16	0.001653	-	

Analysis of variance (MCS standard deviations)

<u>Source</u>	<u>d.f.</u>	<u>M.S.</u>	<u>F</u>	<u>Significance</u>
Sessions	8	0.150040	1.29	NS
Times	2	0.039054	0.34	NS
Residual	16	0.116254	-	

are determined by a set of the subject's responses, so that their variance might be considered as a sort of sampling distribution variance.

ML standard deviations show a slight fall as the threshold falls, MCS standard deviations a somewhat greater rise. However in neither case does the analysis of variance show a significant effect of Times on the standard deviations, so that there is no evidence that the standard deviations are not constant for the different intervals. The failure of the fairly large changes in the MCS standard deviations to be significant reflects the considerable residual variability. A possible reason why this experiment should have produced no significant effect of Times might be that the repeated changes from one procedure to the other were disturbing to the subjects and increased the variability of their performance. Another reason for this variability may have been the rather long sessions. These were necessary to obtain all the threshold determinations, but they may have been rather fatiguing to the subjects, and so may have induced variability. The use of only two stimulus strengths in the MCS series is also unsatisfactory; more reliable results might have been given if a larger number of stimulus strengths had been used.

The false positive rates provide a test of the hypothesis that the increased thresholds at the shorter intervals are due not to a change in a signal-detection procedure but to an increased rate of 'guessing'. If this were true it would be shown by an appropriately

higher false positive rate at the shorter time intervals, yet the false positive rate actually decreases as the interval shortens. The total number of positive responses to 'weak' or 'strong' stimuli is 317 out of 720 presentations at 9 seconds, 519 out of 720 at 1 second. Guessing would account for this difference if the false positive rate were 0 per cent at 9 seconds and 50.1 per cent at 1 second. From the results obtained, this explanation can clearly be rejected.

#### Experiment 10(b)

In the first half of the experiment it was found that the method of constant stimuli could be used to show the fall in threshold at short time intervals and that when it was used a corresponding increase in the threshold variability was shown, which, however, was not significant. The second half of the experiment was designed to decrease the variability of the results by using only one method (MCS) and only two time intervals, and so being able to use more stimulus strengths and to present more stimuli, avoiding possibly disturbing changes of procedure, and somewhat shortening the sessions.

#### Apparatus and procedure

The same apparatus was used as in the first part of the experiment. The method of constant stimuli was employed, and only two time intervals were used in each session. Each session consisted of 5 series at each time interval, presented alternately. For each time interval four stimulus strengths, spaced 0.1 db. apart, were used. A series of

stimuli consisted of 4 stimuli at each of the four stimulus strengths and 4 'blanks', randomly ordered with the aid of random number tables. A 50 msec. increment in the intensity of the 500 c.p.s. tone served as the stimulus. Four subjects each did 3 sessions, with time intervals of 1 and 8 seconds and a pre-warning of 2 seconds. Three of the subjects were experienced, the fourth, who was unpractised, was given a preliminary practice session. In addition two of the well-practised subjects each did two sessions with time intervals of 3 and 6 seconds and a pre-warning of 1 second.

### Results

The results of each subject for one pair of inter-stimulus intervals were pooled and subjected to probit analysis (Finney, 1952). The parallelism of the probit regression lines for the two time intervals was tested by performing an analysis of  $\chi^2$  in each case; a significant departure from parallelism would indicate a significant difference between the standard deviations. These results are given in Table 14. In each case  $\chi^2$  for departures from parallelism has 1 degree of freedom. In no case did the analysis of  $\chi^2$  show significant evidence of heterogeneity of the departures of the experimental results from the regression line, confirming that the normal distribution provides a good fit for these data. The standard deviations are the reciprocals of the probit regression coefficients. Each false positive rate is based on the pooled 'blanks' at that time interval: 240 for 1, and for 8 seconds, and 80 for 3 and 6 seconds.

TABLE 14

<u>Subject</u>	<u>Time Interval</u>	<u>Threshold(db)</u>	<u>S.E.</u>	<u>S.D.</u>	<u><math>\chi^2</math></u>	<u>Significance</u>
C	1 second	0.393	$\pm 0.013$	0.136		
	8 seconds	0.546	$\pm 0.012$	0.135		
	Combined			0.135	0.006	NS
M	1 second	0.505	$\pm 0.012$	0.125		
	8 seconds	0.552	$\pm 0.012$	0.113		
	Combined			0.119	0.370	NS
J	1 second	0.420	$\pm 0.008$	0.067		
	8 seconds	0.455	$\pm 0.009$	0.093		
	Combined			0.075	3.154	NS (p<0.1)
V	1 second	0.412	$\pm 0.013$	0.159		
	8 seconds	0.545	$\pm 0.013$	0.101		
	Combined			0.129	5.693	<u>p &lt; 0.02</u>
C	3 seconds	0.422	$\pm 0.032$	0.177		
	6 seconds	0.471	$\pm 0.025$	0.221		
	Combined			0.199	0.535	NS
H	3 seconds	0.547	$\pm 0.014$	0.127		
	6 seconds	0.583	$\pm 0.015$	0.105		
	Combined			0.117	0.792	NS

Time Intervals (seconds):      1      3      6      8

False positives (per cent):    2.5      2.5      2.5      4.2

## Discussion

In each of the six comparisons the threshold is lower at the shorter inter-stimulus interval. The false positive rates again provide no evidence that this could be accounted for by 'guessing'. In 5 of the 6 cases there is no significant difference between the standard deviations at the two time intervals; in the sixth case the standard deviation is significantly greater at 1 second than at 8 seconds.

This experiment gave each subject either two or three sessions with only two time intervals in an effort to reduce variability; and for most of the subjects the standard deviations are smaller than the mean values obtained with the method of constant stimuli in Experiment 10(a). These results can therefore probably be taken as a fairly reliable indication that the fall in threshold may be accompanied by substantial constancy of the threshold variance.

However one subject did show a significant rise in the standard deviation at the shorter interval. This could be disregarded, the significance of this one result 'disappearing' when the five other non-significant results are taken into consideration, though it invites comparison with the similar rises shown in the earlier experiments. In Experiments 1 and 2 there was a non-significant rise in the standard deviation; in Experiment 4 there was a similar, highly significant rise; in the first half of the present experiment there was a slight fall with the method of limits, and a somewhat larger rise with the method of

constant stimuli, neither of them significant, and the present result shows substantial constancy of the standard deviation. The differences in procedure cannot be excluded as a possible source of the different results, since the earlier rises were obtained with the method of limits, and the present constancy with the method of constant stimuli. Though there was a fairly large rise with the MCS in Experiment 10(a) it was not significant. However no reason presents itself why the two methods should differ in this respect. Another explanation may lie in the differing experience of the subjects. In Experiments 1, 2, and 4, the subjects were for the most part unpractised in threshold determinations of this sort; but in the present experiment three of the four subjects were experimentally sophisticated, having taken part in most of the preceding experiments. Subject V, the only subject who had no previous experience of this sort, was also the only one to show a significant rise in standard deviation as the interval decreased. This suggests that the results should be taken to show that there may be either a rise in the standard deviation, or constancy, as the interval decreases, and that the first is possibly more likely in unpractised, and the second in practised subjects.

How do these results relate to the alternative hypotheses discussed in Section 5.1? They appear to be incompatible with alternatives (a) and (d), which predict a fall in the standard deviation as the threshold falls. The constant variances are compatible with (b), (c),

(e), and (f), but the increase in variance found in Experiment 4 and shown by Subject V in the present experiment was not predicted by any of these alternatives and requires explanation. Two auxiliary hypotheses can be suggested which might explain this finding, one based on the 'warning' and the other on the 'arousal' hypothesis.

In deriving the effects of threshold change on the threshold variance for alternatives (d), (e), and (f), the assumption was tacitly made that all the critical stimuli would fall within the range of expectation, so that the value of  $c$  would be essentially constant for them. However, if the range should begin too late, or end too early, some of the subject's decisions would be determined by the 'asymptotic' level of the criterion,  $c_a$ , and others by the reduced level within the range,  $c_r$ , and this would add to the variance of the response a component derived from the variation of  $c$  between these two values, whose size would depend on the proportion of the responses determined by  $c_a$  and  $c_r$ , and on the difference between them. Thus if the range of the subject's expectations is too narrow in relation to his actual inaccuracy in estimating the most likely time of arrival of the stimulus this should add a small component to the variance which should increase as the threshold falls, resulting in a small increase in variance at the shorter intervals instead of constancy. In such cases occasional detections of stimuli just beyond the range

might lead subjects gradually to extend their range, so that a constant threshold variance would become more likely with practice.

On the 'arousal hypothesis it could be supposed that the fall in  $c$  produced by the mechanisms suggested in alternatives (b) and (c) might not be of the same extent on successive trials, so that there would be some variation in the value of  $c$  during the course of a session. It might also be supposed that this variation would be greater when the average fall in  $c$  is greater. This would provide a source of variance that might cause an increase in threshold variance to be shown as the inter-stimulus interval shortens, instead of constancy.

Though either of these hypotheses could account for a small increase in variance being superimposed on a mechanism tending to produce constancy of the variance with change in threshold level, it can be shown that they would not ordinarily be likely to produce increases big enough to compensate for the decreases in  $\sigma_H$  necessary to give the falls in threshold on hypothesis (a) or (d). Thus, unless extra assumptions are made, this experiment provides evidence tending to exclude the hypothesis that the fall in threshold is due to a decrease in the variance of the 'noise' distribution.

CHAPTER 6

SHORT INTER-STIMULUS INTERVALS

6.1 Introduction

In Section 5.1 the arousal and warning hypotheses were discussed, and it was shown that these might operate in a number of ways: changes in the afferent paths or central sensory mechanisms might alter the parameters attributed to the 'noise' and 'signal + noise' distributions in the threshold decision procedure, or the criterion used to determine the threshold decisions might be altered. There is some evidence of physiological changes occurring in the afferent and central pathways that could correspond to the first alternative. Of particular interest are the descriptions of centrifugal effects on the responses of sense organs to stimuli, or on the spontaneous discharges which may occur in afferent fibres. Kerr & Hagbarth (1955) examined the olfactory centrifugal fibres first described by Cajal. They found that stimulating the basal rhinencephalic area depressed the spontaneous activity of the olfactory bulb. Loewenstein (1956) stimulated the sympathetic nerve supply to isolated frog's skin and found that this lowered the threshold to the mechanical stimulation of touch receptors, and might cause spontaneous firing to appear. It is also well established that the discharge of muscle spindles is controlled by supraspinal centres acting through the gamma motor neurons (Kuffler, Hunt & Quilliam, 1951; Granit, 1955a). Hagbarth & Kerr (1954) showed that the afferent volley produced

by stimulating a dorsal root in the cat could be depressed by stimulating certain central structures, such as the midbrain reticular formation. Hagbarth & Fex (1959) examined post synaptic units in the dorsal grey horn and dorsolateral funiculus of the cat, and found that their spontaneous discharge could be increased or reduced by central stimulation. They noted that the centrifugal outflow might reduce spontaneous activity while at the same time the response to peripheral stimulation increased, and they suggested that spontaneous activity may be maintained by a tonic centrifugal outflow. Galambos (1956) showed that the discharge in the cat's auditory nerve which is evoked by a click can be reduced by stimulating the medulla in the region of the decussation of the olivo-cochlear path. He demonstrated that this was not due to intratympanic reflex action but was mediated by the olivo-cochlear bundle. Hernández-Peón, Scherrer, & Jouvét (1956) used clicks to evoke responses in cats with electrodes chronically implanted in the cochlear nucleus. They found that the responses were reduced if the animals were exposed to visual, olfactory, or painful stimulation, or if the midbrain reticular formation was stimulated, and they interpret this as an inhibitory effect, 'blocking' a pathway to which the animal was not 'attending'. Hernández-Peón, Jouvét & Scherrer (1957) found that the size of the potential in the cochlear nucleus which was evoked by a click declined to a minimum during the course of several thousand repetitions of the click. This habituation was maintained during sleep but not during barbiturate anaesthesia (in

which the reticular formation is depressed), which they considered suggested that an inhibitory effect was maintained by a tonic centrifugal action of the brain-stem.

Evidence of centrifugal effects on the eye has been obtained by Granit (1955b) who found an increase in the response of the cat's retina to a light flash following stimulation of the midbrain reticular formation, and by Dodt (1956), who recorded a spike in the rabbit's retina, which did not appear to be antidromic, on stimulating the contralateral optic tract. Hernández-Peón, Guzmán-Flores, Alcaraz & Fernández-Guardiola (1957) found that in cats attending to auditory or olfactory stimuli the responses to light flashes were reduced, both in the optic pathways and in the visual cortex. They concluded that the 'blocking' occurred at least partly at the retina. Unfortunately they did not control for the effect of changes in pupil size. This has since been investigated by Naquet, Regis, Fischer-Williams, & Fernández-Guardiola (1960) using curarized, unanaesthetized cats. They produced cortical desynchronization by stimulating the sciatic nerve or midbrain reticular formation, and found that this was accompanied by dilatation of the pupil, an increase in the amplitude of the evoked response in the optic chiasma and lateral geniculate body, and a decrease in the evoked response at the visual cortex. When atropine was used to produce mydriasis the evoked potentials at the chiasma and lateral geniculate body became large and invariant, but the cortical response varied as before with changes in

the degree of synchronisation. Horn & Blundell (1959) and Horn (1960) have questioned the assumption of Hernandez-Peón and his co-workers that reduction of evoked potentials implies 'blocking'; they suggest that it may occur in a modality being used to examine the sensory field, and Horn (1960) suggests it may be attributable to suppression of random neural noise in the afferent path. He describes reduction of the cortical response to a flash of light when the cat is looking at a mouse, which agrees with the results of Naquet *et al.*

Although the exact significance of some of these findings may be in doubt, and the meaning of a change in evoked responses may vary at different levels, and may be produced by different underlying neural changes in different situations, these studies do serve to demonstrate physiological mechanisms which might underlie changes in sensory thresholds. Alterations in these mechanisms might correspond to the changes in  $\sigma_N$  or  $M_N$ , and since they are physiological processes they would have some latency, however short. If the assumption is made that some such mechanism underlies an arousal or preparatory effect it follows that as the inter-stimulus interval decreases the fall in threshold will decrease to a minimum at an interval determined by the latency of the physiological effect, and the threshold will then rise, and will again be at the asymptotic level when the two stimuli are simultaneous. This prediction is tested in the next experiment.

## 6.2 Short inter-stimulus intervals

### Experiment 11.

In Experiments 1 - 6 the subject was forewarned of each trial by the experimenter saying "coming", but in Experiment 10 this was replaced by a pre-warning light which preceded the accessory stimulus by a fixed short interval. It seemed possible that this feature of the procedure might itself have an effect on the thresholds recorded. To see if this was the case the experimental design included two pre-warnings, a short one which was judged, on the basis of the earlier results, likely to show any effect that might be present, and a long one which it was thought would be ineffective.

### Apparatus and procedure

The experimental situation was similar to that in Experiment 10, the same apparatus being used to provide the pre-warning, accessory stimulation and critical stimulus, the latter being a 50 msec. increment in the constant tone. Two pre-warning intervals were used, 1 second and 8 seconds, and 5 inter-stimulus intervals, 0.0, 0.3, 0.6, 0.9, and 1.5 seconds, in that order. Thresholds were determined by the method of limits, descending series of stimuli at intervals of 0.1 db being used. A block of five threshold determinations, one at each inter-stimulus interval, was repeated four times in the course of a session, which lasted about an hour. Only one pre-warning was used in any session. Eight subjects were used, each doing two sessions. Four subjects had the

8 second pre-warning for their first session, and the 1 second pre-warning for their second session, and the other 4 subjects had the reverse order. Subjects responded 'Yes' or 'No' verbally; they were in telephonic communication with the experimenter.

### Results

The mean thresholds, standard deviations, and standard errors were calculated as in Experiment 1, and are given in Table 15. An analysis of variance of the subjects' thresholds was made, and the regressions defined by equations (21) and (22) were calculated for the inter-stimulus intervals 0.3 - 1.5 seconds, and their significance tested by analysis of variance of the thresholds at those intervals.

### Discussion

The most notable feature of these results is that there is no evidence of any minimum in the curve, nor does the threshold return to the asymptotic level at the zero interval; instead the curve falls smoothly to a minimal value at zero. This is not in accord with the hypothesis that the accessory stimulus sets up a change in the sensory mechanisms effectively altering the parameters of the 'noise' and 'signal + noise' distributions; and that the decision procedure applied to the afferent input is altered pari passu with this change. The failure to find any evidence that the reduction in threshold has the latency that might be expected of a physiological response cannot be attributed to the latency being considerably shorter than 0.3 seconds;

TABLE 15

Mean Thresholds (decibels)

<u>Time Intervals (seconds):</u>	0.0	0.3	0.6	0.9	1.5
<u>Pre-warnings</u>					
<u>1 second</u>					
Means:	0.541	0.585	0.658	0.667	0.707
S.D.:	0.142	0.075	0.086	0.139	0.093
S.E.:	0.025	0.013	0.015	0.025	0.016
<u>8 seconds</u>					
Means:	0.615	0.633	0.641	0.660	0.712
S.D.:	0.091	0.121	0.134	0.100	0.130
S.E.:	0.016	0.021	0.024	0.018	0.023
<u>All</u>					
Means:	0.578	0.609	0.650	0.663	0.709
S.D.:	0.119	0.100	0.113	0.121	0.113
S.E.:	0.015	0.013	0.014	0.015	0.014

Analysis of variance (Thresholds)

<u>Source</u>	<u>d.f.</u>	<u>M.S.</u>	<u>F</u>	<u>Significance</u>
Subjects (S)	7	100,767	22.26	<u>p &lt; 0.001</u>
Times (T)	4	40,936	9.04	<u>p &lt; 0.001</u>
Pre-warnings (P)	1	13,690	3.02	NS (p < 0.1)
S x T	28	6,351	1.40	NS (p < 0.2)
T x P	4	4,643	1.03	NS
S x P	7	30,459	6.73	<u>p &lt; 0.001</u>
<u>Residual</u>	28	4,527	-	

Regression Equations

$$L = 0.713 - 0.033(1/T) \text{ (p < 0.001)}$$

$$L = 0.855 - 0.030(1/\Delta T) \text{ (p < 0.001)}$$

in that case a minimum should have been found at the latter interval, and the threshold at 0 seconds should have been higher than that at 1.5 seconds. This is clearly not so, and the experiment therefore would seem strongly to support the hypothesis that the change is in the computation and application of the criterion, rather than in the parameters of stimulus and noise.

The thresholds and standard deviations are high when compared with those in Experiment 4. This is probably attributable to the difference in the stimulus, which was 300 msec. long in the earlier experiment, and 50 msec. here. The thresholds can be related equally well to  $(1/T)$  or to  $(1/\Delta T)$ : in each case the regression is highly significant on analysis of variance, and in neither was there a significant deviation from linearity of regression.

The finding that the threshold fall continues to zero removes one doubt about the proper explanation of the results of Experiment 6, in this way increasing its significance as evidence against the arousal hypothesis.

The standard deviations show no suggestion of consistent change as the threshold falls; however the range used is small as compared with earlier experiments.

The thresholds for the two smallest inter-stimulus intervals are lower with the 1 second than with the 8 second pre-warning, but on analysis of variance neither the effect of pre-warnings, nor that of the

interaction between Pre-warnings and Times are significant. The absence of an effect of the pre-warning would be in keeping with the warning hypothesis, since, if the only effect of an earlier stimulus is to mark the beginning of a time interval, then the most reliable temporal information would be given by the stimulus most closely preceding the critical stimulus by a regular interval, and there would be little contribution from any earlier stimulus to this information. If, however, the accessory stimulus causes arousal, then if two successive stimuli are given at an appropriate interval their arousal effects might be expected to sum, and so to give a greater decrease in the threshold. However ineffectiveness of the pre-warning could be reconciled with the arousal hypothesis by supposing that the accessory stimulus causes maximal arousal, and so does not allow of summation. The highly significant interaction between Subjects and Pre-warnings indicates that the difference between the two pre-warnings did have an effect on the performance of some of the subjects. Examination of the individual results showed that the threshold for the zero inter-stimulus interval was lowered by the 1 second pre-warning by 0.3 - 0.4 db for two subjects, but for the remaining six subjects it was lowered by 0.02 to -0.07 db. The simplest explanation for this difference might be that in these two subjects the short pre-warning functioned by ensuring that they were looking at the right spot when the accessory stimulus was switched on, but that after 8 seconds their gaze had sometimes wandered. Alternatively,

we could suppose that for these subjects the short pre-warning served to mark the beginning of an interval, or caused a degree of arousal, and that this reduced the thresholds at the shortest inter-stimulus intervals. This makes it possible to suppose that the minimal threshold at the zero interval might be artefactual, due to the pre-warnings acting to a greater or lesser extent for various subjects as an 'accessory stimulus', and so obscuring the rise in threshold that might otherwise have been found. This possibility is tested in the next experiment.

### 6.3 Randomized pre-warnings.

To exclude the possible effect of the pre-warning it was decided to randomise the pre-warning intervals over a large range, as in Experiment 6, since in that experiment this was found to prevent the monotonic fall in threshold appearing.

#### Experiment 12

In this experiment two extreme conditions of accessory stimulation were compared: an accessory stimulus at zero interval, and no accessory stimulus. The latter condition was considered likely to correspond in its effect to an extremely long inter-stimulus interval, and so perhaps to afford the best estimate of the asymptotic level of the threshold.

#### Apparatus and procedure

The same accessory and critical stimulation were used as in Experiment 10(b), but the length of the pre-warning interval was controlled

by the two timing circuits described in Experiment 5. The 'minimum' was 2 seconds, and the 'range' 5 seconds, so that a rectangular distribution of pre-warning times ranging from 2 - 7 seconds was produced. Each trial began with the upper neon light coming on, and this was followed after an interval of 2 - 7 seconds by a critical stimulus, a 45 msec. increment in the intensity of the 500 c.p.s. tone. In one condition this was given alone, that is, there was no accessory stimulus ('NAS'). In the other a simultaneous accessory stimulus ('SAS') was given, the second neon being switched on at the same moment as the critical stimulus. Both lights stayed on until the subject had responded. The length of the pre-warning interval was recorded on a bank of Dekatrons. The experimental procedure was similar to that of Experiment 10(b). The method of constant stimuli was used, each series consisting of four randomly ordered stimuli at each of four stimulus strengths (0.22, 0.33, 0.44, and 0.55 db) and 4 'blanks'. In each session series of 20 stimuli under one condition were alternated with series under the other until 5 of each had been given, the session lasting about an hour. Two experienced subjects each did three sessions.

### Results

The results were treated by probit analysis and analysis of  $\chi^2$  in the same way as in Experiment 10(b), and are given in Table 16. In the analysis of  $\chi^2$  there is 1 degree of freedom for each of the values of  $\chi^2$  given for departures from parallelism; no significant evidence of

TABLE 16

<u>Subject</u>	<u>Condition</u>	<u>Threshold(db)</u>	<u>S.E.</u>	<u>S.D.</u>	<u><math>\chi^2</math></u>	<u>Signifi- cance</u>
M	SAS	0.394	$\pm 0.011$	0.084		
	NAS	0.474	$\pm 0.014$	0.107		
	Combined			0.095	1.341	NS
C	SAS	0.331	$\pm 0.012$	0.114		
	NAS	0.439	$\pm 0.012$	0.109		
	Combined			0.111	0.052	NS

Condition: SAS NAS

False positives (per cent): 6.7 8.3

Pre-warnings

(seconds): 2.00-2.99 3.00-3.99 4.00-4.99 5.00-5.99 6.00-7.00

Mean thresholds

(db): 0.395 0.388 0.383 0.383 0.388

Analysis of variance:

<u>Source</u>	<u>d.f.</u>	<u>M.S.</u>	<u>F</u>	<u>Significance</u>
Conditions(C)	1	0.0423	52.88	$p < 0.01$
Subjects(S)	1	0.0039	4.88	NS ( $p < 0.1$ )
Pre-warnings(P)	4	0.0001	0.13	NS
C x S	1	0.0024	3.00	NS ( $p < 0.2$ )
C x P	4	0.0017	2.13	NS ( $p < 0.2$ )
P x S	4	0.0002	0.25	NS
<u>Residual</u>	4	0.0008	-	

residual heterogeneity was found in the analysis. To see whether the pre-warning had had any effect a threshold was determined by probit analysis for each fifth of the range of pre-warnings for each condition and subject, and an analysis of variance was performed on these 20 thresholds. Each false positive rate is based on 120 'blanks'.

### Discussion

The fall in threshold found in the previous experiment is clearly present. For each subject a t-test showed the two threshold values to be significantly different ( $p < 0.001$ ), and on analysis of variance the difference is significantly present when the thresholds are calculated separately for each sub-range. The mean thresholds at each sub-range show to a slight extent the fall in threshold at the middle of the range found in Experiment 6, but this is not significant in the analysis of variance, confirming that the experimental design has been successful in preventing the pre-warning contributing to the results. It is also of interest to note that the results of these two subjects confirm their showing in Experiment 10(b): the standard deviation shows no significant change under the two conditions of accessory stimulation.

The fall in threshold cannot be accounted for by an increase in 'guessing' since the false positive rate is lower with the simultaneous stimulus than in the NAS condition. This result, which is similar to the tendency found in Experiment 10 for the false positive rate to be lower at the shorter inter-stimulus intervals, agrees with the analysis in Section 5.1.

It has now been shown, both here and in the previous experiment, that the threshold does not return to the asymptotic level at the zero inter-stimulus interval. This should be strong evidence against any explanation which attributes the threshold fall to a change in the afferent pathways; it tends to exclude alternatives (a) and (b), and (d) and (e), discussed in Section 5.1. There is, however, one difficulty in accepting this conclusion. The accessory stimulus was the switching on of a neon light, which took less than a millisecond to reach its maximum brightness, and the critical stimulus was an intensity increment lasting 45 or 50 msec. It is necessary to consider whether there was time for impulses resulting from the visual stimulus to arrive centrally and initiate a response with a very short latency that could have affected the auditory pathways before the critical stimulus was concluded.

It is usually found that the reaction time to a visual stimulus is longer than that to an auditory stimulus, a common difference being 40 msec. (Woodworth & Schlosberg, 1955; Davis, 1957), and there is evidence suggesting that the difference may be mainly due to a longer time elapsing before the visual signal reaches the central mechanisms. Marshall, Talbot & Ades (1943) found that in the cat the appearance of a cortical response to a flash had a latency of 17 - 25 msec. This has been confirmed by Monnier (1952) who finds a time of 23 msec. for the cat, and a time of 33 - 42 msec., with an average value of 35 msec., in man. He divides this time into a retinal time of 30 msec. and a retino-cortical time of 5 msec.

The peripheral conduction time for auditory stimuli appears to be considerably shorter. In animals the first cortical response appears about 7 - 10 msec. after the stimulus is applied to the ear (Tunturi, 1945, 1946; Galambos, 1954), and the time in man can probably be taken to be of the same order.

Since the peripheral conduction time is so much less for auditory stimulation, a short auditory stimulus might reach the central mechanisms before the visual stimulus could bring about any change in the auditory pathways or central mechanisms. However if we take the peripheral auditory conduction time to be 10 msec. and the visual time 35 msec., the auditory input from the 45 msec. critical stimulus used here would have continued to arrive for some 20 msec. after the visual stimulus had reached the cortex, even if the exponential 'tail' of the pulse is not allowed for, so that any physiological effect of the accessory stimulus not requiring many synaptic crossings could have affected the input from the last part of the tone pulse. This possibility is increased by the fact that strong visual and weak auditory stimuli were used. It is known that reaction times lengthen as the magnitude of the stimulus decreases (Woodworth & Schiffsberg, 1955), and there is some evidence that this could be at least partly due to an increase in the peripheral conduction time as the flash or auditory stimulus decreases in intensity (Monnier, 1952; Bartley, 1934; Kemp, Coppée & Robinson, 1937). The visual peripheral conduction time of 35 msec. was obtained with a very intense stimulus, and

so is unlikely to have been shorter in the present case. However, since the auditory stimulus was at threshold we can probably suppose its conduction time to have been several milliseconds longer than 10 msec. Thus it seems that the experiments in this chapter do not provide conclusive evidence against a preparatory or reflex action on the sensory channels. This point will be examined further in Chapter 8.

CHAPTER 7

RANDOMIZED INTER-STIMULUS INTERVALS

7.1 Introduction

In Chapter 6 evidence was obtained suggesting that the fall in threshold at short inter-stimulus intervals is due not to the accessory stimulus inducing a change in the sensory mechanisms, but to an alteration in the criterion determining the threshold decisions, a change of types (c) or (e). This could be a change of fixed time course, as suggested by the arousal hypothesis, or it could depend on the ability of the subject to estimate time, as required by the warning hypothesis. However a third possibility has not as yet been discussed. In Section 5.1 it was shown that, for a constant value of  $P_N(A)$ , the false positive rate would be less the shorter the inter-stimulus interval. If the subject maintained a constant criterion throughout the full length of the inter-stimulus interval on each trial, but the particular value taken by  $c$  was determined by the requirement that the false positive rate must be kept constant, then a decrease in the inter-stimulus interval would allow  $P_N(A)$  to increase. In this case there would be an inverse relation between the threshold fall and the length of the 'waiting time', the interval from the accessory stimulus to the moment that the subject terminates the trial if he has not received the critical stimulus. This period is usually a little longer than the inter-stimulus interval, since if a subject responds 'No' and then receives the critical stimulus he tends to lengthen his

waiting time in later trials. This hypothesis would be in accord with the relation between threshold and time interval found in Section 4.6, since as the number of 'moments' in the waiting time decreases a greater risk of a false positive at each can be accepted. It would imply a constant threshold variance, and would be compatible with the results of Experiments 11 and 12. In the next section an experiment is designed to test these hypotheses.

## 7.2 Randomized inter-stimulus intervals

If the subject maintains his criterion at a constant, lowered level, throughout the waiting time, then varying the inter-stimulus intervals in a given series over a small range about the end of the inter-stimulus interval might increase the 'waiting time' slightly, and so give a slightly higher threshold, but it should not alter the general form of the threshold curve.

In the discussion of Experiment 6 it was suggested that an argument against accepting that this experiment had provided conclusive evidence against the arousal hypothesis was the possibility that habituation might have occurred. Having the critical stimulus arrive at about the expected time might be necessary to prevent a decrease and disappearance of the arousal provoked by repetitions of the accessory stimulus, which might otherwise normally occur. If the randomization is over a small range only, then the accessory stimulus will still have some relation to the time of arrival of the critical stimulus, which should preclude this

hypothetical habituation. Therefore the arousal hypothesis would lead us to expect that the monotonic fall in threshold found earlier would not be changed by this randomization.

The warning hypothesis, however, leads to a different prediction. At time intervals long enough for the range of randomization to be considerably less than the range of expectation the randomization would have little or no effect. It would cause little increase in the number of occasions on which the critical stimulus falls outside the range of expectation; consequently the threshold level for a given time interval should be little affected. On the other hand, when the time interval is short enough for the range of randomization to be considerably larger than the range of expectation, two effects might occur. The subject might extend the range of expectation to embrace the range of randomization. In this case the range of expectation would become approximately constant in length, instead of decreasing, with further decrease in the inter-stimulus interval, so that the threshold would tend to remain constant for these short times. Alternatively, the range of expectation might be unaffected by the randomization, decreasing as the inter-stimulus interval shortens. However, this too would have little effect on the threshold. When the range of expectation is mainly within the range of randomization the threshold judgments will be determined by two criteria:  $c_a$  for critical stimuli that fall outside the range of expectation, and  $c_r$  for stimuli falling within it. If the range of expectation shortens  $c_r$  will be reduced,

and more of the stimuli falling within this range will give positive responses. However, since the range is shorter, more stimuli will fall outside it and have their effect determined by  $c_a$ . These two factors would be opposed, and the threshold recorded would therefore tend to be constant. Thus, in either case, the threshold would fall in more or less its accustomed fashion as the longer time intervals decrease, then flatten out and become constant for the shorter intervals at which, previously, the steepest falls have been found. This prediction is in contrast to those suggested by both the arousal and the 'waiting time' hypotheses.

### Experiment 13

In this experiment there were two conditions, one in which the inter-stimulus interval was constant in a series, as in the earlier experiments, and one in which the intervals were randomized. To provide an adequate test of the predicted flattening of the curve, a range of inter-stimulus intervals was chosen, 0-3 seconds, which the previous experiments suggested should show a steep fall. A range of randomization was chosen to equal the range of expectation at the middle time interval; since the length of the latter range was not known it was guessed. It seemed plausible to suppose it would be approximately equal to 5 - 6 times the standard deviation; taking the latter as the value of  $\Delta T$  given by equation (17) for  $T = 1.5$  seconds, this suggested a range of approximately 1 second.

### Apparatus and procedure

The critical stimulus was a 50 msec. increment in the intensity of a constant 500 c.p.s. tone at 60-70 db SL produced by the same apparatus as in Experiment 10, and with the same pre-warning and accessory stimulation.

There were two conditions: in the 'Constant Condition' the apparatus was set up as in Experiment 10, and the timing circuits were adjusted to give a pre-warning of 1 second and inter-stimulus intervals of a length constant for each series. Six intervals were used: 0.25, 0.75, 1.25, 1.75, 2.25 and 2.75 seconds. The method of constant stimuli was employed, each series consisting of 21 randomly ordered stimuli: 10 'strong', 10 'weak', and 1 'blank'. On the basis of a preliminary threshold determination at the beginning of each session values were chosen for the 'strong' and 'weak' stimuli which seemed to be just above and just below the subject's threshold. These were always 0.1 db apart. Twelve series, two at each inter-stimulus interval, constituted a session of about an hour; these series were given in a random order different for each subject. The subject was told the length of the inter-stimulus interval, and given a demonstration of it with a stimulus well above threshold before each series.

In the 'Randomized Condition' the apparatus of Experiment 12 was employed. A 1 second pre-warning was followed by an inter-stimulus interval produced by the two randomizing timing circuits. The first timer

produced the 'minimum' and the second the 'range'. The latter was set at 0.5 seconds throughout the experiment, so that when the 'minimum' was fixed a rectangular distribution of stimuli over a half second range was produced. By adding a half second to the 'minimum' on half the trials of a series (in a randomly ordered sequence) a 1 second rectangular distribution could be produced with exactly half the stimuli in each half of the range. Three ranges were used: 0-1, 1-2, and 2-3 seconds. In this condition a series consisted of 42 stimuli: 10 'strong', 10 'weak', and 1 'blank' in each half range. Before each series the subject was told what the range would be. A session consisted of six such series, two at each range, in a counter-balanced order different for each subject. Thus in each condition a session consisted of 252 trials. Subjects responded 'Yes' or 'No' verbally; they were in telephonic communication with the experimenter. Eight subjects were used, each doing two sessions, one under each condition. Four subjects began with the 'Randomized Condition' and four with the 'Constant Condition'.

### Results

In each condition the percentage of positive responses to the 'strong' stimulus ( $P_S$ ), the 'weak' stimulus ( $P_W$ ), and for both together ( $P_M$ ), and their standard errors were calculated, as in Experiment 5, for each inter-stimulus interval, or half-second range, for all the sessions taken together. The false positive rates are each based on 16 'blanks' in the Constant Condition, and 32 in the Randomized Condition. The

proportions of positive responses for each series, subject, stimulus strength, and session had an angular transformation and analysis of variance performed on them. In this, all first and second order interactions with a probability of more than 0.2 on the first analysis were combined with the residual to give an improved estimate of the remainder variance. The results are given in Table 17. Since there was no significant interaction between Times and Intensities the values of  $P_M$  for each condition are plotted in Figure 9.

### Discussion

For both weak and strong stimuli, and in both conditions, the percentage of stimuli evoking positive responses increases as the inter-stimulus interval, or the midpoint of the half-second range, decreases. This produces a significant effect of Times in the analysis of variance. Since there was no significant interaction between Times and Intensities the values of  $P_M$  can be taken as best representing the effects of the two conditions. They show the effect predicted by the warning hypothesis: the more or less continuous rise in the percentage of positive responses as the interval decreases that is shown under the Constant Condition is present only for the longest time intervals in the Randomized Condition. At the shorter intervals there is little change in threshold. The difference between the conditions produces a significant interaction between Times and Conditions in the analysis of variance. A 1 second range was selected in the hope that it would produce a point of inflexion at about

Percentage Response

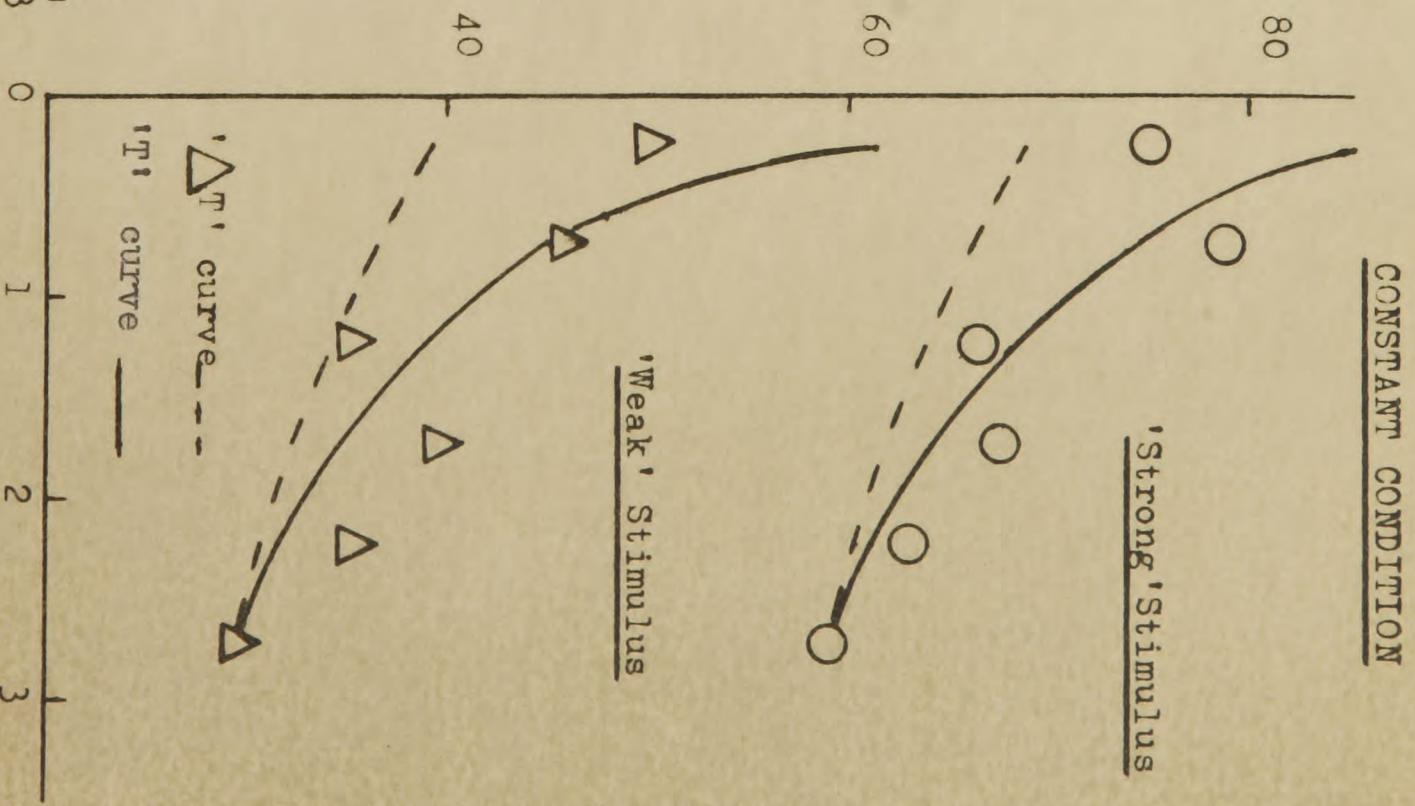
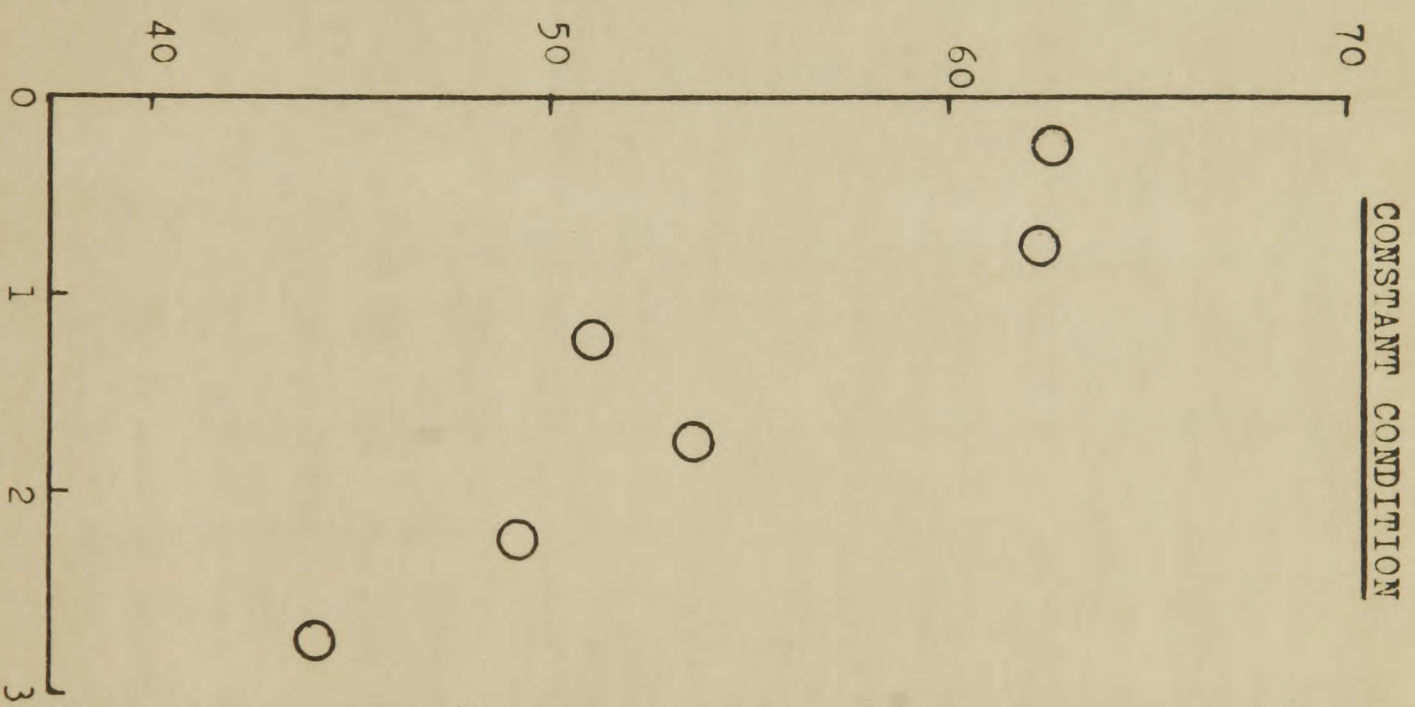
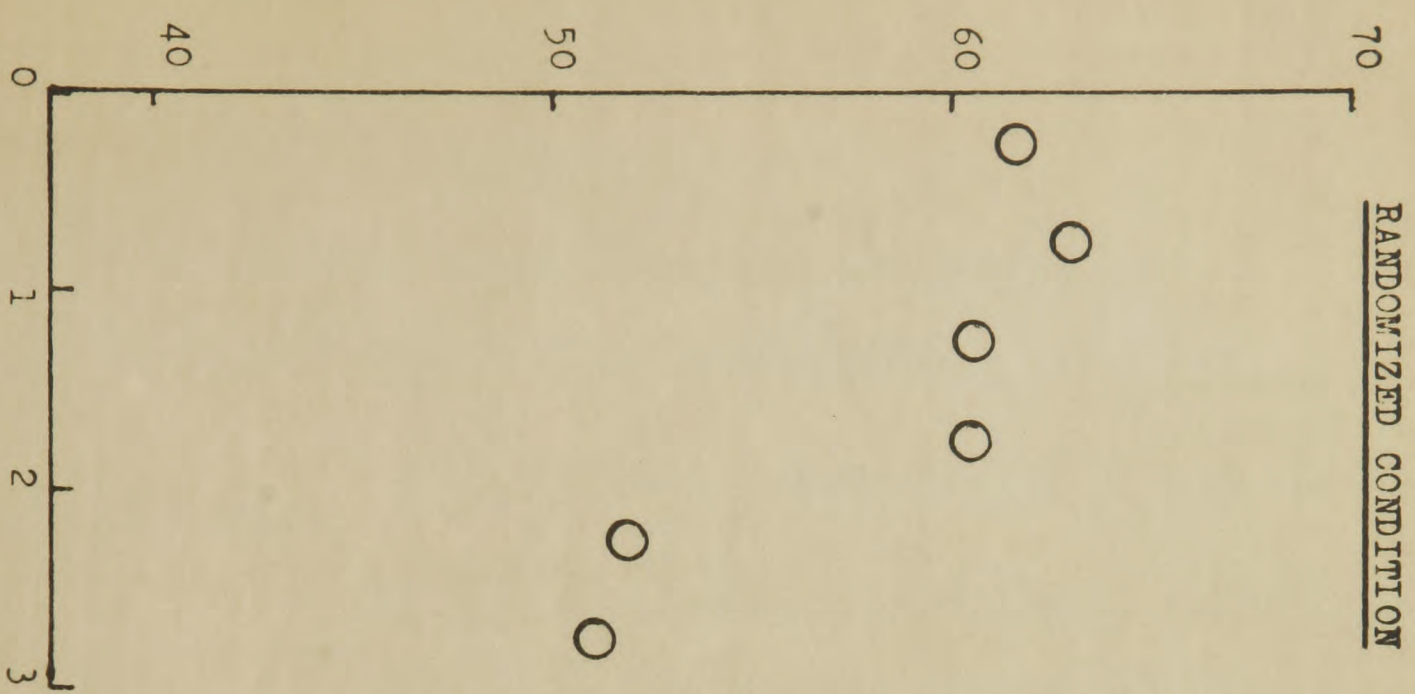


FIGURE 2

Seconds

TABLE 17

Percentages of Positive Response (and Standard Errors)Inter-stimulus interval, or sub-range midpoint:

	<u>0.25</u>	<u>0.75</u>	<u>1.25</u>	<u>1.75</u>	<u>2.25</u>	<u>2.75</u>
<u>Constant Condition</u>						
$P_W$	50.0 <sup>+4.0</sup>	45.6 <sup>+3.9</sup>	35.0 <sup>+3.8</sup>	39.4 <sup>+3.9</sup>	35.0 <sup>+3.8</sup>	29.4 <sup>+3.6</sup>
$P_S$	75.0 <sup>+3.4</sup>	78.8 <sup>+3.2</sup>	66.3 <sup>+3.7</sup>	67.5 <sup>+3.7</sup>	63.1 <sup>+3.8</sup>	58.8 <sup>+3.9</sup>
$P_M$	62.5 <sup>+2.6</sup>	62.2 <sup>+2.6</sup>	50.7 <sup>+2.7</sup>	53.5 <sup>+2.7</sup>	49.1 <sup>+2.7</sup>	44.1 <sup>+2.7</sup>
<u>Randomized Condition</u>						
$P_W$	50.0 <sup>+4.0</sup>	51.3 <sup>+4.0</sup>	42.8 <sup>+3.9</sup>	48.9 <sup>+4.0</sup>	38.5 <sup>+3.9</sup>	40.0 <sup>+3.9</sup>
$P_S$	73.1 <sup>+3.5</sup>	74.4 <sup>+3.5</sup>	78.1 <sup>+3.3</sup>	71.9 <sup>+3.6</sup>	65.0 <sup>+3.8</sup>	61.9 <sup>+3.8</sup>
$P_M$	61.6 <sup>+2.6</sup>	62.9 <sup>+2.6</sup>	60.5 <sup>+2.6</sup>	60.4 <sup>+2.7</sup>	51.8 <sup>+2.7</sup>	51.0 <sup>+2.7</sup>
<u>False positives (per cent):</u>						
Constant Condition:	35.7	14.3	7.1	0.0	14.3	7.1
Randomized Condition:	9.4		0.0			6.3

TABLE 17 (contd)

Analysis of variance

<u>Source</u>	<u>d.f.</u>	<u>M.S.</u>	<u>F</u>	<u>Significance</u>
Subjects (S)	7	509.36	13.00	<u><math>p &lt; 0.001</math></u>
Times (T)	5	480.59	12.27	<u><math>p &lt; 0.001</math></u>
Conditions (C)	1	322.14	8.22	<u><math>p &lt; 0.01</math></u>
Intensities (I)	1	14,789.82	377.58	<u><math>p &lt; 0.001</math></u>
S x T	35	66.67	1.70	<u><math>p &lt; 0.05</math></u>
S x C	7	378.00	9.65	<u><math>p &lt; 0.001</math></u>
S x I	7	196.03	5.00	<u><math>p &lt; 0.001</math></u>
T x C	5	99.46	2.54	<u><math>p &lt; 0.05</math></u>
S x T x C	35	119.15	3.04	<u><math>p &lt; 0.001</math></u>
S x C x I	7	106.26	2.71	<u><math>p &lt; 0.05</math></u>
Residual	81	39.17	-	

the middle of the range of intervals; in these results an inflexion appears at about 1.75 seconds. The percentage rise in  $P_M$  between the 2.75 and 1.75 second points is exactly the same for both conditions; from there to the 0.25 second point there is a further rise of 9.0 per cent in the Constant Condition, but of only 1.2 per cent in the Randomized Condition. To test this feature of the curves the significance of the differences between the value of  $P_M$  at 1.75 seconds and at the two shortest and longest intervals in each condition can be calculated and the probabilities combined (Fisher, 1954). In the Constant Condition the 1.75 second value is significantly higher than the values at 2.25 and 2.75 seconds ( $p < 0.01$ ) and significantly less than those at 0.25 and 0.75 seconds ( $p < 0.001$ ). In the Randomized Condition the 1.75 second is significantly greater than the values at the two longest intervals ( $p < 0.001$ ), but the probability of the difference from the values at the two shortest intervals occurring by chance is fairly large ( $p > 0.2$ ).

Another feature of these results which is compatible with both the 'waiting time' and warning hypotheses but not with the arousal hypothesis is the similarity of the values of  $P_M$  in the pairs of half-second ranges that jointly constituted the 1 second ranges presented to the subjects. On the warning hypothesis this procedure should have led to both half-second ranges showing the same threshold. When the range of randomization is less than the range of expectation the responses to all stimuli would be determined by the same value of  $c_r$ , and this should also be the case

when it is greater if the range of expectation expands to include it. Even if it does not,  $c_r$  and  $c_a$  should contribute in the same proportion to determining the responses in each half-second range, so in this case too the two sub-ranges should show the same threshold. However, this finding, which is not so clearly shown by  $P_W$  and  $P_S$ , might well be fortuitous.

More positive responses were given in the Randomized than in the Constant Condition, which accounts for the significant effect of Conditions. However this has little importance, since the two conditions were given at different sessions, and the values of the 'weak' and 'strong' stimuli were chosen independently at each session on the basis of the experimenter's rough preliminary assessment of the subject's threshold, so that an absolute difference between the two conditions could readily be accounted for by supposing that the stimuli that were chosen were on the average somewhat higher, in relation to the subjects' thresholds, in one condition than in the other.

The false positive rate shows a rise at the 0.25 second point in the Constant Condition. However it is a defect of the experimental design that far too few 'blanks' were given, so that little weight attaches to this variation in false positive rate. The difference between the 0.25 and 0.75 second rates is not significant at the 5 per cent level by the Fisher exact probability one-tailed test (Siegel, 1956).

These results are not in agreement with the 'waiting time' hypothesis, which predicts that little difference should be produced by

randomizing the inter-stimulus intervals over a small range. Thus they provide fairly strong evidence against it. Further evidence against it is the significant variation in threshold over the range of randomization found in Experiment 6. Nor do the results seem readily reconcilable with an arousal hypothesis. The changes produced by the randomization do not seem to be accountable for by habituation. Since the effect of Times is significant complete habituation cannot have occurred. Even if 'partial habituation' were assumed it would not be expected to have altered the form of the curve in the way that has been found. Thus this experiment is evidence for the warning hypothesis.

input is put through a decision procedure determined by the experimentally modified values of  $F_T(x)$  and  $F_{TT}(x)$ . On the second hypothesis the criterion computation is affected directly, and threshold change is a consequence of the employment of the new estimated values in the decision procedure. Decisions need not be made immediately the afferent input arrives externally; the letter could be stored periodically for a short period first. Thus a change in criterion could occur the same change in the response even to a stimulus whether the criterion change preceded or followed the arrival of the stimulus, provided it preceded the application of the decision procedure to the afferent input resulting from the critical stimulus. This would imply that the change in criterion is a 'judgmental' effect dependent on intervening time, or an 'abstract' level of processing.

CHAPTER 8

## THE EFFECT OF AN ACCESSORY STIMULUS SUBSEQUENT TO THE CRITICAL STIMULUS

8.1 Introduction

In Chapter 6 an attempt was made to see whether the fall in threshold at short intervals was due to a change in the 'noise' distribution, i.e. in  $\sigma_N$  or  $M_N$ , or to an alteration in the criterion,  $c$ , determining threshold decisions. For the first type of change to affect the threshold the changed values of the parameters of the 'noise' distribution resulting from the accessory stimulus must apply at the time the critical stimulus arrives, so that the corresponding afferent input is put through a decision procedure determined by the appropriately modified values of  $f_N(x)$  and  $f_{SN}(x)$ . On the second hypothesis the criterion computation is affected directly, and threshold change is a consequence of the employment of the new critical value in the decision procedure. Decisions need not be made immediately the afferent input arrives centrally; the latter could be stored peripherally for a short period first. Thus a change in criterion could cause the same change in the response made to a stimulus whether the criterion change preceded or followed the arrival of the stimulus, provided it preceded the application of the decision procedure to the afferent input resulting from the critical stimulus. This would apply whether the change in criterion is a 'judgmental' effect dependent on estimating time, or an 'alarm'

effect, in which the accessory stimulus causes a rapid reduction of criterion, with a subsequent slower recovery. In either case of 'criterion change' the exact inter-stimulus interval would not be critical, and a fall in threshold might be present even at zero. However a 'distribution change' produced by a 'reflex' or 'preparatory' effect must take some time, however short, to modify the distributions, so that this hypothesis would predict no threshold fall at inter-stimulus intervals less than some latency. Experiments 11 and 12 showed a fall in threshold at the zero interval; however this was not conclusive evidence for the 'criterion change' hypothesis because it remained possible that a 'distribution change' produced by the accessory stimulus might have taken effect before the tail of the critical stimulus had arrived at the central processes.

This possibility could be tested by administering the accessory stimulus at times slightly later than the critical stimulus. (When this is done the interval from the beginning of the first stimulus to the beginning of the second will, for convenience, be referred to as a 'negative' inter-stimulus interval). If the 'distribution' hypothesis is valid the threshold should rise rapidly as the accessory stimulus is retarded, and should have reached the asymptotic level when the effect of the latter no longer overlaps the input of the critical stimulus. It is difficult to estimate the length of the negative interval at which no effect would be shown, since the durations of the effects involved are

uncertain. However, an attempt will be made to make a very rough estimate, erring in the direction of over-estimation.

If the critical stimulus is a 45 msec. intensity increment, and the peripheral conduction times suggested in Section 6.3 apply, then we could suppose that the afferent input resulting from the stimulus would continue to arrive centrally for at least 55 msec. after the beginning of the stimulus. But there are also possible effects of the exponential 'tail' of the pulse, after-discharge from the auditory receptors, and temporal dispersion of the afferent impulses to be considered. The intensity increment decays with a time constant of 25 msec.; its rapid initial decline must soon weaken it to a point at which the reflex or preparatory effect could not make the afferent message it produces equivalent to the unmodified effect of the peak value of the pulse. After-discharges lasting for 20-30 msec. have been recorded in auditory nerve fibres following tone pips (Tasaki, 1954; Davis, 1959), though these were obtained with strong stimuli (50-70 db above threshold), and after-discharges do not appear to have been recorded when near-threshold stimuli were used. There is some temporal dispersion of the impulses arriving at the cortex, probably due to some of the impulses travelling along polysynaptic pathways. Since these impulses are likely to contribute less to the amplitude of the central effect and to be more affected by random noise than the impulses which travel directly, they are probably not important in determining the

threshold level. If, nevertheless, we allow 25 msec. for the decay of the pulse, 25 msec. for after-discharge, and 60 msec., the duration of the primary response at the auditory cortex (Brazier, 1951), to allow for temporal dispersion of the afferent impulses (even though a considerable part of the primary response should be attributed to the activity of intracortical neurons (Chang, 1959)), we get a figure of 165 msec. for the possible duration of an effect of the critical stimulus on the afferent input. If the neon light is regarded as having an instantaneous onset, 35 msec. peripheral conduction time, and a very short latency, say 5 msec., in producing its 'distribution' effect, then it should be unable to affect the threshold if it is retarded by 125 msec. or more. This figure is, of course, not much more than a guess. The times allowed are probably too generous, and it may well be a considerable over-estimation; however it serves to indicate that if falls in threshold are obtained with negative intervals much in excess of 125 msec. these will not be easily accounted for by the hypothesis of a 'distribution change'.

If the fall in threshold is the result of a change in the value of  $c$ , then, for 'positive' intervals, the modified criterion can be applied as soon as the afferent input is received. When, however, the accessory stimulus follows the critical stimulus the modified criterion must be applied to afferent inputs which have been stored. If the warning hypothesis applies retroactively, i.e. the accessory stimulus functions

as a temporal reference point, then the subject must retain sufficient temporal information to allow him to identify traces of the appropriate age when the accessory stimulus is administered. This may be more difficult than estimating a temporal interval whose beginning is clearly indicated. We may also suppose that the 'retroactive range of expectation' will be larger the greater the negative inter-stimulus interval, so that the fall in threshold will be less as this interval increases. The short-term memory store will also be important: the period during which a fall in threshold may be found will be limited to the duration for which stimulus traces can be preserved. If traces decay, so that the most recent inputs are the best preserved, then the latter will benefit most from a fall in the criterion; this, too, will tend to make the fall in threshold greater for the shorter intervals. If the arousal hypothesis is correct, decay in the memory store could result in greater falls in threshold being shown at shorter intervals. Thus in either version of the 'criterion change' hypothesis the threshold would be expected to return to the asymptotic level more quickly with negative than with positive inter-stimulus intervals.

## 8.2 Negative inter-stimulus intervals.

### Experiment 14(a)

This experiment was designed to investigate the threshold at three negative inter-stimulus intervals, using the method of limits. To ensure that any effect found would be an effect of the accessory stimulus,

and not an artefact produced by the pre-warning, the intervals between the pre-warning and the critical stimulus were randomized over a wide range.

#### Apparatus and procedure

The same situation and apparatus were used as in Experiment 12. The upper neon light, which served as the pre-warning, came on at the beginning of each trial. This was followed, after an interval which varied randomly between 2 and 7 seconds, by the critical stimulus, a 45 msec. increase in the intensity of a continuous 500 c.p.s. tone at an intensity of 60-70 db SL. The accessory stimulus was provided by the lower neon light which came on at a fixed interval after the critical stimulus. The subject responded by depressing one tapping-key for 'Yes' or another for 'No'.

There were five conditions of accessory stimulation: no accessory stimulus ('NAS'), -3 seconds (i.e., the accessory stimulus was switched on 3 seconds after the critical stimulus had been given), -1 second, -0.5 seconds, and 0 seconds. A threshold was determined for each inter-stimulus interval by the method of limits, descending series of stimuli being used. A block of five threshold determinations, one at each inter-stimulus interval, was repeated four times in a session, which lasted  $1\frac{1}{2}$ -2 hours. Two subjects each did two sessions; in one the series in each block were given in the order above and in the other they were given in the reverse order.

## Results

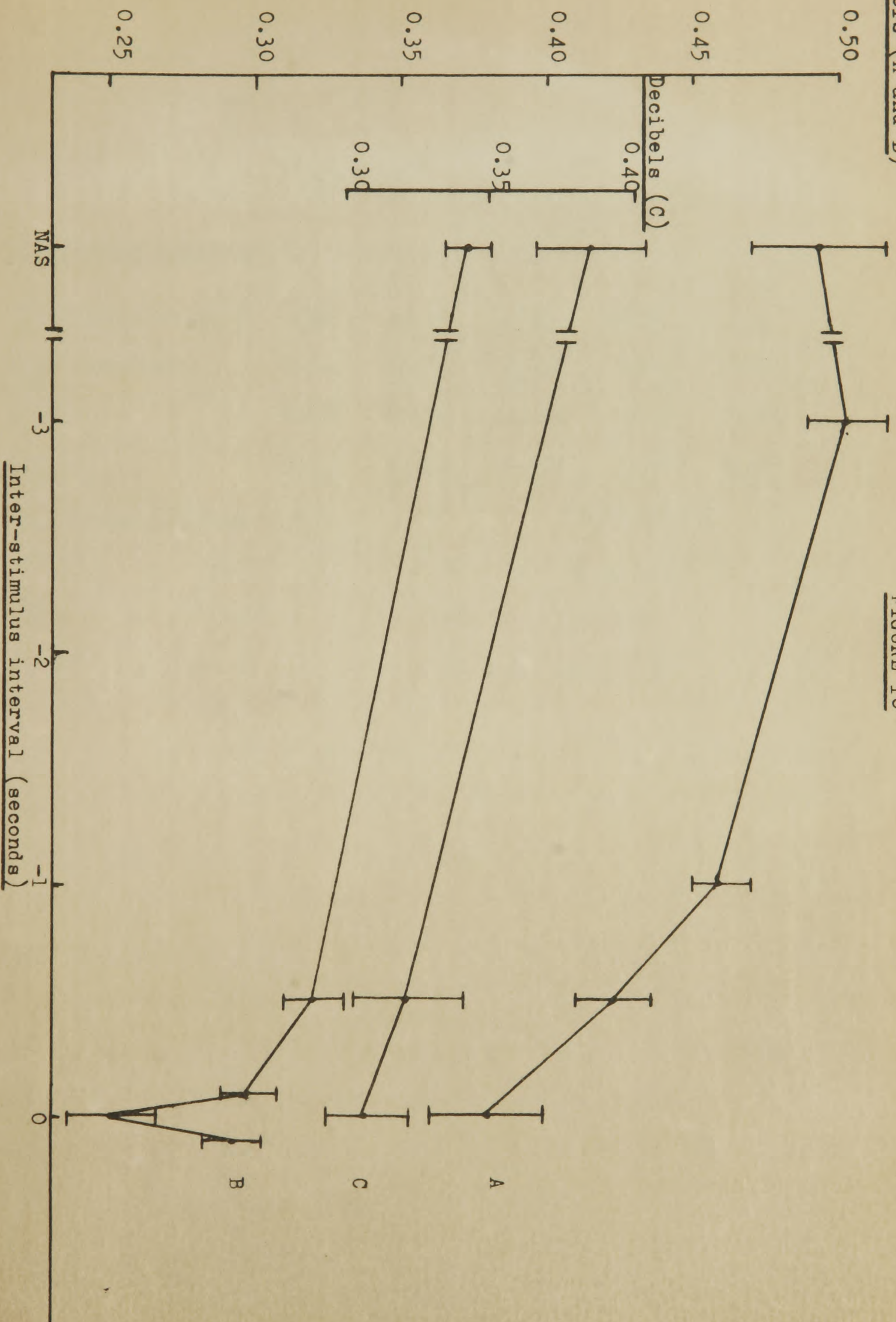
The mean thresholds, standard deviations, and standard errors are given in Table 18, and in Figure 10 the mean thresholds, with their standard errors, are plotted as Curve A. Student's *t* was calculated to give the (one-tailed) significance levels of certain of the differences between mean thresholds.

## Discussion

The significant fall in threshold at the zero interval, as compared with the threshold found in the absence of an accessory stimulus, confirms this finding in Experiments 11 and 12. If the fall in threshold were a 'distribution' effect the threshold should be approximately the same for the negative intervals used here and the NAS condition; but this is not shown. Instead there is a smooth rise from a minimum at the zero interval to a level at -3 seconds not significantly different from the 'asymptotic' threshold given by the NAS condition, which is consistent with the 'criterion' hypothesis. However the fall in threshold is only significant at one of the negative intervals, -0.5 seconds. Though there is a fall at -1.0 second it does not attain significance. The fall at the half-second interval might be produced by a tendency for the subjects to give false positives at that interval; though previous threshold falls have not proved explicable by 'guessing', this cannot be excluded at the negative interval since in this case the threshold was determined by the method of limits, which did not allow the false positive rate to be

Decibels (A and B)

FIGURE 10



NAS

Inter-stimulus interval (seconds)

attached. Therefore the results require qualification.

### Experiment 2(b)

TABLE 18

In this experiment the method of constant stimuli was used.

### Thresholds (decibels)

<u>Intervals (seconds):</u>	<u>NAS</u>	<u>-3.0</u>	<u>-1.0</u>	<u>-0.5</u>	<u>0</u>
Mean Threshold:	0.493	0.503	0.460	0.423	0.379
S.D.:	0.093	0.050	0.041	0.051	0.076
S.E.:	0.023	0.013	0.010	0.013	0.019

<u>Differences between means</u>	<u>t</u>	<u>Significance</u>
'NAS' - '-1.0'	1.30	NS (p = 0.104)
'NAS' - '-0.5'	2.64	<u>p &lt; 0.01</u>
'NAS' - '0.0'	3.79	<u>p &lt; 0.0005</u>

a series of 20 stimuli were given at each interval, 2 at each of 10 equally spaced stimulus strengths, and a 'block', in random order. A block of three series, repeated five times, had two sessions to complete. In half the sessions the three series in each block were given in the order above, in the other half in the reverse order. The length of the pre-warning interval was recorded as was that of the stimulus, and the reaction time, from the presentation of the critical stimulus to the subject's response, as usual. The subjects were not told that reaction times were being recorded nor was any effect made to influence their speed of response. Four subjects had all 5 sessions, 2 of them being under Condition (a) and 3 under Condition (b).

estimated. Therefore the result requires confirmation.

#### Experiment 14(b)

In this experiment the method of constant stimuli was used, allowing the false positive rate to be determined.

#### Apparatus and procedure

The same situation and apparatus were employed as in part (a) of this experiment. The interval from the pre-warning to the accessory or critical stimulus (whichever came first) was again randomized between 2 and 7 seconds. There were two conditions; in each session of each condition three intervals were used. In Condition (a) these were: NAS, -0.5 seconds, and -0.1 seconds, and in Condition (b): -0.1 seconds, 0.0 seconds, and +0.1 seconds. The method of constant stimuli was used; a series of 20 stimuli were given at each interval, 4 at each of 4 equally spaced stimulus strengths, and 4 'blanks', in random order. A block of three series, repeated five times, took two sessions to complete. In half the sessions the three series in each block were given in the order above, in the other half in the reverse order. The length of the pre-warning interval was recorded on one bank of Dekatrons, and the reaction time, from the presentation of the critical stimulus to the subject's response, on another. The subjects were not told that reaction times were being recorded nor was any effort made to influence their speed of response. Four subjects each did 8 sessions, 4 of these being under Condition (a) and 4 under Condition (b).

## Results

The responses to the 80 stimuli given at each interval in each pair of sessions were treated by probit analysis (Finney, 1952) to give thresholds, standard deviations and standard errors, and mean values were calculated for all the sessions. These results are given in Table 19, and the mean thresholds are also plotted in Figure 10 as Curve B. Two values are given for  $-0.1$  seconds, as this interval was employed in each condition. The one-tailed significance levels of some of the differences between mean thresholds are also given; these were determined by calculating the Critical Ratio (difference/S.E. of the difference) and referring it to the normal distribution. Each false positive rate is based on 160 'blanks', and the reaction times are averaged from 2139 positive, and 1354 negative reactions. In calculating these averages, all reaction times more than 3 seconds in length were excluded, as probably representing false positive responses. There were 11 of these, all in the NAS series. In Figure 11A combined distributions of the positive reaction times of two of the subjects who had similar median times are given. All NAS reaction times less than 1.7 seconds are plotted (they are taken to the nearest tenth of a second) as a continuous line. This distribution was subtracted from the corresponding distribution for positive reactions in the  $-0.5$  seconds series, and the distribution of the excess is graphed as a dotted line. The  $-0.5$  seconds distribution, which is not plotted, would equal the sum of these two distributions.

FIGURE 11

Number of Reactions

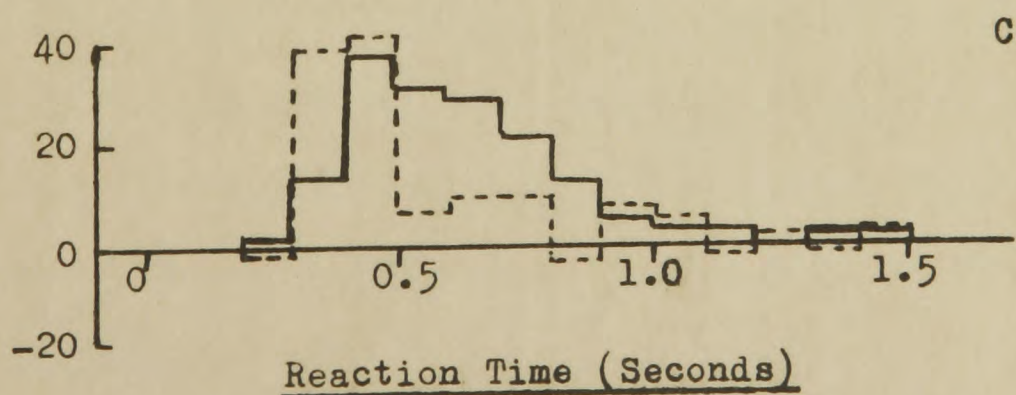
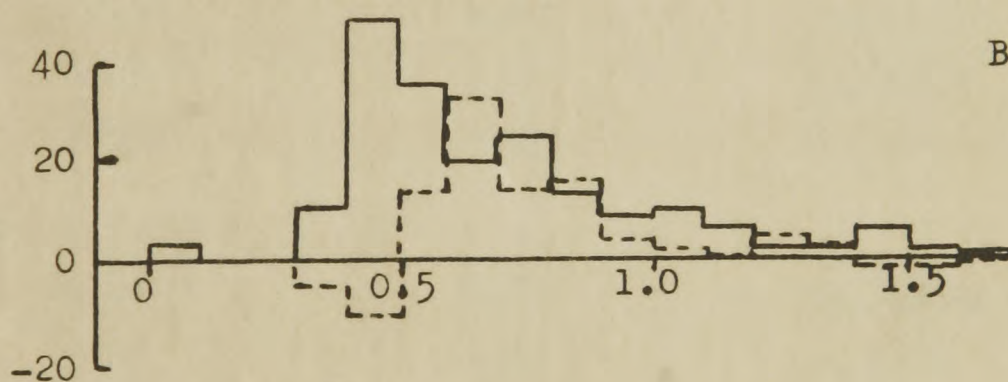
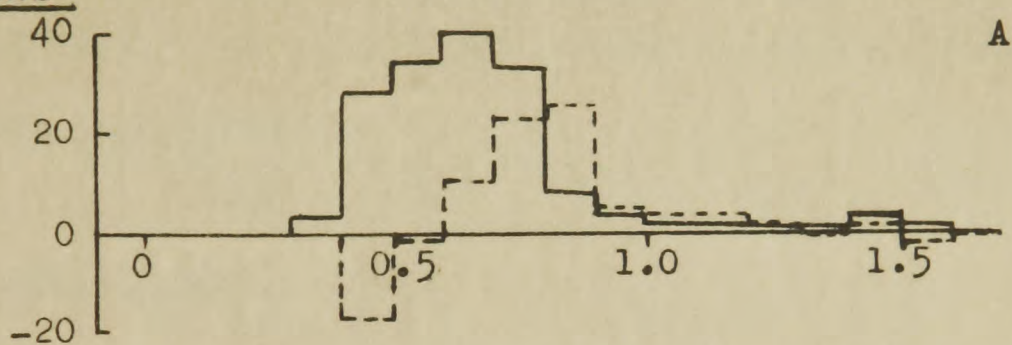


TABLE 19

	<u>Condition (a)</u>			<u>Condition (b)</u>		
	<u>NAS</u>	<u>-0.5</u>	<u>-0.1</u>	<u>-0.1</u>	<u>0.0</u>	<u>+0.1</u>
<u>Intervals (seconds):</u>						
<u>Thresholds (db):</u>	0.373	0.320	0.298	0.297	0.251	0.292
<u>S.D.:</u>	0.114	0.110	0.126	0.112	0.117	0.105
<u>S.E.:</u>	0.008	0.010	0.015	0.008	0.015	0.010
<u>Reaction Times (seconds):</u>						
Positive responses:	0.91	0.86	0.82	0.89	0.83	0.76
Negative responses:	-	1.39	1.08	1.13	1.03	1.04
<u>False positives (per cent):</u>						
	10.0	6.3	8.1	8.8	7.5	11.9

Differences between means:

<u>Condition (a)</u>	<u>Critical Ratio</u>	<u>Significance</u>
'NAS' - '-0.5'	4.14	<u>p &lt; 0.0005</u>
'NAS' - '-0.1'	4.42	<u>p &lt; 0.0005</u>
'-0.5' - '-0.1'	1.24	NS (p = 0.108)
<u>Condition (b)</u>		
'-0.1' - '0.0'	2.74	<u>p &lt; 0.005</u>
'+0.1' - '0.0'	2.35	<u>p &lt; 0.01</u>

Discussion

These results confirm the findings in part (a). The threshold is lowest at the zero interval and rises for both positive and negative intervals. The most interesting comparisons are with the asymptotic level given in the NAS condition; this is significantly higher than both the -0.5 and -0.1 second levels. The latter interval falls within the rather generous limit of 125 msec. suggested in Section 8.1 as the possible extent of the 'negative threshold curve' that might be produced by a 'distribution change'. However the significant fall in threshold with the negative interval of 500 msec., found both here and in part (a) of this experiment, does not appear to be compatible with the assumption that a reflex or preparatory effect of the accessory stimulus exerted on the afferent pathways at a time before the message resulting from the critical stimulus has arrived at the central processes is essential to the fall in threshold. Thus the experiment shows a retroactive effect which appears to be best explained by the hypothesis that the accessory stimulus produces a change in the criterion. There is no suggestion of consistent variation in the standard deviation, as would be expected of an effect which only alters the critical value. A fall in criterion produced by the accessory stimulus could only have an effect on the threshold if threshold decisions are delayed sufficiently long for the new criterion to be applied in making them. It is in accord with this that all the mean reaction times are above 0.5 seconds. It would, of

course, not be necessary to delay all the responses when the inter-stimulus intervals are negative. Stimuli producing a central effect greater than  $c$  might give an immediate positive response; some or all of the others might be held in store for a short time. If the accessory stimulus followed, lowering the criterion, some of these stored traces might then give rise to positive responses. This hypothesis is supported by Figure 11A. It can be seen that there are more positive responses with the  $-0.5$  second interval than under the NAS condition, and that the excess reaction times have a mean greater than the mean for the NAS condition or for the complete  $-0.5$  second distribution (when, of course, reaction times greater than 1.7 seconds are not considered). None of the extra responses have a reaction time less than 0.5 seconds; although there are more positive responses with the  $-0.5$  second interval than under the NAS condition, there are fewer responses under the first condition giving reaction times less than  $-0.5$  seconds. This would occur if some afferent inputs are held in store, when the accessory stimulus is expected, which, under the NAS condition, would give immediate positive responses. The significance of the difference between the number of positive responses with reaction times less than 0.5 seconds under the two conditions was tested using the normal approximation to the binomial distribution (Siegel, 1956); it was found to be significant ( $p=0.004$ , one-tailed). In Table 19 the mean NAS reaction time for positive responses is longer than that for the  $-0.5$  second interval; this is probably due to the inclusion

of a number of reaction times between 1.7 and 3.0 seconds in calculating the former, despite the possibility of their being due to false positive responses.

The reaction times for negative responses are uniformly longer than for positive responses. When the inter-stimulus interval is negative it contributes to the reaction time for negative responses; however the difference is also present at the 0.0 and +0.1 second intervals. Possibly a sample of the afferent input which is greater than the critical value sets the response in motion immediately, but if it is less a further one or two 'moments' or 'sampling times' must be examined before a negative response is made, even when the time of arrival of the stimulus is as exactly indicated as with the zero interval.

The false positive rates show that the fall in threshold with negative inter-stimulus intervals cannot be accounted for by an increase in guessing: the rate is higher for the NAS condition than for either of the negative intervals. However it is still possible to offer an explanation for these results which does not depend on the assumption that the occurrence of the accessory stimulus produces a change in the value of the criterion. The different intervals might have induced different 'sets' with effects on the threshold levels. Thus the NAS condition might have seemed dull, with a consequent soporific effect, but the expectation of an accessory stimulus might have been conducive to greater alertness. Increased alertness, or interest, or effort, in

one condition as compared with another, might correspond to a reduction in  $\sigma_N$  or in  $M_N$  for all trials in that condition. This could have the effect of lowering the threshold for that condition without causing a rise in false positive rate, although the criterion would not vary during the course of individual trials. To some extent this is contradicted by the evidence in Figure 11A that an increase in positive response is associated with an increase in delayed responses, since this implies that the change in threshold on a given trial depends on the changes occurring when the accessory stimulus arrives. However this explanation requires experimental testing.

#### Experiment 14(c)

In part (b) the stimuli were given in series of 20 at a given interval. As well as these 'simple' series, 'complex' series are also used in the next part of the experiment. In a 'complex' series 20 stimuli are given under two conditions of accessory stimulation, which apply to successive trials in a random sequence, so that if, for instance, the NAS condition is combined with accessory stimulation at an interval of -0.5 seconds, then on a given trial the subject does not know, at the moment that the critical stimulus is administered, whether the accessory stimulus will follow in half a second on that trial or not at all, only that its probability, on any trial, is 0.5. If a significant difference in threshold were found between the two conditions in a 'complex' series this could

not be attributed to the subjects' 'set' during the series, which would be the same for the different trials in the series.

#### Apparatus and procedure

The same apparatus was used as in part (b), with the same randomized pre-warning intervals, and critical and accessory stimuli. In each session there were 10 'complex' series of stimuli, each series consisting of 20 stimuli: 2 at each of 3 equally spaced stimulus strengths, and 4 'blanks', under each of two conditions of accessory stimulation, given in random order. In 5 'A Series' NAS and '-0.5 seconds' were combined, and in 5 'B Series' NAS and '0.0 seconds'. In 6 of the 10 sessions 'simple' series were also given. These were one-and-a-half NAS 'Y Series', and one-and-a-half '-0.5 seconds' 'Z Series'. The different series were given in the following order and its counter-balanced equivalents: A, B, Y, A, B, A,  $\frac{1}{2}$ Z,  $\frac{1}{2}$ Y, B, A, B, Z, A, B. A sequence such as this took two sessions, each of about an hour, to administer. Three subjects did 10 sessions, two doing 4 each and one doing 2.

#### Results

Each pair of counter-balanced sessions was combined and thresholds, standard deviations and standard errors calculated by probit analysis. The results of this experiment are given in Table 20, and the mean thresholds for the complex series are plotted as Curve C in Figure 10. Each false positive rate is based on 200 'blanks' for the complex series

TABLE 20

Series:	<u>A</u>		<u>B</u>		<u>Y</u>	<u>Z</u>
<u>Intervals (seconds):</u>	<u>NAS</u>	<u>-0.5</u>	<u>NAS</u>	<u>0.0</u>	<u>NAS</u>	<u>-0.5</u>
<u>Mean thresholds (db):</u>	0.385	0.322	0.385	0.308	0.414	0.375
<u>S.D.:</u>	0.158	0.115	0.101	0.119	0.166	0.079
<u>S.E.:</u>	0.031	0.019	0.023	0.014	0.017	0.016
<u>Reaction Times (seconds):</u>						
Positive responses:	0.74	0.69	0.63	0.62	0.62	0.67
Negative responses:	-	1.25	-	0.89	-	1.21
<u>False positives (per cent):</u>	4.5	4.0	4.5	7.0	0.0	2.4
<u>Differences between means</u>	<u>Critical Ratio</u>			<u>Significance</u>		
<u>A Series</u>						
'NAS' - '-0.5'	1.74			<u>p &lt; 0.05</u>		
<u>B Series</u>						
'NAS' - '0.0'	2.84			<u>p &lt; 0.005</u>		
<u>Y and Z Series</u>						
'NAS' - '-0.5'	1.63			p = 0.052		
<u>A and B Series</u>						
'-0.5' - '0.0'	0.60			NS		

and 40 for the simple series. In calculating reaction times any times greater than 3 seconds were omitted; there were 7 of these, all in the A Series NAS condition. The positive reaction times are derived from 934 and the negative from 243 responses. As all three subjects had similar median reaction times their results were combined to give the distributions in Figures 11B and 11C. In Figure 11B the NAS reaction times for the A Series are given as a continuous line, and the excess when this distribution is subtracted from that for the  $-0.5$  second interval as a dotted line. In Figure 11C the corresponding distributions for the B Series are given, the NAS reaction times as a continuous line, and the excess of the '0.0 second' distribution over the NAS distribution as a dotted line. Critical Ratios are calculated and (one-tailed) significance levels given for some of the differences between mean thresholds. No differences between NAS thresholds in the A, B, and Y series were significant, so these differences are not given in the table.

### Discussion

These results confirm the findings of parts (a) and (b). The threshold falls with the  $-0.5$  second inter-stimulus interval in the complex series as well as in the simple series. Since, however, the difference between the two thresholds in the A series is only significant at the one-tailed 5 per cent level three further sessions were obtained, one from each of three subjects (two of whom had served in the experiment), each session consisting of ten A Series only. When the results calculated

from these are combined with the A Series thresholds given in Table 20 the mean thresholds become: 'NAS' - 0.395 db, '-0.5 seconds' - 0.315 db, C.R. = 2.79 (p 0.005). Thus the effects of negative inter-stimulus intervals which were found earlier cannot be attributed to changes of 'set' between series. Nor can the falls in threshold be attributed to increased guessing, in view of the small changes in false positive rates.

The threshold also falls for the zero interval in the B Series, and in the Z Series. The Y and Z thresholds are higher than the corresponding thresholds in the A Series, but this is an artefact due to the inclusion of simple series in only 6 of the 10 sessions. The thresholds for the A Series for these 6 sessions taken alone are: 'NAS' - 0.420 db, '-0.5' - 0.364 db. These are clearly very similar to the Y and Z thresholds, so that there is no evidence that the use of complex rather than simple series affects the threshold fall. This is further borne out by Figure 11. The results with the complex A Series shown in Figure 11B are very similar to those in Figure 11A, based on the simple series used in part (b). The excess of positive responses with the -0.5 seconds inter-stimulus interval again has a mean reaction time greater than 0.5 seconds, and greater than the mean times for the NAS and '-0.5 seconds' distributions. This is in accord with the hypothesis that some responses which  $c_a$  would make negative are delayed, and positive responses are given instead if the criterion is lowered sufficiently by the administration of the accessory stimulus. Since, in the complex series, the subject

does not know in advance on which trials the accessory stimulus will occur, as many responses should be delayed on NAS trials as on '-0.5 second' trials, though the former will not become positive while the latter may. Figure 11B shows fewer '-0.5 second' responses with reaction times below 0.5 seconds, than NAS responses, but this is not significant at the (one-tailed) 5 per cent level, applying the normal approximation to the binomial theorem. It might be argued that reaction time lengthens as stimulus strength decreases, that the extra positive responses produced by a threshold fall will tend to be responses to stimuli which have been brought just above threshold, and that these should therefore be expected to have a greater mean reaction time than the total distribution, which will include responses to somewhat stronger stimuli as well. However the B Series, which is plotted in Figure 11C, affords a control for this possibility. In this case the extra positive responses produced with the zero interval show no tendency whatever to give greater reaction times than the complete NAS or '0.0 second' distributions. Thus the lengthening of the reaction time must be related to the expected accessory stimulus, and is not a consequence of the fall in threshold as such.

### 8.3 The retroactive effect of accessory stimulation

Experiment 14 has shown a retroactive effect of the accessory stimulus on threshold level which extends backwards for at least 500 msec., extra positive responses being made, after the accessory stimulus has been

administered, which are not false positives. Even allowing the rather generous limit of 125 msec., as calculated in Section 8.1, for the possible retroactive effect, with stimuli such as those used here, of a change in distribution parameters, such a change could not account for these results; they appear to require the assumption that the effect of the accessory stimulus is to change the criterion applied to a stored stimulus trace. However the experiment does not allow us to decide between the 'alarm' and 'judgmental' hypotheses; it is consistent with either. In the former we suppose that the criterion falls as an immediate response to the accessory stimulus, and that the extent of this fall has no relation to the length of the inter-stimulus interval. The new critical value is applied to any decisions not yet made, so that any trace still present in the short-term memory store may now give rise to a positive response, provided its central effect is greater than the lowered critical value. The limit to the retroactive effect would depend on the rate of decay in the short-term store (which might be taken to correspond to Broadbent's (1958) 'S system'), and the upward slope of the retroactive threshold curve in Figure 10 could be explained by this decay. If the traces decay steadily older traces will be less likely than more recent ones of equal initial magnitude to evoke a positive response when the criterion is lowered. The curve would become horizontal at the interval at which the decay of the traces fully cancels out the advantage given by the fall in criterion. To account for the delayed

responses shown in Figures 11A and 11B we could suppose that a stimulus that receives an immediate negative decision nevertheless remains in store and gets a 'second chance' if the criterion falls before it has disappeared.

On the judgmental hypothesis the arrival of the accessory stimulus alters the subjective probabilities attached to the occurrence of signal and noise at different past times. The effect of this would depend on the length of the negative inter-stimulus interval, and the accuracy with which the subject can identify traces of the appropriate age. The 'range of raised signal probability' resulting would be smaller for shorter inter-stimulus intervals, in analogous fashion to the relation between the range of expectation and the length of positive intervals, so that the reduction in  $c$  for inputs falling within the range would be greater for shorter than for longer negative intervals, with a relatively constant false positive rate. Both this effect, and the decay of traces in the store, would here contribute to producing a threshold curve rising rapidly to the resting level.

A few experiments have shown retroactive effects of one stimulus on the perception of another which do not, like metacontrast, appear to be specific to a particular modality. Child & Wendt (1938) measured the frequency of positive response to an auditory stimulus (a 165 msec. tone pip) presented at one of five intensities in the threshold

range, with a flash of light acting as an accessory stimulus. Five values of the inter-stimulus interval, ranging from -0.5 seconds to 2 seconds, were given in randomized order. They found significant rises in the frequency of response, as compared with their 'Tone Alone' condition, at intervals of 0, 0.5, and 1 seconds, with the greatest rise at 0.5 seconds, a "questionable" ( $p < 0.05$ , one-tailed) rise at -0.5 seconds (which was absent on the first two experimental days but "reliably present" on the second two days), and no significant effect at 2 seconds. In view of the results obtained in Experiment 14 the effect at the negative interval should probably be accepted. A point of interest is the similarity between the general form of these results and the curve obtained in Experiment 6. The latter was taken to provide evidence against the arousal hypothesis and it was suggested, in explanation of its form, that the subject might come to attach a greater value of  $P(SN)$  to moments towards the midpoint of the range, and lower his criterion accordingly. In Child's & Wendt's experiment subjective probabilities could have varied in a similar fashion for the different intervals, so as to be highest at 0.5 seconds and lower at the extremes. The increased frequency of response at the -0.5 seconds interval could then be explained in a similar way to the corresponding finding in Experiment 14: stored stimulus traces benefited by a fall in the criterion when the light-flash was administered. The improvement at this interval on successive days might have been a reflection of the course of the subject's probability

learning. The particular interest of this experiment is that though the negative effect, considered alone, could be given an 'arousal' or 'temporal reference' explanation, the former, as in Experiment 6, cannot easily be applied to the positive portion of the curve: there is no monotonic fall in threshold as the positive intervals decrease. Instead the threshold falls from 2 seconds to 0.5 seconds, and then rises again. The experiment as a whole is more easily explained by the 'temporal reference' hypothesis, and in this way constitutes evidence for this account of the effect at negative intervals.

Piéron & Segal (1939) investigated tactile sensitivity using condenser discharges to the base of the proximal phalanx of the middle finger. They compared the effect of a single shock with that of double shocks of equal intensity given at various inter-stimulus intervals, and found that the threshold was lower when the stimulus was repeated than when it was given singly. This was shown with intervals as large as 0.4 seconds, but not at 0.8 seconds, with most subjects. The fall in threshold was not simply due to summation; not only did the second shock become detectable at an intensity at which single shocks could not be felt, but so did the first. For intervals above 10-40 msec. both shocks were separately detectable at strengths below the threshold for single shocks. Thus there was evidence of a process of retroactive facilitation as well as summation. The effect appears to have a similar range of action to the

retroactive fall in threshold in Experiment 14; however it is difficult to account for it in terms of 'arousal'. We would have to suppose that subliminal stimuli may be arousing, and sufficiently markedly so for the residual effect of the first stimulus to bring the second stimulus above threshold, and for the effect of the latter on the criterion to allow the trace of the former to evoke a positive response. But if the changes in criterion produced by these subliminal stimuli have such marked residual or retroactive effects, over periods as long as 400 msec., it is puzzling that the falls in criterion do not bring the stimuli themselves similarly above threshold when they are given singly. If they have such effects on the criterion why are they subliminal when given alone? In contrast, the assumption that the subject takes account of temporal relations in determining the subjective probabilities of signal and noise, and chooses his criterion accordingly, can readily be applied here. We can suppose that when single shocks are given a threshold decision is made each 'moment' for which  $P_N(A) = \epsilon$ , where  $\epsilon$  is a constant. If a series of double shocks at a fixed inter-stimulus interval are given instead, and we suppose the subject can estimate this interval perfectly, he would do best to change his strategy to that of making decisions for pairs of inputs separated by the given interval. If, for a positive decision, he required each of the inputs to exceed the same value of  $c$  as before, he would now have  $P_N(A) = \epsilon^2$  for each decision. Since  $\epsilon^2 < \epsilon$  he could afford to reduce  $c$  to a value at which the false positive

rate would be back at its previous level. If the efficiency with which he can in fact make use of the interval decreases as it gets larger the fall in threshold will be greater at the shorter inter-stimulus intervals. This explanation provides a more plausible fit to the results of the experiment than does the arousal hypothesis, and to this extent supports the application of the 'temporal reference' hypothesis to similar experimental results.

Halliday & Mingay (1961) have made an interesting attempt to extend the findings of Experiment 14, using another modality of stimulation, and varying the strength of the accessory stimuli, and obtain results which they believe are inconsistent with the hypothesis that the accessory stimulus functions only as a temporal reference point or 'warning'. For their critical and accessory stimuli they used condenser discharges administered by silver electrodes applied to the back of each of the subject's forearms, the shock to one forearm acting as the accessory and that to the other as the critical stimulus. They examined the frequency of positive response at 5 inter-stimulus intervals, -0.5, -0.1, 0.0, +0.1, and +0.5 seconds, with 5 strengths of the accessory stimulus, ranging from +2 to +6 db above the "100 per cent threshold". The pre-warning interval was randomized between 2 and 7 seconds, and the different intensities of the critical and accessory stimuli were given in randomized order in a series. Each series had a fixed inter-stimulus interval. They found that the accessory stimulus affected the threshold of the critical

stimulus, and that this effect was shown not only at the zero inter-stimulus interval but also with positive and negative intervals. However the effect was in the opposite direction from that which has been found here: the accessory stimulus raised the threshold, rather than lowering it. They also found that the rise in threshold increased as a function of the strength of the accessory stimulus.

Halliday & Mingay consider their effect relevant to and a source of information about the mechanism producing the fall in threshold at positive and negative intervals which is shown in some of the experiments which have been described above. They argue that if the accessory stimulus acts merely as a "time marker" its strength should be irrelevant to its effect, provided it is perceived. But the intensity of the accessory stimulus did affect the magnitude of the rise in threshold, which pointed to "some active excitatory or inhibitory process of interaction between the two stimuli in the central mechanisms"; since the effect occurred retroactively it could not "be explained in terms of attenuation of the afferent volley".

Since they conclude that the accessory stimulus does not act as a warning, and accept that the retroactive effect excludes an effect on the afferent input, their explanation seems to imply either an effect on the trace, weakening it or tending to prevent it reaching the threshold decision mechanism, or an effect on the latter equivalent to a rise in the

criterion. The latter explanation would be analogous to that given as alternative (c) in Section 5.1. However, it may be doubted whether their effect can rightly be regarded as of the same sort as the falls in threshold described in Experiment 14 and in the earlier experiments. Not only do they find that the accessory stimulus induces a rise in threshold rather than a fall, but the time course is very different from that which has been found above. In Halliday's & Mingay's results the rise in threshold is greatest at the zero interval and at  $-0.1$  seconds, and it is present though less marked at  $+0.1$  seconds. But they find no significant difference in threshold level between series with the  $-0.5$  second and with the  $+0.5$  second interval, and between the  $+0.5$  second interval and the "No warning" (NAS) condition, indicating that the total range of the effect is less than 0.5 seconds in each direction, which is very different from the extent of the threshold changes described above. If their result is not clearly due to the mechanism under investigation here it does not necessarily constitute evidence about it. Nevertheless the problem remains of how their interesting finding can best be explained.

If the "excitatory or inhibitory process" suggested by Halliday & Mingay is interpreted, in terms of the signal-detection hypothesis, as corresponding to a rise in criterion, why should this occur, and why should its effect be retroactive? A possible explanation can be derived from the fact that they used accessory and critical

stimuli differing only in intensity (and not very greatly in this respect: the two weakest accessory stimuli were of equivalent strength to the two strongest critical stimuli used) and in their two symmetrically placed sites of application. The afferent pathways for two such stimuli might be so related that a spread of excitation could occur from one to the other. If this were so a stimulus at one site, say the left forearm, could sometimes cause 'cross-talk' in the corresponding pathway from the symmetrically opposite part of the right forearm. This hypothesis gains some support from the occurrence of disturbances such as alloaesthesia and synchiria in certain lesions of the nervous system (Critchley, 1953). With stimuli such as these shocks, which can perhaps be regarded as differing only in a simple 'one-dimensional' way, 'cross-talk' from the left pathway might not be distinguishable from random firing arising directly on the right, producing an equivalent effect to a temporary change in the 'noise' distribution for right-sided afferent inputs. (To facilitate further discussion the pathways affected by the critical and accessory stimuli will be referred to as the 'critical' and 'accessory pathways'). If, in this case, the probability of false reports of critical stimuli is not to increase when an accessory stimulus is given, the decision mechanism for the critical pathway must adjust its criterion for an appropriate period, which will correspond to the usual duration of the change in its 'noise' parameters. It would follow from this that the duration of the effect found by Halliday & Mingsay should be related to the

time course of the afferent changes resulting from the accessory stimulus. It is in accord with this hypothesis that the maximal effect was at about the zero interval, and that there was no effect at a positive or negative inter-stimulus interval of 0.5 seconds.

The stimuli used by Halliday & Mingay were very short: condenser discharges with a time constant of 10-20 microsec. Yet an effect was found with positive and negative inter-stimulus intervals as large as 100 msec. A number of factors can be suggested, some or all of which could have contributed to this: (i) In terms of the signal-detection model the subject's criterion should be adjusted in a way that approximately parallels the time course of the changes in the parameters of the noise distribution for the critical pathway. Even with an almost instantaneous accessory stimulus temporal dispersion of impulses traveling at different rates, and after-discharge from the stimulated receptors, would lengthen the duration of the corresponding afferent message, giving a corresponding increase in the period during which 'cross-talk' might appear in the critical pathway. This might provide a basis for an effect at positive inter-stimulus intervals. (ii) The time course of the change in the noise parameters which is allowed for by the threshold decision mechanism might be expected to be an average based on some or all of the subject's experiences with similar brief and fleeting stimuli. Since stimuli are rarely as abrupt in onset and termination as those used in this experiment, signal detections might often seem to be preceded and

followed by relatively high noise levels, due in fact to the subliminal effects of gradually applied or removed stimuli. This would contribute to an expectation of noise at both positive and negative intervals.

(iii) If the 'moment' or 'sampling time' hypothesis is correct, even if the afferent messages could be supposed to begin and end abruptly the time course of the change in the noise parameters would not appear to be correspondingly 'rectangular', since the beginning or end of the afferent message would sometimes fall within a sufficient proportion of a 'sampling time' to increase the apparent level of activity at that 'moment' without bringing it above the criterion. This would contribute to positive and negative results.

These considerations suggest that the 'criterion computer' would do best to raise the criterion for the critical pathway not only at the time that a message can be detected in the accessory pathway but also at times before and after this. One way in which the sensory discrimination mechanism could approximate to this performance would be to feed the afferent input in the accessory pathway into the criterion computer for the critical pathway in such a way that it continuously affects the critical value. This would result in the criterion being high when the accessory pathway is active and low when it is inactive. If a rise in the criterion corresponded to the induction of some degree of inhibition or refractoriness which would take some small period of time to disappear, this would have the effect of raising the criterion for a short time just

after the accessory stimulus is detectable. Some neurophysiological studies are of interest in this respect. Towe & Amassian (1958) examined the discharge patterns of individual units in the somatosensory cortex of the rhesus monkey in response to electrical and mechanical stimulation of the digits. They found that the discharge evoked in many of the units by a shock to the fingerpads could be prevented by prior or simultaneous stimulation at a nearby point; in these cases a brief shock in the inhibitory receptive field produced a rapidly developing depression of the evoked discharge lasting for about 150 msec. Jabbur & Towe (1959, 1960) examined the effect of stimulation of the cat's motor cortex on the discharge evoked in single units in the cuneate nuclei by cutaneous stimulation. In some cases a single shock or a train of shocks reduced their excitability for periods of 100-200 msec. Stoiber & Barnes (1960a, 1960b) similarly examined the effect of cerebellar stimulation on evoked responses in cuneate neurons, and in some cases found an inhibitory effect which disappeared in 150-200 msec. Thus it is neurophysiologically plausible to suggest that the accessory stimulus might have an inhibitory or depressing effect on the mechanism corresponding to the criterion, producing a rise in the critical value which might take 100-200 msec. to disappear.

Towe & Amassian (1958) did not obtain any effect with negative inter-stimulus intervals greater than 2 msec. The negative effect found at -0.1 seconds could however be accounted for if the latency of the

effect of the accessory stimulus on the criterion were less than the time taken for the afferent input from the critical stimulus to reach the decision mechanism. A relative delay of this sort might be produced in a number of ways: (i) A shorter or quicker pathway might be employed in bringing the afferent input in the accessory pathway to bear on the determination of the criterion than in bringing the input from the critical stimulus to its decision mechanism. (ii) The peripheral conduction time to the cortex increases as the intensity of the stimulus decreases (Towe & Amassian, 1958). This would delay the afferent input from the critical stimulus as compared with the accessory stimulus. (iii) If the sampling time hypothesis is correct, the afferent input from the critical stimulus will frequently be delayed while the sample is completed; it is assumed however that the activity in the accessory pathway does not undergo sampling for this purpose but acts continuously on the criterion computer.

Another, perhaps simpler, model which would also have the effect of producing a rise in threshold at short negative and positive intervals can be derived from the sampling time hypothesis. This is based on the assumption that there is a reciprocal inhibitory effect between inputs in the two pathways which fall in the same time-sample. The effect of this on the weaker input would be equivalent to a rise in its criterion, and the rise in threshold of the critical stimulus resulting from this would be a function of the probability of the peak effects of the two stimuli falling within the same sampling time. This probability would be unity

corpus callosum. Ender (1951) has shown that in normal cortex, when the two peaks arrive centrally simultaneously, and zero when they are separated by more than one sampling time, and would be a linear function of the interval between them between these limits. The effect would not depend on the order in which they arrive, as long as they fall in the same sampling time. Since the input from the weaker stimulus will travel more slowly than that from the strong stimulus it will arrive centrally somewhat later when the stimuli are administered simultaneously. If the weak stimulus is administered first this will tend to reduce the interval between the times at which the two inputs arrive centrally, if it is given second the interval will tend to be increased. Thus, in the experiment by Halliday & Mingay, the zero interval would correspond to a small interval between the central times of arrival of the two inputs,  $-0.1$  seconds would correspond to an interval less than 100 msec., and  $+0.1$  seconds to an interval correspondingly greater than 100 msec. It is in keeping with this that they found no significant difference between the thresholds obtained with intervals of  $-0.1$  seconds and zero, but that the threshold was significantly lower at the  $+0.1$  second interval.

The hypothesis that the 'insulation' of symmetrically opposite pathways from one another may demand a fairly high level of central nervous organization finds some support in an old observation made by Pavlov (1927). He noted that dogs found great difficulty in developing differential conditioned reflexes to stimulation of symmetrically opposite points on their bodies, but this difficulty disappeared on section of the

corpus callosum. Bender (1952) has shown that in normal humans tactual stimuli affecting symmetrical or related pathways may mask one another. An advantage of the 'cross-talk' hypothesis is that it suggests a link between the experimental findings and certain abnormal phenomena which can, perhaps, best be understood as exaggerations of features of nervous functioning which are normally present. Halliday & Mingay refer to the phenomenon of 'inattention', sometimes found with lesions of the parietal lobe, in which the patient is unable to perceive a stimulus on the affected side if there is simultaneous contralateral stimulation. It is possible that in such cases the lesion has produced a loss of 'insulation' of the affected pathways, with a corresponding increase of cross-talk from related sources, and the disorder is in fact the expression of the large compensatory rise in the criterion necessary when there is an input on a related pathway. Giblin's (1960) finding that in such a case there was no reduction in the cerebral potentials evoked by stimuli when 'inattention' to them was present supports the hypothesis that the change is in the central mechanisms. If the criterion did not rise in compensatory fashion stimulation on the sound side should produce an hallucination of similar symmetrical stimulation, and this symptom, synchiria, is sometimes reported in such cases (Critchley, 1953).

It appears, then that the findings of Halliday & Mingay do not constitute a reason for rejecting the 'temporal reference' hypothesis for Experiment 14. They appear to provide evidence of a different phenomenon,

very possibly critically dependent on their use of stimuli identical but for intensity differences and given at symmetrical locations, and which can be given an explanation from which the greater threshold rises produced by the stronger accessory stimuli follow naturally. Experiment 14 could be explained by either the 'arousal' or the 'temporal reference' hypothesis; but the experiments both of Child & Wendt (1938) and Piéron & Segal (1939) can be regarded as showing retroactive effects which in each case the former hypothesis, but not the latter, meets with difficulty in explaining. Taken together these experiments provide evidence of some weight for the temporal reference hypothesis.

#### 8.4 Short-term information storage

The results of Experiment 14 can be regarded as providing information about a short-term sensory information store. In each part of the experiment there was a significant fall in threshold at  $-0.5$  seconds, and in part (a) an effect was also present, though not significantly so, at  $-1$  second, so that the trace appears to be stored peripheral to the threshold decision mechanism for a period of the order of  $0.5 - 1$  seconds. A similar period is suggested by the threshold fall at  $-0.5$  seconds found by Child & Wendt (1938), and by the experiment of Piéron & Segal (1939), who found retroactive facilitation with subliminal shocks separated by  $0.4$  seconds, but not with a difference of  $0.8$  seconds. A short-term storage effect of similar duration has also been described by Sperling (1960) for vision.

Sperling presented a visual array of letters or numbers to his subjects as a tachistoscopic exposure, usually 50 msec. long. He compared the information in immediate memory when the subject attempted to report the whole of what had been exposed (the span of apprehension), with the total information available to the subject at various times after the exposure. The 'information available' was estimated by a sampling technique, the subject being required to make partial reports consisting of only one of the rows of letters in the display, a coded auditory signal (a tone of variable frequency) being given at a fixed delay after the flash to indicate which row was to be reported. In this way Sperling found that two to three times as much information was available to the subject immediately after the exposure as was contained in a whole report. The extra information disappeared rapidly, the accuracy of the partial reports closely approaching that of whole reports at a delay of 1 second. Thus the experiment appears to demonstrate in a somewhat different way a short-term 'S system' (Broadbent, 1958) with a duration of effect very similar to that found in Experiment 14.

Sperling's explanation for his findings is that information is stored as a persisting "subjective image of sensation", and that the high accuracy of partial report results from the "ability of the observer to read a visual image that persists for a fraction of a second after the stimulus has been turned off". An explanation couched in phenomenological

terms, such as this, has the grave disadvantage that we are thereby limited, in extending its application, to instances which are phenomenologically similar. To attempt to apply it to Experiment 14 would lead to odd results; we would have to say, in those instances where response was delayed until the arrival of the accessory stimulus, that the subject then 'read' or perhaps 'listened to' a persisting auditory sensation of image. For the A Series of Experiment 14(c) it would be necessary to ascribe the extra positive responses given on trials on which the accessory stimulus was administered to the effect of a persisting sensation which, if the accessory stimulus had not been administered, we would have said had never been present. It would seem desirable to have an account of short-term peripheral information storage that could be satisfactorily applied to both types of experimental situation.

The explanation has the further disadvantage that it does not cover all the phenomena which Sperling himself found. When the coded signal is made to indicate which row to report first, this row is considerably worse reported than when it is to be reported alone; yet the persisting visual image must be in much the same state when the coded signal is given in each case. If the array consists of both letters and digits, and the coded signal indicates that the subject must report one or other of these categories, no more information is available with partial than with whole report; phenomenological explanation does not account for this. And it offers no suggestion on what the limitation

may be that prevents the whole report transmitting all the information available for partial reports; in each case the process of making the report has much the same temporal relation to the exposure.

If, however, phenomenological temptations are scouted, an explanation can be offered which covers a number of features of the two sets of experiments by adapting Broadbent's (1958) hypothesis that the sensory input is stored for a short time in an 'S system' from which it may pass to a limited capacity 'P system', the 'immediate memory' storage system, entrance to which is protected by a 'filter'. We suppose that the S system corresponds to the short-term storage of a fading neural trace of the afferent input resulting from the stimulus, and that passage into the P system corresponds, in a situation such as that of Experiment 14, to the action of the threshold decision mechanism. Such a mechanism performs the simplest type of categorization, into one of two classes. A similar account would be given of Sperling's experiment: a neural effect of the tachistoscope flash is preserved as a stimulus trace which decays within about a second. The categorization of the trace, however, which corresponds to entry into the P system, is much more complex: it involves the identification of each of the presented symbols as one of a number of alternatives (with the preservation of their relative location).

This simple model suggests an explanation for the limit - the 'span of apprehension', or of immediate memory for simultaneously presented material - which prevents the whole reports transmitting all the

information available for sampling by partial report. It is not altogether satisfactory to attribute this to a limit to the capacity of the P system, since the immediate memory span for sequentially presented material such as digits or letters is usually of the order of 6 - 8 items, whereas the span of apprehension for unconnected letters is 4 - 5 (Miller, 1956; Crossman, 1961; Woodworth & Schlosberg, 1955; Sperling, 1960). It could, however, be attributed to the categorization system: if this required a minimum time to deal with each symbol, this, in conjunction with the rate of decay of the trace, would determine a limit to the number of symbols which could be correctly reported. Thus if the symbols were categorized sequentially, the categorization time for each being about 0.15 seconds, and the effective life of the stimulus trace was 0.75 seconds, the span of apprehension would have an upper limit of 5 symbols, even though the span of immediate memory might be greater than this. The system might be able to deal with symbols simultaneously, taking a correspondingly longer time. If this time were equal to or longer than the effective life of the trace in the S system, then the limit would be determined by the maximum number of letters that could be handled simultaneously, which we can take to be 4 - 5.

With the addition of some further assumptions the model becomes capable of explaining a number of findings. These assumptions are: (a) Mutual interference occurs between items in the S system, increasing their rate of decay. Some support for this assumption is provided by the old

observation that pairs or groups of letters may be more difficult to read than single letters (Woodworth & Schlosberg, 1955). (b) To the extent that items in the S system, or undergoing categorization, can be identified without having to go through the full categorization process first, the subject can partly or completely 'suppress' or 'delete' them. Thus items could be eliminated on the basis of the modality of stimulation, or retinal location, but not on the basis of meaning. (c) The subject can similarly select categories of items from the S system for categorization. This is, of course, an application of Broadbent's (1958) 'selective filter'.

Some of the consequences of this model are:

- (i) Partial report will be more accurate than whole report because the sample indicated by the coded signal will lie within the limit determined by the categorization system, and suppression of the other rows will reduce their interference with the items in the sample.
- (ii) Sperling (1960) found that when the coded signal indicated which of two rows of four letters each was to be reported, 90 per cent of the letters were correct; when the signal indicated which of the two rows was to be reported first, only 69 per cent of the letters in the first row were correct. This can be explained by supposing that in the first case the signal allowed the rejected row to be suppressed, thus removing its interfering effect on the selected row, but this could not

be done in the second case.

(iii) When the coded signal was used to indicate which of two rows was to be reported first, different subjects showed different patterns of performance. There was a general tendency to report the first row more accurately than the second, and the top row more accurately than the lower one, but some subjects showed no position preference, or no effect of order of report, or both, without appreciable decrement in accuracy, so that accuracy did not appear to be accountable for by order of report.

This, too, can be explained as a consequence of selective suppression. If the subject partially suppresses one row of the array it will interfere less with the items in the other row, producing a secondary enhancement of the latter, so that the former row would be expected to be worse and the latter better reported than if there had been no suppression, and this will have little effect on the span of apprehension for the whole display. When no order of report was prescribed the subjects all reported the top row more accurately than the lower one. When the order of report was indicated by the coded signal some subjects may have changed to producing relative enhancement of the first row, or other patterns of selective suppression. These strategies would have resulted in the differences in performance found while having little effect on the span of apprehension.

(iv) When the subjects were asked to give partial reports of

the letters or numbers alone in a mixed array there was no increased information available as compared with whole reports. This would follow from the impossibility (assumption (b)) of using selection and suppression to facilitate the categorization of certain items if more or less complete categorization is necessary before the items to be selected or suppressed can be identified. This is comparable with the difference in the 'selective filter's' ability to select 'channels' defined by simple physical characteristics or by meaning (Broadbent, 1958; A. Treisman, 1960).

(v) If the decrease in available information which is shown by partial reports as the coded signal is delayed is to some extent a function of interference from the rejected symbols, this interference should be greater the longer the delay, and the more letters there are besides those which are to be reported. Using matrices of 9 letters or of 12 letters, each consisting of three equal rows, Sperling found that for delays greater than 0.3 seconds partial reports showed more available letters in the 9-letter than in the 12-letter matrices. (The reports themselves were, of course, shorter (3 letters) for the former than for the latter (4 letters), though both are within the span of apprehension).

(vi) Performance should improve if interference between items in the S system is reduced. One way of achieving this would be to administer simultaneous material by functionally separate S systems. Some support for this prediction is provided by the finding that if material is

presented dichotically, or on two frequency bands, the amount remembered is less than the immediate memory span, but if the items are given simultaneously to the eye and ear, retention is at least as good as the ordinary memory span (Broadbent, 1958).

(vii) These consequences can be derived whether the categorization system is assumed to handle items sequentially or simultaneously. A test to distinguish between these two possibilities might be provided by the effect on the whole report of an increase in the rate of decay in the S system. Thus if the life of the stimulus trace were halved, and symbols are treated sequentially, each requiring, say, 0.15 seconds to be categorized, the span of apprehension would be halved; but if 4 - 5 symbols pass immediately into the categorization system an increase in the rate of decay of the traces remaining in the S system would not affect the span. Sperling (1960) showed that the rate at which the information available for partial report decayed could be increased by using a light rather than a dark post-exposure field. However the effect of using a light post-exposure field on the whole report is not clear. Sperling (1960) reports that the effect of the light post-exposure field on 5 subjects was to reduce the accuracy of their whole reports to about half. But Averbach & Sperling (1961) present results for "one representative subject" who shows almost identical whole reports with light and dark post-exposure fields. Possibly the system may operate in either way, and

either performance may be obtained; this problem needs experimental clarification.

The model that has been proposed here is an attempt to provide an explanatory framework that can be applied both to retroactive effects on threshold level, and to the short-term information storage shown by the experiments on the span of apprehension. Its agreement with the experimental findings on a number of points suggests that it may prove possible to relate the results of these different experimental procedures in this way.

This reviewer can suggest from the second observation that the intensity of a stimulus is important for attracting attention and evoking responses (Woodrow, 1914; Boring, Langfeld & Wald, 1940), and from findings such as that habituation occurs more slowly to an intense than to a weak stimulus, and may not occur at all if the stimulus is very intense (Sampson, 1933; Glanfield, 1937). If, on the other hand, the entire stimulus served only as a temporal reference point, then its intensity should be irrelevant to its effect. This argument was used by Saffrey & Klogy (1961), who found that the intensity of their auditory stimulus affected the amount of the threshold change in a way that is a reason for rejecting the temporal reference hypothesis in favour of "some active excitatory or inhibitory process ... in the central auditory system".

Although, on the wording hypothesis, there should be no effect of stimulus intensity for most of the range, this might not hold for very intense stimuli, which might evoke responses with a characteristic

## CHAPTER 9

### VARIATION IN THE INTENSITY OF THE ACCESSORY STIMULUS

#### 9.1 Introduction

Stimuli that evoke a simple reflex response may give a larger response if they are made stronger (Sherrington, 1947). On the hypothesis that the accessory stimulus affects the threshold by inducing a central change which then gradually dies away, it seems plausible that a similar relation might hold between the intensity of the accessory stimulus and the degree of this effect. This receives some support from the common observation that the intensity of a stimulus is important for attracting attention and evoking responses (Woodrow, 1914; Boring, Langfeld & Weld, 1948), and from findings such as that habituation occurs more slowly to an intense than to a weak stimulus, and may not occur at all if the stimulus is very intense (Humphrey, 1933; Oldfield, 1937). If, on the other hand, the extra stimulus served only as a temporal reference point, then its intensity should be irrelevant to its effect. This argument was used by Halliday & Mingay (1961), who found that the intensity of their accessory stimulus affected the extent of the threshold change it induced, as a reason for rejecting the temporal reference hypothesis in favour of "some active excitatory or inhibitory process ... in the central mechanisms".

Although, on the warning hypothesis, there should be no effect of accessory stimulus intensity for most of the range, this might not hold for very intense stimuli, which might evoke responses with a distracting

effect and so reduce the subject's efficiency, or for very weak stimuli, which might not always themselves be above threshold. The simplest assumption in the latter case is that if the accessory stimulus is within the range of uncertainty (for convenience such stimuli will be referred to as 'intraliminal') then it will provide temporal information on those occasions when it is itself above threshold, but not when its central effect is below the criterion, and it will on the average, therefore, produce a smaller fall in threshold than a supraliminal stimulus. The same result would be expected on the arousal hypothesis, since the degree of central arousal produced by intraliminal stimuli will be less than that evoked by stronger stimuli.

The intensity and time relations of an intraliminal accessory stimulus have been sufficiently well controlled in two experiments by Brogden and his co-workers to make their results of relevance here. Gregg & Brogden (1952) investigated the effect of a simultaneous visual accessory stimulus on an absolute auditory intensity threshold. The visual stimulus was an increase in the brightness of a fixation circle 0.25 inches in diameter. They studied the effects of a zero increase, a weak increase, and a somewhat stronger increase in its brightness, both the last two being within the range of uncertainty. The weak increase was detected by the subjects on 37 per cent, and the strong stimulus on 80 per cent of presentations. The critical stimulus was a 1,000 c.p.s. tone, and both stimuli were given simultaneously for 2 seconds, the auditory threshold

being measured by a modification of the method of limits. One second before each stimulus presentation a buzzer was sounded for 0.3 seconds.

They found that the results varied with the instructions given to the subjects. With subjects who had been informed about the visual stimulus at the beginning of the experiment, and were asked to report the presence of both stimuli, the auditory threshold was raised by both the brightness increments. The rise was greater for the weaker increment, though the difference was not significant. With subjects who were asked only to report the auditory stimulus, and were not informed about the visual stimulus, the auditory threshold fell, the fall being significantly greater with the stronger brightness increment.

Thompson, Voss & Brogden (1958) extended these findings by using eight intensities of the accessory visual stimulus, including the zero value. These intensities, which were numbered from 1 to 8, evoked 0, 4, 25, 50, 75, 100, 100, and 100 per cent. of positive responses from a normative group of 30 subjects. One group of experimental subjects was instructed to report the presence of the visual stimulus after they had responded to the auditory stimulus. Their results were different for stimuli 1-4 and 5-8. Stimuli 2 and 3 (i.e. those giving 4 and 25 per cent positive response) caused a rise in threshold, the rise being greater for stimulus 2 than for 3. There was no change in threshold with stimulus 4. This is in agreement with the finding of Gregg & Brogden for intraliminal stimuli given under this instruction. With stimuli 5-8 a significant fall

in threshold was obtained. With the group of subjects who were not asked to report, nor even told about the visual stimulus, there was no change in threshold for stimuli 1-4, and a significant fall for stimuli 5-8, this fall being significantly greater than that produced by the same stimuli under the previous instruction.

These experiments provide little support for the expectation suggested by both the warning and arousal hypotheses. In both experiments intraliminal accessory stimuli caused a rise in threshold when the subject was required to report the presence of both stimuli, though it cannot be said whether this was a consequence of having to report the presence of the accessory stimulus, or merely of being informed about its possible occurrence. When the subjects were not told about the accessory stimulus, it produced a fall in threshold which increased as its intensity increased in the experiment of Gregg & Brogden, but Thompson, Voss & Brogden found no effect of stimuli 1-4 in their experiment. Though the former result agrees with the prediction suggested here, the latter authors, comparing their result with the earlier finding of Gregg & Brogden, suggest that "the present results probably reflect more adequately the effect of the experimental variables". There are further reasons for doubting the earlier finding: (1) Subjects might have discovered during the course of the experiment that a light stimulus was occurring simultaneously with the auditory stimulus, and this might have produced a change in 'set' with an effect on the threshold, as discussed in Section 8.2. Since this discovery

was more likely when the stronger visual stimulus was being given than during series with the weaker intensity, and since all possible orders of the three accessory stimulus conditions were given equally often, the subjects were more likely to discover or to have discovered the accessory stimulus when the threshold was being determined with the stronger visual stimulus, so that an effect of set would have occurred more frequently in the latter case. This in itself could account for the difference in the threshold levels with the different accessory stimulus strengths.

(ii) Whether or not the subjects 'discovered' the visual stimulus, it might have had the effect of causing an increase in 'guessing', or outputs from the visual detection mechanism might have caused positive responses to be made. Since the auditory threshold was measured by the method of limits no false positive rates were recorded, so this possibility cannot be excluded.

Since a threshold-lowering effect of intraliminal stimuli appears likely on both the arousal and the warning hypotheses, and yet is only dubiously present in these experiments, which in some circumstances provide evidence of a rise instead, the effect of intraliminal stimuli is investigated in the next section to see whether or not a fall in threshold can be unequivocally demonstrated.

## 9.2 'Intraliminal' accessory stimuli

### Experiment 15

In order to avoid the possibility of 'set', or changes in

'guessing', affecting the results, 'complex' series, in which the different intensities of the accessory stimulus were given in random order, rather than separate series for each intensity, were used. The subjects were informed of, but not required to respond to the accessory stimulus. If the rises in threshold sometimes found with intraliminal accessory stimuli were due to both stimuli competing for some mechanism of limited capacity, with consequent delays or mutual interference, the intraliminal stimulus might cause a rise if given simultaneously, but not if given a short while before the critical stimulus; therefore an interval of 0.2 seconds was used as well as the zero interval.

#### Apparatus and procedure

The subject, who was alone in the experimental room, sat facing a white cardboard screen, 50 cm. by 60 cm. in size, 1 metre in front of him. At the centre of the screen was a small cross which he was instructed to fixate, and 25 cm. to the right of this (i.e. about  $14^{\circ}$  from the point of fixation) was a neon bulb, shielded by a sheet of white paper, which provided the pre-warning. The same apparatus was used as in Experiment 14 to provide the critical stimulus, a 40 msec. increment in the intensity of a constant 500 c.p.s. tone. The accessory stimulus was produced by passing a 10 msec. constant current square pulse, obtained from the constant current generator used in Experiment 1, through a neon bulb, the light from which was focused onto the centre of the fixation cross, so as to give a square spot of light, 4 mm. by 4 mm.,

subtending about 14 min. of arc at the subject's retina. The intensity of the light spot could be varied by altering the strength of the constant current pulse. The brightness of the walls of the room was 0.20 foot lamberts, of the screen 3.5 foot lamberts, and of the pre-warning light 10 foot lamberts. Electronic timing circuits were used to determine the intervals between the pre-warning coming on (it stayed on until the subject had responded on each trial) and the accessory stimulus, and between the latter and the critical stimulus. The subject responded 'Yes' or 'No' by depressing one of two tapping keys.

At the beginning of each session approximate determinations of the subject's thresholds for the auditory and visual stimuli were made by the method of limits, and on the basis of these four equally spaced values of the auditory stimulus and two values of the light stimulus were chosen. 'Weak' and 'strong' values of the accessory stimulus were selected which, it was hoped, were capable of evoking about 50 per cent and about 95-100 per cent of positive responses respectively. To check the accuracy of this selection, two series, each of 40 trials, were given in each session, an Initial Series at the beginning and a Terminal Series at the end of the session, in which the subject was instructed to respond to the light, and the 'weak' and 'strong' accessory stimuli were each given 20 times in random order. In the main portion of each session the method of constant stimuli was used to determine auditory thresholds. Twenty series of auditory stimuli were given, each series consisting of

15 trials, the 5 strengths of the auditory stimulus (the 4 chosen above, and zero) and 3 strengths of the visual stimulus ('weak', 'strong', and zero) being given so that each combination occurred once, in a random order different for each series. Each trial began with the pre-warning lighting up. This was followed, after an interval which varied randomly between 2 and 4 seconds, by the accessory stimulus. In half the series the critical stimulus was given simultaneously with the accessory stimulus and in half there was an inter-stimulus interval of 200 msec. The '0.0 second' series were alternated regularly with the '0.2 second' series.

The subjects were instructed to fixate the cross when the pre-warning was given, as a warning light spot might occur with the pip, or 200 msec. before it. They were required to report the presence or absence of the auditory stimulus alone. Before each series they were informed what inter-stimulus interval to expect, and given a preliminary trial with supraliminal values of both stimuli, the response to which was not recorded. Each session lasted about  $1\frac{1}{2}$  hours. Two well-practised subjects did four sessions each, and a third did eight.

### Results

The results of each set of four sessions were combined, and probit analyses performed to determine the thresholds with the 'strong' (S), 'weak' (W), and zero (NAS) accessory stimuli at each inter-stimulus interval. One subject did 8 sessions, but as these were done in two

groups of four separated by a long interval the two sets were analysed separately. The mean thresholds for each condition, standard deviations, and standard errors, are given in Table 21. Critical ratios are also given for some of the differences between means; for differences at the same inter-stimulus interval one-tailed significance levels are given, for differences between inter-stimulus intervals they are two-tailed. Each false positive rate is based on 160 'blanks'.

### Discussion

The effect of the 'strong' accessory stimulus is in full agreement with the results of previous experiments: at each interval it causes a fall in threshold, and the threshold at the zero inter-stimulus interval is lower than at the 0.2 second interval.

The effect of the 'weak' stimulus agrees with the prediction made by both the warning and arousal hypotheses: it produces a significant fall in threshold which is nevertheless less than that produced by the 'strong' stimulus. This confirms the result of Gregg & Brogden (1952), but not that of Thompson, Voss & Brogden (1958), for the condition in which the subject did not report the accessory stimulus. However this result cannot here be attributed to an effect of set, since all three accessory stimulus strengths were given in each series. The false positive rate appears to increase with increase in the strength of the accessory stimulus; however this does not account for the threshold falls found. Average false positive rates of 52.4 per cent at 0 seconds

TABLE 21

Inter-stimulus intervals (seconds):

	<u>0.0</u>			<u>0.2</u>		
<u>Accessory stimuli:</u>	<u>NAS</u>	<u>W</u>	<u>S</u>	<u>NAS</u>	<u>W</u>	<u>S</u>
<u>Threshold (db):</u>	0.668	0.559	0.473	0.608	0.558	0.522
<u>S.D.:</u>	0.244	0.190	0.218	0.155	0.213	0.161
<u>S.E.:</u>	0.020	0.011	0.024	0.010	0.012	0.017
<u>False positives (per cent):</u>	0.6	3.1	6.3	1.9	3.1	8.1

Responses to the accessory stimuli in the four sets of sessions:Initial series

W:	48.8	43.8	51.3	71.3
S:	88.8	96.3	98.8	98.8

Terminal series

W:	47.5	50.0	27.5	45.0
S:	93.8	96.3	85.0	95.0

All sessions and series

W:	48.3
S:	94.1

Differences between mean thresholdsCritical ratioSignificance  
(one-tailed)0.0 seconds

NAS - W	4.78	<u>p &lt; 0.0005</u>
NAS - S	6.22	<u>p &lt; 0.0005</u>
W - S	3.23	<u>p &lt; 0.005</u>

TABLE 21 (contd)

<u>Differences between mean thresholds</u>	<u>Critical ratio</u>	<u>Significance</u> <u>(one-tailed)</u>
<u>0.2 seconds</u>		
NAS - W	3.10	<u>p &lt; 0.005</u>
NAS - S	4.33	<u>p &lt; 0.0005</u>
W - S	1.72	<u>p &lt; 0.05</u>
<u>'0.0 seconds' - '0.2 seconds'</u>		
NAS	2.67	<u>p &lt; 0.01</u>
W	0.06	NS
S	1.66	NS (p < 0.1)

However the reliability of the standard deviation in this test was too great to provide evidence about the differences between results obtained in the two experiments. However they also show an unexpected feature. Although at the zero interval the threshold with the 'strong' stimulus is lower, though not quite significantly so, than the corresponding threshold at 0.2 seconds, the thresholds with the 'weak' necessary stimuli are almost identical at the two intervals, while on the trials on which no necessary stimulus was given the threshold is significantly higher at the zero interval. Like the other in threshold found by Wang & Wright, and Thompson, Voss & Wright, this is not explainable by the hypotheses put forward so far.

However a model can be suggested, similar to that of Section 3.4, which appears capable of explaining some of these findings. It

and 34.2 per cent at 0.2 seconds would be required to produce the increased response rates found with the 'strong' accessory stimuli as compared with the NAS condition. Moreover the thresholds given in Table 21 were calculated after the results had been corrected for false positives.

On the hypothesis that the 'weak' accessory stimulus has an intermediate effect because it is sometimes detected and sometimes not, and is only effective in the first case, the standard deviation should be greater for the 'weak' threshold than for the zero or 'strong' thresholds. However the variability of the standard deviations in this experiment was too great to provide evidence about this.

These results confirm in the main the expectations deriving from both hypotheses. However they also show an unexpected feature. Although at the zero interval the threshold with the 'strong' stimulus is lower, though not quite significantly so, than the corresponding threshold at 0.2 seconds, the thresholds with the 'weak' accessory stimuli are almost identical at the two intervals, while on the trials on which no accessory stimulus was given the threshold is significantly higher at the zero interval. Like the rises in threshold found by Gregg & Brogden, and Thompson, Voss & Brogden, this is not explicable by the hypotheses put forward so far.

However a model can be suggested, similar to that of Section 8.4, which appears capable of explaining some of these findings. Its

main assumptions are: (i) There is a limit to the discriminatory capacity available for detecting signals. When only one threshold decision must be made most or all of this capacity is available to the corresponding threshold decision mechanism. But if more than one decision must be made simultaneously, the capacity available for each will be reduced. A reduction in the capacity of a threshold decision mechanism may result in a rise in threshold, or an increase of  $P_N(A)$  for a given value of  $P_{SN}(A)$ , perhaps because of more rapid decay or more frequent disappearance of the afferent inputs before or during categorization, which can be considered equivalent to a decrease in  $M_{SN}$ . (ii) The instructions, or the subject's expectations, can have the effect of increasing or decreasing the capacity available to a particular decision mechanism. The effect of this would be equivalent, in extreme cases, to 'switching in' or 'switching out' that decision mechanism. (iii) With increase in the strength of the stimulus awaited the capacity required to determine its presence or absence decreases, so that intraliminal stimuli will require considerable capacity, but in order to be able to detect stimuli well above threshold the decision mechanisms may draw upon the available capacity to a minor extent only. (iv) More capacity is required for categorizing a stimulus when a report on it is required than when this is not so. (v) It is further assumed that an accessory stimulus that has been detected provides temporal information, so that it can have the effect of lowering  $c$  for the critical stimulus. If a

negative decision has already been reached on the afferent input from the critical stimulus, the fall in criterion resulting from the detection of the extra stimulus may give the critical trace a 'second chance', as discussed in Section 8.2.

A consequence of this model is that the effect of an accessory stimulus on the threshold will depend on the balance between the competition for the limited capacity that it produces, and the advantage resulting from the temporal information that it provides. Thus when a supraliminal accessory stimulus is employed it will require relatively little capacity and its effect will be mainly due to the change in  $c$  which it may produce; but with intraliminal stimuli these factors may be more evenly balanced.

This model appears to offer explanations for a number of findings: (a) If two stimuli must be categorized simultaneously, the capacity available for each will be reduced. If there is no increase in temporal information when they are given together, the reduction in capacity will not be compensated for and both thresholds will rise. An experiment that appears to fit this description was carried out by Pauli (Woodworth & Schbsberg, 1955). He administered a touch stimulus to a finger of each of the subject's hands, requiring him to say which pressure was stronger, or a brief visual exposure of 3-6 short lines to be counted. Either task alone was so easy that nearly 100 per cent of correct responses were given; but when they were presented simultaneously performance on both deteriorated markedly. Here each task appears likely to have

required considerable capacity, but in each case the concomitants of the stimulus, the sum of the pressure stimuli or the total briefly exposed field, very possibly constituted a supraliminal stimulus conveying adequate temporal information of the moment at which the comparison or the count was to be made. If this was so, presenting the stimuli simultaneously would not have increased the temporal information available, so that only the effects of the sharing of the capacity would have been observed, an expectation which is in agreement with the result obtained.

(b) Gregg & Brogden (1952) and Thompson, Voss & Brogden (1958) found a rise in threshold when the subject was required to report simultaneous intraliminal accessory stimuli. A difficulty in considering these two experiments is that each stimulus was presented for 2 seconds. However it is possible that, due to adaptation, the main neural effect of the stimulus was at its onset, and perhaps at the end. No attempt appears to have been made to avoid clicks when the tone was switched on or off, and these could have contributed to increasing the detectability of the stimulus at those moments. A buzz lasting 0.3 seconds preceded the stimulus by 1 second, so that the onset occurred at an interval of 0.7 seconds after the end of the buzz, whereas the end of the stimulus came 2 seconds later. The threshold-lowering effect would have resulted in the threshold for the former being lower than that for the latter. Thus it seems possible that their stimulus was equivalent to a short

stimulus given at an inter-stimulus interval of 0.7 seconds. Giving an intraliminal accessory stimulus at the same time as the critical stimulus would have reduced the capacity available for the latter, tending to raise the auditory threshold. When the extra stimulus was detected it would have provided temporal information tending to lower the threshold. With the weaker brightness increment used by Gregg & Brogden detection of the accessory stimulus, with a consequent fall in the auditory criterion, may have occurred on 37 per cent of trials; with their stronger stimulus it may have occurred on 88 per cent. In both cases a reduction in capacity would have been present for every judgment, so that the greater rise in threshold with the weaker brightness increment would be expected. Similarly Thompson, Voss & Brogden found the greatest rise in threshold with their stimulus 2, a lesser rise with stimulus 3, and no change with stimulus 4. With their stronger stimuli, which were likely to have been detected almost every time, and would require less capacity to be detected, the threshold fell. In both experiments there was no rise in threshold with the zero brightness increment, which suggests that if the accessory stimulus is not detected some minimal proportion of times the capacity allocated to detecting it is reduced.

(c) When no report on the accessory stimulus was required Gregg & Brogden found a fall which was a function of the intensity of the intraliminal accessory stimulus, Thompson, Voss & Brogden found no effect with their stimuli 1-4, and in Experiment 15 a fall in threshold

was found which was greater with the stronger accessory stimulus. As has been discussed above, the first result may have been an effect of 'set' or of indiscriminate responding. However these could not account for the present result, which was obtained using complex series. Since no report on the accessory stimulus was required, the capacity allocated to detect it may have been less than if reports on it had been asked for, so that the tendency for an intraliminal accessory stimulus to cause a rise in threshold would have been less. The criterion would have been lowered more often by the 'strong' stimulus, which would also have made less demand on the subject's signal detection capacity, which would account for its effect being significantly greater than that of the 'weak' stimulus. The absence of an effect of stimuli 1-4 in the experiment of Thompson, Voss & Brogden may be a consequence of a peculiarity of their experimental design. Their subjects were not informed of the occurrence of the light stimulus, and thresholds were determined for every one of stimuli 1-4 before any of stimuli 5-8 were presented, although in each group of four stimuli the order in which they were used was random, to postpone their subjects discovering the light stimulus. Their result suggests that they were successful in this, and that unless the subject has discovered or been informed about the accessory stimulus he will not 'switch in' the appropriate threshold detection mechanism by increasing its capacity to the level required for the intraliminal stimuli, so that the consequences of categorizing

the accessory stimuli will not occur.

That an opportunity for subjects to 'discover' a relation between the stimuli can influence their expectation and affect their further performance is well illustrated by an experiment by Brogden (1950). He measured subjects' auditory thresholds with and without a simultaneous visual stimulus (a small increase in the illumination of the room) accompanying each tone. The thresholds were measured at the end of an experimental sequence which commenced with the presentation of 10 pairings of a tone of about 15 db SL and the visual stimulus for the experimental group and 10 presentations of the tone alone for the control group. In each case the experiment was presented to the subjects as an investigation of auditory acuity. The subjects, in the experimental group, but not the controls, showed a significant fall when the threshold was measured with the light stimulus present.

(d) If simultaneous intraliminal accessory and critical stimuli are assumed to compete for a limited discriminatory capacity, an explanation can be offered for the disparity between the results at the zero interval and at 0.2 seconds in Experiment 15. When the two stimuli are given successively at a sufficient interval there will be no reduction of capacity for the critical stimulus, and the only effect on the threshold will be that produced by the temporal information provided by the accessory stimulus. At 0.2 seconds we can suppose most or all of the effects on the threshold to be due to the latter

source. At the zero interval there will be, as well, the effect of the reduction in the capacity available for the critical stimulus. This will be opposed by the threshold-lowering effect of a simultaneous accessory stimulus on those trials on which it is detected. If the advantage, as a temporal reference point, of a simultaneous accessory stimulus over one at an inter-stimulus interval of 0.2 seconds is greater than the effect of the reduction in capacity for the critical stimulus in the former case, then the threshold at the zero interval will be lower than that at 0.2 seconds for trials with an accessory stimulus which is almost always detected. This advantage will decrease as the proportion of times that the accessory stimulus is detected decreases, so that for some weak intraliminal stimulus the two effects should balance out and the thresholds should be the same at both intervals. Finally, in the NAS condition, only the effect of the reduction in capacity will be present, and this will produce a higher threshold for the complex 0.0 second series than for the 0.2 second series.

Perhaps the least well supported of the assumptions in the model is that having to report the accessory stimulus increases the capacity required for detecting it. Where conditions with and without report have been compared, this factor has been confounded with the effect of informing or not informing the subject about the accessory stimulus. We do not know what effect the two intraliminal stimuli in the experiment of Gregg & Brogden, or stimuli 1-4 in that of Thompson,

Voss & Brogden, would have had on the thresholds if the subjects had been informed in advance about their possible occurrence but not required to report them. Stimuli 5-8 in the latter experiment caused a greater fall in threshold when no report was required, but the two groups of subjects may have had different 'sets' induced by their differing experience in the first half of the experiment. In Experiment 15, in which the accessory stimulus was not reported, the 'weak' stimulus, which itself gives 48 per cent response, causes a fall in threshold as compared with both NAS conditions. In Thompson's, Voss' & Brogden's experiment stimulus 4 (50 per cent response) causes no change in threshold when the accessory stimulus is reported, and in Gregg's & Brogden's experiment stimuli evoking 37 and 88 per cent of response both cause a rise in threshold. However in both the latter experiments the reference level is the threshold given with no accessory stimulus and a fixed pre-warning interval of 0.7 seconds; in Experiment 15 the pre-warning was large and randomly varied, so that the NAS threshold was determined in the absence of usable temporal information, and the accessory stimulus should give a correspondingly greater fall.

Other models could be suggested, differing in various ways, which might account for these experiments equally well. Thus it could be supposed that the two threshold decision mechanisms operate sequentially, with decay of the trace categorized second. The model has been

described in terms of the temporal reference hypothesis, but it could as well have been supposed that detection of an accessory stimulus caused an 'alarm' effect with a fall in criterion, provided it is supposed that in an experiment such as Pauli's the arousal effects of the separate stimuli do not sum. However, in its present form, the model serves to show that the various findings on the effects of intraliminal accessory stimulation are not incompatible with the hypotheses that have been advanced to explain the threshold-lowering effect.

Experiment 15 has confirmed a prediction of both the arousal and the warning hypotheses, but it has not been of much help in discriminating between them. In the next section the effect of supraliminal stimuli differing in intensity will be examined to see whether this will prove more useful for the latter purpose.

### 9.3 Supraliminal accessory stimuli

On the arousal hypothesis it appears plausible that the fall in threshold should be greater for the stronger accessory stimuli, but on the temporal reference hypothesis the intensity of a supraliminal stimulus should be irrelevant to its effect. In the experiment by Thompson, Voss & Brogden (1958) similar falls in threshold were obtained with three stimulus strengths each of which evoked 100 per cent of positive responses. However these stimuli covered a range of less than 0.4 millilamberts; had a greater range of stimuli been used a difference in their effects might have been found.

### Experiment 16

This experiment was designed to examine the effects on the threshold of two accessory stimuli, one just above the threshold range, and the other considerably greater. It seemed desirable to attempt to obtain some index of the arousing effects of the two stimuli, other than their effect on the threshold. An old view is that reaction time can be used as an indicator of clearness of attention (Woodrow, 1914). Recently Fuster (1958) has provided evidence for an association between the sensory threshold, the arousal mechanism, and reaction time: he stimulated the mesencephalic brain stem area in rhesus monkeys engaged in discriminating between a pair of objects exposed by a tachistoscopic flash, and found that this stimulation increased the proportion of correct choices for a given flash duration, and reduced the mean reaction time. The area of the reticular system stimulated has been shown to evoke electroencephalographic and behavioural arousal (Magoon, 1958), and Fuster concluded that its excitation "induces general activity of the cortex (arousal), presumably facilitating its receptivity to the sensory impulses ascending over the classic sensory paths". In the present experiment the reaction times of the subjects' responses are also measured.

### Apparatus and procedure

The pre-warning and critical stimulus were provided by the same apparatus as in Experiment 15. The stimulus was a 40 msec. increment

in the intensity of a constant 500 c.p.s. tone. The pre-warning interval was varied randomly between 2 and 5 seconds from trial to trial. The accessory stimulus was a 40 msec. light flash produced by passing a current through a small fluorescent tube of the type used in the tachistoscope employed in Experiment 1 (Humphrey, Dave & Mandell, 1955). The duration of this light flash was determined by an electronic timer.

The subject sat facing a white cardboard screen 50 cm. by 60 cm. in size, in the centre of which was a circular aperture 7 cm. in diameter filled with a 0.2 neutral density Ilford Filter. The subject was 85 cm. from the screen, so that the aperture subtended about  $4^{\circ}40'$  of arc. The fluorescent tube was contained in a box placed behind the aperture and some distance from it, so that the flash illuminated the fixation circle diffusely. A neutral density filter could be inserted between the fluorescent tube and the screen to reduce the intensity of the flash. The brightness of the screen, measured with an SEI photometer, was 3.2 foot lamberts, the unilluminated fixation circle was 0.40 foot lamberts, the 'weak' (W) flash raised its brightness to 0.56 foot lamberts, and the 'strong' (S) flash to 180 foot lamberts.

Each trial began with the pre-warning coming on. This was followed, after an interval varying at random between 2 and 5 seconds, by the accessory stimulus. The critical stimulus was given either simultaneously with or 1.5 seconds after the accessory stimulus. The

subject responded by depressing one tapping key for 'Yes' or another for 'No'. His reaction time was recorded on a bank of Dekatrons.

Each session began with an approximate estimation of the subject's threshold by the method of limits. The main part of the session consisted of 15 simple series, each of 20 trials. Each series consisted of 4 critical stimuli at each of 4 equally spaced intensities, and 4 'blanks', randomly ordered, all given with one strength of the accessory stimulus and at one inter-stimulus interval. There were 5 combinations of accessory stimulus and inter-stimulus interval: NAS; Strong accessory stimulus-0 seconds interval (S-0); Strong-1.5 seconds (S-1.5); Weak-0 seconds (W-0); and Weak-1.5 seconds (W-1.5); and a session consisted of 3 series under each condition given in a random order different for each subject. Each session lasted 1-2 hours. Five subjects each did one session. They were not informed that their reaction times were being recorded.

### Results

The threshold for each subject under each condition was determined by probit analysis. The mean thresholds, standard deviations, and standard errors for the experiment are given in Table 22. Each false positive rate is based on 60 'blanks'. The mean reaction times given are based on 647 positive reactions; responses to 'blanks' were not used in calculating these means. The (two-tailed) significance levels of some of the differences between mean thresholds and reaction times are also given.

TABLE 22

<u>Condition:</u>	<u>NAS</u>	<u>W-0</u>	<u>S-0</u>	<u>W-1.5</u>	<u>S-1.5</u>
<u>Threshold (db):</u>	0.576	0.511	0.518	0.536	0.500
<u>S.D.:</u>	0.136	0.197	0.160	0.126	0.172
<u>S.E.:</u>	0.013	0.019	0.016	0.013	0.029
<u>False positives (per cent):</u>					
	6.7	3.3	1.7	0.0	6.7
<u>Reaction Times (seconds):</u>					
	1.108	1.010	0.849	0.723	0.804
<u>S.D.:</u>	0.758	0.484	0.315	0.303	0.392
<u>S.E.:</u>	0.077	0.041	0.027	0.028	0.032
<u>Differences between mean thresholds: Critical ratio      Significance</u>					
<u>0 second interval</u>					
NAS - W			2.77		p < 0.01
NAS - S			2.77		p < 0.01
W - S			0.28		NS
<u>1.5 second interval</u>					
NAS - W			2.18		p < 0.05
NAS - S			2.38		p < 0.02
W - S			1.14		NS
<u>Both intervals</u>					
'W-1.5' - 'W-0'			1.08		NS
'S-1.5' - 'S-0'			0.54		NS

TABLE 22 (contd)

Differences between mean reaction times: Critical ratio      Significance

0 second interval

NAS - W	1.13	NS
NAS - S	3.20	<u>p &lt; 0.01</u>
W - S	3.31	<u>p &lt; 0.001</u>

1.5 second interval

NAS - W	4.73	<u>p &lt; 0.001</u>
NAS - S	3.66	<u>p &lt; 0.001</u>
W - S	1.91	NS (p < 0.1)

Both intervals

'W-1.5' - 'W-0'	5.85	<u>p &lt; 0.001</u>
'S-1.5' - 'S-0'	1.07	NS

## Discussion

At both intervals both the 'weak' and 'strong' accessory stimuli produce significant falls in threshold, though the difference between the two intervals is not significant. Despite this there is no significant evidence, at either interval, of an effect of the intensity of the accessory stimulus. This is in contrast to the results of Experiment 15, in which a very much smaller difference in the intensity of the accessory stimuli produced significant differences in threshold level. As in Experiment 15, the false positive rate increases as the accessory stimulus gets stronger at the 1.5 second interval, but the reverse is found at the zero interval.

If reaction times are regarded as an index of 'arousal', their changes here are of interest. Reaction times usually increases as the intensity of the stimulus evoking the response decreases (Guilford, 1954; Woodworth & Schlosberg, 1955). As the same critical stimuli were used in the different conditions, and the variation in threshold does not appear to be sufficient to affect reaction time (cf. Figure 11C and Experiment 6), and as the reaction time varies even when there is no significant change in threshold, the effects on reaction time which are found can be attributed to the intensity of the simultaneous or preceding accessory stimulus. That the presence of this stimulus does produce an effect on the subject's reaction time is shown by the significant fall in reaction time with both accessory stimuli at the 1.5 second interval,

and with the 'strong' stimulus at the zero interval. That the two stimulus strengths can produce significantly different effects is shown by the difference between the mean reaction times at the zero interval, the 'weak' stimulus having no apparent effect while the 'strong' stimulus produces a significant fall.

Thus using reaction time as an index of arousal suggests that the accessory stimuli do produce arousal, but that its occurrence and extent is not closely related to the changes in threshold. At the zero interval, despite a significant difference in the arousing effect of the two stimuli, there is no difference in the corresponding thresholds, the 'weak' threshold falling significantly in the absence of any evidence of arousal. At the 1.5 second interval a significant degree of arousal develops with the 'weak' stimulus, without a corresponding effect on threshold level.

This result is not in keeping with the association between threshold, reaction time, and brain stem stimulation found by Fuster (1958). However, over and above the differences in subjects and techniques employed in the two experiments, it must be noted that Fuster is not justified in concluding that in his monkeys brain stem stimulation produced a central state altering "receptivity to the sensory impulses", since he did not exclude the possibility that the stimulation produced the change in threshold by a peripheral effect, such as mydriasis. Naquet, Regis, Fischer-Williams, & Fernandez-Guardiola (1960) have shown

that this may be produced by direct stimulation of the brain stem reticular formation. In Fuster's experimental situation an effect of this sort could have caused the fall in threshold.

The failure to find an effect of accessory stimulus strength, despite the presence of a clear effect on threshold, further confirms that the effect under investigation here is different from that of Halliday & Mingay (1961). It affords evidence that the threshold-lowering effect of accessory stimulation is not due to arousal, and this is strengthened by finding that arousal, as indicated by the reaction time, was produced in some conditions, but in a different pattern from the threshold changes. It is not in itself conclusive, however, since it could be argued that the threshold change is produced by an effect which is all-or-none in nature, results in one of the changes given as alternatives (a) to (c) in Section 5.1, and is independent of and different from those effects of a sudden stimulus usually referred to as 'arousal'. If the hypothesis is maintained in this form, however, it loses the plausibility which is otherwise given to it if the common physiological effects of an arousing stimulus can be used as analogues of the change hypothesized.

The changes in the reaction time are also of some interest. Since the reaction time is here disjunctive, depends on a poorly discriminable stimulus, the responses to only one alternative are recorded, and there is no pressure on the subject to respond quickly,

the results can be expected to differ from those obtained using simple reactions to clearly discriminable stimuli. It can be seen that the results with the 'weak', but not with the 'strong' accessory stimulus, are in keeping with Woodrow's (1914) finding that with fixed fore-periods the reaction time is least at an interval of about 2 seconds.

A possible explanation for the result found here is that an effect of the accessory stimulus tends to reduce the subject's decision or response time, and that the latency of this effect is shorter the more intense the accessory stimulus; when the latter is weak the latency may be too long for the response to a simultaneous critical stimulus to be affected. The effect of the accessory stimulus at the zero interval is in contrast with Todd's finding that adding a light stimulus to a sound, or to an electric shock, did not shorten the reaction time, as compared with reactions to the sound or shock alone. (Woodworth & Schlosberg, 1955). It seems that responding to very weak stimuli, with the long reaction times produced, may allow some effects on reaction time to be shown which are not revealed by the more rapid simple reactions.

in sequences of varying length, as in the method of Squires, the subject cannot really be expected to know  $f_{ij}(x)$  until he has gone through the experimental procedure. Furthermore it follows from equation (10) that if the subject knows that  $f_{ij}(x) < 1$ , as in the method of Squires, or the

## CHAPTER 10

### CONCLUSION

#### 10.1 A further test of the warning hypothesis

Although the subject is usually assumed to employ the 'maximum expected value' criterion (equation (10)) in applications of signal-detection theory to the threshold (Tanner & Swets, 1954; Licklider, 1959), there is an advantage in supposing that he uses the Neyman-Pearson criterion (equation (23)) instead. Less knowledge is required to make use of this criterion, so that it is plausible in more situations.

The application of equation (10) requires the subject to know  $P(N)$  and  $P(SN)$ , as well as the 'values' and 'costs', before he can calculate the critical likelihood ratio,  $\beta$ . In order to identify the critical value of  $x$ , corresponding to  $\beta$ , that is to be used as the cut-off point he must know both  $f_N(x)$  and  $f_{SN}(x)$ . These requirements restrict this model to situations where the subject has considerable information about the stimulus, so that it is not easily generalisable to account for sensory discriminations made in relation to possible stimuli which are imperfectly known or unexpected. Even in the laboratory it meets with difficulties: when a number of different stimulus strengths are used in sequences of varying length, as in the method of limits, the subject cannot really be expected to know  $f_{SN}(x)$  until he has been through the experimental procedure. Furthermore it follows from equation (10) that if the subject knows that  $P(SN)=1$ , as in the method of limits, or the

method of constant stimuli if no 'blanks' are given, the threshold should become extremely low. This is not found in practice; the method of limits usually gives thresholds not very dissimilar from those given by the method of constant stimuli (in Experiment 10(a) the former were higher). A consequence is that procedures such as the method of limits are not usually taken account of in signal-detection theory.

Although the Neyman-Pearson criterion will correspond to a value of likelihood ratio it is not necessary to suppose that the subject calculates this value, nor that he knows  $P(SN)$  or  $f_{SN}(x)$ . Since the criterion only specifies the limiting rate of false positive decisions that the subject will tolerate, it can be applied equally well to analysing the detection of familiar and unfamiliar signals. In order to determine the cut-off point,  $c$ , the subject needs only to be able to make some estimate of  $P_N(A)$  (which he should ordinarily be well able to do, since daily life affords many opportunities to make difficult discriminations which can be later verified), and to vary the value of  $c$  until  $P_N(A) = \epsilon$ , where  $\epsilon$  is the acceptable limiting rate of false positive decisions. The criterion requires less foreknowledge on the subject's part, and is correspondingly more general, than equation (10).

The finding that subjects can be got to alter their criterion by varying monetary values and costs associated with their responses in an experiment (Tanner & Swets, 1954) or by altering the announced value of  $P(SN)$  (Tanner, Swets & Green, 1956) does not necessarily

imply that these procedures have affected essential stages in the process of determining the criterion. Information of this sort might well be used by the subject to alter the preferred value of  $\epsilon$  rather than in the way prescribed by equation (10).

If the subject is assumed to employ the Neyman-Pearson criterion, this does not imply that the false positive rate in experiments using different inter-stimulus intervals must be constant. Though on each trial the signal is presented only once, there are a number of opportunities to make a decision leading to a false positive response. If, as in Section 5.1, the waiting time on each trial is assumed to be sampled by the subject as a succession of 'moments', at each of which a decision is made, and for each of which  $P_N(A) = \epsilon$ , then the probability of a positive response being made on a trial on which no stimulus has been administered ( $P_{FP}$ ) is the probability of the decision being positive at one or more of these moments. If the trial is taken to contain  $m$  moments, then the relation between the recorded false positive rate,  $P_{FP}$ , and the probability of a false positive decision at any 'moment',  $\epsilon$ , is given by

$$P_{FP} = 1 - (1 - \epsilon)^m \quad \dots\dots\dots (25)$$

If the subject is concerned, when the stimulus is not presented more than once on any trial, to maintain his rate of false positive overt responses at or below a limiting value, then he will not necessarily keep  $\epsilon$  constant; the number of moments on each trial will be relevant

to determining its value.

If it is assumed that the subject attempts to keep  $P_{FP}$  constant or nearly constant by varying  $c$ , then, by applying the warning hypothesis, it is possible to predict the threshold changes that will occur as the length of the inter-stimulus interval,  $T$ , varies. For this purpose the waiting time on each trial can be considered to consist of the 'range of expectation' during which the subject expects the stimulus, and a residual period. The subject might attempt to keep  $P_{FP}$  constant for the whole trial, or to keep the probability of his making a false positive response to an input during the period when he expects the critical stimulus at or below a limiting value, i.e. to keep the portion of the false positive rate arising during the range of expectation,  $P_{FP}(r)$ , constant and the contribution of the residual period low. In the former case the value of  $P_{FP}$  would be constant despite variation in  $T$ , in the latter it would tend to rise to a small extent as  $T$  increases.

Some information on this point is supplied by the false positive rates recorded in Experiments 10 to 16. Making all possible comparisons between false positive rates recorded under conditions differing only in the length of the inter-stimulus interval, there are 17 cases in which the rate is greater at the longer interval, 7 in which there is no change, and 12 in which it is less. Though the difference is in the direction that would be expected on the hypothesis

that the subject keeps  $P_{FP}(r)$  constant, it is not significant. However it will be assumed that this is what the subject attempts to do.

If an inter-stimulus interval  $T_0$  is changed to an interval  $T_1$ , so that the range of expectation  $R_0$ , containing  $r_0$  moments, is replaced by the range  $R_1$ , containing  $r_1$  moments, then  $P_{FP}(r)$  can be kept constant by altering  $\epsilon_0$  to a new value  $\epsilon_1$  which is given by

$$\epsilon_1 = 1 - (1 - \epsilon_0)^{r_0/r_1} \dots \dots \dots (26)$$

If the range of expectation is taken to be a constant proportion of the inter-stimulus interval, then

$$r_0/r_1 = T_0/T_1 \dots \dots \dots (27)$$

If the range of expectation is taken to be a constant number of j.n.d.s, where the j.n.d.,  $\Delta T$ , is given by  $\Delta T = k(T+a)$ , then

$$r_0/r_1 = (T_0+a)/(T_1+a) \dots \dots \dots (28)$$

If  $z(p)$  is the standardized normal deviate for which the normal distribution function  $\Phi(z) = 1-p$ , then

$$z(P_{SN}(A)_1) = z(P_{SN}(A)_0) - (z(\epsilon_0) - z(\epsilon_1)) \dots \dots (29)$$

provided that the N and SN distributions are normal with equal variance.

In this case, if the frequency of response to a given stimulus at  $T_0$  is known, then it is possible to predict the frequency of response to the same stimulus at  $T_1$  provided  $\epsilon_0$  is known and equations (26), (27), and (29) are applied, or  $\epsilon_0$  and 'a' are known and we apply equations (26),

(28), and (29).

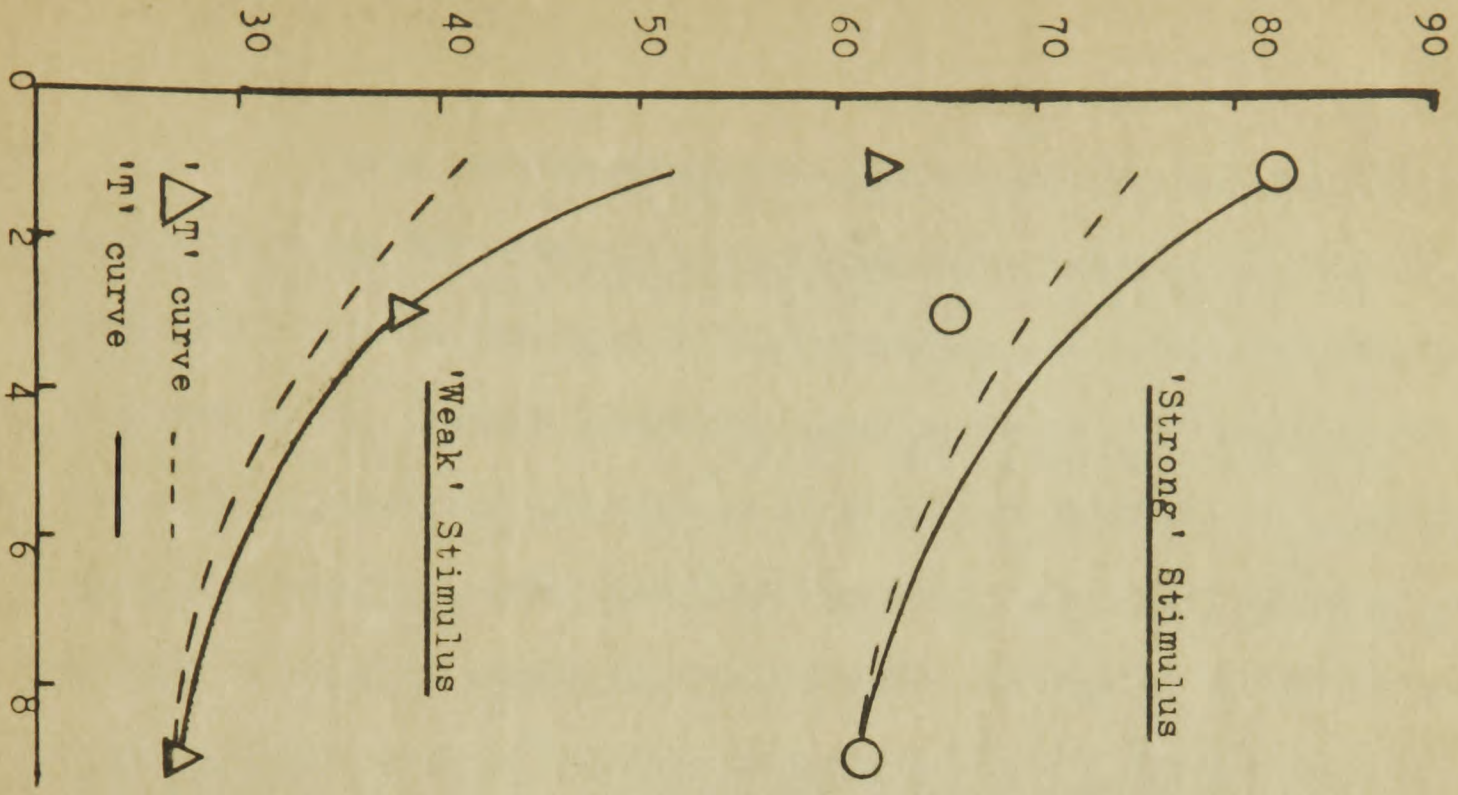
In studies of the sensory threshold the psychometric function has usually been found to be approximately normal (see Section 1.6); the use of probit analysis in determining thresholds in the present series of experiments allows the hypothesis that the subjects' responses were normally distributed in relation to the decibel scale used here to be tested. This was done by summing the values of  $\chi^2$  for residual heterogeneity about the probit regression line found in each analysis. In Experiments 10(b), 12, 14(b), 14(c), 15 and 16, 145 probit analyses were performed; these give a combined value of  $\chi^2 = 284.10$  (d.f. = 258) which is not significant ( $p < 0.15$ ), so that there is no evidence that the present data depart significantly from normality. It is also plausible to suppose that in Experiments 10 to 16  $\sigma_N \doteq \sigma_{SN}$ , since in these experiments the former represents the variability of the central response to a 500 c.p.s. tone 60-70 db above threshold, and the latter the response to a tone a fraction of a decibel stronger.

Experiments 10(a), 10(b), and 13 (Constant Condition) offer data suitable for checking predictions made using equations (26) to (29), since in these experiments the stimuli were presented by the method of constant stimuli, at more than one positive inter-stimulus interval, in order to measure incremental thresholds. (At zero and very short intervals equation (27) would be expected to break down as the range of

expectation approaches the 'sampling time'). In order to obtain an estimate of  $\epsilon_0$  it was assumed that at 10 seconds the difference between the values of  $\epsilon$  in the range of expectation and in the residual period is negligible and Stroud's (1955) estimate that a 'moment' is approximately 0.10 seconds long was applied to give  $m = 100$ . 4.1 per cent, the average false positive rate in Experiments 10(a), 10(b) and 13 (Constant Condition), was used as an estimate of  $P_{FP}$ , and these two quantities were substituted in equation (25) to give  $\epsilon_0 = 0.00043$ .

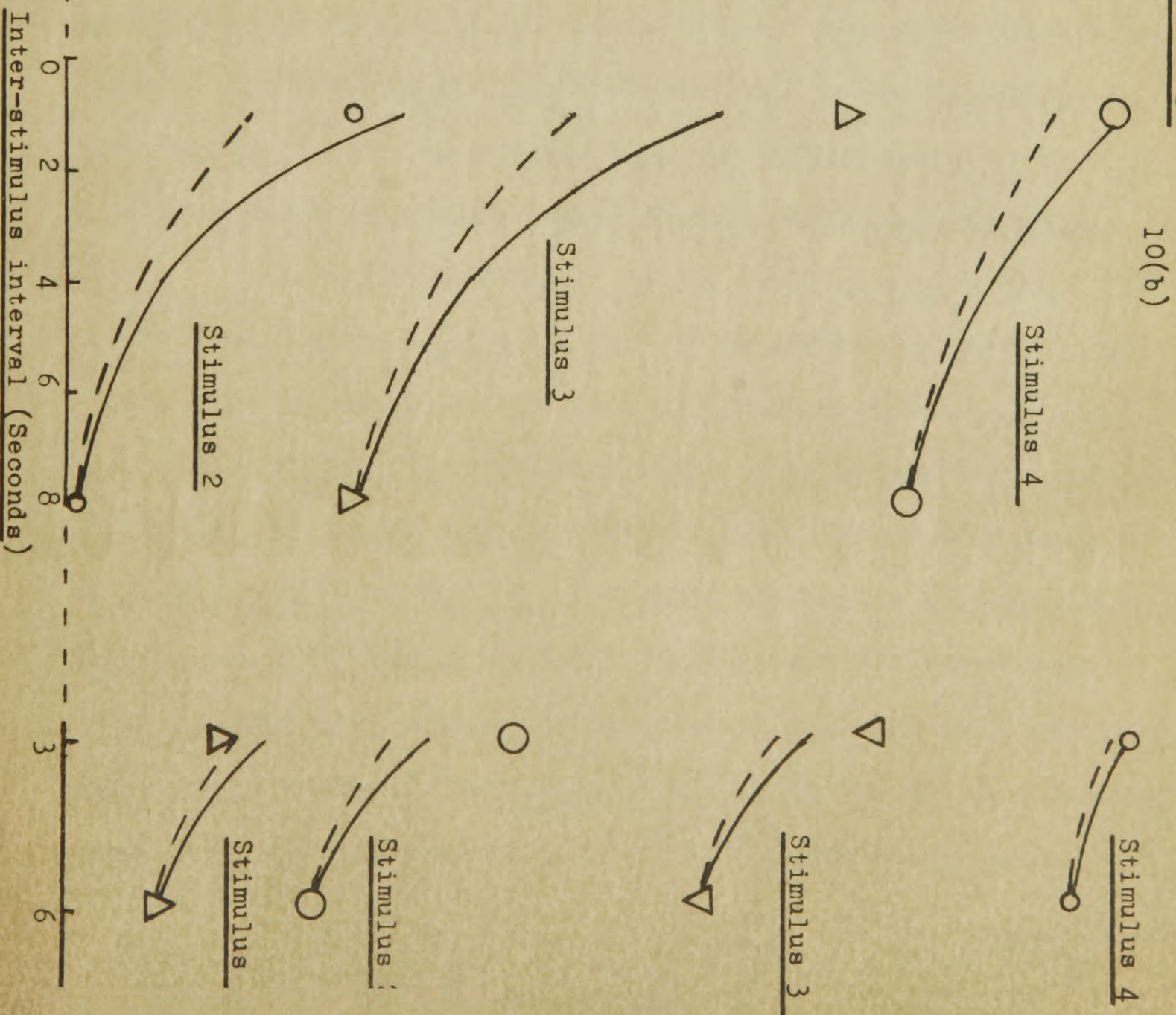
For 'a' the value of 1.57, given in equation (17), was taken, and values of  $\epsilon$  were calculated for the different inter-stimulus intervals in Experiments 10 and 13, applying equation (27) or (28). Using the frequency of response to each stimulus strength at the longest inter-stimulus interval in the experiment as an estimate of  $P_{SN}(A)_0$ , the predicted values of  $P_{SN}(A)$  for each shorter inter-stimulus interval were obtained from equation (29). These results are given in Table 23, predictions obtained using equation (27) being indicated by 'T', and those obtained with equation (28) by ' $\Delta$ T'. Each of the 'experimental' percentages of positive response was based on 360 responses in Experiment 10(a), 240 for the 1 and 8 second intervals, and 80 for the 3 and 6 second intervals in Experiment 10(b), and 160 in Experiment 13. The 'T' and ' $\Delta$ T' curves for the 'weak' and 'strong' stimuli in Experiment 13 are shown in Figure 9, and those for Experiment 10 in Figure 12.

Percentage Response



10(a)

FIGURE 12



10(b)

Stimulus 4

Stimulus 3

Stimulus 2

Stimulus 1

Inter-stimulus interval (Seconds)

These results show quite close agreement between the 'T' predictions and the obtained percentages of response. In view of the error variation in the experimental results, and in the frequencies of response at the longest intervals, used in making the predictions, and the somewhat arbitrary assumptions necessary to estimate  $\epsilon_0$ , better agreement could not well have been expected. The exact value of  $\epsilon_0$  is not critical in making the predictions: it can be varied between 0.001 - 0.0001 without markedly impairing their agreement with the observed results. The success of the predictions tends to support the assumptions on which they were based, notably the warning hypothesis. Although the 'judgmental' hypothesis has been assumed in making the predictions, it could be supposed that the subject varies  $\sigma_N$  or  $M_N$  in a fashion geared to produce changes in  $c$  of about the same order.

The ' $\Delta T$ ' predictions are not quite as good as the 'T' predictions. Of 21 comparisons, the 'T' prediction is closer to the experimental result in 16, and the ' $\Delta T$ ' prediction in 5 ( $\chi^2 = 5.76$ , d.f. = 1,  $p < 0.02$ ). This inferior result is in line with the tentative conclusion, in Section 4.6, that the thresholds obtained by the method of limits in Experiments 1, 2, and 4 were better described as a linear function of  $(1/T)$  than of  $(1/\Delta T)$ . It is possible to improve the ' $\Delta T$ ' predictions by using a larger value of  $\epsilon_0$ . However it must be considerably increased, to a range of 0.05-0.10, for the predictions given to be approximately as good as the 'T' predictions, and these

TABLE 23

Obtained and predicted percentages of response ( $P_{SN}(A)$ )

Experiment 10(a)

<u>Stimulus</u>	<u>Source</u>	<u>Inter-stimulus interval (seconds)</u>		
		<u>1</u>	<u>2</u>	<u>2</u>
'Weak'	Experimental:	62.2	38.3	26.9
	T:	51.6	37.8	-
	$\Delta T$ :	41.3	34.8	-
'Strong'	Experimental:	81.9	65.6	61.1
	T:	82.6	72.2	-
	$\Delta T$ :	75.2	69.5	-

Experiment 10(b)

		<u>1</u>	<u>2</u>
		2	Experimental:
	T:	42.5	-
	$\Delta T$ :	32.6	-
3	Experimental:	72.5	38.8
	T:	64.1	-
	$\Delta T$ :	54.0	-
4	Experimental:	90.4	76.3
	T:	91.3	-
	$\Delta T$ :	86.4	-
		<u>3</u>	<u>6</u>
1	Experimental:	30.0	26.3
	T:	33.4	-
	$\Delta T$ :	31.2	-
2	Experimental:	50.0	36.3
	T:	44.0	-
	$\Delta T$ :	41.7	-

TABLE 23 (contd)

3	Experimental:	73.8	62.5
	T:	69.9	-
	$\Delta T$ :	67.7	-
4	Experimental:	91.3	87.5
	T:	91.2	-
	$\Delta T$ :	90.2	-

Experiment 13 (Constant Condition)

		<u>0.25</u>	<u>0.75</u>	<u>1.25</u>	<u>1.75</u>	<u>2.25</u>	<u>2.75</u>
'Weak'	Experimental:	50.0	45.6	35.0	39.4	35.0	29.4
	T:	61.4	45.2	38.6	34.5	31.6	-
	$\Delta T$ :	39.4	36.7	34.5	32.3	30.9	-
'Strong'	Experimental:	75.0	78.8	66.3	67.5	63.1	58.8
	T:	85.3	73.9	68.1	64.1	61.0	-
	$\Delta T$ :	68.8	66.3	64.1	61.8	60.3	-

values of  $\epsilon_0$  would correspond to false positive rates of 99.41 to 99.997 per cent. The difference could also be made less marked by reducing the value used for 'a'. These results therefore tend to suggest that in central applications of the 'internal clock' the additive constant 'a' is not as large as when Weber's law is determined from overt responses, and may even be equal to zero.

The most marked disparity between the results predicted using equation (27) and those obtained occurs at 0.25 seconds in Experiment 13, where the response rates are lower than predicted. This may be an accidental variation, or it may be due to equation (27) ceasing to apply at this interval, perhaps because the range of expectation is nearing the sampling time.

In Figure 8 the 'T'  $z(\epsilon)$  values used in making the predictions for inter-stimulus intervals between 1 and 9 seconds are plotted. They represent the values of  $c$  determining the threshold at these intervals. It is of interest to note their similarity to the threshold curves given by Experiments 1, 2, and 4. The best curve of the form

$$(A-z)T = k' \dots\dots\dots (30)$$

has been fitted to these values. It shows somewhat greater curvature than do the values of  $z$ . In contrast, the threshold values determined in Experiments 1, 2, and 4, and plotted in Figures 7 and 8, do not appear to be systematically less curved than the best curves fitted to

them by equation (22). This suggests the possibility that in varying the value of  $c$ , as the inter-stimulus interval changes, in order to keep  $P_{FP}(r)$  constant, the subject does not go through a process equivalent to calculating equations (26), (27) and (29), but instead applies the simpler approximation afforded by equation (30), at least for intervals above about 1 second. For intervals somewhat below this the equation would be modified. The subject's experience of the changes in  $P_N(A)$  as the value of  $z$  is altered could be used to determine the appropriate value of  $k'$  to keep  $P_{FP}(r)$  approximately constant. If the values of  $z(\epsilon)$  used in making the 'T' predictions in Table 23 are replaced, for intervals of 1 to 9 seconds, by the corresponding values given by equation (30), the predicted frequencies of response agree approximately as well with the experimental data.

### 10.2 Conclusion

This thesis has been mainly concerned with the proper explanation of the change in threshold induced by variation in the inter-stimulus interval that was found in the first experiment, and whose generality was demonstrated in Experiments 2, 3, and 4. This finding posed two questions: How did the subject vary his threshold? What determined the relation between the threshold level and the inter-stimulus interval? In the attempt to define the processes underlying the subjects' performance in these experiments two sets of hypotheses were proposed.

In terms of the signal-detection model of sensory discrimination

alternative ways in which the threshold could be varied are: (a) A reduction in the variance of the afferent inputs. (b) An increase in the separation between the mean values of the 'noise' and 'signal + noise' inputs, i.e. a reduction of  $M_N$  relative to  $M_{SN}$ . These were referred to as 'distribution' effects since in each case the distributions of the afferent inputs corresponding to 'noise' or 'signal + noise' would be altered and a change in the critical value,  $c$ , would occur as a consequence of this. (c) A 'criterion' effect: a change in the computation of the criterion,  $c$ , produced directly, i.e. not consequent on a change in  $f_N(x)$  or  $f_{SN}(x)$ .

Processes suggested by the question about the relation between threshold and time interval were: (a) An 'arousal' or 'alerting' effect producing a rapid fall in threshold with a slower return to the resting level. (b) A 'temporal reference' or 'warning' effect, the accessory stimulus affecting the threshold by virtue of temporal information possessed by the subject. This was associated with two subsidiary hypotheses: (i) The threshold is reduced for the complete 'waiting time', to an extent inversely related to its length, or (ii) it is reduced for the duration of the period when the subject expects the critical stimulus to arrive, and the length of this 'range of expectation' is a function of the inter-stimulus interval.

These two sets of alternatives are independent in the sense that any of the signal-detection hypotheses can be combined with any of

the hypotheses about the form of the threshold curve; most of the experiments have provided information on one or the other set alone.

Experiment 10 provided the first main finding relevant to the signal-detection hypotheses. The constancy of (or rise in) the threshold variance as the threshold fell made hypothesis (a) difficult to maintain but left the other two alternatives equally plausible. The 'distribution' and 'criterion' effects were tested further by Experiments 11 and 12. Since no minimum was found in the threshold curve it did not seem necessary for the accessory stimulus to have acted on the afferent system at the time that the critical stimulus was administered; any possibility of a very rapid action of the accessory stimulus affecting a slowly conducted portion of the afferent input from the critical stimulus was excluded by the finding in Experiment 14 that the threshold could be lowered by an accessory stimulus given as much as 500 msec. after the critical stimulus.

By showing that a fall in threshold as the inter-stimulus interval decreased was not shown if the intervals were randomized over a large range Experiment 6 threw considerable doubt on the arousal hypothesis, though it remained possible that temporal information was important not because it acted directly on the threshold decision mechanism but because it prevented habituation of the arousal response. However an effect on the threshold was shown when the intervals were randomized over a small range, in Experiment 13, but the curve flattened

out at short intervals, a finding which was predicted by the 'range of expectation' hypothesis but not by either the 'waiting time' or the 'arousal' hypotheses. Further evidence for rejecting the arousal hypothesis was provided by Experiment 16, in which no effect of change in the intensity of a supraliminal accessory stimulus on the extent of the threshold fall produced by it could be found.

The hypothesis of a simple inverse relation between the degree of threshold fall and the range of expectation was supported by finding that the thresholds in Experiments 1, 2, and 4 were related to the inter-stimulus intervals in this way. In Section 10.1 it was further shown that the assumption of a 'criterion effect' combined with the temporal reference hypothesis allowed changes in the frequency of response to a stimulus to be predicted, and that the predictions showed satisfactorily close agreement with the obtained results. The relation of the changes in criterion required, to the inter-stimulus intervals, could be fairly closely approximated by an inverse function similar to those that were fitted to the results of Experiments 1, 2, and 4. Though it is possible to suppose that an arousal response might decay in such a fashion that the reduction in critical value would be inversely related to the time interval, it would be somewhat surprising if the change in  $c$  produced in this way were at the rate which approximates that calculated to maintain the probability of a false positive arising during the range of expectation constant.

Thus it seems possible to conclude, in view of these findings and other points mentioned in the discussion of particular experiments, that when the subject is attempting to detect a stimulus he reduces his threshold about the expected time of arrival of the stimulus for a period related to the inter-stimulus interval, and that the threshold fall is produced by a reduction in the critical value applied to afferent inputs arising during that period which is not secondary to a change in the parameters of the 'noise' and 'signal + noise' distributions. Furthermore it appears possible that the range of expectation may be an approximately constant proportion of the inter-stimulus interval (except for very short intervals), rather than a function of the differential threshold for duration, and that in reducing the criterion the change may be calculated to prevent the probability of a false positive response arising when the stimulus is expected exceeding a limiting value.

- WICKENS, C. (1948) Variation of the mean in a signal-detection task. *Psychol. Monographs*, 61, 749-767.
- WICKENS, C. (1967) *Psychological Methods for the Behavioral Sciences*. New York: McGraw-Hill.
- WICKENS, C. (1969) *Psychological Methods for the Behavioral Sciences*, 2nd ed. New York: McGraw-Hill.
- WICKENS, C. (1971) The theory of signal-detection. In *Psychological Methods for the Behavioral Sciences*, 3rd ed. New York: McGraw-Hill, pp. 21-54.
- WICKENS, C. (1981) The influence of time collection procedures upon psychological measurement of the sensory threshold. *Psychol. Monographs*, 94, 306-35.

REFERENCES

- ATTNEAVE, F. & McREYNOLDS, P.W. (1950) A visual beat phenomenon. *Amer. J. Psychol.*, 63, 107-110.
- AVERBACH, E. & SPERLING, G. (1961) Short Term Storage of Information in Vision. In C. Cherry(ed): Information Theory. Fourth London Symposium. London: Butterworths, 196-211.
- BAKER, K.E. & DUDEK, F.J. (1955) Weight scales from ratio judgments and comparisons of existent weight scales. *J. exp. Psychol.*, 50, 293-308.
- BARLOW, H.B. (1956) Retinal Noise and Absolute Threshold. *J. opt. soc. Amer.*, 46, 634-639.
- BARLOW, H.B. (1957a) Increment thresholds at low intensities considered as signal/noise discriminations. *J. Physiol.*, 136, 469-488.
- BARLOW, H.B. (1957b) Purkinje Shift and Retinal Noise. *Nature*, 179, 255-256.
- BARLOW, H.B. (1957c) Noise and the Visual Threshold. *Nature*, 180, 1405.
- BARLOW, H.B. (1958) Intrinsic Noise of Cones. In: Visual Problems of Colour, Vol. II. National Physical Laboratory Symposium No. 8, London: H.M.S.O. pp. 615-630.
- BARTLEY, S.H. (1934) Relation of intensity and duration of retinal stimulation by light to the electrical response of the optic cortex of the rabbit. *Amer. J. Physiol.*, 108, 397-408.
- BEKÉSY, G. von (1948) Vibration of the head in a sound field and its role in hearing by bone conduction. *J. Acoust. Soc. Amer.*, 20, 749-760.
- BEKÉSY, G. von (1960) Experiments in hearing. New York: McGraw-Hill.
- BENDER, M.B. (1952) Disorders in Perception. Thomas: Springfield, Ill.
- BIRDSALL, T.G. (1955) The Theory of Signal Detectability. In: H. Quastler (ed) Information Theory in Psychology. The Free Press: Glencoe, Ill. pp. 391-402.
- BLACKWELL, H.R. (1952) The Influence of Data Collection Procedures upon Psychophysical Measurement of Two Sensory Functions. *J. exp. Psychol.*, 44, 306-15.

BLACKWELL, H.R. (1953a) Evaluation of the Neural Quantum. Theory in Vision. Amer. J. Psychol., 66, 397-408.

BLACKWELL, H.R. (1953b) Psychophysical Thresholds. Experimental Studies of Methods of Measurement. Engineering Research Bulletin No. 36. Engineering Research Institute, University of Michigan.

BLACKWELL, H.R., PRITCHARD, B.S., & OHMART, J.G. (1954) Automatic Apparatus for stimulus presentation and recording in visual threshold experiments. J. opt. soc. Amer., 44, 322-326.

BLODGETT, H.C., JEFFRESS, L.A., & TAYLOR, R.W. (1958) Relation of masked threshold to signal-duration for various interaural phase-combinations. Amer. J. Psychol., 71, 283-290.

BORING, E.G. (1926) Auditory theory with special reference to intensity, volume, and localization. Amer. J. Psychol., 37, 157-188.

BORING, E.G. (1942) Sensation and Perception in the History of Experimental Psychology. New York: Appleton-Century-Crofts, Inc.

BORING, E.G., LANGFELD, H.S., & WELD, H.P. (1948) Foundations of Psychology. New York: Wiley.

BOUMAN, M.A., & VAN DER VELDEN, H.A. (1947) The Two-Quanta Explanation of the Dependence of the Threshold Values and Visual Acuity on the Visual Angle and the Time of Observation. J. opt. soc. Amer., 37, 908-919.

BRAZIER, M.A.B. (1951) The Electrical Activity of the Nervous System. London: Pitman.

BRINDLEY, G.S. (1960) Physiology of the Retina and the Visual Pathway. London: Arnold.

BROADBENT, D.E. (1958) Perception and Communication. London: Pergamon Press.

BROADBENT, D.E., & LADEFOGED, P. (1959) Auditory Perception of Temporal Order. J. acoust. soc. Amer., 31, 1539.

BROGDEN, W.J. (1950) Sensory conditioning measured by the facilitation of auditory acuity. J. exp. Psychol., 40, 512-519.

BULLER, A.J., NICHOLLS, J.G., & STRÖM, G. (1953) Spontaneous fluctuations of excitability in the muscle spindle of the frog. J. Physiol., 122, 409-18.

- BULMER, M.G., & HOWARTH, C.I. (1957) Noise and the Visual Threshold. *Nature*, 180, 1403-1404.
- CALLAWAY, E. III, & ALEXANDER, J.D.JR. (1960) The Temporal Coding of Sensory data: An Investigation of Two Theories. *J. gen. Psychol.*, 62, 293-309.
- CANE, V.R. (1956) Some Statistical Problems in Experimental Psychology. *J. Roy. Statist. Soc. B.*, 18, 177-194.
- CANE, V., & GREGORY, R.L. (1957) Noise and the Visual Threshold. *Nature*, 180, 1404-1405.
- CAUSSE, R., & CHAVASSE, P. (1947) Études sur la fatigue auditive. *l'Année Psychol.*, 43-44, 265-298.
- CHANG, H-T. (1959) The Evoked Potentials. In: J. Field, H.W. Magoun & V.E. Hall (eds). Handbook of Physiology. Section 1: Neurophysiology. Vol. 1, 299-313.
- CHEATHAM, P.G., & WHITE, C.T. (1952) Temporal Numerosity: I. Perceived Number as a function of flash number and rate. *J. exp. Psychol.*, 44, 447-451.
- CHEATHAM, P.G., & WHITE, C.T. (1954) Temporal Numerosity: III. Auditory Perception of Number. *J. exp. Psychol.*, 47, 425-428.
- CHILD, I.L., & WENDT, G.R. (1938) The Temporal Course of the Influence of Visual Stimulation upon the auditory threshold. *J. exp. Psychol.*, 23, 109-127.
- CLAUSEN, J. (1950) An Evaluation of Experimental Methods of Time Judgment. *J. exp. Psychol.*, 40, 756-61.
- CORSO, J.F. (1951) The Neural Quantum in Discrimination of Pitch and Loudness. *Amer. J. Psychol.*, 64, 350-368.
- CORSO, J.E. (1956) The neural quantum theory of sensory discrimination. *Psychol. Bull.*, 53, 371-393.
- CRITCHLEY, M. (1953) The Parietal Lobes. Arnold: London.
- CROSSMAN, E.R.F.W. (1955) The measurement of discriminability. *Quart. J. exp. Psychol.*, 7, 176-195.

- CROSSMAN, E.R.F.W. (1961) Information and Serial Order in Human Immediate Memory. In: C. Cherry (ed): Information Theory, Fourth London Symposium. London: Butterworths, 147-161.
- DAVIS, R. (1957) The Human Operator as a Single Channel Information System. Quart. J. exp. Psychol., 2, 119-129.
- DAVIS, H. (1959) Excitation of auditory receptors. In: J. Field, H.W. Magoun, & V.E. Hall (eds). Handbook of Physiology, Section 1: Neurophysiology. Vol. 1, 565-584.
- DENTON, E.J., & PIRENNE, M.H. (1954) The Absolute Sensitivity and functional Stability of the Human Eye. J. Physiol., 123, 417-442.
- DODT, E. (1956) Centrifugal Impulses in Rabbit's Retina. J. Neurophysiol., 19, 301-7.
- EGAN, J.P. (1948) The Effect of Noise in one ear upon the loudness of speech in the other ear. J. acoust. soc. Amer., 20, 58-62.
- EKMAN, G. (1958) Two generalized ratio scaling methods. J. Psychol., 45, 287-295.
- EKMAN, G. (1959) Weber's law and related functions. J. Psychol., 47, 343-352.
- EKMAN, G., & KUENNAPAS, T. (1960) Note on direct and indirect scaling methods. Psychol. Reps., 6, 174.
- FINNEY, D.J. (1952) Probit Analysis. Second Edition. Cambridge University Press.
- FISHER, R.A. (1954) Statistical Methods for Research Workers. Twelfth Edition. London: Oliver & Boyd.
- FISHER, R.A., & YATES, F. (1948) Statistical Tables for Biological, Agricultural and Medical Research. London: Oliver & Boyd.
- FITZHUGH, R. (1957) The Statistical Detection of Threshold Signals in the Retina. J. gen. Physiol., 40, 925-948.
- FRAISSE, P. (1948) Les Erreurs Constantes dans la Reproduction de Courts Intervalles Temporels. Arch. de Psychol., 32, 161-176.
- FRAISSE, P. (1957) Psychologie du Temps. Paris: Presses Universitaires de France
- FUSTER, J.M. (1958) Effects of Stimulation of Brain Stem on Tachistoscopic Perception. Science, 127, 150.

- GALAMBOS, R. (1954) Neural mechanisms of audition. *Physiol. Rev.* 34, 497-528.
- GALAMBOS, R. (1956) Suppression of auditory nerve activity by stimulation of efferent fibers to cochlea. *J. Neurophysiol.*, 19, 426-37.
- GALAMBOS, R., & DAVIS, H. (1944) Inhibition of Activity in Single Auditory Nerve Fibers by Acoustic Stimulation. *J. Neurophysiol.*, 7, 287-303.
- GARNER, W.R. (1952) An equal discriminability scale for loudness judgments. *J. exp. Psychol.*, 43, 232-238.
- GARNER, W.R. (1958) Advantages of the discriminability criterion for a loudness scale. *J. acoust. soc. Amer.*, 30, 1005-1012.
- GARNER, W.R. (1959) An argument for the use of discriminability scaling procedures in scaling sensory intensities. *Acta Psychol.*, 15, 94-7.
- GARNER, W.R., & HAKE, H.W. (1951) The amount of information in absolute judgments. *Psychol. Rev.* 58, 446-59.
- GEBHARD, J.W. (1953) Motokawa's studies on electric excitation of the human eye. *Psychol. Bull.*, 50, 73-111.
- GIBLIN, D.R. (1960) The effect of lesions of the nervous system on cerebral responses to peripheral nerve stimulation. *EEG Clin. Neurophysiol.*, 12, 262.
- GILBERT, G.M. (1941) Inter-sensory facilitation and inhibition. *J. gen. Psychol.*, 24, 381-407.
- GILLILAND, A.R., & HUMPHREYS, D.W. (1943) Age, Sex, Method, and Interval as Variables in Time Estimation. *J. genet. Psychol.*, 63, 123-130.
- GOLDSTONE, S., BOARDMAN, W.K., & LHAMON, W.T. (1959) Intersensory comparisons of temporal judgments. *J. exp. Psychol.*, 57, 243-248.
- GOLDSTONE, S., JERNIGAN, C., LHAMON, W.T., & BOARDMAN, W.K. (1959) A Further Note on intersensory differences in temporal judgment. *Percept. & Motor Skills*, 9, 252.
- GOODFELLOW, L.D. (1934) An Empirical Comparison of audition, vision, and touch in the discrimination of short intervals of time. *Amer. J. Psychol.*, 46, 243-258.
- GRANT, R. (1955a) Receptors and Sensory Perception. Yale University Press: New Haven.

- GRANT, R. (1955b) Centrifugal and antidromic effects on ganglion cells of retina. *J. Neurophysiol.*, 18, 388-411.
- GREEN, D.M. (1960) Psychoacoustics and detection theory. *J. acoust. soc. Amer.*, 32, 1189-1203.
- GREGG, L.W. (1951) Fractionation of temporal intervals. *J. exp. Psychol.*, 42, 307-12.
- GREGG, L.W., & BROGDEN, W.J. (1952) The effect of simultaneous visual stimulation on absolute auditory sensitivity. *J. exp. Psychol.*, 43, 179-186.
- GREGORY, R.L. (1956) An experimental treatment of vision as an information source and noisy channel. In: E.C.Cherry(ed): Information Theory, Third London Symposium. London: Butterworths Scientific Publications, pp. 287-299.
- GREGORY, R.L., & CANE, V. (1955) A Statistical Information Theory of Visual Thresholds. *Nature*, 176, 1272.
- GREGORY, R.L., & CANE, V.R. (1958) A Theory of Visual Thresholds. *Nature*, 181, 1488.
- GUILFORD, J.P. (1936) Psychometric Methods. First edition. New York and London: McGraw-Hill.
- GUILFORD, J.P. (1954) Psychometric Methods. Second edition. New York: McGraw-Hill.
- GUILFORD, J.P., & DINGMAN, H.F. (1954) A validation study of ratio-judgment methods. *Amer. J. Psychol.*, 67, 395-410.
- HAGBARTH, K-E., & FEX, J. (1959) Centrifugal influences on single unit activity in spinal sensory paths, 22, 321-38.
- HAGBARTH, K-E., & KERR, D.I.B. (1954) Central Influences on Spinal Afferent Conduction. *J. Neurophysiol.*, 17, 295-307.
- HALLIDAY, A.M., & MINGAY, R. (1961) Retroactive Raising of a Sensory Threshold by a contralateral stimulus. *Quart. J. exp. Psychol.*, 13, 1-11.
- HANES, R.M. (1949) The Construction of Subjective brightness scales from fractionation data: a validation. *J. exp. Psychol.*, 39, 719-728.
- HARPER, R.S., & STEVENS, S.S. (1948) A psychological scale of weight and a formula for its derivation. *Amer. J. Psychol.*, 61, 343-351.

- HARRIS, J.D. (1950) Some relations between vision and audition. Springfield, Ill: Thomas.
- HECHT, S. (1942) The Quantum Relations of Vision. *J. opt. soc. Amer.*, 32, 42-49.
- HECHT, S., SHLAER, S., & PIRENNE, M.H. (1942) Energy, Quanta, and Vision. *J. gen. Physiol.*, 25, 819-840.
- NELSON, H. (1959) Adaptation level theory. In: S. Koch (ed). Psychology: A Study of a Science. Study 1. Conceptual and Systematic. Vol. 1. New York: McGraw-Hill. pp. 565-621.
- HERNÁNDEZ-PEÓN, R., GUZMAN-FLORES, C., ALCARAZ, M., & FERNANDEZ-GUARDIOLA, A. (1957) Sensory transmission in visual pathway during "attention" in unanaesthetized cats. *Acta Neurol. Latinoamer.*, 3, 1-8.
- HERNÁNDEZ-PEÓN, R., JOUVET, M., & SCHERRER, H. (1957) Auditory potentials at cochlear nucleus during acoustic habituation. *Acta Neurol. Latinoamer.*, 3, 144-156.
- HERNÁNDEZ-PEÓN, R., SCHERRER, H., & JOUVET, M. (1956) Modification of electrical activity in cochlear nucleus during "Attention" in unanaesthetized cats. *Science*, 123, 331-2.
- HICK, W.E. (1952) On the rate of gain of information. *Quart. J. exp. Psychol.*, 4, 11-26.
- HIRSH, I.J. (1959) Auditory perception of temporal order. *J. acoust. soc. Amer.*, 31, 759-67.
- HIRSH, I.J., BILGER, R.C., & DEATHERAGE, B.H. (1956) The effect of auditory and visual background on apparent duration. *Amer. J. Psychol.*, 69, 561-574.
- HOAGLAND, H. (1932) The Physiological Control of Judgments of Duration: Evidence for a Chemical Clock. *J. gen. Psychol.*, 2, 267-287.
- HORN, G. (1960) Electrical Activity of the Cerebral Cortex of the Unanaesthetized Cat during Attentive Behaviour. *Brain*, 83, 57-76.
- HORN, G., & BLUNDELL, J. (1959) Evoked potentials in visual cortex of the unanaesthetized cat. *Nature*, 184, 173-4.
- HOWARTH, C.I. (1954) Strength duration curves for electrical stimulation of the human eye. *Quart. J. exp. Psychol.*, 6, 47-61.

- HOWARTH, C.I., & BULMER, M.G. (1956) Non-random sequences in visual threshold experiments. *Quart. J. exp. Psychol.*, 8, 163-171.
- HUMPHREY, G. (1933) The Nature of Learning. London: Kegan Paul, Trench, Trubner & Co., Ltd.
- HUMPHREY, G., DAWE, P.G.M., & MANDELL, B. (1955) New high-speed electronic tachistoscope. *Nature*, 176, 231-234.
- JABBUR, S.J., & TOWE, A.L. (1959) Blocking and excitation of cuneate neurons by sensori-motor cortical stimulation in the cat. *Fed. Proc.*, 18, 73.
- JABBUR, S.J., & TOWE, A.L. (1960) Effect of pyramidal tract activity on dorsal column nuclei. *Science*, 132, 547-548.
- KEMP, E.H., COPPEE, G.E., & ROBINSON, E.H. (1937) Electric responses of the brain stem to unilateral auditory stimulation. *Amer. J. Physiol.*, 120, 304-315.
- KERR, D.I.B., & HAGBARTH, K.-E. (1955) An investigation of Olfactory Centrifugal Fiber System. *J. Neurophysiol.*, 18, 362-74.
- KOESTER, T., & SCHOENFELD, W.N. (1947) Some comparative data on Differential Pitch Sensitivity under Quantal and Non-Quantal Conditions. *J. gen. Psychol.*, 36, 107-112.
- KOHATA, T., KOMATSU, M., & MOTOKAWA, K. (1956) After-Effects of Electrical Test Shocks as a factor affecting retinal color effects. *Japanese J. Physiol.*, 6, 236-248.
- KORZUN, P.A. (1957) Changes in the electrical sensitivity of the eye after illumination with intense flashes of different colours. *Biophysics*, 2, 428-437.
- KOWALSKI, W.J. (1943) The effect of delay upon the duplication of short temporal intervals. *J. exp. Psychol.*, 33, 239-246.
- KUFFLER, S.W., FITZHUGH, R., & BARLOW, H.B. (1957) Maintained activity in the cat's retina in light and darkness. *J. gen. Physiol.*, 40, 683-702.
- KUFFLER, S.W., HUNT, C.C., & QUILLIAM, J.P. (1951) Function of medullated small nerve-fibers in mammalian ventral roots: efferent muscle spindle innervation. *J. Neurophysiol.*, 14, 29-54.

- KÜLPE, O. (1901) Outlines of Psychology. New York: Macmillan.
- LEWIS, D., & BURKE, C.J. (1949) The use and misuse of the Chi-square Test. *Psychol. Bull.*, 46, 433-489.
- LICKLIDER, J.C.R. (1948) The influence of interaural phase relations upon the masking of speech by white noise. *J. acoust. soc. Amer.*, 20, 150-159.
- LICKLIDER, J.C.R. (1959) Three Auditory Theories. In S. Koch (ed). Psychology: A Study of a Science. Study I. Conceptual and Systematic. Vol. 1. New York: McGraw-Hill. pp. 41-144.
- LIVINGSTON, R.B. (1959) Central control of receptors and sensory transmission systems. In J. Field (ed): Handbook of Physiology. Section 1: Neurophysiology. Vol. 1. pp. 741-60.
- LOEWENSTEIN, W.R. (1956) Modulation of cutaneous mechanoreceptors by sympathetic stimulation. *J. Physiol.*, 132, 40-60.
- LONDON, I.D. (1954) Research on sensory interaction in the Soviet Union. *Psychol. Bull.*, 51, 531-68.
- LUCE, R.D. (1959a) On the possibility psychophysical laws. *Psychol. Rev.*, 66, 81-95.
- LUCE, R.D. (1959b) Individual Choice Behavior. New York: Wiley.
- LUCE, R.D., & EDWARDS, W. (1958) The derivation of subjective scales from just noticeable differences. *Psychol. Rev.*, 65, 222-37.
- MACCORQUODALE, K., & MEEHL, P.E. (1948) On a distinction between hypothetical constructs and intervening variables. *Psychol. Rev.*, 55, 95-107.
- MAGOON, H.W. (1958) The Waking Brain. Springfield, Ill: Thomas.
- MARRIOTT, F.H.C. (1956) Frequency-of-seeing curves and spontaneous excitation. *J. opt. soc. Amer.*, 46, 661.
- MARRIOTT, F.H.C. (1958) A Theory of Visual Thresholds. *Nature*, 181, 1487-1488.
- MARSHALL, W.H., TALBOT, S.A., & ADES, H.W. (1943) Cortical response of the anesthetized cat to gross photic and electrical afferent stimulation. *J. Neurophysiol.*, 6, 1-15.

- MICHAELS, R.M. (1957a) The electrical phosphene threshold as a measure of retinal induction and visual organization. *J. exp. Psychol.*, 54, 21-27.
- MICHAELS, R.M. (1957b) Personal communication.
- MICHAELS, R.M. (1960) Anisotropy and interaction of fields of spatial induction. *J. exp. Psychol.*, 60, 235-41.
- MILLER, G.A. (1956) The Magical Number Seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.*, 63, 81-97.
- MILLER, G.A., & GARNER, W.R. (1944) Effect of random presentation on the psychometric function: implications for a quantal theory of discrimination. *Amer. J. Psychol.*, 57, 451-467.
- MONNIER, M. (1952) Retinal, cortical and motor responses to photic stimulation in man. *J. Neurophysiol.*, 15, 469-86.
- MOTOKAWA, K. (1949a) Retinal processes and their role in color vision. *J. Neurophysiol.*, 12, 291-303.
- MOTOKAWA, K. (1949b) Physiological Studies on Mechanisms of Color Reception in Normal and Color-Blind Subjects. *J. Neurophysiol.*, 12, 465-474.
- MOTOKAWA, K. (1949c) Physiological induction in human retina as basis of color and brightness contrast. *J. Neurophysiol.*, 12, 475-488.
- MOTOKAWA, K. (1949d) Visual function and the electrical excitability of the retina. *Tohoku J. exper. Med.*, 51, 145-153.
- MOTOKAWA, K. (1950) Field of retinal induction and optical illusion. *J. Neurophysiol.*, 13, 413-426.
- MOTOKAWA, K. (1951a) Propagation of retinal induction. *J. Neurophysiol.*, 14, 339-351.
- MOTOKAWA, K. (1951b) Some remarks on measurements of electrical excitability of the human eye. *Tohoku J. Exper. Med.*, 54, 385-392.
- MOTOKAWA, K., EBE, M., ARAKAWA, Y., & OIKAWA, T. (1951) Studies on the physiological color blindness of the human fovea with the polarization method. *Japanese J. Physiol.*, 2, 50-59.

- MOTOKAWA, K., & IWAMA, K. (1949a) Electric stimulation of the human retina with exponentially increasing currents. *Tohoku J. exper. Med.*, 50, 25-37.
- MOTOKAWA, K., & IWAMA, K. (1949b) Color receptors in frog eyes as revealed by electrical stimulation. *Tohoku J. exper. Med.*, 50, 240 only.
- MOTOKAWA, K., & IWAMA, K. (1949c) The three colour processes in the retina of frog. *Tohoku J. exper. Med.*, 50, 258 only.
- MOTOKAWA, K., & IWAMA, K. (1951) Color processes and physiological induction in Frog's retina. *Tohoku J. exper. Med.*, 53, 341-349.
- MOTOKAWA, K., IWAMA, K., & EBE, M. (1952) Retinal colour processes in cats. *Japanese J. Physiol.*, 2, 198-207.
- MOTOKAWA, K., IWAMA, K., & TUKAHARA, S. (1951) Color processes in single retinal elements. *Tohoku J. exper. Med.*, 53, 399-406.
- MOWRER, O.H. (1940) Preparatory set (expectancy) - some methods of measurement. *Psychol. Monogr.* 52, no. 233.
- MOWRER, O.H. (1941) Preparatory set (expectancy) - further evidence of its 'central' locus. *J. exper. Psychol.*, 28, 116-133.
- MOWRER, O.H., RAYMAN, W.N., & BLISS, E.L. (1940) Preparatory set (expectancy) - an experimental demonstration of its 'central' locus. *J. exp. Psychol.*, 26, 357-372.
- MUELLER, C.G. (1950) Quantum concepts in visual intensity-discrimination. *Amer. J. Psychol.*, 63, 92-100.
- MURDOCK, B.B. Jr. (1960) The distinctiveness of stimuli. *Psychol. Rev.*, 67, 16-31.
- NAQUET, R., REGIS, H., FISCHER-WILLIAMS, M., & FERNANDEZ-GUARDIOLA, A. (1960) Variation in the responses evoked by light along the specific pathways. *Brain*, 83, 52-6.
- NEISSER, U. (1957) Response-sequences and the hypothesis of the neural quantum. *Amer. J. Psychol.*, 70, 512-527.
- NEWHALL, S.M. (1923) Effects of attention on the intensity of cutaneous pressure and on visual brightness. *Arch. Psychol.*, N.Y., No. 61, 1-75.

- NEYMAN, J., & PEARSON, E.S. (1933) On the problem of the most efficient tests of statistical hypotheses. *Phil. Trans. Roy. Soc. A.*, 231, 289-337.
- OBERLIN, K.W. (1936) Variation in intensitive sensitivity to lifted weights. *J. exp. Psychol.*, 19, 438-455.
- OLDFIELD, R.C. (1937) Some recent experiments bearing on 'Internal Inhibition'. *Brit. J. Psychol.*, 28, 28-42.
- PAVLOV, I.P. (1927) Conditioned Reflexes. Oxford: Oxford University Press.
- PETERSON, W.W., & BIRDSALL, T.G. (1953) The Theory of Signal Detectability. Tech. Report no. 13, Electronic Defense Group, University of Michigan.
- PIERON, H. (1952) The Sensations. Their Functions, Processes and Mechanisms. London: Frederick Muller Ltd.
- PIERON, H., & SEGAL, J. (1939) Sur un phénomène de facilitation rétroactive dans l'excitation électrique de branches nerveuses cutanées (sensibilité tactile). *J. Neurophysiol.*, 2, 178-191.
- PIRENNE, M.H. (1943) Binocular and unocular threshold of vision. *Nature*, 152, 698-699.
- PIRENNE, M.H. (1956) Physiological mechanisms of vision and the quantum nature of light. *Biol. Rev.*, 31, 194-241.
- PIRENNE, M.H., & MARRIOTT, F.H.C. (1955) Absolute threshold and frequency-of-seeing curves. *J. opt. soc. Amer.*, 45, 909-912.
- REESE, T.S., & STEVENS, S.S. (1960) Subjective intensity of coffee odor. *Amer. J. Psychol.*, 73, 424-8.
- RIGGS, L.A., CORNSWEET, J.C., & LEWIS, W.G. (1957) Effects of light on electrical excitation of the human eye. *Psychol. Monogr.* 71, Whole No. 434, 1-45.
- ROSE, A. (1948) The sensitivity performance of the human eye on an absolute scale. *J. opt. soc. Amer.*, 38, 196-208.
- RYAN, T.A. (1940) Interrelations of the sensory systems in perception. *Psychol. Bull.*, 37, 659-98.

- SIEGEL, S. (1956) Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill.
- SHERRINGTON, C.S. (1947) The Integrative Action of the Nervous System. Cambridge: Cambridge University Press.
- SMITH, M., & WILSON, E.A. (1953) A model of the auditory threshold and its application to the problem of the multiple observer. *Psychol. Monogr.* 67, Whole no 359, pp. 1-35.
- SPELTING, G. (1960) The information available in brief visual presentations. *Psychol. Monogr.* 74, Whole no. 498, 1-29.
- STEVENS, J.C., MACK, J.D., & STEVENS, S.S. (1960) Growth of sensation on seven continua as measured by force of handgrip. *J. exp. Psychol.*, 59, 60-7.
- STEVENS, J.C., & STEVENS, S.S. (1960) Warmth and cold: dynamics of sensory intensity. *J. exp. Psychol.*, 60, 183-92.
- STEVENS, S.S. (1936) A scale for the measurement of a psychological magnitude: loudness. *Psychol. Rev.*, 43, 405-416.
- STEVENS, S.S. (1951) Mathematics, Measurement, and Psychophysics. In S.S. Stevens (ed). Handbook of Experimental Psychology. New York: Wiley. pp. 1-49.
- STEVENS, S.S. (1955) The measurement of loudness. *J. acoust. soc. Amer.*, 27, 815-829.
- STEVENS, S.S. (1956) The direct estimation of sensory magnitudes - loudness. *Amer. J. Psychol.*, 69, 1-25.
- STEVENS, S.S. (1957) On the psychophysical law. *Psychol. Rev.*, 64, 153-181.
- STEVENS, S.S. (1958) Problems and methods of psychophysics. *Psychol. Bull.*, 55, 177-196.
- STEVENS, S.S. (1959a) On the validity of the loudness scale. *J. acoust. soc. Amer.*, 31, 995-1003.
- STEVENS, S.S. (1959b) Cross-modality validation of subjective scales for loudness, vibration, and electric shock. *J. exp. Psychol.*, 57, 201-9.

- STEVENS, S.S. (1959c) Tactile vibration: dynamics of sensory intensity. *J. exp. Psychol.*, 57, 210-8.
- STEVENS, S.S. (1959d) On the measurement of sensation. *Acta Psychol.*, 15, 91-4
- STEVENS, S.S. (1960a) Ratio scales, Partition Scales and Confusion Scales. In H. Gulliksen & S. Messick (eds) Psychological Scaling: Theory and Applications. New York: Wiley. pp. 49-66.
- STEVENS, S.S. (1960b) The psychophysics of sensory function. *Amer. Scientist*, 48, 226-53.
- STEVENS, S.S. (1960c) On the new psychophysics. *Scand. J. Psychol.*, 1, 27-35.
- STEVENS, S.S., & DAVIS, H. (1938) Hearing. Its Psychology and Physiology. New York: Wiley.
- STEVENS, S.S., & GALANTER, E.H. (1957) Ratio scales and category scales for a dozen perceptual continua. *J. exp. Psychol.*, 54, 377-411.
- STEVENS, S.S., MORGAN, C.T., & VOLKMANN, J. (1941) Theory of the Neural Quantum in the discrimination of loudness and pitch. *Amer. J. Psychol.*, 54, 315-335.
- STEVENS, S.S., & VOLKMANN, J. (1940) The relation of pitch to frequency: a revised scale. *Amer. J. Psychol.*, 53, 329-53.
- STOIBER, A.M., & BARNES, C.D. (1960a) Excitatory interaction of cerebellar and peripheral inputs to cuneate units. *Physiologist*, 3, 524.
- STOIBER, A.M., & BARNES, C.D. (1960b) Modification of peripherally evoked cuneate responses by cerebellar stimulation. *Fed. Proc.*, 19, 288.
- STROUD, J.M. (1955) The Fine Structure of Psychological Time. In H. Quastler (ed). Information Theory in Psychology. Glencoe, Ill: Free Press.
- SWETS, J.A. (1959) Indices of signal detectability obtained with various psychophysical procedures. *J. acoust. soc. Amer.*, 31, 511-513.
- SWETS, J.A., SHIPLEY, E.F., McKEY, M.J., & GREEN, D.M. (1959) Multiple Observations of signals in noise. *J. acoust. soc. Amer.*, 31, 514-521.

- SWETS, J.A., TANNER, W.P.Jr., & BIRDSALL, T.G. (1955) The evidence for a decision-making theory of visual detection. Technical Report No. 40., Electronic Defense Group, Dept. of Electrical Engineering, University of Michigan.
- SYMONS, J.R. (1954) An investigation of intersensory relationships. D.Phil. Thesis, University of Reading.
- TANNER, W.P.Jr. (1955) On the Design of Psychophysical Experiments. In H. Quastler (ed) Information Theory in Psychology. Glencoe, Ill: Free Press. pp. 403-414.
- TANNER, W.P.Jr., & SWETS, J.A. (1954) A decision-making theory of visual detection. *Psychol. Rev.*, 61, 401-409.
- TANNER, W.P.Jr., SWETS, J.A., & GREEN, D.M. (1956) Some general properties of the hearing mechanism. Technical Report No. 30. Electronic Defense Group, Dept. of Electrical Engineering, University of Michigan.
- TASAKI, I. (1954) Nerve impulses in individual auditory nerve fibers of guinea pig. *J. Neurophysiol.*, 17, 97-122.
- THOMPSON, R.F., VOSS, J.F., & BROGDEN, W.J. (1958) Effect of brightness of simultaneous visual stimulation on absolute auditory sensitivity. *J. exp. Psychol.*, 55, 45-50.
- THURSTONE, L.L. (1927a) Psychophysical Analysis. *Amer. J. Psychol.*, 38, 368-389.
- THURSTONE, L.L. (1927b) A law of Comparative Judgment. *Psychol. Rev.*, 34, 273-286.
- THURSTONE, L.L. (1927c) A Mental Unit of Measurement. *Psychol. Rev.*, 34, 415-423.
- THURSTONE, L.L. (1927d) Three psychophysical laws. *Psychol. Rev.*, 34, 424-432.
- THURSTONE, L.L. (1928) The phi gamma hypothesis. *J. exp. Psychol.*, 11, 293-305.
- THURSTONE, L.L. (1948) Psychophysical Methods. In T.G. Andrews (ed) Methods of Psychology. New York: Wiley. pp. 124-157.
- THURSTONE, L.L. (1954) The Measurement of values. *Psychol. Rev.*, 61, 47-58.

- TORGERSON, W.S. (1958) Theory and Methods of Scaling. New York: Wiley.
- TORGERSON, W.S. (1960) Quantitative Judgment Scales. In H. Gulliksen & S. Messick (eds). Psychological Scaling: Theory and Applications. New York: Wiley. pp. 21-31.
- TOWE, A.L., & AMASSIAN, V.E. (1958) Patterns of activity in single cortical units following stimulation of the digits in monkeys. *J. Neurophysiol.*, 21, 292-311.
- TREISMAN, A.M. (1960) Contextual cues in selective listening. *Quart. J. exp. Psychol.*, 12, 242-8.
- TUNTURI, A.R. (1945) Further afferent connections of the acoustic cortex of the dog. *Amer. J. Physiol.*, 144, 389-394.
- TUNTURI, A.R. (1946) A study of the pathway from the medial geniculate body to the acoustic cortex in the dog. *Amer. J. Physiol.*, 147, 311-319.
- VERPLANCK, W.S., COLLIER, G.H., & COTTON, J.W. (1952) Nonindependence of successive responses in measurements of the visual threshold. *J. exp. Psychol.*, 44, 273-282.
- WARREN, R.M. (1958) A basis for judgments of sensory intensity. *Amer. J. Psychol.*, 71, 675-87.
- WARREN, R.M., & POULTON, E.C. (1960) Basis for lightness-judgments of grays. *Amer. J. Psychol.*, 73, 380-7.
- WARREN, R.M., SERSEN, E.A., & PORES, E.B. (1958) A basis for loudness-judgments. *Amer. J. Psychol.*, 71, 700-9.
- WARREN, R.M., & WARREN, R.P. (1956) Effect of the relative volume of standard and comparison-object on half-heaviness judgments. *Amer. J. Psychol.*, 69, 640-3.
- WARREN, R.M., & WARREN, R.P. (1958) Basis for judgments of relative brightness. *J. opt. soc. Amer.*, 48, 445-50.
- WEBER, A.O. (1933) Estimation of time. *Psychol. Bull.*, 30, 233-252.
- WEVER, E.G. (1949) Theory of Hearing. New York: Wiley.
- WHITE, C.T., & CHEATHAM, P.G. (1959) Temporal Numerosity: IV. A comparison of the major senses. *J. exp. Psychol.*, 58, 441-444.

- WOODROW, H. (1914) The measurement of attention. Psychol. Monogr. 17, Whole No. 176, 1-158.
- WOODROW, H. (1930) The reproduction of temporal intervals. J. exp. Psychol., 13, 473-499.
- WOODROW, H. (1934) The temporal indifference interval determined by the method of mean error. J. exp. Psychol., 17, 167-188.
- WOODROW, H. (1951) Time Perception. In S.S. Stevens (ed) Handbook of Experimental Psychology. New York: Wiley. pp. 1224-1236.
- WOODWORTH, R.S. (1938) Experimental Psychology. New York: Holt.
- WOODWORTH, R.S., & SCHLOSBERG, H. (1955) Experimental Psychology. London: Methuen.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor R.C. Oldfield, for the provision of research facilities, for his advice and interest and for assistance in many other ways which has contributed to making the completion of this thesis possible. I would also like to thank Dr. C.I. Howarth, Dr. J.A. Deutsch, Dr. Martin Halliday, Dr.F.H.C. Marriott, Dr. D. Drazin, and my wife, Anne Treisman, for many useful comments and conversations, and to express my gratitude to Dr. Howarth for his cheerful and efficient collaboration in carrying out Experiments 1-6, 10, 12, 14 and 15. My thanks are due to Mr. J.F. Scott and Dr. M. Bulmer for their kind help with many statistical problems, to Mr. P. Coughlin, Mr. R. Shrimpton, and Mr. P. Dawe for their willing assistance on many occasions with the construction of apparatus, to Mr. B. Mandell for the loan of the tachistoscope used in the earlier experiments, to Miss R. Williams for the care and efficiency with which she has typed this thesis, to Miss J. Burnett Brown for her assistance with the figures, and to the many subjects whose contributions, in however disguised a form, figure in these pages.

Part of the work reported here was performed while in receipt of a grant from the Medical Research Council and a Senior Hulme Scholarship from Brasenose College; I would like to express my gratitude to these two bodies for their support.