1978

# Cross-modality temporal resolution for auditory, vibrotactile, and visual stimuli

Donal Gene Sinex

## Recommended Citation

Sinex, Donal Gene, "Cross-modality temporal resolution for auditory, vibrotactile, and visual stimuli" (1978). *Independent Studies and Capstones.* Paper 465. Program in Audiology and Communication Sciences, Washington University School of Medicine. http://digitalcommons.wustl.edu/pacs_capstones/465

WASHINGTON UNIVERSITY

Department of Psychology

Dissertation Committee:

Charles S. Watson, Chairman
Ira J. Hirsh
James D. Miller
Thomas T. Sandel
James A. Simmons

# CROSS-MODALITY TEMPORAL RESOLUTION FOR AUDITORY, VIBROTACTILE, AND VISUAL STIMULI

by

Donal Gene Sinex

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August, 1978

St. Louis, Missouri

# ABSTRACT

The sensory environment outside the laboratory presents us with stimuli constantly changing in time, that may be perceived in more than one sensory modality. We hear speech and can at the same time see the talker's lip movements. A number of studies of the perception of temporal relations in two modalities are limited by the ability to resolve that relation, but they are subject to other limitations as well. There are large effects due to attentional mechanisms, as indicated by the phenomenon of prior entry. The present set of experiments attempted to assess the resolution of temporal intervals without regard for the percepts elicited by those intervals. The intervals were marked by brief stimuli delivered to the auditory, the tactile, and the visual modalities. Three questions were asked. First, how does cross-modality temporal resolution compare to within-modality temporal resolution? Second, are there differences in cross-modality temporal resolution that are related to the particular pair of modalities stimulated? Third, does the presence of perceptual boundaries in the temporal dimension, such as that between perceived simultaneity and perceived order, affect discrimination performance, as it does in cases of categorical perception?

Experiments 1-3. Within- and cross-modality temporal resolution was measured with the same-different psychophysical procedure. Standard intervals, designated Ts, ranged from 0 to ±480 msec, and comparison intervals longer and shorter than Ts were tested.

Three conclusions were drawn from the results. First, within-modality temporal resolving power is generally greater than cross-modality temporal resolving power. That is, smaller changes are discriminable at a fixed performance level when the stimuli are presented to a single modality. Second, for Ts=0 msec, auditory-tactile temporal resolution is greater than visual-auditory or visual-tactile temporal resolution. For larger Ts, differences between cross-modality conditions are less pronounced. Third, for certain cross-modality conditions, the just-resolvable change is quite large, relative to predictions based on Weber's Law and on performance in other conditions of the same experiment. The difficulty is hypothesized to be related to the subjects' tendency to respond categorically when the stimuli are within the range that corresponds to perceptual simultaneity.

Experiment 4. Cross-modality temporal resolution for the conditions found to be poorly resolved in Experiments 1-2 was remeasured with a criterion-free forced-choice procedure. However, the hypothesis that smaller changes might be resolved with this procedure was not supported. The form of stimulus tracks produced by the adaptive procedure suggested that fluctuations in attention or in subject strategy result in performance that is alternately much better or much worse than the long term average. To evaluate the variability of the tracks, the data were simulated with a Monte Carlo procedure, under the assumption that an invariant psychometric function governs performance. The simulated runs were significantly less variable than the actual runs, supporting the hypothesis that attentional factors are not constant throughout an experimental session.

Experiment 5. Within-modality and cross-modality temporal resolution over the range of Ts from 0 to ± 120 msec was measured with the method of adjustment. With this procedure, smaller and more reliable estimates of the just-resolvable change were obtained. The major conclusions suggested by Experiments 1-3 were confirmed. First, within-modality temporal resolution is greater than cross-modality temporal resolution. Second, auditory-tactile temporal resolution is better than that observed with cross-modality pairs involving vision. For the latter cases, the just-resolvable change is approximately constant over the range of Ts tested. Auditory-tactile temporal resolution is consistent with Weber's Law.

A model that assumes (1) that within-modality temporal discriminations are mediated at more peripheral sites in the nervous system than are cross-modality discriminations, and (2) that at least three independent sources of variance combine to limit performance was proposed. The three sources are associated with the encoding of the two marker stimuli and with the measurement of the interval itself. The model predicts the major findings, at least on a qualitative level, including the difference between within- and cross-modality performance; the differences between cross-modality conditions involving vision and the auditory-tactile case; and the near-constancy of the just-resolvable change in cross-modality cinditions including the visual modality.

OK
CSWatson

## TABLE OF CONTENTS

Page

# LIST OF TABLES

LIST OF FIGURES

# CHAPTER ONE. Introductory remarks and review of the related literature.

## I. INTRODUCTION

The study of the perceived temporal relation between events in two sensory modalities is older than experimental psychology itself. In the nineteenth century, astronomers tried to write "personal equations" because they had already discovered that idiosyncratic biases enter into the perception of and description of ordered events (Boring, 1959, p. 134). Experimental investigations of the perception of nearly-simultaneous events have continued to be characterized by ambiguity of results and great variation in individual performance. In the experiments reported here, an attempt was made to understand a different aspect of the sensory representation of paired events. Here, the capability for resolving differences in the onset time of the elements of a cross-modality stimulus pair was measured. The experiments were designed to provide the first complete and systematic investigation of temporal resolution for cross-modality stimuli. In addition, they were further motivated by the possibility that this new information would add to the understanding of the interaction between perceptual reports and resolving power. This interaction is most evident in the phenomenon of "categorical perception", a mode of sensory processing once thought to be unique to speech stimuli. There are now a number of reports that indicate that other stimulus configurations will induce subjects to treat the sensory continuum as if it were broken into discrete regions. These are reviewed below. There are two reasons for considering the relation between resolving power and perceived qualities of stimuli in the temporal domain. First, when paired events are separated by only a few milliseconds, observers may report that they experience simultaneity or order, or under some conditions apparent motion. If discrimination performance is enhanced at the boundary between regions of differing perceptual quality, one might expect to find more than one temporal relation for which resolution is better or worse than the discrimination performance expected under Weber's Law. Second, there are currently very few studies that have directly measured temporal resolving power for cross-modality stimulus pairs. There are no studies that examine combinations other than visual-auditory. If one wished to compare temporal resolution for other combinations, the most relevant data are to be found in those experiments that determine the precision with which the order of events can be labeled. While those results certainly reflect the limitations imposed by temporal resolving power, they are confounded by other factors, which are discussed below. It is impossible to extract the component due to simple resolving power from those studies. Discrimination must be measured directly, and that is the contribution of the current experiments.

The remainder of this chapter is divided into six sections, reviewing experiments that are relevant to the consideration of temporal processing and temporal aspects of processing in each of three sensory modalities: the auditory, the visual, and the tactile. The first section defines the terms that will be used throughout the dissertation and distinguishes between the two experimental strategies that have been applied to the study of temporal phenomena. It also presents the rationale behind the current experiments. A historical review follows. The next section examines temporal processing within the three modalities mentioned. Experiments dealing with cross-modality stimuli are presented next, with special consideration given to the experimenters' thoughts about the role of peripheral and central mechanisms. In the last section, theories of temporal processing are discussed.

## II. CONCEPTS and TERMINOLOGY

The Theory of Signal Detectability (Swets, Tanner, and Birdsall, 1961) provides a way to separate the effects of the subject's criterion for responding from his actual sensitivity or resolving power. The possibility of bias in psychopysical judgments had long been recognized, and procedural variations such as extensive training and the use of catch trials had been used in an attempt to stabilize criterion. With detection-style analysis, a quantitative appraoch to the criterion problem is possible.

The fact that sensitivity is independent of criterion is of interest, but it is perhaps less interesting than the observation that for a given stimulus configuration, subjects are likely to place their response criteria in a non-random manner. The fact that a particular response criterion is selected when the subject has a number of alternative choices is an important aspect of sensory behavior that should be studied separately from resolving power. Watson (1973) has pursued the distinction between two aspects of sensory functioning: sensory capabilities and response proclivities. When the limits of a sensory capability are estimated, the proper experimental question is, how much information about the stimulus sequence can the subject transmit. The experiments reported here are intended to measure cross-modality temporal resolving power, a sensory capability. By temporal resolving power it is meant the capability to distinguish between two intervals of differing temporal extent. Temporal resolving power, under this definition, is a general capability, the measurement of which is not limited to certain lengths of interval or to intervals marked with special stimuli. There are several psychophysical tasks for which performance is limited by temporal resolving power but that do not themselves produce uncontaminated estimates of resolving power. In these tasks, response proclivities are studied. A response proclivity is a tendency for an observer to respond to a stimulus in a certain way, when other responses might have been possible. Such a response might be a verbal description of a percept. For example, if an observer indicates that of two sounds differing in intensity by 10 dB, one is twice as loud as the other, he is giving information about a response proclivity. A 3-dB increase in intensity could have been associated with a doubling of loudness; the fact is that it is not. To study response proclivities, the methods for the study of sensory capabilities are inappropriate, for those procedures typically restrict the subject's opportunity to express his perceptual experience.

If a subject observes a stimulus pair and is asked to describe the temporal relation between its elements, a response proclivity is being measured. A stimulus pair with the onsets of elements separated by a few milliseconds could result in the experience of simultaneity or of order. One cannot say which of the two occurs without conducting the experiment. And having determined the label assigned to such a pair, one cannot predict the discriminability of that pair and another in which the relative onset time of elements is slightly different. The relation between the proclivity to label a stimulus and the capability to resolve it is complex. In some instances, a response proclivity, a tendency to assign a label to a particular stimulus configuration, is accompanied by enhanced resolution (see the discussion of categorical perception below). In others, the stimulus may be resolved with the same precision as adjacent stimuli that would be labeled differently. In each case, sensory capability must be estimated independent of the measurement of response proclivities.

The definition of temporal resolution allows the possibility that modality differences exist; however, it requires that statements about those differences be based upon the results from experimental methods used with all the stimuli to be compared. For example, cross-modality temporal resolution estimated using procedures appropriate to the study of sensory capability can be contrasted to within-modality temporal resolution estimated using similar methods. These results can be compared to the results of order-labeling experiments with the understanding that the two procedures measure different aspects of temporal processing.

Two conventions for distinguishing between physical stimuli and their associated perceptual responses are followed. Almost all of the stimuli described in the literature review consist of pulses delivered to one of the three major sensory modalities. Following Bartley's (1958) suggestion, the term "pulse" refers to a stimulus. The response to a pulse is called a click, a flash, or a tap, depending on the modality stimulated. The stimuli in many cases are paired pulses, presented to the same or to different modalities. The pulses are also described as being the elements in the pair. The perceptual correlates of the elements are called events. Thus if a subject observes a stimulus pair whose elements are simultaneous, it may be the case that he experiences an ordered pair of events.

The temporal relation between elements of a pair is measured from onset to onset, unless noted, and is symbolized by T. In describing experiments that use single-interval psychophysical procedures, no subscripts are used. For multiple-interval procedures, subscripts are used to distinguish between standard and comparison stimuli. The difference in onset times in a standard stimulus is indicated by Ts. The difference in onset times in the stimuli to be discriminated from the standard is indicated by Tc. Note that Tc denotes the absolute temporal relation in the comparison stimulus. The algebraic difference between Ts and Tc is the quantity that is resolved, and it is symbolized by $\Delta T$. For all pairs formed from two elements, the sign of Ts indicates the order of elements. The sign convention is established as specific cases are presented. The sign of $\Delta T$ indicates the direction in time in which a change is made. An example is given. A pair of elements A and B has Ts=+100 msec. A change in the positive direction, $\Delta T=+50$, produces the pair with Tc=+150 msec. A change in the negative direction, i.e. $\Delta T=-50$ msec, produces the pair with Tc=+50 msec. When Ts is negative, a positive $\Delta T$ forms Tc with a smaller absolute value; that is, for all cases,

$$Tc = Ts + \Delta T \qquad \text{(eq. 1.1).}$$

The experiments reviewed employed two basic procedures. One class of experiment is used to determine the stimulus configuration that will produce a specified perceptual report. The most common experiment is one that requires the subject to label the order of a pair of events, such as a flash and a click. As was emphasized above, this procedure is suitable for investigating a response proclivity. The subject is usually limited to two choices; for a pair consisting of element A and element B, he must say that "A" or "B" came first. In physical time, a pair of elements may be simultaneous or have one of two orders. In perceptual time, no fewer than five relations may be reported, according to Fraisse (1963, pp. 99-115): unity; simultaneity; non-simultaneity without the percept of distinct events; succession; and finally, identifiable order. For a given stimulus pair, the nature of the elements will limit the possible percepts. For two identical elements, four reports are possible as increasing delay between onsets (T) is introduced. For the smallest T, the two events will fuse and seem to be a single event (unity). As T increases, the quality of the percept will be changed, such that the pair is discriminable from a single event. For still larger T, two events will be perceived as successive; perceived simultaneity, however, does not occur. While two separate events will eventually be perceived, it is apparent that a subject cannot be asked which of two identical events came first. For similar but not identical elements, unity is not perceived but simultaneity is; two tones that form a chord are easily experienced as simultaneous. Fraisse (1963) suggests that the perception of simultaneity is most easily achieved when the two events are related as if from a single source, when they

"form a figure with one unit of meaning" (p. 108). Non-simultaneity without succession occurs when T is increased slightly. In this region, the quality of the complex changes; in vision, apparent movement between two positions might be reported, for example. Both succession and ordering will occur as T is made larger. For two stimuli that are very different (this includes the case in which the elements of a pair are presented to different modalities), unity is never perceived, and true simultaneity, according to Fraisse, is difficult. In this case, the zone between simultaneity and succession may not be marked by a change in perceptual quality, as it is for within-modality events. The values of T at which these various perceptual boundaries are crossed are generally determined by the choice of elements. An exception is seen in the case of order identification; nearly any stimulus pair can be labeled in the same way (Hirsh and Sherrick, 1961).

Perception of simultaneity, succession, and finally order is reported as T increases. But the reports are not necessarily veridical with physical time; simultaneous elements may not produce the experience of simultaneity, and over a range of non-simultaneous values, T may be called simultaneous. Often, a subject presented with stimulus pairs for order labeling produces a monotonic labeling function whose central tendency is displaced from T=0 msec. It is inferred that in forced-choice order identification experiments, a stimulus perceived as simultaneous will be assigned each order equally often. Two parameters of the labeling function are usefully considered: the point of subjective simultaneity (PSS) and the difference limen (DL). The PSS is the median, or sometimes the mean, of the labeling function, the value of T for which the probability that each label is assigned equals 0.5. The DL is the quartile deviation of the function. The DL represents the amount of change from subjective simultaneity that is required to produce the pair that is given one of the labels on 75% of the trials. When the PSS differs from physical simultaneity, the DL is a better indicator of the precision of labeling than a statistic based on some level of correct responses (Sternberg and Knoll, 1973). The term precision is used to refer to the consistency of judgments, and thus is given by the DL. The PSS is a measure of the accuracy of the central tendency of those judgments.

The precision of labeling is related to resolving power, but as experiments generated by the Theory of Signal Detectability (Swets et al., 1961) have shown, observers can sometimes extract information from stimuli independent of the labels they would apply to those stimuli. In the fundamental detection experiment, for example, subjects are often surprised to learn that they perform above the chance level with weak signals. Asked to produce the threshold signal, they typically adjust to a slightly higher intensity than can be shown to be detectable at a performance level greater than chance in a forced-choice procedure. To estimate the limits of sensory capability, it is necessary to develop psychophysical procedures that eliminate or compensate for the effects of decision processes or criterion on performance. Estimates of temporal resolution deduced from the precision of identification are influenced by decision variables. Particularly clear examples of the effects of subjects' criteria arise in the study of the two-pulse threshold, which is discussed below. The second class of experiments attempts to measure temporal resolution with minimal reliance on the labeling of stimuli. In considering the distinction made here between experimental strategies, it is helpful to recall the difference between a sensation and a perception, proposed by the philosopher Thomas Reid (Boring, 1942). The sensations were said to be primary experiences that are immediately generated by stimuli. Sensations are available to introspection. A perception arises when the observer makes an inference about the source of the sensations. A perception, while simpler for the observer to have, is a more complex cognitive event, involving learning, expectation, and prediction as well as sensation. The goal in perception is to reduce all of sensory input to a manageable amount of information, and it is to be expected that the fine grain of sensation will be blurred as the decision about the source is made. This was made clear by Boring (1921) in his discussion of the stimulus error. When the subject directs his attention to the underlying sensation, he may preserve the small differences between events that would be lost if his task were to categorize the experience in terms of a small number of possible objects that could produce it. The same distinction is appropriate in the case of the dimension of time. There are experiments that elicit from the subject a report of some relation - a description of a percept. These are the labeling and identification experiments, which, it has been suggested, do not provide an unbiased estimate of temporal resolution. The second class of experiments measures discrimination directly. Ideally, they avoid the stimulus error, but it will be shown that this is not always the case.

Temporal resolution has, of course, been measured with experimental procedures of the second type. If the dimension of time is similar to other sensory continua, the resolution of a standard interval can be predicted from Weber's Law; that is, the just-detectable change in Ts, $\triangle T$, will be a constant proportion of Ts. There are a number of time discrimination experiments for which this prediction is reasonably well upheld. When Ts is defined by brief "marker" stimuli at the interval's beginning and its end (an empty interval), the Weber fraction, $\triangle T/Ts$, is found to be approximately 0.10 (Woodrow, 1951). For values of Ts that are large (greater than approximately 2000 msec) or very small, the fraction departs from the value of 0.10, but departures from Weber's Law are found in other sensory domains as well, when the standard stimuli approach the limits of the dynamic range of the sensory system. Most time discrimination experiments have used two identical pulses as marker stimuli. Most often, they have been delivered to the auditory modality. The auditory modality is believed to be the best equipped for resolving temporal relations (Hirsh, 1959; Julesz and Hirsh, 1972; Geldard, 1970). However, at large values of Ts, differences between modalities are less pronounced. The differences appear to be greatest when the subject's task is to resolve one pulse from two.

To summarize, two classes of experiment have been contrasted, one based on the subject's analysis of percepts, or response proclivities, and the other based on the analysis of a more basic level of stimulus representation, or sensory capabilities. At the present time, it is not possible to comment on the relation between the two kinds of data, for with few exceptions, experimenters have restricted their efforts to one class of investigation. In the auditory modality, both order labeling and temporal resolution have been well studied. The results of each type of experiment are reviewed in detail below. Larger stimulus changes are needed to change the label of temporal order assigned to a pair than are required for two stimulus pairs to be discriminable. Cross-modality stimulus pairs that differ in relative onset time have been studied from the standpoint of response proclivities. However, the understanding of simple temporal resolution, a sensory capability, is incomplete. The auditory example emphasizes that both types of experiments must be done.

The results of the experiments reported here describe cross-modality temporal resolution more completely. Estimates of temporal resolving power were obtained using cross-modality stimulus pairs and using several psychophysical procedures. The cross-modality stimulus is often used in temporal identification experiments, but it has rarely been used in systematic investigations of temporal resolving power. The discrimination procedures used proved to vary in their suitability; that is, greater resolving power is indicated by some of the procedures than by others. The differences between procedures are attributed to differences in the extent to which they induce the subject to categorize stimulus events. The interpretation of the results borrows heavily from the literature on categorical perception, although the experiments here were not conducted within the framework of the original demonstrations of that phenomenon (Liberman, Harris, Kinney, and Lane, 1961). When discrimination performance differs from predictions based on Weber's Law, the results are described in terms of the cues gained or ignored by subjects who might be responding according to their tendency to categorize the stimuli.

## III. HISTORICAL OVERVIEW

As was mentioned above, the first inquiries into the nature of cross-modality temporal processing were made by astronomers, who had a practical problem to solve (Boring, 1959, pp. 134-142). It was necessary for the astronomers to calibrate a clock by noting the exact time at which a star passed over a telescope cross-hair. The "eye and ear" method was typically used; the observer listened to the beat of a metronome and tried to estimate the fraction of an interval that had elapsed when the star just reached the designated position. The method was reliable and was presumed to be accurate as well until 1796 when, as every psychologist knows, Maskelyne, the astronomer at Greenwich, fired his assistant Kinnebrook. The reason for Kinnebrook's dismissal was the lack of agreement in the two men's observations of star crossings; Kinnebrook's judgments were found to be about one-half second later than Maskelyne's. In spite of Kinnebrook's attempts to bring his judgments into line, the discrepancy eventually increased to about 800 msec.

It was another astronomer, Bessel, who later wondered whether this discrepancy should be called an error. Over the course of several years, he had opportunities to compare the judgments of various astronomers, and he determined that two observers rarely recorded identical times. Thus was developed the personal equation, which was intended to allow one astronomer's observations to be scaled in terms of another's. With the introduction of better time-measuring equipment in the 1850's it became possible not only to compare two persons but also to determine the absolute error of each.

Though the existence of the personal equation ceased to be a problem for astronomy, it had captured the attention of experimental psychology. Experiments revealed that the equation was subject to a number of influences, including the brightness and rate of motion of the object being judged. These kinds of variables were better studied under simpler conditions, and so experimentation began with apparatus designed to produce visual and auditory stimuli with controllable T. Experiments involving stimulation of two modalities had been termed "complications" by Herbart in 1816 (Boring, 1959, p. 142), and one of the first usable devices was Wundt's "complication clock". The clock was used by von Tchisch in 1885 for controlled determinations of the conditions for perceptual simultaneity (Boring, 1959, p. 146). From these experiments and those involving the personal equation, it was soon concluded that temporal judgments were subject to two kinds of influence. There are unalterable physiological delays in each modality; latencies and conduction times introduce a delay between the time of occurrence of the stimulus and the time at which it is perceived. By unalterable, it is meant that once the stimulus parameters such as intensity are fixed, the delay in each channel will be relatively constant over successive presentations. The amount of delay differs across modalities. But there are also effects of attention or expectation. One of these is the prior entry effect. Early attempts to account for prior entry were ridiculed by William James (1890, p. 411), for some authors had described the phenomenon as if a percept could occur before its associated physical event. The prior entry effect was discussed by Titchener (1908, pp. 251-259) and was included as one of his laws of attention. He noted that when an observer's attention is directed toward one of two nearly-simultaneous events, or if his expectation is that one of the events will occur first, then the element for which the observer is predisposed will be first in perceptual time. The implications of this fact for the labeling of temporal relations are obvious, and the attentional factor cannot be overlooked in any analysis of temporal processing (Sternberg and Knoll, 1973). A clear demonstration of the magnitude of the prior entry effect in order labeling was given by Stone (1926). Her stimuli were paired clicks and mild shocks; the subjects' task was to report which event occurred first. Two sets of instructions were given, which induced the subjects to anticipate that one or the other event would in fact be leading. In each condition, the DL was constant for individual subjects. But the PSS was shifted such that at the PSS, the expected element actually trailed the other. The average shift was approximately 50 msec, which is nearly twice the DL.

Displacement of the PSS from physical simultaneity is consistent with the assumptions that physiological simultaneity (or order) at a central comparator is the necessary condition for perceptual simultaneity (or order) and that one modality transmits its information to the site for comparison more slowly. Reaction time experiments had revealed a difference between vison and audition that was assumed to be in the sensory channels, rather than in the motor component. Helmholtz estimated the speed of conduction in the tactile modality from the difference in reaction times to stimuli delivered to different parts of the body (Boring, 1959, p. 144). Prior to Helmholtz's experiment, most scientists had assumed that the conduction of nervous activity was essentially instantaneous. The variability in the PSS might also have been explained <u>post hoc</u> by the observation that, given the great difference in the rates of transmission of light and sound, physical events that are generated simultaneously will never arrive at the receptor organs simultaneously. The delay in the arrival of the acoustic information may vary over a range of tens of milliseconds, the range over that the PSS seems to vary. It can be hypothesized that mechanisms have evolved to compensate for these physical asynchronies. Such a mechanism could affect judgments of simultaneity or order. There are many explanations of the first type, but none of the second.

There were other order labeling experiments in the late nineteenth and early twentieth centuries. The results of these studies have in all respects been replicated with more sophisticated technology. The originals remain of interest, however, because of the conclusions drawn from them. Three experiments, by Exner (1875), Whipple (1899), and Smith (1933), have been selected because when considered as a group, they anticipated many of the contemporary findings. They showed that the PSS can vary across experiments, a result that still characterizes these kinds of studies. They first reported that certain procedural options, such as the use of repeated presentations, will increase the precision of order labeling. Finally, each of these authors explained his results in one of two ways. Some suggest that the PSS is fixed by temporal characteristics of the peripheral sensory systems; others postulate a central mechanism that operates independent of the modalities.

Exner (1875) conducted the first order labeling experiment. He found that the PSS was obtained when a visual preceded an auditory stimulus. Because of a very conservative psychophysical method, Exner obtained unduly large values for the DL (Hirsh and Sherrick, 1961). Exner was the first to interpret the displacement of the PSS from physical simultaneity in terms of latency differences; he proposed that auditory transmission delays were smaller than visual and that perceptual simultaneity was determined by simultaneous arrivals at some center in the brain.

Some time later, Whipple (1899) found the opposite displacement of the PSS from physical simultaneity. Since neural transmission times are similar across subjects, Whipple favored an explanation involving attention, or variance introduced at a central processing stage. Central processing, in this and later usages, refers to analysis that takes place beyond the primary sensory pathways. It typically refers to mechanisms that are as yet unidentified. Whipple's second significant finding was the increased precision of labeling when the observers were allowed repeated observations of the stimulus prior to making the decision about order. The DL was reduced by approximately one-half with repeated presentations.

Smith (1933), who also favored the central explanation for the displacement of the PSS, varied the intensities of visual and auditory stimuli in an attempt to manipulate their salience to the subjects. A brighter flash should be more compelling and should perhaps capture the attention of the subject; Smith would then predict that an intense flash would trail a click at the PSS. His results were consistent with the hypothesis. The predicted shifts were observed with intensity manipulations in each modality.

To summarize the study of the perception of temporal relations prior to 1959, the most obvious characteristic of the several studies was the disagreement as to the exact value of the PSS. After considering the eye-and-ear method and the personal equation, and the early order labeling results, it is still not possible to specify the physical conditions that will elicit a report of perceived simultaneity or succession or a reliable identification of order. In spite of this, some investigators (Exner, 1875, and others cited by Smith, 1933) contended that perceived temporal relations between events are determined by the temporal relations between volleys of afferent activity elicited by the stimulus elements (the latency-difference hypothesis). They theorized that the PSS is predictable from the difference in conduction times to the place in the brain at which the arrival times are compared. The alternative explanation (Whipple, 1899; Smith, 1933; Stone, 1926) is that afferent impulses may not necessarily be processed in the order in which they arrive, that attention or an imperfect decision mechanism may control the order in which events register.

It is important to keep in mind that the instability of the PSS may in part reflect the manner in which it is determined. Fraisse (1963, p. 109) suggests that dissimilar stimulus events, such as flashes and clicks, may never be experienced as occurring simultaneously. As Sternberg and Knoll (1973) have pointed out, the PSS is usually inferred from forced-choice labeling, and thus has never been identified as eliciting the experience of simultaneity. It may actually be the value of T that is most ambiguous without actually producing the integrated percept that Fraisse considers to be perceptual simultaneity. If the PSS is to be studied, it should be estimated directly. When the subject's task is to label the order of two events, conclusions about the perception of order can be drawn. Conclusions about perceived simultaneity must be made with caution.

The discrimination literature has been considered in less depth because there are few surprises in it. Woodrow's review notes that Weber's Law describes most of the data, with the value of $\Delta T/Ts$ generally about 0.10. There are slight differences in the resolution of intervals marked by pulses to one of the three major modalities. Goodfellow (1934) measured temporal resolution for a 1000-msec standard interval marked by pairs of auditory, tactile, or visual stimuli. He also compared estimates of the Weber fraction obtained using the three classical psychophysical procedures: the methods of adjustment, limits, and constant stimuli. He reported averaged Weber fractions of 0.07, 0.09, and 0.12 for the auditory, tactile, and visual cases. The Weber fractions did not differ greatly across methods, and the modalities were ordered the same way using each. As noted, there are no discrimination experiments that use the kinds of stimuli favored for the studies of perceptual phenomena.

## IV. REVIEW OF CONTEMPORARY LITERATURE: WITHIN-MODALITY TEMPORAL PHENOMENA

More recent reports of temporal processing are considered next. The first step will be to examine the characteristics of the auditory, visual, and tactile modalities in the time domain. Both physiological and psychophysical data will be considered. One aspect of sensory physiology has already been mentioned; the modalities differ with respect to transmission times. However, these three modalities also differ in the precision with which an onset time can be encoded in each. The variance in single neuron latencies at various levels of the projection pathways in each modality has rarely been considered in discussions of temporal resolution. Within each modality, the review considers the following topics: physiological delays; transmission time variance; simple reaction time; temporal acuity and two-pulse thresholds; apparent movement; and perceptual simultaneity and temporal order labeling.

### A. Physiological delays.

The characteristic delay between the onset time of a stimulus and the time at which its occurrence is represented in neural activity has been given great weight by some theorists. This delay will be called transmission time. By transmission time, it is meant the total time that elapses between the presentation of a stimulus and the time at which it has an effect at the neural location under consideration. Transmission time will include components that arise due to stimulus propagation velocity; receptor latency; synaptic delays; and fiber conduction velocities. Strong conclusions about transmission times to centers where temporal judgments are assumed to be mediated cannot be drawn from the kinds of data that are available. We have only tentative bases for drawing inferences about the locus of a temporal processing center. For example, the literature contains reports that aphasic patients are unable to make normal temporal order judgments (Efron, 1963c; Swisher and Hirsh, 1970) that would seem to imply that the left hemisphere is essential to the judgment. But these results must be interpreted with caution. They do not provide sufficient information to place a temporal organizing system in the part of the brain in which the speech areas are found. Since the location of a hypothetical center for temporal processing is not known, the data most often cited in latency difference arguments cannot reflect the transmission time to the appropriate cortical receiving area. It is not clear what additional delays are imposed as the train of impulses is reprocessed in secondary areas.

Typical transmission delays to primary projection areas are known, and the data support the contention that the visual system is "sluggish" (Pieron, 1952). In the human auditory system, Davis (1973) reports that a component of the evoked response with a latency of 16 msec may be of cortical origin. That value may be slightly high. In recordings from single neurons in the auditory cortex of cats, deRibaupierre, Goldstein, and Yeni-Komshian (1972) reported that some units responded to clicks with a latency of 9-11 msec.

The tactile system seems not to differ greatly from the auditory. Pagni (1967) recorded directly from the surface of somatosensory cortex in human patients undergoing surgery to control seizures and reports that the latency of cortical response to stimulation of the face is also about 9 msec. In the tactile system, the distance between the receptor and the brain must be considered; the same stimulus applied to the foot excites the cortex with a delay of 27-32 msec. The auditory system and the tactile system are similar with respect to latency, at least for tactile stimuli applied to the upper body.

In the visual system, the data for humans are less definitive, but greater delays seem certain. Latencies to the primary receiving area based on evoked potentials vary between 20-25 msec (Cobb and Dawson, 1960) and 35 msec (Vaughan and Hull, 1965). But McKay and Jeffreys (1973) indicate that recordings from electrodes on the cortical surface reveal the earliest component at 50 msec. That the longer times may be more correct is suggested by comparisons to data from the cat. Stone and Hoffman (1971) distinguish between two afferent transmission systems on the basis of their conduction velocities. The retina prior to the ganglion cell layer imposes a 30-50 msec delay, while the conduction time from ganglion cell to cortex is an additional 5 msec (X-system) or 10 msec (Y-system). Ikeda and Wright (1975b) confirm those values, reporting cortical latencies of 40 msec and 60 msec for the two systems. The human visual system is unlikely to be faster than the cat's.

In one experiment that compared the waveform of auditory and visual evoked potentials (EP's), it was reported that within a subject, the later components of the waveforms had similar appearance but that the auditory EP reliably occurred 10-15 msec earlier (Ciganek, 1965). Absolute latencies were not reported.

Generally, the visual system is the slowest of the three modalities, but it must be reemphasized that the physiological delays that are measured are delays to sites that are relatively early in the sequence of processing that eventually results in a perceptual judgment. The transmission time to the primary cortical receiving areas in each modality can be regarded only as a lower bound on the transmission time to areas in which interactions between modalities can occur.

B. Transmission time variance.

An aspect of physiological encoding that has been overlooked by psychophysically-oriented investigators concerns the reliability of latencies and conduction times. Emphasis has been placed on the average latency in an effort to account for such findings as the displacement of the PSS. A more successful venture might be an attempt to predict the consistency of performance in temporal discrimination experiments by considering the variablity with which relative onset times are recorded in the central nervous system. Variance measures are not always directly available, but they may be deduced from data that are available. When the discharge pattern of a neuron is locked to the phase of the stimulating waveform, it may be concluded that the temporal information in the stimulus is being preserved. Thus, the highest frequency for which phase-locking is observed provides an estimate of the neuron's temporal capability. When the response of cortical neurons is considered, there are more reports describing the capability of the visual system to follow the time waveform of the stimulus than there are for the auditory or tactile systems. This may reflect the popularity in visual psychophysics of measures of the critical fusion frequency (CFF), the frequency of amplitude modulation that is indistinguishable from a steady light. If a neuron responds in a phase-locked manner to a flickering stimulus, it can be concluded that the range of arrival times to that neuron is less than the period of the stimulating waveform. As a conservative estimate, the duration of one-half cycle of the CFF may be taken to indicate the time window within which the majority of impulses are generated. In the visual system, the CFF of single neurons decreases systematically as one ascends toward the cortex (van de Grind, Grusser, and Lunkenheimer, 1973). At the ganglion cell level, CFF is typically within the range from 20 to 40 Hz (Maffei, Cervetto, and Fiorentini, 1970; Fukada, Motokawa, Norton, and Tasaki, 1966; Fukada and Saito, 1971). At 40 Hz, one half-period is 12.5 msec. Levick (1973) reported the results of detailed measurements of the variability in the latency of the response of retinal ganglion cells to brief light pulses. He reported that the standard deviation of the onset of an identifiable response ranged from 2-6 msec. That result suggests that the half-period criterion provides a fair estimate of the interval within which most latencies will fall. An interval equal to the duration of one quarter-cycle of the CFF appears to estimate the standard deviation quite well. In the lateral geniculate nucleus (LGN), neurons still follow up to frequencies of 10 to 30 Hz (Spekreisje, van Norren, and van den Berg, 1971). These authors noted that the characteristics of psychophysically-measured temporal modulation transfer functions (de Lange, 1954) were already present in the response of LGN neurons. At the cortical level, these properties are no longer present in single neuron response; Ikeda and Wright (1975a) report that neurons classified as transient have the best response to time-varying stimulation, but that even they reach CFF at 5 Hz. Van de Grind et al. (1973) confirm this finding and add that in the secondary areas (18 and 19) of the cat cortex, there is no response to diffuse flicker, and that patterned stimuli are again followed only at frequencies lower than 5 Hz.

The response of neurons in primate somatosensory cortex has been described by Mountcastle, Talbot, Sakata, and Hyvarinen (1969). They reported that one population of neurons, called "thin-spike" units because of the waveform of the action potentials, responded with phase-locked discharges for vibration frequencies in excess of 100 Hz. Cortical neurons with regular spikes could not follow the same frequencies. All thalamo-cortical fibers responded in a phase-locked manner to high-frequency vibration. The authors compared the precision of phase-locking to psychophysically-measured frequency discrimination. They concluded that the coding was sufficiently precise to predict the psychophysical results under the assumption that the periods of the stimulus frequencies were being compared.

In auditory cortex, deRibaupierre, Goldstein, and Yeni-Komshian (1972) described the response of neurons to click stimulation. A subpopulation of neurons, called "lockers", responded to clicks with a short burst of spikes, at a fixed latency of 9-11 msec. The total range of variation in these latencies was on the order of 2-3 msec. The lockers responded to click trains at low repetition rates but were not excited by sinusoidal stimuli at the same frequencies. Thus they seem specialized for encoding temporal, rather than spectral information.

The modalities are ordered with respect to the consistency of transmission time in the same way as they are ordered with respect to the magnitude of transmission times. The former information seems more relevant to temporal discrimination, as Mountcastle et al. (1969) suggested. The absolute value of transmission time should be irrelevant to temporal resolution provided that it is relatively constant.

### C. Simple reaction time.

The reaction time experiment is one of the oldest in psychology and one that has often been used to make inferences about the time course of internal events. While the simple reaction experiment will not reveal the latency of a sensation, many investigators have used it to estimate the change in latency with changes in stimulus intensity

(Roufs, 1963). Also of interest here is the absolute value of reaction time (RT) in each modality. Woodworth (1938; Woodworth and Schlosberg, 1954) gives 180 msec as the generally-accepted RT for vision and 140 msec as the RT for touch and hearing. He also lists 167 and 105 msec as the "irreducible minimum" RT's for vision and audition. RT can be altered by manipulating the criterion for reacting, of course, but when criterion is held constant, the visual system is again suggested to be slower than the other senses. More recent determinations of RT in the visual and tactile modalities have emphasized what has been known almost from the first RT experiment, that RT decreases with increases in stimulus intensity. Using a weak tactile stimulus, Lele, Sinclair, and Weddell (1954) found that simple RT decreases by 60 msec for a factor of ten increase in applied force. In the visual system, various investigators (Raab and Fehrer, 1962; Roufs, 1963; Keitzman and Gillam, 1972) report similar results, about a 15-20 msec decrease in RT per log unit increase in stimulus intensity. Cross-modality comparisons of the effects of intensity are to be taken with caution, however; Woodworth notes that the outcome depends upon whether the stimuli were matched on physical or subjective scales. Raab, Fehrer, and Hershenson (1961) have found that visual RT is less affected by changes in brightness (manipulated by changes in duration) than by changes in intensity.

### D. Temporal acuity and two-pulse thresholds

Thus far, the time course of sensory events following the delivery of a single stimulus element has been considered. The visual system, it was found, takes longer to transmit information than the auditory and tactile systems, and its response to a pulse is spread out in time, relative to the other modalities. The latter characteristic will affect the two-pulse threshold, the smallest interval at which a pair of identical elements is discriminable from a single pulse. Some authors refer to this measure of temporal capability as temporal acuity (Pieron, 1952), since it is the analog in time of visual spatial acuity, the just-resolvable separation between two lines. If the modalities differed by only a characteristic delay, then two-pulse measures of acuity would not be different across the modalities. Since the visual system seems to lose temporal precision as well, a difference should be expected. It was noted above that in these measures, one finds excellent examples to illustrate the difference between tests of labeling and tests of discrimination.

Keitzman (1967; Keitzman and Sutton, 1968) has considered the distinction between the two-pulse threshold and the two-flash threshold. The difference is important; recall that pulses refer to stimuli and flashes to responses. Thus in the former case, the subject must resolve a change in the stimulus, independent of his labeling of it; in the latter, the T at which the percept is identified as resulting from two events is determined. The former is a discrimination experiment, and the subject may base his response on a brightness or a flicker cue, rather than a clear time difference. But the result is valid, since the temporal information in the stimulus is being transmitted. In the latter case, the observer himself must set a criterion for "twoness"; what is measured in this case is something more complex than resolving power. Keitzman measured labeling by presenting a single pair of pulses. The two-flash threshold was interpolated from the resulting labeling functions and found to be approximately 70 msec. The two-pulse threshold, in contrast, was determined in a three-alternative forced choice procedure. The non-signal alternatives were temporally contiguous elements (Ts=0 msec, measured from offset to onset), and the signal alternative contained elements separated by Tc msec. The thresholds from this procedure were only 20-30 msec.

When the observer is asked to commit the stimulus error, he may divide a continuum differently than he would if asked to indicate the smallest detectable change. The smallest discriminable temporal changes may not be perceived as changes in time. Thus a labeling experiment may underestimate resolving power.

Analogous experiments in the other modalities have been reported by Gescheider (1966,1967). In these single-interval experiments, the stimulus error may have occurred, although the instructions apparently were intended to relax the subjects' criteria for "twoness"; they were encouraged to use any available cue, such as a "rough quality" to guide their decisions. For tactile stimuli, 10 msec had to intervene between pulses; for auditory stimuli, 1.6 msec onset asynchrony was sufficient for identification. The possibility that the value of 10 msec for the tactile modality is a conservative estimate must be considered, given the psychophysical method employed.

Though only an approximate comparison between modalities can be made, the result seems consistent with the previous evidence, which suggested that the visual system is less acute.

The experiments reviewed above used as stimuli consecutive identical pulses delivered to the same receptor site. It was appropriate in those experiments to ask whether the source of stimulation was one element or two. At the small values of T, the subject's label indicated that only a single event was experienced. The next level of stimulus complexity is achieved by stimulating two different receptor sites. Recalling the classification of perceptual experiences given by Fraisse, in this case there will be no perception of unity, since the elements have individual identities. That is, the subject is aware that there are two elements; he cannot be asked whether one or two were presented if the perception of a single event never will occur. However, the distinction in labeling experiments may now be made between simultaneity and succession. The pair can be labeled, but a new set of labels must be used. Discrimination experiments may be conducted as before, using forced-choice procedures.

Sweet (1953) conducted several experiments of the first type, using visual stimuli and varying such factors as retinal eccentricity and spatial separation of elements. The method of limits was used to determine the values of T to which the labels of simultaneity or succession were applied. With one pulse presented foveally and the other peripherally, T for identifiable successiveness increased from 5 to 12 msec with increasing eccentricity. These values are measured from perceptual simultaneity, which was displaced from physical simultaneity. In a second experiment, spatially-contiguous pulses were presented at various retinal eccentricities. With this configuration, perceptual simultaneity always accompanied physical simultaneity, and once again the threshold for succession increased with eccentricity, to a maximum of 35 msec. When subjects were advised that a cue in the form of apparent movement might be available, the threshold was reduced to 5 msec in all cases. This last finding emphasizes once again that temporal resolution cannot be estimated from a labeling procedure, because of the confounding introduced by the subject's criterion. The subject is likely to be able to resolve changes that he believes do not place the stimulus in the next perceptual class.

There are a number of reports of experiments that measure the discriminability of stimulus pairs for which a unique perceptual quality emerges from changes in T. A naive subject will not be aware that changes in time are responsible for the perceptual differences, but since time is the only variable, it must in these cases be acknowledged that the temporal information is preserved in the sensory system.

A number of experiments using acoustic stimuli have established that T=2 msec is sufficient to distinguish between certain kinds of paired stimuli. Green (1971) summarizes a series of experiments in which temporal acuity was measured independent of spectral acuity. Brief acoustic stimuli are broadband signals, and often changes in time result in discriminable spectral changes. Leshowitz (1971) had found that the two-pulse threshold experiment for audition is confounded by the accompanying spectral changes. Thresholds determined in a two-alternative forced choice procedure were approximately 10 microsec. But control procedures established that changes in the energy distribution in the 8-10 kHz band carried the information required to perform the task. The stimuli selected by Green were Huffman sequences, a class of short-duration waveforms which differ in the time domain but can be equated for bandwidth. Acuity was measured for pairs of Huffman waveforms with a phase delay in a frequency band of one element. Discrimination based on the delay was possible with 2-msec pulses. Wier and Green (1975) found the discriminability of waveforms consisting of two sinusoidal components differing in their order was still high when total waveform duration was only 2 msec. These waveforms also have identical spectra, since the same components appear in each stimulus.

Efron (1973; Yund and Efron, 1974) has studied the discriminability of "micropatterns" in three modalities. A micropattern consists of a pair of elements presented simultaneously; the elements are chosen such that a unique perceptual quality is associated with the micropattern. Visual micropatterns, for example, were formed from mixtures of monochromatic light (510 and 620 nm) that had a neutral hue. With changes in Ts, the pattern took on the hue of the trailing element, enabling the discrimination to be made. Auditory and tactile standard micropatterns were perceived as chords; changing Ts produced a pitch change. Discrimination varied with $\Delta F$ in the auditory experiment;

at $\Delta F=160$ Hz, a 2-msec $\Delta T$ was discriminable. For the same $\Delta F$, a 20-msec $\Delta T$ was discriminable in the tactile modality. In vision, $\Delta T=12$ msec was required for discriminability. These thresholds are not directly comparable, in the absence of a rationale for equating differences in frequency (acoustic and tactile elements) to differences in wavelength (photic elements). Yund and Efron (1974) extended the experiment to include dichoptic and dichotic presentation of elements. Performance in these conditions was slightly poorer than in monocular and monaural controls.

### E. Apparent movement

The next group of experiments deals with an illusion, apparent movement, that is dependent upon the temporal relation between two stimulus elements and that has been reported in each of the three modalities studied in the current experiments. It has been suggested that the determinants of apparent movement are central rather than in the sensory projection systems (Geldard, 1970; Sherrick and Rogers, 1966) and that there may be a link between these phenomena and the perception of ordering (Hirsh and Sherrick, 1961; Westheimer and McKee, 1977).

Sherrick and Rogers (1966) compared their results with apparent haptic movement to those of Neuhaus (1930) for visual apparent movement. The stimuli were sinusoidal bursts from two vibrators; element duration was an independent variable in each experiment. Each investigation found that the optimal T for perceived movement increased with increases in element duration. Three increasing functions of duration were shown by Sherrick and Rogers. The parameter was the report given by the observers: simultaneity, apparent movement, or order. For a given duration, small values of T were called simultaneous; larger T elicited the illusion; and the larger values were identified as ordered. The precision of each judgment decreases with duration. The similarity of results between the tactile and visual modalities is striking; for durations up to 200 msec, it is impossible to distinguish between the two sets of data. The agreement is found for both vibrotactile and electrocutaneous stimuli. The optimal T for perceived movement is a non-monotonic function of duration, with the minimum T, 75 msec, appearing when duration equals 25 msec. Smaller values of T are reported to be resolvable under other experimental conditions.

Auditory apparent movement can be elicited, but smaller intervals are effective in that modality (Briggs and Perrott, 1972; Perrott, 1972). Also, the optimal T is a monotonic function of the duration of the stimulus components. Stimulus elements were noise bursts, presented dichotically. Movement was perceived at a slightly smaller T if the noise sources were independent. One difference between these experiments and those of Sherrick and Rogers (1966) and Neuhaus (1930) is in the nature of the stimulus elements. The latter two used narrow-band stimuli, sinusoidal vibrotactile bursts or steady pulses of light. Visual stimuli may be characterized in the frequency domain by the Fourier transform of the envelope of luminance changes (deLange, 1954; Kelly, 1961). Simmons (1973) has shown that echolocation, which is limited by temporal resolution, may be modeled in terms of the autocorrelation function of the orientation sound. Broadband signals are therefore better suited for resolving temporal intervals. This line of thinking has not been applied to human temporal resolution, although the apparent movement data would not be inconsistent with predictions based on stimulus bandwidth. In each modality, the range of T that elicits the percept increases with increases in duration or decreases in signal bandwidth. In addition, the shortest effective range of T was found in audition using broadband noise as the stimulus.

### F. Labeling of simultaneity and temporal order.

The last topic for which within-modality experiments are reviewed is perceived simultaneity and order. These experiments often use the single-interval labeling procedure mentioned above. Two discriminably-different elements are presented, and the subject judges their temporal relation. The elements may be qualitatively different, in which case they may be presented at the same place, or they may be identical elements separated in space. When perceived simultaneity is under investigation, the PSS is the statistic of interest. The stimulus requirements for perceived simultaneity can also be determined directly.

The first contemporary labeling experiments were done by Hirsh (1959) using auditory stimuli. A total of five experiments were reported, testing various combinations of clicks, noise bands, and pure tones. Within each experiment, frequency, duration, and level were varied. In all cases, stimuli were presented monaurally in earphones, and listeners were permitted multiple observations of the pair before reporting its order. In every case, the labeling functions could be approximated by the same straight line (in normal-linear coordinates) such that the PSS=0 msec and the DL=17-20 msec. Only for some combinations of clicks were slightly steeper functions seen. Hirsh noted the insensitivity of the DL to stimulus features and the large difference between the DL for order and the report by others that a T of 2 msec is adequate to distinguish two events from one. His suggestion was that a central processor or a time-organizing system would be required to account for the finding.

Two years later, Hirsh and Sherrick (1961) reported additional experiments with auditory perception of order and added experiments with visual and tactile stimuli as well (their cross-modality experiments are discussed below). In the auditory experiments, damped sinusoidal pulses were presented dichotically. In one case, identical tones were led to each ear; in another, different frequencies were used. The DL was again about 20 msec, and the entire psychometric function was reasonably well fit by the same line.

For visual stimuli, Hirsh and Sherrick used 5-msec pulses; parameters included the spatial separation of the elements, the orientation of the array, and retinal eccentricity. The results for thse conditions were also summarized by the line from the 1959 auditory experiments. For order-labeling, the performance of the visual system is no worse than that of the auditory system.

Filtered tactile pulses from Goodmans V-47 vibrators were delivered to fingertips to compare the somatosensory system to the other modalities. No difference in labeling performance was seen; the DL remained at approximately 20 msec. As in the other experiments, the PSS was not displaced from physical simultaneity. Feedback was not given in these experiments, but the subjects were allowed to observe as many pairs as necessary prior to responding. The PSS is believed to be sensitive to procedures, as is discussed below (Gengel and Hirsh, 1970). Sternberg and

Knoll (1973) suggest that it may tend to be placed in the middle of the stimulus range.

Other experiments involving the labeling of the order of two stimuli have generally upheld Hirsh's findings. The value of 20 msec for the DL is frequently cited. The PSS is much less stable, however. The order labeling procedure in later studies was used primarily to track changes in the PSS under different stimulus conditions.

Rutschmann (1966; 1973; Gibbon and Rutschmann, 1969) has applied the order-labeling procedure to the estimation of visual latencies. Recall that from the time of the first experiments (Exner, 1875), one explanation for shifts in the PSS has been in terms of the relative physiological delays in the two stimulus channels. In Rutschmann's view, perceptual simultaneity occurs when those delays are just offset by T. In this interpretation, central factors are considered to be negligible; the decision mechanism introduces no variance (Sternberg and Knoll, 1973). In the labeling functions shown in her various publications, the slope is similar to Hirsh and Sherrick's, but retinal position influences the PSS in a systematic way. When one pulse is presented foveally, a peripherally-presented pulse must precede it at the PSS. Measures of simple RT to the same stimulus elements lead to the same conclusion, that the transmission delays associated with the peripheral retina are longer. The PSS can also be shifted by intensity changes, in the direction expected if higher intensities reduce latency (Rutschmann, 1973). These conclusions are in agreement with those of Sweet (1953) and Roufs (1963).

Efron (1963a, b) investigated the perception of temporal order and simultaneity, under the assumption that the relevant processing occurs in the dominant hemisphere and that simultaneity of neural arrivals at that center determines perceptual simultaneity. In addition to visual stimuli, Efron also used brief shocks. The stimulus elements were presented such that the initial neural representation of each was confined to one hemisphere: right or left fingertip, or right or left temporal retina. In each modality, it was found that a stimulus relayed to the non-dominant hemisphere had to precede one relayed to the dominant hemisphere at the PSS. Increasing intensity to decrease transmission times had the predicted effect on the PSS. It shifted back toward physical simultaneity if the leading element was made more intense.

Corwin and Boynton (1968) further tested the notion that there is a single center that mediates the perception of temporal relations by examining the transitivity of simultaneity for various visual stimuli. Latency was manipulated by stimulating foveally or at one of two peripheral locations. It was found that the PSS determined for two pairs would predict the third. The result was regarded as supporting the arguments for a common center, as Efron had proposed. However, Sternberg and Knoll (1973) have clarified the implications of transitivity for the existence of such a center. They note that a single center is neither a necessary or a sufficient condition for transitivity. An arrangement of several centers that compensate for latency differences can be envisioned; in addition, transitivity of arrival latencies could exist but not be preserved by a decision mechanism at a single center.

A slightly different experimental manipulation was shown to shift the PSS by Matteson (1970). It was found that the addition of an annular surround to one of the elements in a visual stimulus pair had the effect of moving that element forward in time. For example, if the PSS had ocurred at physical simultaneity without the surround, it would occur with the center-surround field trailing in the test condition. It was suggested that the effect of the surround is to reduce background noise, which defines the onset of the event more sharply. The effect on the PSS is similar to that obtained by increasing intensity of one element, although in the presence of the surround, brightness decreases.

Oatley, Robertson, and Scanlon (1969) used a modified order-labeing procedure, and the authors proposed a different interpretation of the results than is found elsewhere. The method of presentation is reminiscent of Stevens' (1961) quantal method. A fixed value of T was used throughout a block of trials; the subjects labeled each presentation. This should not be confused with the repeated-presentation method used by Hirsh (1959) in which T varies within the block. The stimulus elements were visual, but only one order was tested. The labeling function had slope of approximately zero for T less than 40 msec, then increased monotonically as T increased further. The authors attributed the flat portion of the function to the existence of a psychological moment, which was triggered by the leading event. Events occurring within a moment cannot be perceived as ordered. Moment theories are reviewed in more detail below. This version is unique in postulating that the moment must be triggered by a stimulus element. Other such theories assume that the moments elapse independent of external events. With respect to the method used in this experiment, it has been noted by Barlow (1961) that the quantal method advocated by Stevens may bias the psychometric function toward the shape predicted by the quantum theory. Subjects develop a set for responding after several presentations of a stimulus that they know is always the same. Thus if the first few presentations of a small T were labeled as simultaneous, the observer may tend to perseverate in that response on the later trials. This tendency could produce the flattened psychometric functions shown by Oatley et al. (1969).

Auditory labeling of order was examined further in two experiments (Babkoff and Sutton, 1963; Babkoff, 1975). The case of dichotic clicks was studied, since Hirsh (1959) had reported that the labeling function for clicks was slightly steeper than for other stimuli. Babkoff and Sutton (1963) reported that clicks of identical waveform interacted at interaural T between 2-20 msec, producing a cue in the form of a loudness difference. When subjects were advised that the louder-sounding click was the earlier-occurring event, their labeling functions did steepen. Babkoff (1975) examined the labeling of T for values small enough to produce a lateralized image. Not surprisingly, labeling functions are strongly non-monotonic when those values are included. Labeling performance is perfect for T less than 2 msec, decreases as T increases slightly, and returns to perfect as T is increased to 100 msec. Lateralization is mediated by specialized neural mechanisms, and as Sternberg and Knoll (1973) have pointed out, within-modality temporal judgments may be influenced by peripheral interactions among the stimulus elements. They propose that cross-modality pairs be used if the central component is to be studied.

A tactile order-labeling experiment was conducted by Halliday and Mingay (1964). As were several of the visual labeling experiments, this study was concerned with the effects of transmission time on the PSS. The authors had noted that the latency of evoked potentials following stimulation of the toe was 20 msec greater than that evoked by stimulation of the finger. The authors asked whether this difference would be compensated for in order-labeling, or whether the same 20-msec difference would be reflected in finger-toe simultaneity. For two subjects, the PSS occurred with the more-distant stimulus leading by 9 and 17 msec. In control judgments of order for stimuli presented to fingertips on the two hands, the PSS was nearly zero. The slight displacement that was seen was interpreted as support for Efron's hypothesis that comparison occurs in the dominant hemisphere. Although the PSS was not displaced by the full 20 msec, the authors concluded that compensation did not occur, that order judgment reflects the order of arrival of afferent volleys. They commented that for motor tasks, compensation for conduction delays does occur; a subject instructed to move a finger and a toe simultaneously can do so. Apparently, this was an informal observation for which no data were presented.

Variations in the PSS across experiments or across conditions within experiments are common. The other parameter of the labeling function, the DL, seems to be nearly constant. The value obtained by Hirsh, 20 msec, has been widely cited and is not observed to change with the

PSS. There are a few cases in which comparable values have not been found. Of the three studies that do propose smaller values, one could not be replicated. Robinson (1967) presented data to indicate that the order of visual forms, a square and a triangle, presented dichoptically to homologous retinal locations could be labeled perfectly when T was only 5 msec. Thor (1967) attempted an exact replication of Robinson's experiment but could not demonstrate the same high level of performance. Recently, Westheimer and McKee (1977) obtained order judgments for pairs of lines, presented at various orientations. They relied heavily on the experience of apparent movement to guide the judgments. The DL was as low as 3 msec in the best cases. Binocular presentation yielded lower estimates of the DL than did dichoptic presentation. In the latter condition, the DL was 10-15 msec.

Patterson and Green (1970), using auditory stimuli, reexamined order-labeling in that modality and concluded that labeling could be as precise as temporal acuity measures would predict. However, in these results, the cue may have been spectral, rather than temporal. It is also not likely that naive subjects would spontaneously describe the difference between pairs as being an order difference. A subject can be taught to apply two arbitrary labels to any pair of stimuli that are discriminably different. Discrimination procedures require that at least minimal labels be used. However, the goal of experiments such as Hirsh's is to determine the way in which subjects can apply a non-arbitrary set of labels. The Patterson and Green experiment is not a study of order-labeling but a discrimination experiment.

Three additional experiments involving the labeling of paired auditory stimuli differing in onset time are discussed here; each of these takes the further step of comparing labeling to discrimination. These studies present examples of categorical perception (Liberman, Cooper, Shankweiler, and Studdert-Kennedy, 1967), a hypothesized mode of sensory processing that makes different predictions about the discriminablity of pairs of stimuli than does Weber's Law. The latter predicts that the just-discriminable change will be a constant proportion of the standard stimulus. For categorical stimulus dimensions, discrimination depends upon the relative distance of the stimuli from the categorical boundary, the stimulus value at which the assigned label changes. Three requirements for a categorically-perceived stimulus dimension were set down by Studdert-Kennedy, Liberman, Harris, and Cooper (1970): 1) There must be regions of the stimulus range within which the discrimination of stimulus pairs of equal physical difference is enhanced; that is, peaks in the discrimination function must be seen. 2) There must be regions of the stimulus range within which the discrimination of stimulus pairs of equal physical difference is at or near chance; that is, troughs in the discrimination function must be seen. 3) The peaks and troughs must bear a characteristic relation to the labeling function. Peaks must occur when the pairs to be discriminated are taken from two sides of the identification boundary; troughs must occur when the boundary does not fall between the pairs. Liberman, Harris, Kinney, and Lane (1961) held that this mode of perceiving was unique to speech sounds and presented data to support the claim. The stimuli were synthetic consonant-vowel (CV) syllables, differing in voice onset time (VOT). The sounds were generated by the Pattern Playback, a device that converts a sound spectrogram into a sound. A set of synthetic consonant-vowel syllables with VOT varied in 10-msec steps was used. The syllables were labeled by subjects, with the break occuring when VOT was 20-30 msec. At the boundary, the label changed from /do/ to /to/. Pairs of the same stimuli with differences in VOT of 10, 20 or 30 msec were then presented for pairwise discrimination. In most cases, the predicted discrimination performance was obtained. Performance for a given size of change was greatest when the pairs presented had been given different labels in the first experiment. To test the hypothesis that non-speech stimuli would not be perceived in the same manner, the authors measured the discriminability of the pairs of sounds produced by playing the inverted spectrograms. These were rarely discriminable at a level above chance. However, the authors did not obtain labeling data for these sounds; the possibility that they were not labeled categorically was not considered.

More recent evidence indicates that non-speech stimuli can also be perceived categorically. Lane (1965) reviewed the experiments using synthesized speech and concluded that the results were similar to what is found when other classes of stimuli are investigated under the same conditions. Miller, Wier, Pastore, Kelly, and Dooling (1976) have presented data that indicate that a set of stimuli differing in onset time formed by pairing a noise burst with a 100-Hz "buzz" will satisfy the requirements listed by Studdert-Kennedy et al. (1970). They can be divided into two categories, using a criterion such as "no noise" and "noise"; the boundary obtained probably reflects the onset time difference at which the noise is just masked by the buzz. The discriminability of noise-buzz pairs can also be predicted from labeling functions, just as was done by Liberman et al. (1961) for synthesized speech-sound pairs. Miller et al. (1976) suggest that perceptual categories arise not because the stimulus is one for which special mechanisms exist, but because the stimulus is a complex combination of elements for which some new perceptual quality emerges at the categorical boundary. In their experiment, that quality was the sudden detectability of the noise at the appropriate T. Enhanced discriminability merely reflects the ability to indicate the presence or absence of the new quality. They also state that the possibility of demonstrating categorical perception can be increased by certain aspects of experimental procedure. For example, subjects can be induced to compare labels, i.e., to commit the stimulus error, rather than to focus on the underlying sensations.

Pisoni (1977) has shown that the familiar temporal order labeling experiment can be used to demonstrate categorical perception of non-speech sounds. In one experiment, subjects labeled the order of paired 500 and 1500 Hz tones. The labeling was similar to the function obtained by Hirsh (1959), with the exception that the PSS, or categorical boundary as it is called in this case, was displaced from zero. When listeners made ABX discriminations of the same stimuli, the functions relating performance to T peaked at the boundary and fell to chance for within-category pairs, as predicted. In a second experiment, the listeners were encouraged to use three labels: two orders plus simultaneous. Three categories will have two categorical boundaries. These subjects were not tested for discrimination, but their ability to use three labels was suggested to be related to a pattern of discrimination observed before, in which two peaks in the discrimination function were seen. The implication is that categorical boundaries are not rigidly fixed; subjects can be induced to change their mode of responding with different instructions or with the availability of different strategies.

### G. Within-modality temporal resolution.

In this section, measures of the general capability for detecting differences in intervals of time are reviewed. Almost all of these experiments use auditory stimuli. The review is restricted to the discrimination of "empty intervals", those marked with a pulse at the beginning and the end. Temporal resolution can also be measured for "filled intervals", such as continuous tones, but since the current set of experiments used cross-modality stimuli, the empty-interval is assumed to be the most similar case for comparison. There is not a large difference between filled and empty interval discrimination within modalities. The resolution of empty intervals can be measured using any of the standard psychophysical procedures for determining sensory capabilities. Some experiments reviewed will be described as "gap discrimination" tasks. In these, the subjects must resolve a silent interval between the elements of the standard stimulus. The just-discriminable gap is usually measured from the offset of the leading element to the onset of the trailing element.

Abel (1972b) measured the detection of changes in gaps between noise-burst markers; the standard gaps, (Ts), ranged from 0.63 msec to 640 msec, measured from the offset of the leading element to the onset of the trailing element. The markers were either 10 or 300 msec in duration; two levels of 10-msec markers were used. The just-discriminable increase, $\triangle T$, in the standard gap increased as Ts increased. For very brief Ts, $\triangle T$ was approximately 2 msec. For larger Ts, the increase in $\triangle T$ was proportional to the square root of Ts. This result was said to be more consistent with the Poisson counter model proposed by Creelman (1962) than with Weber's Law. However, Abel's estimates of $\lambda$, the parameter of the Poisson, differed greatly from Creelman's. Just-discriminable changes were slightly larger for long or low-level markers.

Getty (1975) described the results of measurements of temporal resolving power for intervals bounded by clicks; Ts in this experiment varied from 50-3200 msec. For Ts between 500 and 2000 msec, the just-discriminable change was approximately 0.06 of Ts for each of two observers. For Ts outside of this range, the Weber fraction was larger. Getty compared the results to predictions based on the Creelman model and to predictions based on Weber's Law; the latter prediction more nearly approximated the data. For the conditions of Getty's experiment that are most similar to Abel's (1972b), he reports much smaller changes to be discriminable. For example, Ts of 100 msec was discriminable from Tc of 105 mec according to Getty; the Weber fraction is 0.05. In Abel's data, Ts of 80 msec is discriminable from Tc of 97 msec; in this experiment, the Weber fraction is 0.21. Both experiments employed the two-alternative forced-choice procedure. There are two differences between the experiments. Getty's marker stimuli were clicks of very short duration (100 sec), while Abel used longer noise bursts. Another difference is that Abel's experiment requires the discrimination to be based on offset to onset time, while Getty's subjects were resolving a Ts marked from onset to onset. If the decay of the internal representation of the marker were slower than its rise (see Green, 1971, for a discussion of such a system), then the results of the two experiments would be expected to differ.

Divenyi and Danner (1977) used the same procedure to measure resolution for 25, 80, and 320 msec Ts. Their results are in agreement with Weber's Law and with Getty; $\Delta$T/Ts was approximately 0.06. For Ts=80 msec, Divenyi and Danner find that $\Delta$T of approximately 5 msec is resolvable. For some conditions in which the markers were of different frequencies, the smallest Ts was less well resolved, however. Large frequency differences between the elements in a pair increased the just-discriminable change. They modified the Poisson model in interpreting their results. Parameters estimated from these data are different from

Abel's and Creelman's estimates.

Williams and Perrott (1972) determined the just-detectable gap between two tones, centered around 1 kHz, as a function of the tone duration, D, and the frequency separation between them. This experiment is similar to a two-pulse threshold determination. In addition, there are conditions that are close to conditions described by Abel (1972b); the detectable gap of approximately 5 msec given 10 msec tones and no frequency difference may be compared to the values of 2-3 msec found with noise bursts as elements. For D=300 msec, tones must be separated by 10 msec, measured from offset to onset, and noise bursts by 3 msec. In all cases, larger gaps are required for detection when a frequency difference exists between the marker stimuli. This result is consistent with Divenyi and Danner's finding, although their standard intervals were longer than the interval used here.

For the visual modality, Nilsson (1969) measured the just-resolvable change in Ts for Ts=0 to 90 msec, using a three-alternative forced choice procedure. One-msec pulses from a single light source bounded the intervals. The smallest detectable change was approximately 10 msec, when Ts=45 msec. For Ts smaller or larger than 45 msec, $\Delta T$ increased; the slope of the psychophysical function, the plot of the just-resolvable change as a function of the standard interval, was approximately 0.2 for Ts greater than 45 msec. The just-discriminable change at Ts=75 msec is approximately 20 msec. For auditory discrimination experiments, comparable values were 5 msec (Getty, 1975; Divenyi and Danner, 1977) or 17 msec (Abel, 1972b).

### H. Summary

This concludes the review of within-modality temporal phenomena. In summary, the modalities differ in the time required to transmit afferent information and in the variation in the arrival time of the activity elicited by a pulse or periodic signal. The visual system consistently imposes a longer delay, as indicated by the latencies of evoked potentials and single neuron discharges, and by simple reaction times. In each modality, increases in stimulus intensity will decrease the transmission delay. The visual system cannot follow rapid envelope changes, as indicated by two-pulse thresholds; the tactile and auditory modalities resolve successive events more precisely. This difference was anticipated in the discussion of the variance in arrival times. Tasks for which the differences between modalities are minimized have been inferred to be mediated centrally. Illusory movement and the labeling of temporal order require similar stimulus relations in each modality and have been suggested to involve higher-order processing. These phenomena also require larger temporal separations than do the fairly simple discrimination tasks.

In Section II, a distinction was made between two classes of experiment. One class measures response proclivities (Watson, 1973) or aspects of perceptual experience elicited by a stimulus pair. The second class estimates the limits of sensory capabilities, in this case the resolution of a temporal interval, independent of how that interval might be described. In two cases discussed in this section, it was seen that the first type of experiment does not provide an estimate of the same capability for temporal resolution that the pure discrimination experiment does. In one case, a subject can discriminate two pulses of light from one when the difference in onset times is approximately one-quarter of the difference required to elicit the verbal report of "twoness" (Keitzman, 1967). In another case, comparisons across the two experimental strategies reveal a order-of-magnitude difference in the change in the temporal relation between the elements of an auditory stimulus pair required for threshold-level discrimination (Getty, 1977) or for threshold-level identification of perceived order (Hirsh, 1959). The use of the first type of procedure is valid when the nature of a perceptual experience is to be studied; in fact, it is mandatory. But as indicated, these same procedures will not provide an unbiased estimate of resolving power.

## V. CROSS-MODALITY TEMPORAL PHENOMENA

This section reviews experiments that are more directly relevant to the work reported in later sections. These examine the perception of simultaneity and of temporal order, and temporal resolution when the elements of the stimulus pair excite different sensory modalities. Such experiments are relevant to hypotheses about central processors. If such general-purpose temporal organizing centers exist, then cross-modality stimuli should be used to study the central components in the absence of confounding peripheral interactions (Sternberg and Knoll, 1973).

### A. Labeling experiments

The appropriate starting point is again the paper by Hirsh and Sherrick (1961). The final experiment in that series asked for identification of order in auditory-tactile, auditory-visual, and visual-tactile stimulus pairs. Procedures were as described in Section IV. As with within-modality pairs, the finding was that the PSS conformed to physical simultaneity and that the DL was approximately 20 msec. In summarizing the results of this and the other experiments, the authors expressed the view that such uniform results could not be accounted for by factors operating within the modalities, given the differences in other measures of temporal capability. They postulated the existence of a central organizing process that was accessible by events originating in any modality.

As was noted earlier, the specific procedures used by Hirsh (1959; Hirsh and Sherick, 1961) may account for some aspects of his results, especially the consistent agreement between physical and subjective simultaneity. Gengel and Hirsh (1970) sought to clarify this issue by determining the effect on the parameters of labeling of several variables, including the number of presentations of a click-flash pair to be labeled, the use of feedback, and the amount of training. As had been suspected, the precision of labeling was increased by the use of repeated presentations; the asymptotic DL for single presentations was 30 msec, compared to the value of 15 msec obtained with repeated presentations. Feedback about the actual physical order of the pair surprisingly did not increase the precision of labeling. It did tend to make the PSS agree with physical simultaneity. Even though subjects reported that they were initially aware of a lack of correspondence between their percepts and the physically-correct order, they learned to report the actual order without loss of consistency. Practice, as expected, decreases the DL; it appears to stabilize after 9-16 sessions. With single presentations, the DL after the first session is 56 msec, replicating the earlier study by Hirsh and Fraisse (1964), which had used untrained observers.

Rutschmann and Link (1964) obtained temporal order judgments for pairs whose elements were light pulses and acoustic noise bursts. It was expected that the PSS might be predictable from RT data; the prediction, however, was not confirmed. They were unable to show that the difference in RT, which was 45 msec with auditory RT faster, entailed a PSS of 45 msec with the visual element leading. The PSS for each of two subjects was attained with the noise burst leading. They chose not to abandon the idea that physiological simultaneity at some point in the nervous system determines perceptual simultaneity, however. It was concluded that unequal delays must be interposed between the neural substrates that mediate RT and temporal order determinations.

Roufs (1963) obtained direct measures of auditory-visual simultaneity using the method of adjustment. This experiment was mentioned earlier in the discussion of the relation between RT and stimulus intensity. Observers were asked to set the onset of a 1-kHz tone to be simultaneous with the onset of the visual stimulus, which could be varied in intensity. The mean of simultaneity adjustments was shifted by the intensity manipulation, according to the same trading ratio obtained in the RT phase of the experiment. For two subjects, the mean adjustment had the visual element leading by 20 msec when it was presented at its maximum level. The PSS shifted by approximately 15 msec per log unit decrease in intensity; the light had to be advanced by over 50 msec to achieve perceptual simultaneity at the minimum intensity used. Roufs reports that changing the intensity of the tone had no similar effect, although the tone's level and the range of variation were not provided. The standard deviation of the simultaneity judgments was given as 13 msec. Roufs notes that day-to-day variation in the PSS was common.

Simultaneity judgments were used as a measuring device by Haber and Standing (1970), in this case to estimate the duration of a percept. Their subjects were instructed to make simultaneous the onsets of a click and a letter display. On later trials, they were instructed to set the click at the offset of the display. The visual display led at the PSS of onsets, by 33 and 26 msec for two subjects. The variance in the adjustments was not discussed, but the judgments were reported to range over approximately 40 msec. The assumption was made that the criterion for simultaneity of click and onset was equivalent to that for click and offset; thus, the difference between PSS's would estimate the duration of the percept. For displays lasting more than 300 msec, the duration of the percept was approximately equal to the duration of the physical stimulus. For displays lasting less than 300 msec, perceptual duration exceeded physical duration.

Efron (1970; 1973b) used similar reasoning to measure the duration of sensory events in three modalities. He determined the PSS using the method of limits, rather than the method of adjustment. In each modality, an effect similar to the one reported by Haber and Standing (1970) was observed. The duration of the perceptual event, estimated from the difference in PSS to onset and offset, was equal to the duration of the physical element, for durations greater than 130 msec. For elements whose durations are equal to or less than 130 msec, the duration of the percept seems to be 130 msec always. It should be noted that Penner (1975) and Allan (1976) have raised questions about the assumption that underlies this use of the PSS. The issue concerns the criterion for simultaneity of offset. Penner (1975) and Allan (1976) found that not all subjects produce data of the form described by Efron. Some subjects' judgments provide no evidence that perceptual duration ever exceeds physical duration. Penner (1976) suggests that both patterns of performance can be understood in terms of the subject's choice of a criterion for offset. The effect of the stimulus is assumed to decay gradually after the physical stimulus has ended. The subject must decide at which point the trace has dropped to the level he considers to be offset. Various criterion placements will produce the two types of result seen by Penner. In this review, the precision of the judgment is of primary interest, however, and that aspect of the judgments seems relatively tolerant of criterion fluctuation. Efron (1973b) provides some of the data from which the PSS was determined for two subjects. For one who matched a click to a vibratory standard, the standard deviations of 20-trial blocks were approximately 8 msec. Block means appeared to fluctuate, suggesting that criterion for simultaneity was unstable. For the second subject, whose matches were between a flash and the vibratory standard, block standard deviations were slightly larger, 10 msec. However, this subject was more consistent across blocks; the standard deviation of the block PSS was 7 msec, compared to 16 for the first subject.

Most of the investigators who have studied perceptual simultaneity have commented on the instability of the PSS. Roufs and Efron both suggest that criterion fluctuations can shift the PSS between sessions. One possibility is that the slow rise time of a visual sensation interferes with the ability to specify its onset precisely. But Efron's first subject is characterized by more fluctuation than the subject who made matches using a flash. Unfortunately, these subjects served in only one condition each; the effect due to individual differences cannot be accounted for. These observations bring to mind the early discussion of the prior entry effect and the role of attentional factors in general. In some reports, it is stated that a certain procedure was used to determine the PSS, because the data obtained were less variable. Variance might be reduced in two ways. When criterion fluctuation, which changes the PSS, contributes to the variance that is used to describe resolving power, precision may appear to be low. Judgments made during periods of stable criterion may be more consistent. Not all procedural changes that reduce variance will be desirable, since they may have that effect by imposing stability on performance where it does not normally exist. An example has been mentioned, in the context of the order-labeling experiments. Gengel and Hirsh (1970) showed that feedback can stabilize the PSS. This option would reduce the variance in data averaged across sessions, but it would lead to erroneous conclusions about the requirements for perceptual simultaneity if it were used to determine that property of perception.

### B. Discrimination experiments.

The final experiments to be reviewed are the cross-modality discrimination experiments. These experiments are distinguished from the labeling or identification experiments in that they measure resolving power without regard to a specific response set. Forced-choice methods are often used, which require a decision about some attribute of the stimulus pairs to be made. But the judgment is one of similarity or of the relative lengths of two intervals. The standard for comparison is external, rather than internal, and so seems less likely to vary.

Several reports of auditory-visual temporal resolution have been provided by Kristofferson and colleagues. These experiments are offered as tests of the attention-switching model of temporal processing. The theory was originally presented by Schmidt and Kristofferson (1963) and has been elaborated by Kristofferson (1967) and Allan and Kristofferson (1974). Briefly, the model postulates a periodic central process that quantizes arrival times. Discrimination of events arriving in the same moment should be at chance. The period of the process is one of the parameters of interest. The two-alternative forced-choice experimental procedure is used to generate psychometric functions from which the parameters of the model may be estimated. The subject's task is to identify the trial interval in which the offsets of long-duration stimulus elements are non-simultaneous. In the non-signal interval, the offsets are usually simultaneous. The authors feel that the use of offset rather than onset asynchrony is preferable because the observer's attention is more easily directed to the stimulus channels.

Schmidt and Kristofferson (1963) reported that a 30-50 msec $\Delta T$ was discriminable using the procedure described. Approximately the same result was obtained by Kristofferson

(1967) with eight new subjects. In 1974, Allan and Kristofferson added two parameters to the mathematical model. These were the probability that the subject was attending to the stimulus channel on which the first element was presented and the probability that the observer was in a "two-quantum state", an indicator of the efficiency with which the observer could switch attention between channels. In this experiment, $\Delta T$ required for discrimination was closer to 60 msec. Psychophysical procedures, however, had not been changed. There was no discussion of the decrease in performance, except to say that the added parameters would account for it.

Allan and Kristofferson (1974) and Kristofferson and Allan (1973) described an interesting aspect of their findings. In these experiments, Tc is always positive. When Ts was made negative, no improvement in performance was seen, for values that increased $\Delta T$ by up to 25 msec. That is, the proportion of correct choices of a given Tc was no higher when it was paired with Ts=-25 msec than when it was paired with Ts=0 msec. A similar result has been observed in the current experiments.

Collyer (1973) used a single-interval psychophysical procedure to generate the psychometric function for cross-modality gap detection. On half the trials, a 1500-msec tone was followed immediately (Ts=0 msec) by a light. On the remaining trials, the light followed by Ts + $\Delta$T msec. Performance was expressed in terms of a criterion-tolerant index, analogous to d'. The just-detectable gap was approximately 100 msec. The experiment was repeated, using two tones to bound the gap. When the tones were of the same frequency, a 20-30 msec gap was detected; with frequency disparity, 30-50 msec was required. The effect of frequency difference is comparable to that reported by Divenyi and Danner (1977) and Williams and Perrott (1972). The latter experiment, the detectable gap between two tones increased from 10 to 45 msec as $\Delta$F increased from 0 to 240 Hz.

Rousseau and Kristofferson (1973) examined the resolution of non-zero intervals in a cross-modality case. A single-interval psychophysical procedure was used; subjects indicated whether the pair was Ts or Tc. The tested values of Ts, measured from offset of a light to onset of a tone, were 100, 600, 1200, and 2000 msec. For each Ts, the just-discriminable $\Delta$T was approximately 160 msec. The authors noted that psychometric functions, d' x $\Delta$T, were also similar for all Ts. They were linear and were approximated by the relation,

$$d' = 0.0086(\Delta T) \qquad \text{(eq. 1.2).}$$

Solving for the value of $\Delta$T required for a d' of 1.0 would reduce the just-discriminable change to 116 msec. These results are greatly at odds with Weber's Law; rather than being a constant, the Weber fraction decreases from greater than 1.0 at the smallest value of Ts to 0.08 at Ts=2000 msec. The latter value is as low as the values obtained in within-modality experiments reported by others.

### C. Summary

Cross-modality temporal resolution has not been thoroughly studied with the appropriate psychophysical procedures, e.g., discrimination methods as opposed to labeling or identification procedures. There is no published report that includes measurements at several values of Ts with each element leading. The studies that are available are restricted to the auditory-visual case, and since most of these originate with Kristofferson's group, they are presented as tests of the attention-switching model. Little space is devoted to describing the nature of performance apart from the model. In most cases, interpolations from psychometric functions are left to the reader. The data reviewed do suggest that cross-modality temporal resolution is much poorer than within-modality temporal resolution, especially for nearly-simultaneous values of Ts.

## VI. THEORIES OF TIME PERCEPTION

In this section, a few of the theories advanced to account for aspects of time perception are reviewed. Most of these arise as an explanation of a specific type of experimental result. A truly general model of temporal processing would account for labeling phenomena, including the prior entry effect, and temporal resolution as well. Such a model does not exist at present. In this section, the term theory is used loosely; the review will consider several attempts at explanation that have not been presented as formal theories.

### A. Explanations based on labeling data

The first hypothesis to be considered cannot be called a theory, but it is an assumption that is often made in studies of order labeling. The assumption has been called the latency-difference hypothesis here. Each of the elements in a stimulus pair initiates afferent neural activity, which is propagated towards a hypothesized point in the central nervous system where a comparison of arrival times will occur. It is further assumed that the percept reflects the relative arrival times of the impulses, either simultaneous or ordered. Variance in arrival times determines the DL for order labeling, in this model, and the mean difference in arrival times determines the PSS. The simplicity of the model is appealing, and comparators of this type are thought to exist in the auditory system, where sound localization is cued by a difference in arrival times at the two ears. For example, neurons in the superior olivary complex have been shown to be tuned for small ranges of interaural T (Galambos, Schwartzkopf, and Rupert, 1959). However, the model has obvious difficulty in explaining the wide range of PSS reported in the literature, and the ease with which the PSS can be shifted in individual subjects (Stone, 1926; Gengel and Hirsh, 1970). It has received support from experiments in which factors known to affect latency, such as stimulus intensity or retinal eccentricity, are shown to produce systematic and predictable shifts in the PSS. Several authors have interpreted their results in terms of latency differences (Rutschmann, 1973; Halliday and Mingay, 1964; Roufs, 1963; Efron, 1963a, b; Corwin and Boynton, 1968; Matteson, 1970; Exner, 1875).

To account for the aspects of labeling performance not predictable from the latency-difference hypothesis, others have included components of central origin in their explanations. These models may recognize the existence of transmission delay differences, but they do not accept simple coincidence coding. In this explanation, the decision process accounts for the absolute value of the PSS and the DL also. Hirsh (1959) and Hirsh and Sherrick (1961) argued for a

> "time-organizing system that is both
> independent of and central to the
> sensory mechanisms" (p. 431).

They argued that only a central process could account for the finding that stimulus parameters had no role in determining the precision of order labeling. The limiting factor in order labeling was believed to be the central component. Thus the variance contributed by the sensory channels is insignificant. Two experiments involving order labeling by brain-damaged subjects are relevant to questions about the number and location of central processing centers. As was noted above, caution is always required in the interpretation of the performance of neurologically-impaired subjects. Patients differ with respect to the extent and location of their lesions; the information is usually not available at all. Pre-injury performance is unknown, and long periods of training are not usually feasible. Finally, the loss of a function in the presence of a lesion implicates that part of the brain as participating in the behavior, but it does not pinpoint the role of the damaged area. In the studies that investigate order labeling, aphasic patients are included in the subject groups. For these groups, at least, the site of lesion can be more precisely specified, since the mediation of speech production and speech reception is known to be well localized in Broca's and Wernicke's Areas, respectively. Efron (1963c) found that expressive aphasics, those whose speech production was affected, were severely impaired on an auditory order-labeling task (DL=400 msec compared to 75 for control patients). The DL for receptive aphasics was intermediate, approximately 140 msec. On the visual order labeling task, the receptive aphasics were slightly more impaired than the expressive group. Swisher and Hirsh (1970) administered a larger battery of tests and concluded that the most difficult mode of presentation was one in which the elements came from the same location, a single earphone or a single light source. In their experiment, the fluent patients had the most difficulty, however. The authors conclude that their results are equivocal with respect to the identity of the processing center. Disruptions in temporal order labeling clearly occur in aphasia, but do not always accompany damage to non-speech areas of the brain.

Sternberg and Knoll (1973) attempted to formulate a general model for the perception of temporal order, which they called the independent-channels model. The model incorporates many of the features listed in this section. Under the model, the psychometric function for order judgment reflects two underlying functions: one describing the relative arrival times of stimulus-elicited activity and one describing the decision rule. Each underlying process may be considered separately, for they are independent. As was seen, many investigators will make simplifying assumptions with respect to one process, enabling them to make inferences about the other. In the latency-difference argument, for example, the decision rule is assumed to be deterministic and unbiased. This is equivalent to the statement made earlier, in the discussion of the latency difference argument, that all the variance in the psychometric function arises in afferent transmission. It is apparent that transmission variance could be assumed to be zero, in which case the psychometric function becomes a pure decision function. The processes are additive in any case, and an alternative to fixing one of them would be to seek other sources to estimate only one of the functions. Given the psychometric function and one of two contributing functions, it should be possible to deconvolve the psychometric function to obtain the other component. The independent-channels model provides an elegant way to consider the temporal-order labeling task. The model suggests a strategy for getting at the factors involved in the task, but it remains for ingenious investigators to apply it successfully.

### B. Explanations based on discrimination data.

Kristofferson's attention-switching model proposes a central process which limits temporal resolution (Schmidt and Kristofferson, 1963; Kristofferson, 1967; Allan and Kristofferson, 1974). This is one of a class of theories that postulate that subjective time is divided into discrete periods, or psychological moments (Stroud, 1955; Harter, 1967). The attention-switching theory is based on data from auditory-visual cross-modality forced-choice experiments. It postulates the existence of an internal timing mechanism that generates a succession of equally-spaced points in time, independent of external stimuli. Allan and Kristofferson (1974) state the three basic assumptions of the theory. First, it is assumed that there are independent input channels, which means that they cannot be attended to at the same time and that activity in one channel does not influence activity in another. Second, it is assumed that activity in an unattended channel can signal a switch in attention to that channel. Finally, it is assumed that when attention is to switch, it cannot do so until the occurrence of the next time point. The information in the channels is routed to a display area, where comparisons are made. Since transmission times are expected to vary across channels, simultaneity of arrivals occurs when the physical temporal relation compensates for the difference between channels. Two events will be perceived as successive if after the first is observed, a time point occurs and attention is switched to the other channel before the occurrence of the second event. Since the occurrence of the time points is uncorrelated with the occurrence of stimuli, the probability that a time point intervenes is a simple linear function of Ts; when Ts equals the period of the generator, the probability is 1.0. The decision process is deterministic; thus the psychometric function predicted for the discrimination of a simultaneous standard, Ts=0 msec, from a non-simultaneous pair separated by Tc msec consisted of three segments. Discrimination was perfect for Tc greater than or equal to the clock period, usually estimated to be approximately 50 msec. Discrimination was at chance for Tc between 0 and the value that just compensated for transmission delays, usually estimated to be approximately 10 msec. Between those values, the psychometric function was linear. The later version of the theory allowed the possibility that two periods could elapse before the switch was made, and that attention might initially be directed to the wrong channel. Each of these will make the psychometric function less steep.

To summarize, there are four free parameters in the psychometric function, representing the difference in transmission delays, the period of the internal clock, the probability that the subject attends to the channel on which the first event actually occurs, and the probability that he is in a "two-quantum state". Empirical psychometric functions are used to estimate these parameters, and an excellent fit of the function is obtained. Allan and Kristofferson (1974) made two estimates of the set of parameters for three subjects and found them to be in good agreement.

A psychological moment has some of the properties of a perceptual category. Events arriving at the display area within the moment would be labeled simultaneous, assuming that labeling is unbiased (Allan, 1975), and they would also be indiscriminable. However, if resolution of non-zero Ts is considered, different predictions are made by the two models. The following example predicts the perfectly-discriminable $\Delta T$ assuming either a psychological moment or a categorical model. Stimulus pairs are formed from two arbitrary elements. The form of the psychophysical function can be derived from the predictions. Following Egan (1975), a psychophysical function is a plot of the change in a standard stimulus required for a fixed level of performance as a function of the standard stimulus. Let the period of the moment be 50 msec. For Tc to be perfectly discriminable from Ts, the probability that a different number of time points intervenes must equal 1.0. For an unbiased observer in a single-quantum state, a $\Delta T$ of 50 msec satisfies this requirement and will always be perfectly discriminable. It is assumed that the observer has perfect memory for the number of time points that elapse between events. For a given Ts, Tc that is 50 msec smaller or larger should be resolvable.

This may be compared to a prediction based on the partitioning of the Ts dimension into two categories, corresponding to the two possible orders of elements.

Assume that the categorical boundary is located at Ts=+25 msec, and for simplicity assume that the labeling function is a step. Two pairs, Ts and Tc, are discriminable if they elicit different labels. In this case, the discriminable Tc differs for changes in different directions. For Ts=-100 msec, $\Delta T$ greater than or equal to 125 msec would be discriminable. It follows that no $\Delta T$ which produces Tc more negative than Ts will be discriminable. The psychophysical function, $\Delta T$ by Ts, will have a slope of one in this case; its slope was zero if predicted from a quantal model.

Another way to model temporal resolution originates with Creelman (1962), who attempted to account for the resolution of filled auditory intervals. The Creelman model postulates that the internal representation of duration is in terms of the final value of a counter that is started at the onset of the stimulus. A similar suggestion had been made by Boring (1933), when he suggested that time may be encoded by integration, a process that increases monotonically and that proceeds to a unique value determined by the time at which the integration is stopped. In Creelman's model, the input to the counter is assumed to obey the rules of a Poisson process. Counts may arise from any of several sources and at random times; in such cases, the total number of counts observed over an interval will be Poisson distributed (Parzen, 1962). The single parameter of the Poisson, $\lambda$, is proportional to the rate at which the process proceeds. For a duration T, the observed number of counts is a random variable with mean and variance both equal to $\lambda T$. The parameter, $\lambda$, may be estimated, and a statistic analogous to d' calculated from overlapping Poisson distributions. Predictions from the model generally agree with data. Divenyi and Danner (1977) have used a similar Poisson model to interpret the temporal resolution of empty intervals, and Abel (1972a, b) has examined both filled and empty intervals. All of these investigators have been able to estimate $\lambda$ for their data, but there is no agreement across the studies. As Divenyi and Danner (1977) pointed out, the data from studies of time discrimination have most often been interpreted as reflecting either a quantal or a Poisson process. In spite of the lack of quantitative agreement between versions of the Poisson model, they are supported by the fact that most sensory processes seem to be continuous rather than quantal in nature. In those experiments whose data are said to argue for a discrete representation of temporal extent, the argument is based on the form of the psychometric function. If the underlying process is quantal, the psychometric function will have at least one linear segment. However, the expected difference between a linear and a curvilinear psychometric function is small enough that it is difficult to demonstrate empirically. This question probably cannot be settled until a better understanding of the underlying neural mechanisms is reached.

The models of the final type have in common their application to situations in which two inputs are combined, resulting in a unique event that reflects the temporal relation of the components. These are correlation models. They draw heavily from principles developed in electrical engineering, in the design of radar systems. The use of such models in psychophysics has been limited to cases in which special mechanisms are likely to exist. They have not been applied to the modeling of simple temporal resolution. The models differ from the others considered in two ways. First, they emphasize characteristics of the signal as the major determinants of its utility for carrying temporal information. In addition, they allow the possibility that temporal information may be preserved without being recognized or labeled as such. For example, a time difference may be experienced as a change in position. Models of psychophysical or neural correlation have been proposed by Jeffress for binaural localization (1948) and binaural detection (1972); by Taylor and Clarke (1971) for monaural detection with a contralateral cue; and by Simmons (1973) for echolocation by bats.

The Jeffress (1972) and Taylor and Clarke (1971) models are applied to detection situations. In one case, a binaural signal in noise is made more detectable if the interaural phase of the signal is different from that of masking noise (Jeffress, 1972). The difference between the threshold for homophasic signal and noise and the threshold for antiphasic signal and noise is called the masking-level difference (MLD). In the latter case, Taylor and Clarke (1971) found that when a signal is presented monaurally in one interval of a 2AFC trial, its detectability is enhanced if a contralateral replica of the signal is present in both intervals. It was also found that small asynchronies, up to 1 msec, between cue and signal could be tolerated. The subjective effect of this time shift is to change the perceived location of the signal. However, certain features of their results, notably worse-than-monaural performance when cue leads the signal, are best explained by cross-correlation, rather than by lateralization.

Simmons' (1973) work has shown that the resolution of target range by echolocating bats is limited by the bat's ability to resolve the time at which the echo returns, which in turn is limited by and is predictable from the bandwidth of the orientation sound. For several species of bat, the autocorrelation function of the orientation sound was found to predict the psychometric function for the discrimination of changes in target distance/time.

One consequence of psychophysical cross-correlation seems to be the preservation of a small time difference, and it has been suggested that the mechanism involves encoding of the time difference into a spatial difference in the nervous system (Jeffress, 1948; Boring, 1933). That is, two pairs of signals that differ in T may be diverted to different neural channels, and the organism is apparently capable of discriminating information in one channel from identical patterns of activity in another. It also seems plausible that the process would allow the organism to recognize the similarity between two signals in spite of an onset-time difference. In these models, the physical delay, T, is compensated by an equal neural delay, and the magnitude of the neural delay is preserved by the identity of the tuned channel that transmits the information.

Neural cross-correlation is a mechanism that seems best suited to special cases of time discrimination, such as sound localization or echo ranging. It does not seem to be involved in the more general capability that temporal resolution has been defined to be. There are reports, however, that suggest that a cross-modality correlation mechanism may exist, and for that reason, the concept was included. König (1965) has presented data for one observer who was confronted with a speech intelligibility problem under three conditions: auditory information only, lipreading only, or both cues combined. For single-modality presentations, the information was degraded to make intelligibility approximately 0.20. When both channels were presented simultaneously, intelligibility was considerably greater than simple arithmetic summation of the intelligibility scores for each presented alone; the score here was approximately 0.80. In addition, intelligibility did not decrease when the auditory cue was delayed for over 200 msec. Dodd (1977) has reported a similar result, finding improved intelligibility for masked speech when the acoustic signal trails the visual by 400 msec. Greater improvement was seen for simultaneous visual-auditory presentation. Intermediate values of delay were not tested. These results are remarkable in light of the role played by temporal relations in speech, where the perceived syllable may change as a result of only a few milliseconds' variation in the interval between two phonemes. The result seems to require an interpretation in terms of correlation mechanisms such as were described above. A compensating neural delay imposed on the leading channel would create the proper relative timing between the two channels and allow a point-by-point comparison to take place. This kind of processing would seem to be a necessity, even for the case called simultaneous, since physical simultaneity at the point at which an auditory-visual signal originates is lost when the receiver is separated from the source. Physical asynchrony may be compounded by physiological delay as well.

It seems possible that such a mechanism might preserve information about relative onset times without an ongoing experience of asynchrony; if so, tasks other than the kinds reviewed here will be needed to reveal the capability for using the information in temporal differences.

## VII. SUMMARY AND CONCLUSIONS

At this point, a brief summary of Chapter One is provided. The major conclusions of the literature reviewed are restated and related to the experiments reported in the following sections.

### A. Perceptual qualities of stimulus pairs differing in T.

Many of the studies mentioned used the single-interval order-labeling procedure, which provides estimates of the point of subjective simultaneity, PSS, and the difference limen for order, DL. From these studies, it was concluded that the DL has a lower limit, usually 20 msec (Hirsh, 1959; Hirsh and Sherrick, 1961). Higher values will be obtained when the subjects are untrained or when single presentations of the stimulus pair must be labeled (Hirsh and Fraisse, 1964; Gengel and Hirsh, 1970). Because it is fairly constant in spite of changes in the PSS or in the nature of the stimulus pair, the DL has been said to reflect the limitations of a central processing center, into which any sensory channel might project (Hirsh and Sherrick, 1961). The PSS, on the other hand, is found to fluctuate. It can be shifted by instructions (Stone, 1926); by the use of feedback (Gengel and Hirsh, 1970); or, especially for within-modality experiments, by manipulating aspects of the stimulus elements that are known to affect latency (Rutschmann, 1973; Efron, 1963a, b; Corwin and Boynton, 1968; Matteson, 1970; Halliday and Mingay, 1964). Several authors have mentioned the ease with which the criterion for simultaneity fluctuates. The fact that the precision of order labeling does not change with changes in the PSS suggests that the precision measured by this procedure is of the decision process, rather than of the encoding process. Other psychophysical tasks, those that eliminate the criterion problem, can reveal aspects of temporal processing other than those demonstrated in labeling experiments.

### B. Discrimination of changes in Ts

The perception of order is well understood. The second class of experiment discussed above was that in which the smallest discriminable change from a standard, Ts, is determined. For the most general case, that value, expressed as $\Delta T/Ts$, is approximately 0.06-0.10 (Woodrow, 1951). Most often, auditory marker stimuli have been used. There are differences between modalities, with the visual being least precise (Goodfellow, 1934).

For the special case where Ts=0 msec and two identical elements stimulate the same place, the experiment is called a determination of the two-pulse threshold, or gap detection. The just-detectable change in these experiments is 1.6 msec (auditory stimuli, Gescheider, 1966); 10 msec (tactile, Gescheider, 1966); or 20 msec (visual, Keitzman, 1967). When the two stimulus elements have different identities, the result can be called a successiveness threshold. The successiveness threshold has the same value as the two-pulse threshold for auditory and tactile stimuli (Gescheider, 1966) and is only 5-12 msec for vision (Sweet, 1954). For the visual-auditory cross-modality case, it is said to be 100 msec (Collyer, 1974). This experiment can be confounded by labeling, as was discussed in making the distinction between the two-pulse and the two-flash thresholds.

### C. Discrimination and its relation to labeling and perceived simultaneity

The experiments reported in the following chapters were conducted to determine the limits of temporal resolution for cross-modality stimuli and to clarify the relation between labeling data and discrimination data. The best-known cross-modality result is that of Hirsh and Sherrick (1961) for order labeling. Those data do not suffice to estimate temporal resolution, however, for as those authors stated, more than resolving power must be considered to account for order-labeling. Those experiments that have measured cross-modality temporal resolution have used only one cross-modality stimulus combination and usually have not varied Ts parametrically.

Two possible relations between the precision of temporal resolution as indicated by order labeling and as indicated by criterion-tolerant procedures might exist. One is that the observer may be more sensitive to temporal changes than he will indicate by applying labels to his perceptual experience. In this case, forced-choice discrimination procedures will reveal that pairs of elements that are given the same labels are resolvable. Results such as the following might be obtained. In the auditory modality, assume a Weber fraction, $\Delta T/Ts$, of 0.10. Then when Ts=25 msec, a 2.5-msec change should be discriminable. Abel (1972b), Getty (1975), and Divenyi and Danner (1977) obtain this result with various auditory markers. However, Hirsh (1959; Hirsh and Sherrick, 1961) and Pisoni (1977) both find that the precision of labeling is much less than that. As noted, the precision of order labeling is at best 17-20 msec. In this case, it seems that pairs given the same order label may be shown to be clearly distinguishable by some other psychophysical procedure.

The second possibility is similar to the special relation between labeling and discrimination hypothesized by Liberman as categorical perception (Liberman et al., 1961; Miller et al., 1976). If there are regions of Ts and stimulus combinations over which subjects tend to compare labels rather than absolute features of the stimulus, different results would be anticipated. Assume that the result of an order-labeling experiment places the PSS, or categorical boundary, at +25 msec, with a DL of 20 msec. Considering the same example described above, Ts=25 msec, the discriminable $\Delta T$ would be one which was labeled differently from the standard. If that standard were called simultaneous, then a change of 20 msec, to generate a percept of order, would be discriminable. In Divenyi and Danner's experiment, the discriminable change for Ts=25 msec was 10 msec when a large frequency difference characterized the elements in the stimulus pair. It may be that with the frequency difference, a characteristic percept, perhaps related to pitch, emerges and that the subject tends to base his judgment on the quality, rather than the time difference. For a standard that is not centered in the hypothesized category, an additional prediction could be made. In this case, the just-discriminable change in one direction would be larger than in the other, since Ts is farther from one categorical boundary than the other.

The remaining chapters summarize the cross-modality discrimination experiments that were performed. The results indicate that cross-modality temporal resolution is a capability that may be modeled in either of these two ways, depending on the psychophysical procedure chosen.

## CHAPTER TWO. Results of preliminary experiments.

The experiments described in this chapter were completed prior to those presented as part of the dissertation. The current experiments were suggested by the results of these preliminary studies. They were conducted in the Signal Detection Laboratory of the Central Institute for the Deaf (CID). Brief descriptions of the results have appeared in CID Periodic Progress Reports Nos. 18-21, and in abstract form (Watson, Sinex, and Kelly, 1976).

### A. Cross-modality simultaneity judgments.

In these experiments, which used the method of adjustment (MOA), subjects were presented with an ongoing train of paired stimulus elements. The stimuli were generated by delivering a brief pulse to the appropriate transducer, to produce stimuli for three sensory modalities. The auditory stimuli were produced by delivering a 0.5-msec pulse to a TDH-49 earphone. The tactile stimulus was generated by delivering the same pulse, amplified and filtered, to a ceramic bender. When tactile stimuli were used, masking noise was always present in both earphones. The visual stimulus was produced by triggering a Grass photostimulator or, later, by lighting a yellow, light-emitting diode (LED) for 20 msec. All the stimulus elements were presented at levels well above threshold.

The method-of-adjustment procedure was as follows. Subjects were given control over the temporal relation, T, of the stimulus pair. They were permitted to take as long as they wished to produce a pair that they perceived as simultaneous. Data were collected in five-trial blocks, which lasted approximately five minutes. A feedback scheme was employed, by which the subjects were informed of the trial-to-trial fluctuation in their responses, but not the actual values of T they had selected. It was hoped that this feature of the procedure would help to stabilize the criterion for simultaneity. However, for two subjects who made adjustments with and without feedback, no difference in performance under the two procedures was observed. In the original experiment, auditory-visual pairs were presented to four subjects. Later, two new subjects were tested with auditory-visual (A-V), auditory-tactile (A-T), and visual-tactile (V-T) pairs. Figure 2.1 presents the frequency distributions of all the judgments from the four original subjects. The standard deviation of simultaneity matches for a typical subject is approximately 25 msec. If this figure is taken to indicate a range of uncertainty about the temporal relation of elements that elicit the percept, it may be compared to the DL for order labeling, with one qualification. The standard deviation includes a larger proportion of the underlying distribution function than does the quartile deviation; thus if simultaneity judgments and order labeling reflected the operation of a single process, the standard deviation would still be 1.5

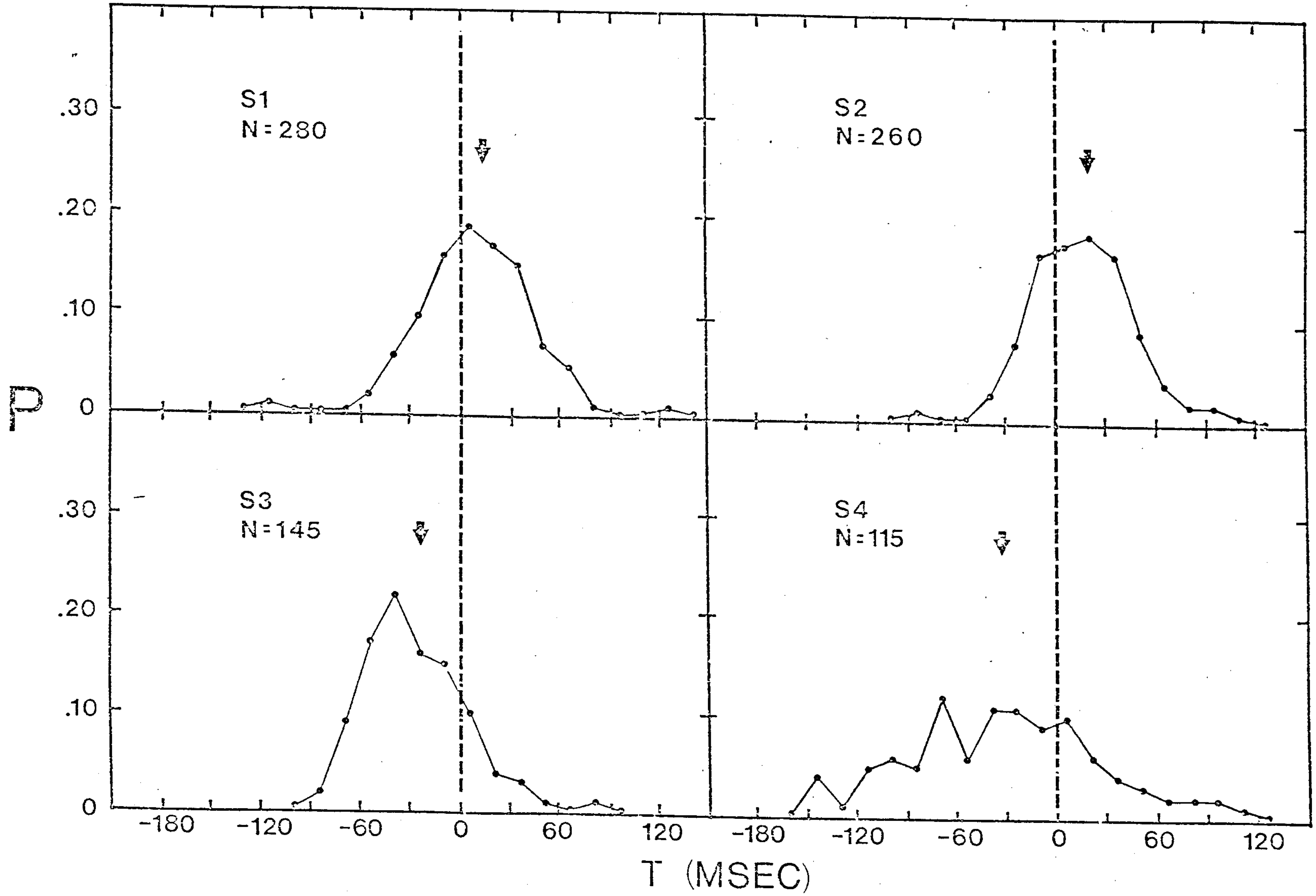

Figure 2.1. Distributions of visual-auditory simultaneity judgments. Frequency distributions for four subjects' simultaneity matches are plotted. The abscissa is T; negative values mean auditory leads. The arrows indicate the mean of each distribution. The mean varies across subjects (from Watson et al., 1976).

times greater than the DL. A 26-msec standard deviation is therefore consistent with a 17-msec DL (Hirsh and Sherrick, 1961). It is also consistent with the observation of Fraisse (1963) that true simultaneity of auditory-visual pairs is difficult to experience, unless they can be associated with a single common source. For the V-A stimulus pairs used in these experiments, it may be concluded that the range of T called simultaneous is the same range of T over which reliable order judgments cannot be made.

The result for auditory-visual matches was replicated with two new observers in the experiment that included three modalities. The distributions of judgments from these observers with V-T stimulus pairs were similar to their V-A matches in precision, but for the auditory-tactile case, the standard deviation of the matches was approximately 10 msec. The distributions are shown in Figure 2.2. This value is too small to be consistent with the result from order-labeling. That is, there is a range of T that is not called simultaneous by the observers in this experiment and over which temporal order could not be identified by the observers of Hirsh and Sherrick (1961). In addition to the difference in the standard deviation of matches, both subjects informally expressed greater confidence in their judgments of auditory-tactile simultaneity. It may be concluded that with A-T pairs, an experience of true simultaneity occurs. This differs from the V-A and V-T

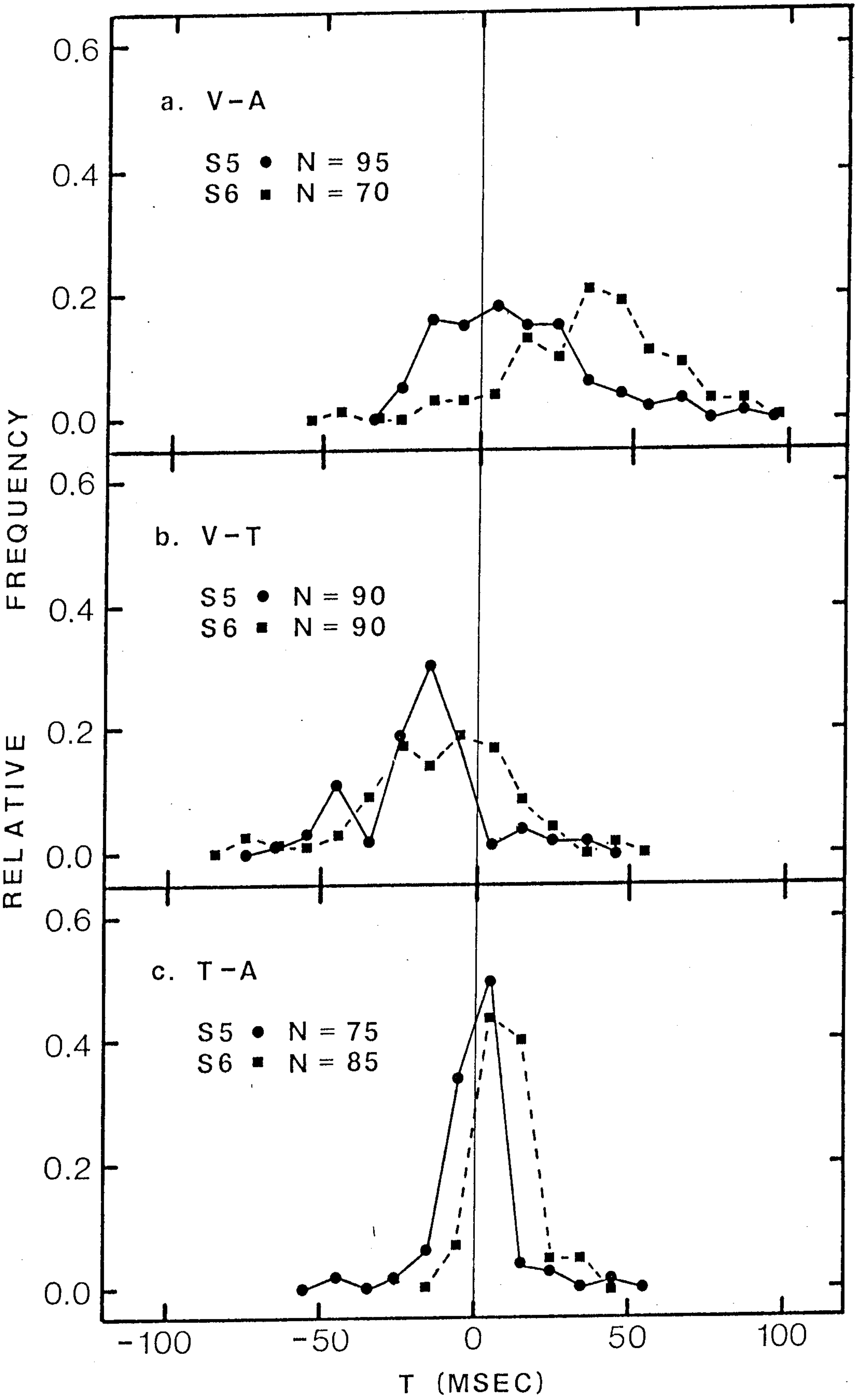

Figure 2.2. Distributions of visual-auditory, visual-tactile, and auditory-tactile simultaneity judgments. Matches from two subjects are shown. Positive T indicates V leads in panels a and b, and T leads in panel c.

cases, in which pairs are sometimes called simultaneous when the perceived temporal relation is ambiguous, to the extent that no other label can be reliably assigned.

The mean of simultaneity matches is variable, as is the PSS inferred from order labeling. For the total of six observers who made A-V matches, the means covered a range of over 60 msec. For the other cross-modality cases, the range is smaller but includes only two values in each case. The mean judgment of auditory-tactile simultaneity is 1 or 10 msec; for V-T it is 15 or 12 msec. The tactile element was leading at the mean value of T in each case.

As noted, the data were obtained in blocks of five trials. The distributions of V-A adjustments of two of the original group of observers were subjected to an additional analysis when it was observed that the criterion for simultaneity might be more stable over the course of the 5-minute block than it was for successive blocks. Given a distribution of judgments for which the mean, variance, and N are known, the standard error (SE) is given by the equation,

$$SE = \sigma / \sqrt{N} \qquad \text{(eq. 2.1)}$$

where $\sigma$ is the standard deviation of the distribution and N is the number of cases. If random samples of size n are drawn without replacement from the distribution, the distribution of sample means will have its mean equal to the distribution mean and its standard deviation equal to the standard error. Such samples were drawn from the distributions of judgments, and the statistics of the sampling distributions were calculated. For the random samples, the relation described above held. However, for non-random samples, the actual blocks, the standard deviation of block means was significantly larger than the standard error. This would occur if the central tendency of individual blocks of judgments was not constant. The hypothesis of a process with a fixed criterion for simultaneity thus appears to be untenable. This result is taken as evidence that while the PSS and the criterion for simultaneity do vary, they are stable at least for moderate periods of time.

### B. Auditory-visual temporal resolution.

This experiment used the same-different psychophysical procedure to measure temporal resolving power for auditory-visual stimulus pairs. The stimulus elements were the same as those used in the latter set of simultaneity judgments, in which the LED was used to produce the visual stimulus. The standard, Ts, ranged from -480 to +480 msec; the visual element leads the auditory for positive values. Comparison stimuli were formed by adding or subtracting $\Delta T$. Thus Tc could be longer or shorter than Ts, or it could have the opposite order. Four subjects participated in the experiment. The just-resolvable change in Ts, △T for which d'=1.0, was determined by interpolation from fitted psychometric functions. These values were displayed as a function of Ts.

Two such functions are shown in Figs. 2.3a and 2.3b. Each point on these functions is the median of the best estimates of the just-resolvable △T for four subjects. Figure 2.3a indicates performance for those conditions in which △T was positive. If Ts is also positive, Tc will be a longer interval than Ts. If Ts is negative, Tc will be a smaller interval or will have the opposite order from Ts. Larger and smaller refer to absolute values, since the signs indicate order only. The just-resolvable change is smallest when Ts=+60 msec, not at Ts=0 msec as might have been expected. For Ts>60 msec, the △T for d'=1.0 increases with increases in Ts. For Ts<+60 msec, performance is poor until Ts=-120 msec. For the remainder of the conditions, the △T for d'=1.0 again increases with increases in the standard. In the region, -120 msec <Ts< +60 msec, an analogy to discrimination of categorically-perceived stimuli can be made. The values of Ts and Tc that are presented for discrimination in these conditions are those that were labeled simultaneous by the observers in the previous experiment. In this region, both Ts and Tc fall into that range, unless △T is large, relative to the values that could be discriminated when Ts was slightly larger or

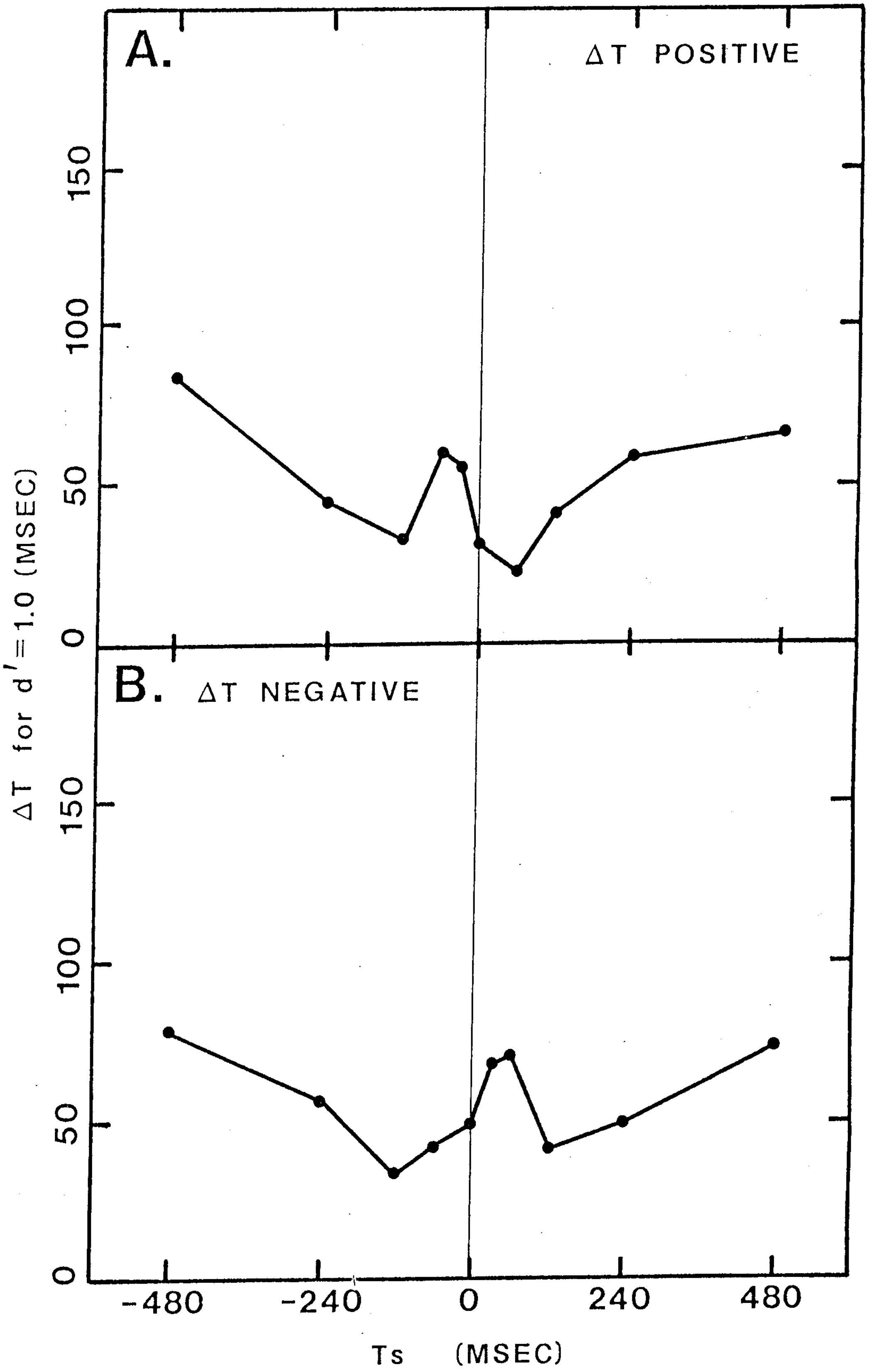

Figure 2.3. Discrimination function for visual-auditory stimulus pairs. The abscissa is Ts. The ordinate is the value of ΔT for which d'=1.0. Panel a is for positive values of ΔT; panel b is for negative values of ΔT. Each function has characteristic peaks and valleys that are explained in the text.

smaller. It is suggested that in these conditions, subjects in the same-different experiment tend to rely more heavily on labels assigned to Ts and Tc, rather than on a fine-grained analysis of the stimulus information. The same explanation is applicable to the observation that the smallest $\Delta T$ for d'=1.0 occurs at Ts=60 msec. In this case, a tendency to label Ts will enhance performance when Tc is longer, since Tc will seldom be called simultaneous.

Figure 2.3b presents the median data for conditions in which $\Delta T$ is negative. In this case, the curve reaches its minimum at Ts=-120 msec; this function is nearly a mirror image of the one shown in Fig. 2.3a. Values of Ts between -60 and +60 msec are most difficult to resolve. Once again, the poorest performance is observed when Ts + $\Delta T$ is approximately 0 msec. That is, the just-discriminable change in these conditions deviates from the line of constant slope predicted by Weber's Law.

An apparent relation between labeling and discrimination was found in this experiment. It is not suggested that the observer is incapable of resolving smaller changes in nearly-simultaneous values of Ts. It does seem that a problem related to criterion or memory may be present when the same-different procedure is used. That is, when Ts is nearly simultaneous, a single presentation of the standard elicits an ambiguous percept for which the best strategy may be to classify the pair as "simultaneous" or "ordered". The response will then be "same" unless Tc is given a different label. Ts and Tc are most likely to be given the same labels when Ts + $\triangle$T is approximately 0 msec. They are most likely to be given different labels if Ts is small but non-zero and $\triangle$T increases the absolute value of Tc. If the observer adopts this strategy, which will be called the "labeling strategy", his discrimination performance will take the form seen in Fig. 2.3a and 2.3b.

Given the apparent influence of perceived temporal relations on A-V temporal resolution and the large difference between A-T and V-A or V-T simultaneity matches, it may be hypothesized that performance in temporal discrimination tasks using A-T stimuli will be qualitatively different from performance in the same tasks using V-A or V-T stimulus elements. To test this, the experiments presented in the following chapters were designed to measure temporal resolution with all three cross-modality stimulus combinations. In addition, two procedural variations were used in an effort to reduce the observer's decision problem. In one case, a three-alternative forced-choice (3AFC) procedure was tried, to reduce the effects of criterion. In the other, a multiple-presentation adjustment procedure was used to reduce the ambiguity associated with the presentation of the standard pair.

# CHAPTER THREE. Results of experiments using the same-different psychophysical procedure.

The first experiments to be discussed used the same-different psychophysical procedure with auditory-tactile (A-T) and visual-tactile (V-T) stimuli. In addition, within-modality control conditions were tested. These experiments complete the series that began with the visual-auditory (V-A) experiment, conducted at CID.

## I. Experiment 1. Same-different judgments using auditory-tactile stimulus pairs.

### A. Methods.

Temporal resolution for A-T stimulus pairs was measured using the same-different psychophysical procedure. The standard intervals, Ts, ranged between -480 and +480 msec. As before, the sign of Ts indicates the order of elements. In this case, auditory elements lead tactile if Ts or Tc is positive. The comparison intervals, Tc, were formed by changing Ts by the amount indicated as $\Delta T$. The sign of $\Delta T$ indicates the direction, in time, in which the change is made. The relation,

$$Tc = Ts + \Delta T \qquad \text{(eq. 1.1)}$$

describes all conditions. Thus, positive values of $\Delta T$ have the effect of delaying the tactile stimulus, relative to its temporal relation to the auditory element in Ts. Two examples illustrate the symbol conventions. Let Ts=+100 msec and $\Delta T$=+50 msec. Then Tc=+150 msec; the acoustic stimulus element leads the tactile by 100 msec in the standard and by 150 msec in the comparison. If Ts=-100 msec, the same $\Delta T$ produces Tc=-50 msec. The tactile element leads by 100 msec in the standard but by only 50 msec in the comparison. Negative values of $\Delta T$ are treated in a similar manner.

A combination of a given Ts and a direction of change is called a condition. For each direction of change, 10 values of Ts were used. The 20 conditions were tested in an arbitrary order that alternated, as nearly as possible, the sign of Ts and $\Delta T$. Following approximately 30 hours of training, the cycle of 20 conditions was tested three times. Each time a condition was tested, three blocks of trials were obtained. Each point on the resulting psychometric function was based on 96 trials.

The stimulus elements were produced by driving the appropriate transducers with 5-msec rectangular pulses. The devices for generating tactile and visual stimulus elements were in separate enclosures located in sound attenuating chambers (IAC). The tactile transducer was an electromagnetically-driven contactor, constructed by the author. The device is shown in Fig. 3.1, which also presents some of its characteristics. The frequency response of this vibrator is adequate for the somatosensory system, which has its greatest sensitivity near 200 Hz

Figure 3.1. Description of the electromagnetic vibrator. Panel a: Schematic diagram of the device. The driving mechanism consists of the magnet and coil from a loudspeaker. The contactor tip is attached to the end of a shaft that projects 2 mm above the surface of the enclosure. The device is mounted on a wooden base (shaded areas). Panel b: Vibrator frequency response. The input signal was a sinusoid, at 0.5 V RMS. The ordinate is the output force in dB re 1 dyne. During the measurements, the vibrator was loaded with a Bruel and Kjaer Artificial Mastoid, model 4930, used in conjunction with a model 2209 Sound Level Meter. Panel c: Vibrator response to a 5.0 msec pulse. Measurement apparatus is the same as in panel b.

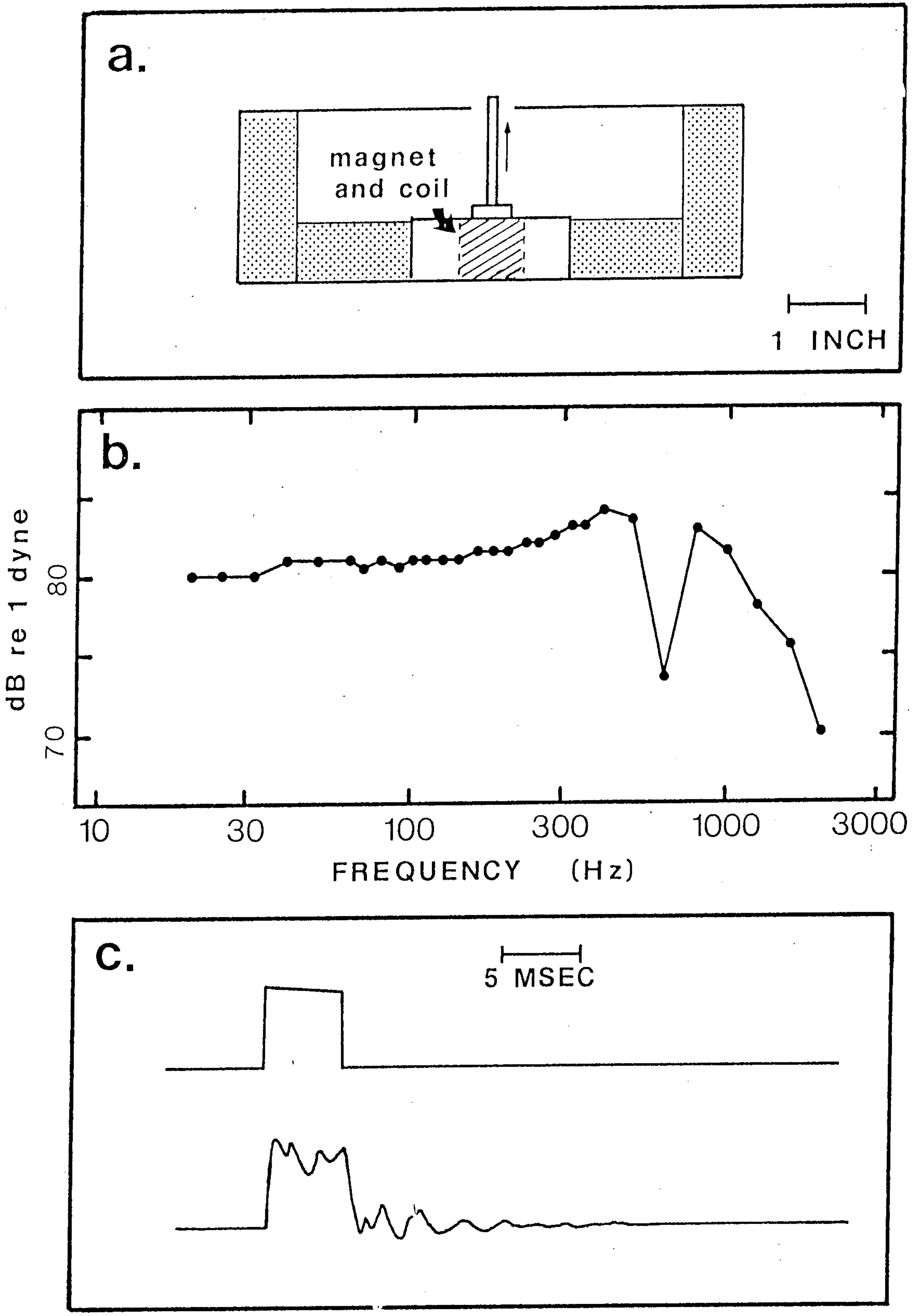
a.
magnet
and coil
1 INCH
b.
dB re 1 dyne
80
70
10
30
100
300
1000
3000
FREQUENCY (Hz)
c.
5 MSEC

Figure 3.1

(Verillo, 1963). The tactile stimulus was delivered to a fingertip on the right hand. The contactor tip was circular, with a diameter of 2 mm. The subject's hand rested on the surface of the enclosure. The vibrator tip projected approximately 2 mm above the surface; the tip slightly depressed the finger when it was not being driven. The pulse displaced the tip toward the finger. The stimulus was presented at 26 dB sensation level (SL). When activated, the vibrator produced an audible click. Consequently, during all experimental sessions, subjects wore earphones (TDH-39 with MX41-AR supraaural cushions). Broadband noise at 70 dB SPL was presented binaurally to mask the vibrator click. This level was at least 20 dB higher than the lowest level required to eliminate this potential auditory cue.

The acoustic stimulus was produced by mixing the 5-msec pulse with the noise waveform going to the right earphone only. The click had a peak SPL of 74 dB. This stimulus was 24 dB above its threshold in the noise.

The experiments were conducted in the Psychoacoustics Laboratory of the Boys Town Institute. The sessions were controlled by a PDP-11/34 computer. The 5-msec pulses were obtained by gating single cycles of a square-wave from a Tektronix FG-501 function generator. A block diagram of the apparatus is shown in Fig. 3.2.

Figure 3.2. Block diagram of experimental apparatus. The diagram shows the equipment used to generate within-modality and cross-modality stimulus pairs with specified temporal relations and to record subjects' responses. The laboratory uses a Digital Equipment Corp. PDP-11/34 computer. Two D-A converters are used to gate independent Tektronix function generators (FG), in square-wave mode (V=visual channel; T=tactile channel; A=auditory channel). The noise generator (NG) provides the masking noise. The noise waveform to one earphone is passed through a Charybdis programmable attenuator (PROG ATT) to generate the feedback intensity increment. Manual attenuators (ATT) and impedence-matching networks ($\Omega$) were used as needed.

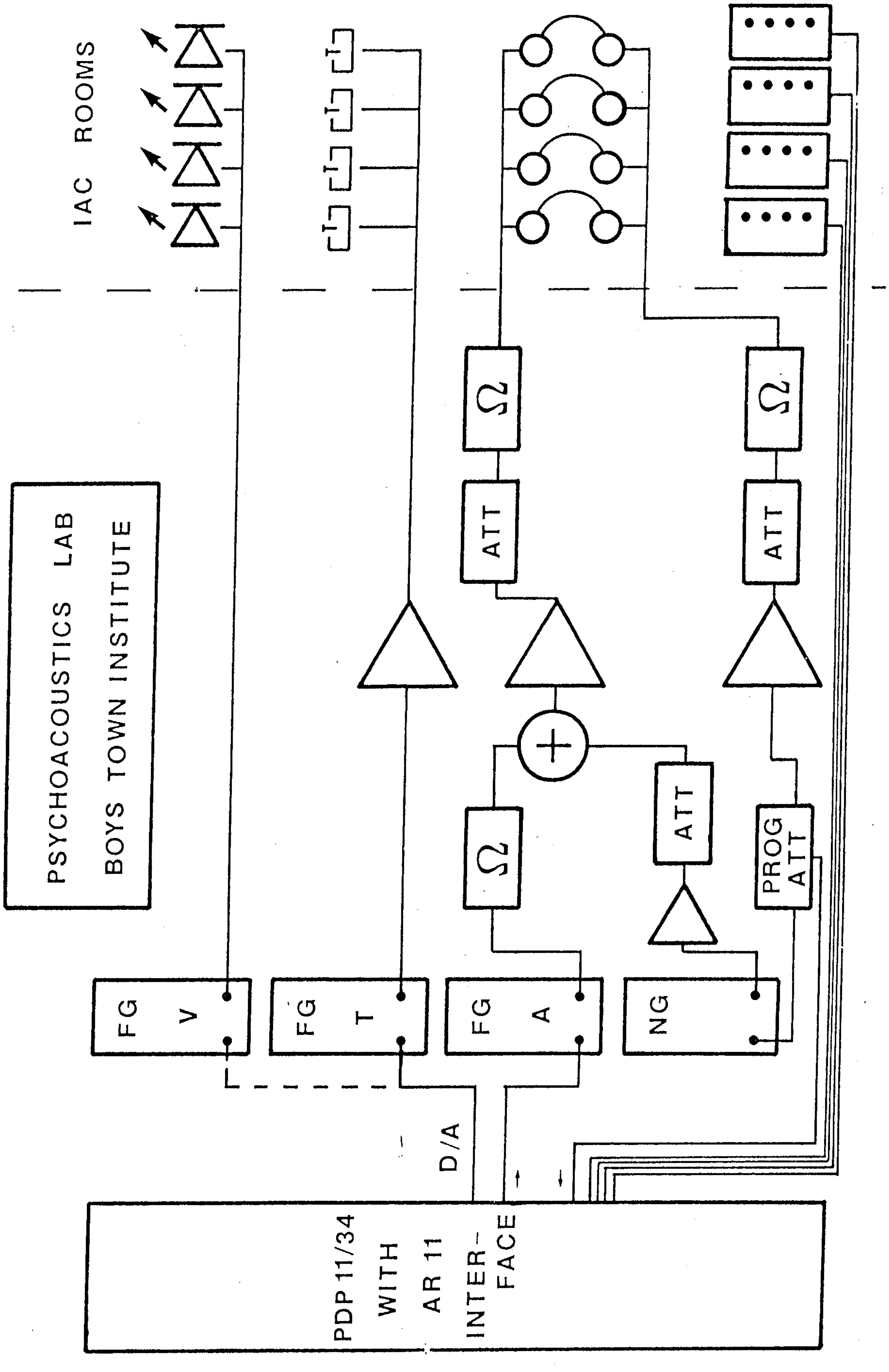
IAC ROOMS
PSYCHOACOUSTICS LAB
BOYS TOWN INSTITUTE
Ω
ATT
Ω
ATT
Ω
ATT
PROG ATT
FG V
FG T
FG A
NG
D/A
PDP 11/34 WITH AR 11 INTER- FACE

Figure 3.2

Data were collected in 96-trial blocks. In each cycle of conditions, each was tested for three blocks, or 288 trials. The trial sequence is illustrated in Fig 3.3. Ts was always presented as the first stimulus pair on a trial. The second pair was Tc on half the trials and Ts on half the trials. Within the 96-trial block, three values of Tc could occur. The three values were changed when the condition was retested, to accomodate changes in performance. Based on the previous data, the three values of Tc were chosen to elicit d' of 0.5, 1.0, and 2.0, on the average. The interpair interval was varied randomly to guarantee that the only temporal information available in the trial was in the stimulus. The base interval was 600 msec for Ts less than 240 msec; 700 msec for Ts equal to 240 msec; and 1100 msec for Ts equal to 480 msec. An additional interval, selected with uniform probability from 30 to 200 msec, was added to the base on each trial. During the response interval, subjects indicated "same" or "different" with a button press. If the stimuli were in fact different, feedback was provided. The feedback signal was a 200-msec 10 dB intensity increment in the masking noise delivered to the left ear. If the stimuli were the same, a 200 msec delay preceded the next trial.

Data were converted to d' and plotted as 3-point psychometric functions, d' by $\Delta T$. Fitted curves of the form,

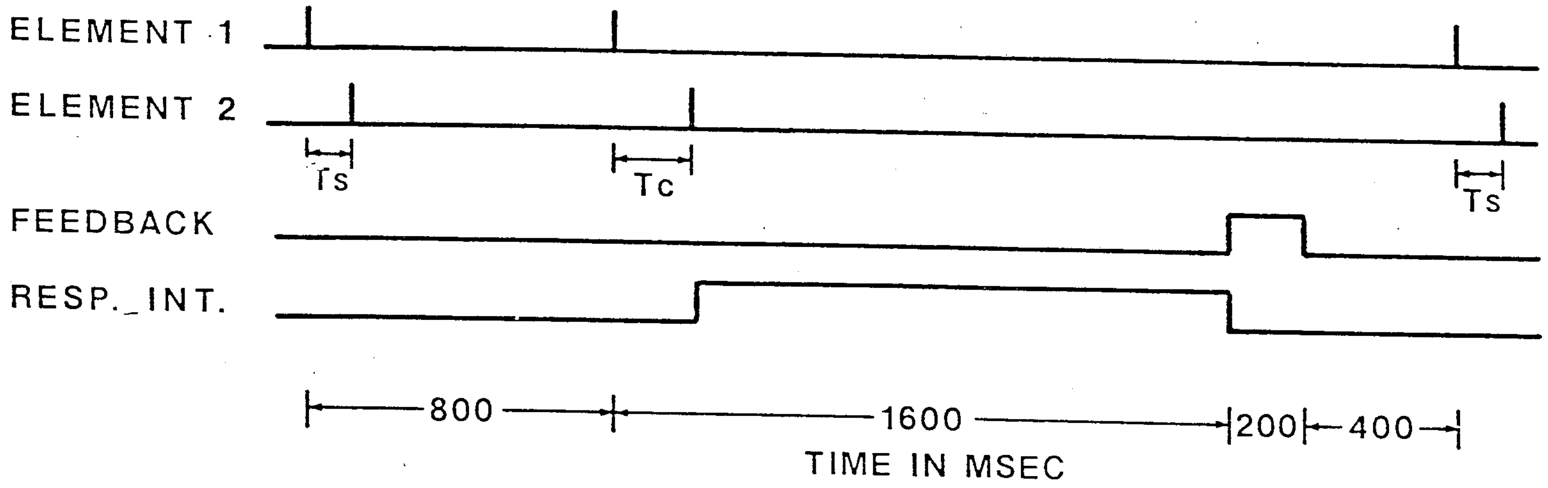

Figure 3.3. Same-different trial sequence. A "different" trial is shown. Ts is presented, followed 630-800 msec later by Tc. In this example, Ts=+120 msec and Tc=+200 msec. The duration of the response interval varies. It always begins 10 msec after the trailing element in Tc, and it always ends 1600 msec after the leading element of Tc. On different trials, the response interval is followed by 200 msec of auditory feedback and 400 msec of delay before the start of the next trial. On same trials, 600 msec of delay elapse.

$$d' = a(\triangle T)^b \qquad \text{(eq. 3.1)}$$

(Egan, Lindner, and McFadden, 1969) were obtained by linear regression, in the coordinates log d' and log $\triangle T$. The value of $\triangle T$ for which d'=1.0 was interpolated from the fitted equation. Good fits were usually obtained, but for certain cases, the data were such that the fits were not acceptable. In some conditions, for example, non-monotonic psychometric functions were obtained. These cases are discussed below. Examples of psychometric functions with the fitted curves are shown in Fig. 3.4.

## B. Results

Data were obtained from four subjects. As in previous cross-modality discrimination experiments (Watson, Sinex, and Kelly, 1976), the range of individual performance was large. Temporal resolution is given by the estimates of $\triangle T$ for which d'=1.0 for each condition. These estimates are listed in Table 3.1. The summary value listed for each subject is the lowest of the three estimates obtained in each condition. The median of the four subjects is also given.

As noted, the experiment was preceded by 30 hours of training, during which the subjects were exposed to the different types of condition: simultaneous and non-simultaneous Ts, and both directions of change. They

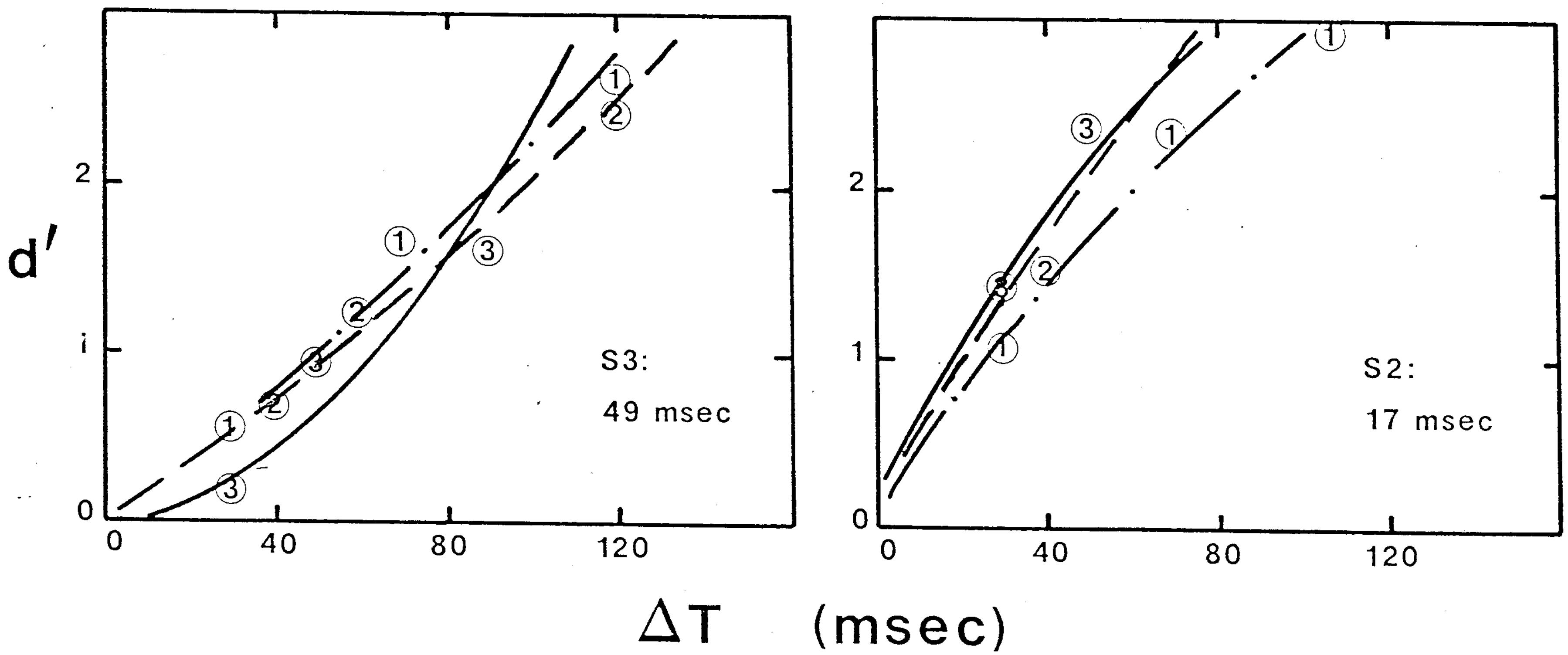

Figure 3.4. Fitted psychometric functions. Examples of the fitted psychometric functions for two subjects are shown. The symbols indicate the cycle of testing from which the data were taken. Center lines: cycle one. Dashed lines: cycle two. Solid lines: cycle three. Ts was +60 msec, and $\Delta T$ was positive. The lowest value of $\Delta T$ for $d'=1.0$ is given in the lower right corner of each panel.

## TABLE 3.1 RESULTS OF AUDITORY-TACTILE SAME-DIFFERENT JUDGMENTS

| Cond: | | ΔT for d' = 1.0 (msec): | | | | |
|---|---|---|---|---|---|---|
| Ts | ΔT | S1 | S2 | S3 | S4 | Median (4 Ss) |
| -480 | + | 79 | 21 | 49 | 89 | 64 |
| -240 | + | 80 | 24 | 75 | 69 | 72 |
| -120 | + | 100 | 14 | 43 | 86 | 65 |
| -60 | + | 166 | 16 | 84 | 103 | 94 |
| -30 | + | 24 | 8 | 48 | 96 | 36 |
| 0 | + | 3 | 4 | 9 | 5 | 4 |
| +60 | + | 84 | 17 | 49 | 36 | 42 |
| +120 | + | 92 | 23 | 68 | 66 | 67 |
| +240 | + | 84 | 28 | 46 | 67 | 56 |
| +480 | + | 144 | 43 | 90 | 66 | 78 |
| -480 | - | 88 | 12 | 64 | 86 | 75 |
| -240 | - | 69 | 28 | 60 | 73 | 64 |
| -120 | - | 94 | 23 | 69 | 60 | 64 |
| -60 | - | 66 | 19 | 56 | 45 | 50 |
| 0 | - | 4 | 4 | 12 | 15 | 8 |
| +30 | - | 49 | 7 | 15 | 88 | 32 |
| +60 | - | 65 | 15 | 25 | 140 | 45 |
| +120 | - | 101 | 14 | 36 | 71 | 54 |
| +240 | - | 83 | 28 | 61 | 52 | 56 |
| +480 | - | 83 | 29 | 39 | 73 | 56 |

also completed the auditory within-modality control condition prior to Experiment 1. In spite of the extent of the subjects' prior experience with the procedure and with related stimulus configurations, performance in a given condition was not stable across the three cycles of the experiment. This fluctuation is not due to learning, because there is no consistent tendency for the $\Delta T$ for d'=1.0 to decrease with time. Therefore, the lowest of all estimates of the $\Delta T$ for d'=1.0 is used to summarize performance. The number of trials run is sufficient to provide a valid estimate of sensory capability. The fact that performance is sometimes worse than the asymptotic level is interesting in itself but is less significant for the consideration of the limits of temporal resolution.

Individual psychophysical functions are shown in Fig. 3.5. The $\Delta T$ for d'=1.0 is plotted as a function of Ts. The top panel includes the 10 conditions in which $\Delta T$ was positive; that means that in panel a, the tactile element always ocurred later in time in Tc than in Ts. In the bottom panel, the 10 conditions for which $\Delta T$ was negative are shown. In these, the tactile stimulus element always occurred earlier in time in Tc than in Ts. The range of individual performance was large, as was pointed out, with the exception of two conditions. The $\Delta T$ for d'=1.0 always reached a minimum at Ts=0 msec. For the two conditions that use the simultaneous standard, all subjects can detect a

Figure 3.5. Auditory-tactile discrimination functions for individual subjects. The best estimate of $\Delta T$ for d'=1.0 is plotted for each observer. Abscissa: Ts. The tactile element leads if Ts is negative. The top panel shows data for conditions in which $\Delta T$ is positive. The lower panel shows data for conditions in which $\Delta T$ is negative. In each panel and for each subject, the $\Delta T$ for d'=1.0 at Ts=0 msec is less than 15 msec. For adjacent Ts, the $\Delta T$ for d'=1.0 increases sharply.

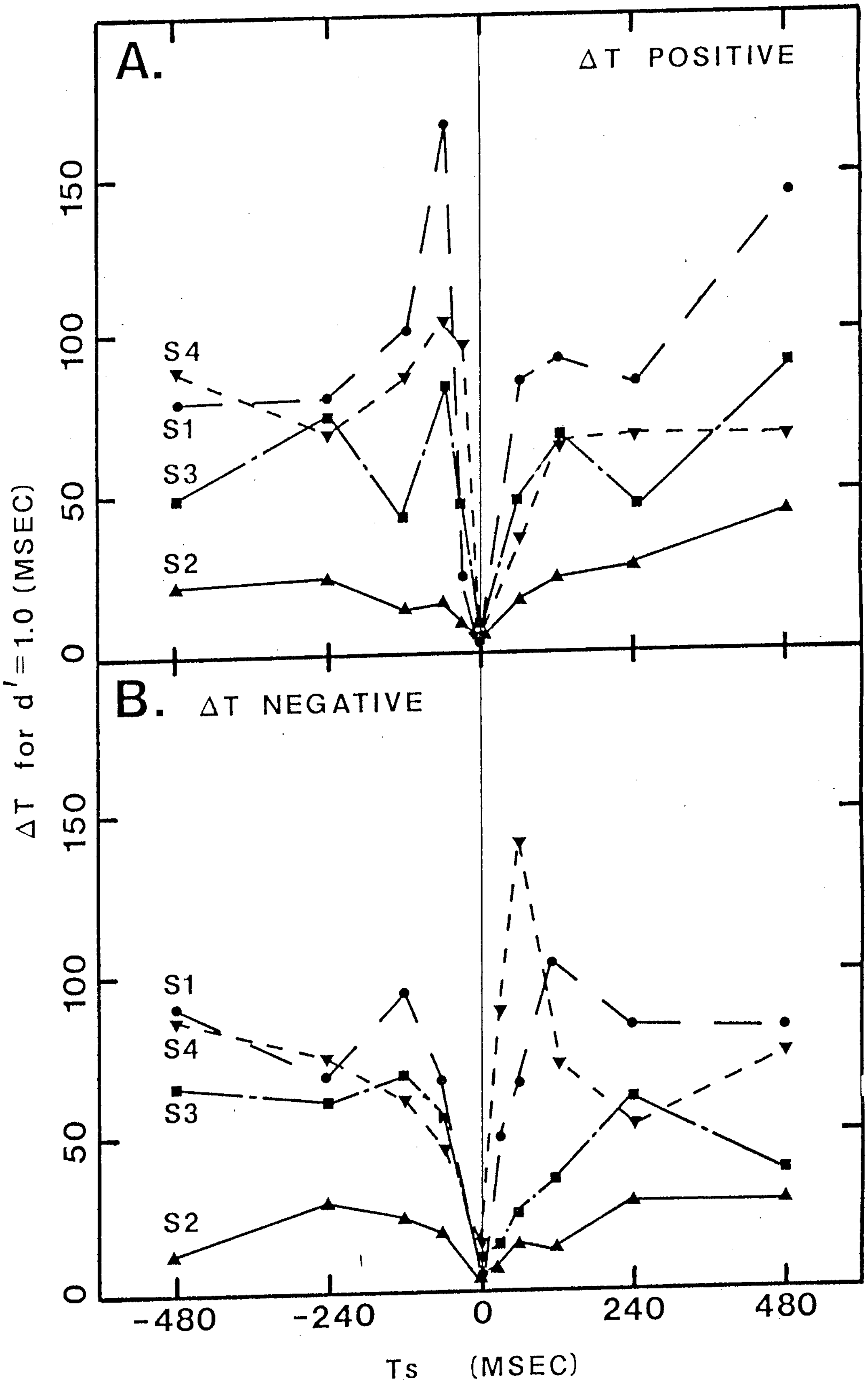
A.
ΔT POSITIVE
S4
S1
S3
S2
B. ΔT NEGATIVE
S1
S4
S3
S2
ΔT for d' = 1.0 (MSEC)
0
50
100
150
-480
-240
0
240
480
Ts (MSEC)

Figure 3.5

change of less than 15 msec. In addition, performance in these conditions was stable over time. This contrasts sharply with the result seen with V-A pairs; in that case, the smallest change was discriminable at a non-zero Ts.

The median $\Delta T$ for d'=1.0 for the four subjects is plotted in Fig. 3.6. The sharp improvement in performance at Ts=0 msec is apparent in each panel. The $\Delta T$ for d'=1.0 increases with Ts until the absolute value of Ts is 120 msec. Outside that range of Ts, the systematic increase is no longer observed. This pattern is most clear in panel b of the figure. The greatest deviation from the pattern is found in panel a when Ts=-60 msec. This condition was the most difficult (required the largest $\Delta T$) of all the conditions for three of the four subjects (see Fig. 3.5). In the individual psychophysical functions, this characteristic is more apparent than in Fig. 3.6. In the V-A experiment, the difficult conditions were those for which both Ts and Tc might be called simultaneous. In this experiment, a simultaneous standard is resolved very well, but for some non-zero values of Ts, the $\Delta T$ for d'=1.0 are large, relative to the size of the just-discriminable change when Ts is slightly larger or smaller. In panel a of Fig. 3.5, maxima in the discrimination functions are seen at Ts=-60 or -30 msec; in panel b, at Ts=+30 or +60 msec. These are the same conditions that were poorly resolved in the V-A experiment described in Chapter Two. Performance in

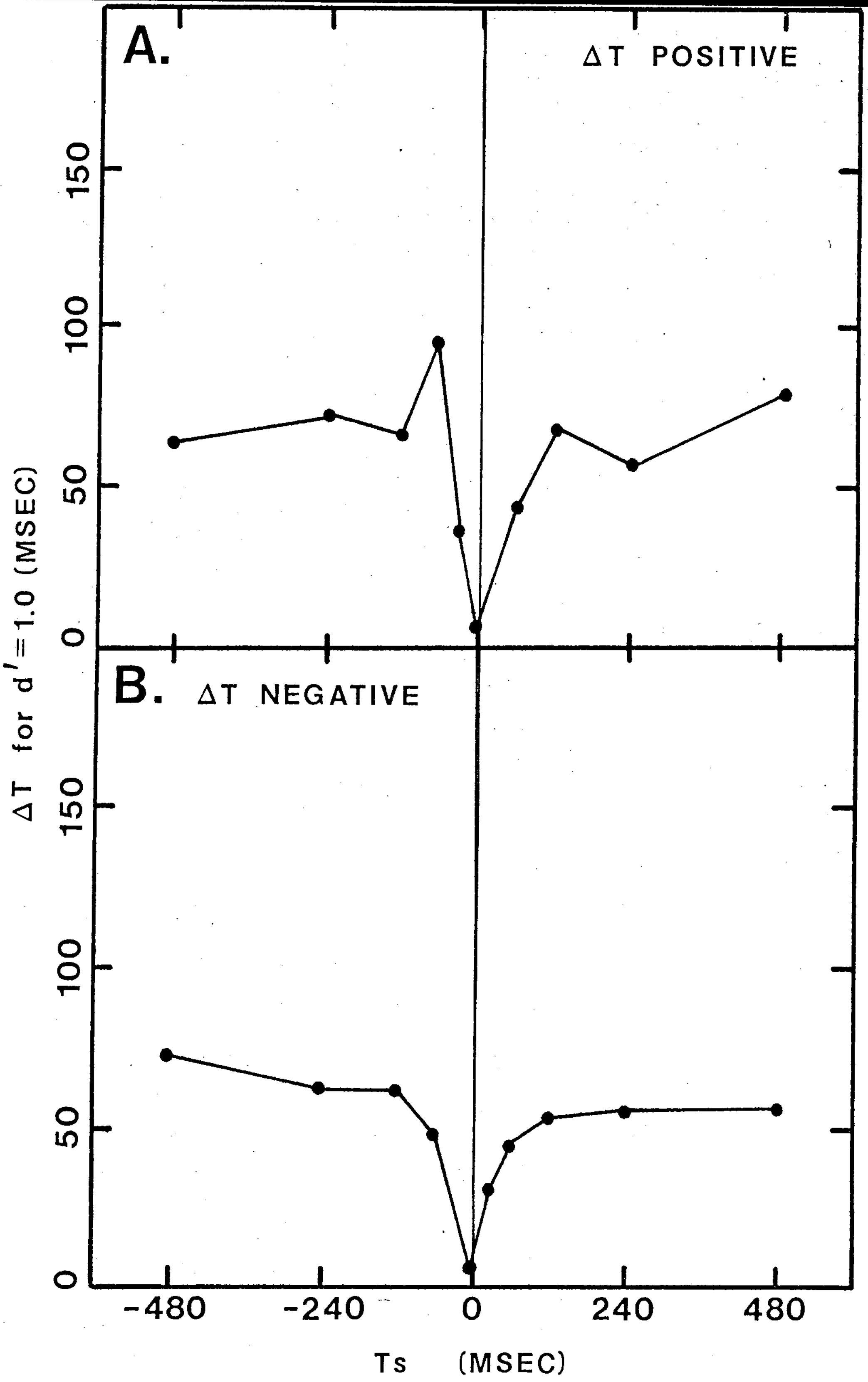

Figure 3.6. Median auditory-tactile discrimination function. The coordinates are the same as in Fig. 3.5. The points are the medians of the four individuals shown in Fig. 3.5. The median increases with Ts up to 120 msec and is approximately constant for larger values. At Ts= 0 msec, the median ΔT for d'=1.0 is less than 10 msec.

these conditions may still be understood in terms of the perceptual qualities of the pairs to be compared.

### C. Discussion.

Three aspects of the A-T conditions examined in Experiment 1 are of interest. For two of these, an explanation for the observed performance in terms of a characteristic quality of the stimulus pairs will be offered. First, when Ts=0 msec, subjects resolve either direction of change very well. In the preliminary experiment using V-A pairs, equivalent performance had never been seen. In that experiment, the smallest $\Delta T$ for d'=1.0 never occurred at Ts=0 msec and was larger than the smallest $\Delta T$ for d'=1.0 from Experiment 1. The results of cross-modality simultaneity judgments suggested that the sensory experience of true simultaneity (Fraisse, 1963) was possible for an A-T pair. By true simultaneity, it is meant that the subject does not call a pair simultaneous as a result of a process of elimination; the experience of simultaneity is compelling for those pairs. As a result, A-T simultaneity judgments are more consistent than either V-A or V-T judgments. That result is consistent with the discrimination result reported here, if it is accepted that an additional cue for discrimination arises when a change in perceived quality is associated with the change from Ts to Tc (Miller, Wier, Pastore, Kelly, and Dooling, 1976). In this experiment, that cue is the same one that sharpens the simultaneity judgments. A simultaneous standard is perceived as such, and Tc that differs by only 10 msec produces a qualitatively different sensory experience. In Chapter Two, the discussion of the interaction between labeling tendencies and the discrimination of V-A pairs considered the case in which discriminability of standard and comparison decreased if the subject tended to classify them. In this case, discriminability is enhanced by the same tendency. The difference between the two effects may be attributed to the certainty with which labels may be applied. The experience of A-T simultaneity is compelling; subjects have confidence in that judgment. For nearly-simultaneous V-A stimulus pairs, no label is applied with confidence. For stimulus elements separated by -60 to +60 msec, perceived temporal relations are ambiguous. In the V-A case, subjects probably adopt the labeling strategy because their ability to generate or store a precise representation of the standard stimulus is poor.

As noted, individual subjects recorded poor performance for certain conditions of the A-T experiment as well. Examples are seen in the data of S1 for Ts=-60 msec, $\Delta T$ positive, and in the data of S4 for Ts=+60 msec, $\Delta T$ negative. For these conditions, some factor appears to interfere with performance, but it does not seem to be related to perceived simultaneity. For these cases, the discriminability of Ts from Tc=0 msec should be high. To some extent that was true, although better discrimination would be predicted from the conditions in which Ts=0 msec. When Ts=0 msec and Tc=+60 msec, the difference is easily resolvable. But when Ts=+60 msec and Tc=0 msec, d' is not as large as in the former case. Psychometric functions for one of these two conditions are shown in Fig. 3.7. Ts=+60 msec for these data; the abscissa has been transformed to Tc, using the relation given in eq. 1.1. The psychometric function is non-monotonic for subjects 1 and 4. Discrimination performance reaches a relative maximum when Tc is approximately 0 msec, falls to a minimum when Tc is approximately -60 msec, and then increases again. The minimum occurs when the comparison pair is an exact reversal of the standard, that is, Tc=-Ts. The form of these functions suggests that these two subjects were confused by the extent of the interval and were not able to detect the order of the elements. The nature of the error was described to the subjects, but it continued to be made. This seems to be another, unanticipated case in which an irrelevent attribute of the stimulus pair interferes with the discrimination task. It was noted in Chapter two that A-T simultaneity judgments were more precise than A-T order judgments (Hirsh and Sherrick, 1961). Poor performance in this case is related to the inability to identify the order of Ts and Tc.

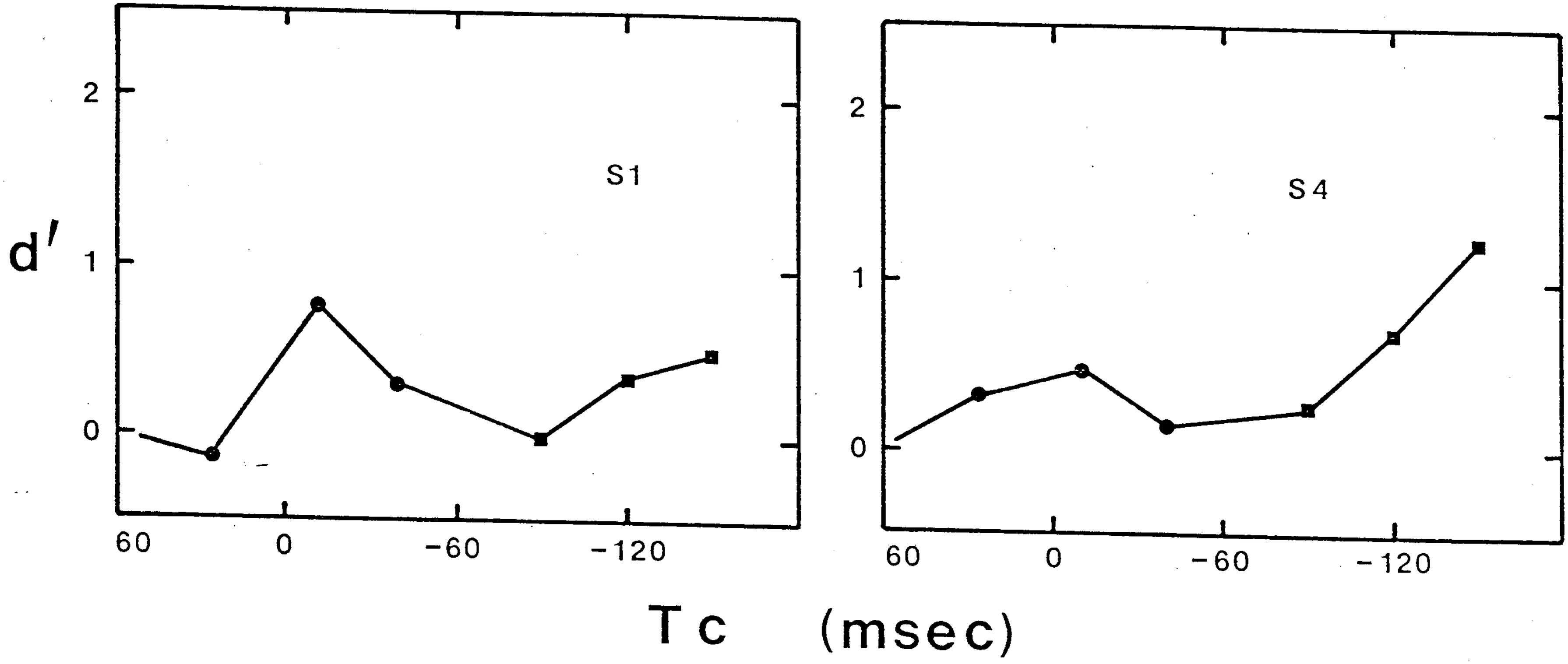

Figure 3.7. Non-monotonic psychometric functions. The abscissa, normally ΔT, has been transformed to Tc, the temporal relation in the comparison pair. Ts was +60 msec, and ΔT is negative. The functions increase initially, decrease to a minimum when Tc is approximately -Ts, then begin to increase again.

Finally, performance for large Ts can be considered. Values of Ts between 120 and 480 msec, both positive and negative, and including both directions of change, account for 12 of the 20 conditions in the experiment. The median $\Delta T$ for d'=1.0 for these conditions always falls between 54 and 78 msec. There is no tendency for the $\Delta T$ for d'=1.0 to increase with Ts over that range. The data are best summarized by stating that the just-resolvable change is a constant.

These data may be compared to the V-A discrimination experiment reported in Chapter Two. In that experiment, the $\Delta T$ for d'=1.0 was not observed to be constant over the analogous set of conditions; it was observed to increase with Ts. The absolute level of the $\Delta T$ for d'=1.0 was similar for the larger Ts in the two experiments, but performance with V-A pairs is slightly better than with A-T pairs for Ts=+120 msec and -120 msec. This difference is considered at the end of the chapter, when the V-T discrimination data have been presented.

## II. Experiment 2. Same-different judgments using visual-tactile stimulus pairs.

### A. Methods.

The psychophysical procedure and experimental design were identical to those of Experiment 1. In place of the acoustic stimulus element, a photic element was used. Positive values of Ts and Tc indicate that the visual stimulus led the tactile. The sign of $\Delta T$ again indicates the direction of change; positive values mean that the tactile element occurs later in time in Tc than in Ts. The same 20 conditions were tested in the same order. The four subjects who served in Experiment 1 also served in Experiment 2. They practiced for 12 hours with the new stimulus pair before the experiment began. It then continued through three complete cycles of the 20 conditions.

The tactile stimulus was the same as that described in section 1.A. The visual stimulus was produced by introducing the 5-msec rectangular pulse to a yellow light-emitting diode (LED). The rise/fall time of output of the LED is measured in nanoseconds. The LED was mounted in a metal enclosure such that it was viewed against a neutral gray background. It was 6 mm in diameter. Viewing distance was not controlled; at 50 cm, the LED subtended approximately 40' of visual angle. The visual stimulus was viewed binocularly. With the normal background illumination in the IAC chamber, the stimulus was presented at approximately 3 log units above its threshold, or at 30 dB SL.

Procedures for data analysis were as before. Fitted psychometric functions did not differ greatly from those of Experiment 1. Non-monotonic functions of the type discussed in section I.C were not observed in this experiment.

## B. Results.

Data were obtained from the four subjects who had previously served in Experiment 1. However, the performance of S1 in this experiment was such that good estimates of the $\Delta T$ for d'=1.0 were rarely obtained. Table 3.2 lists the best estimate of the $\Delta T$ required for d'=1.0 for each subject. For S1, it is often the case that the value listed is the only estimate that could be obtained. The data from S1 are exceptionally unreliable, and they are omitted from the remaining analysis. The median $\Delta T$ for d'=1.0 given in the table reflect the performance of S2, S3, and S4 only.

Individual psychophysical functions are shown in Fig. 3.8. These functions do not have a common minimum as did those from Experiment 1 (Fig. 3.5). In general, the $\Delta T$ for d'=1.0 is larger for this experiment, but the difference between experiments is greatest at Ts=0 msec. The functions in Fig. 3.8 reach minima at different values for different subjects. In panel a, conditions employing positive $\Delta T$ are shown. While no strong trend emerges, there is a tendency for the $\Delta T$ for d'=1.0 to be smaller when Ts is +60 or +120 msec. In panel b, the results of conditions in which $\Delta T$

## TABLE 3.2 RESULTS OF VISUAL-TACTILE SAME-DIFFERENT JUDGMENTS

| Cond: | | ΔT for d' = 1.0 (msec): | | | | |
|---|---|---|---|---|---|---|
| Ts | ΔT | S1 | S2 | S3 | S4 | Median (3 Ss) |
| -480 | + | 170 | 27 | 55 | 143 | 55 |
| -240 | + | 244 | 22 | 190 | 81 | 81 |
| -120 | + | 264 | 28 | 97 | 83 | 83 |
| -60 | + | 184 | 34 | 195 | 104 | 104 |
| -30 | + | 166 | 33 | 161 | 79 | 79 |
| 0 | + | 85 | 34 | 123 | 87 | 87 |
| +60 | + | 121 | 27 | 102 | 76 | 76 |
| +120 | + | 99 | 43 | 74 | 59 | 59 |
| +240 | + | 108 | 22 | 104 | 57 | 57 |
| +480 | + | 141 | 55 | 111 | 88 | 88 |
| -480 | - | 212 | 55 | 138 | 100 | 100 |
| -240 | - | 132 | 35 | 113 | 103 | 103 |
| -120 | - | 130 | 17 | 67 | 82 | 67 |
| -60 | - | 200 | 26 | 76 | 58 | 58 |
| 0 | - | 220 | 38 | 102 | 90 | 90 |
| +30 | - | 203 | 31 | 134 | 74 | 74 |
| +60 | - | 216 | 28 | 180 | 94 | 94 |
| +120 | - | 139 | 35 | 191 | 93 | 93 |
| +240 | - | 80 | 33 | 83 | 80 | 80 |
| +480 | - | 168 | 34 | 122 | 81 | 81 |

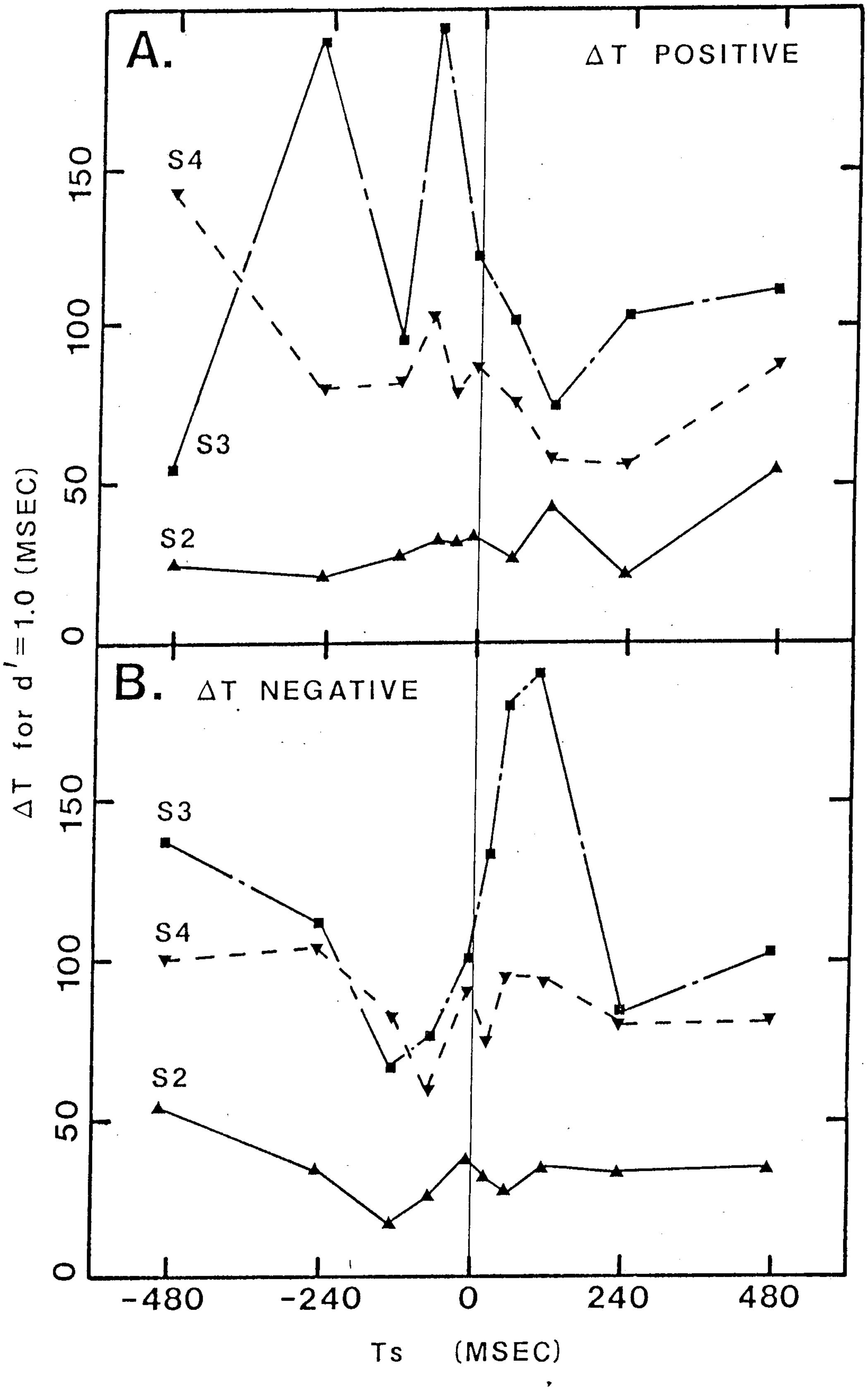

Figure 3.8. Visual-tactile discrimination functions for individual subjects. Similar to Fig. 3.5; however, no sharply-defined minimum exists at Ts=0 msec in the visual-tactile case. As noted in the text, data for one subject are not shown.

was negative, the smallest $\Delta T$ for d'=1.0 is obtained for Ts between -60 and -120 msec. These results are similar to those of the V-A experiment, reported in Chapter Two.

Again, very large changes are required for d' to equal 1.0 in certain conditions. These data are again consistent with the hypothesis that discrimination performance will decline when Ts and Tc are both taken from the range over which a stimulus pair can be called simultaneous. Regions of Ts for which relatively large $\Delta T$ for d'=1.0 are observed are Ts=-60 msec ( $\Delta T$ positive, panel a) and Ts=0 to +120 msec ( $\Delta T$ negative, panel b). The data for S3 show the greatest increase in the $\Delta T$ for d'=1.0. This subject had not had such difficulty in the A-T experiment. The psychometric functions for those conditions do not differ in form from those for other conditions. They are, except for small deviations, monotonic with $\Delta T$. Exact reversals of Ts did not cause undue confusion in this experiment, as they had in Experiment 1.

Median $\Delta T$ for d'=1.0 are plotted in Fig. 3.9. The trends described above are observable in the medians as well. Larger changes are required for d'=1.0 when Ts=-60 msec and $\Delta T$ is positive, or when Ts is +60 or +120 msec and $\Delta T$ is negative. Smaller changes are discriminable if Ts is -60 msec and $\Delta T$ is negative, or if Ts is +60 or +120 msec and $\Delta T$ is positive.

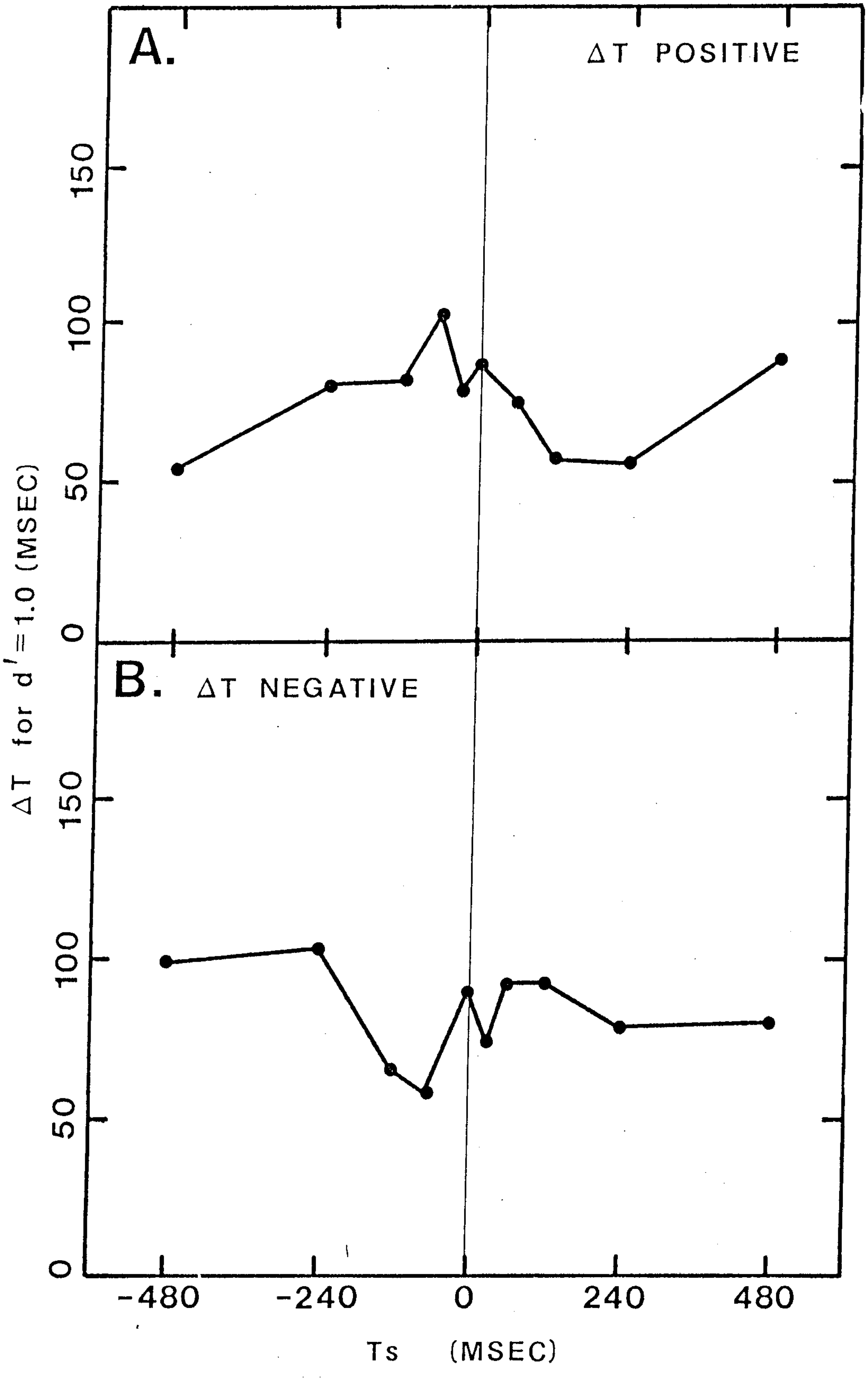

Figure 3.9. Median visual-tactile discrimination function. Similar to Fig. 3.6 for visual-tactile stimulus pairs. As noted, medians are based on three subjects.

## C. Discussion.

The general form of the psychophysical function relating discrimination performance to Ts was similar in Experiment 2 to that obtained in Experiment 1 and in the previous experiment using V-A stimulus pairs. That is, the $\Delta T$ for d'=1.0 does not increase monotonically with increases in the absolute value of Ts. The interpretation of this experimental result is the same as that previously given. When Ts and Tc are both within the range from which simultaneity judgments are made, they are not easily discriminable. When Ts is at the border of that range, smaller changes will be discriminable, provided that the direction of change is such that the absolute value of Tc is larger than that of Ts. Thus the smallest $\Delta T$ for d'=1.0 obtained in Experiment 2 was in conditions for which the absolute value of Ts is 60 to 120 msec and for which the absolute values of Tc are greater than Ts. The poorest discrimination occurs when the absolute values of the Tc are smaller than Ts.

Figure 3.10 compares the median $\Delta T$ for d'=1.0 determined in Experiments 1 and 2. First, consider Ts over the interval from -120 to +120 msec. More precise discrimination performance is observed for the A-T stimulus pairs. For this cross-modality combination, temporal resolution reaches a well-defined maximum when Ts=0 msec.

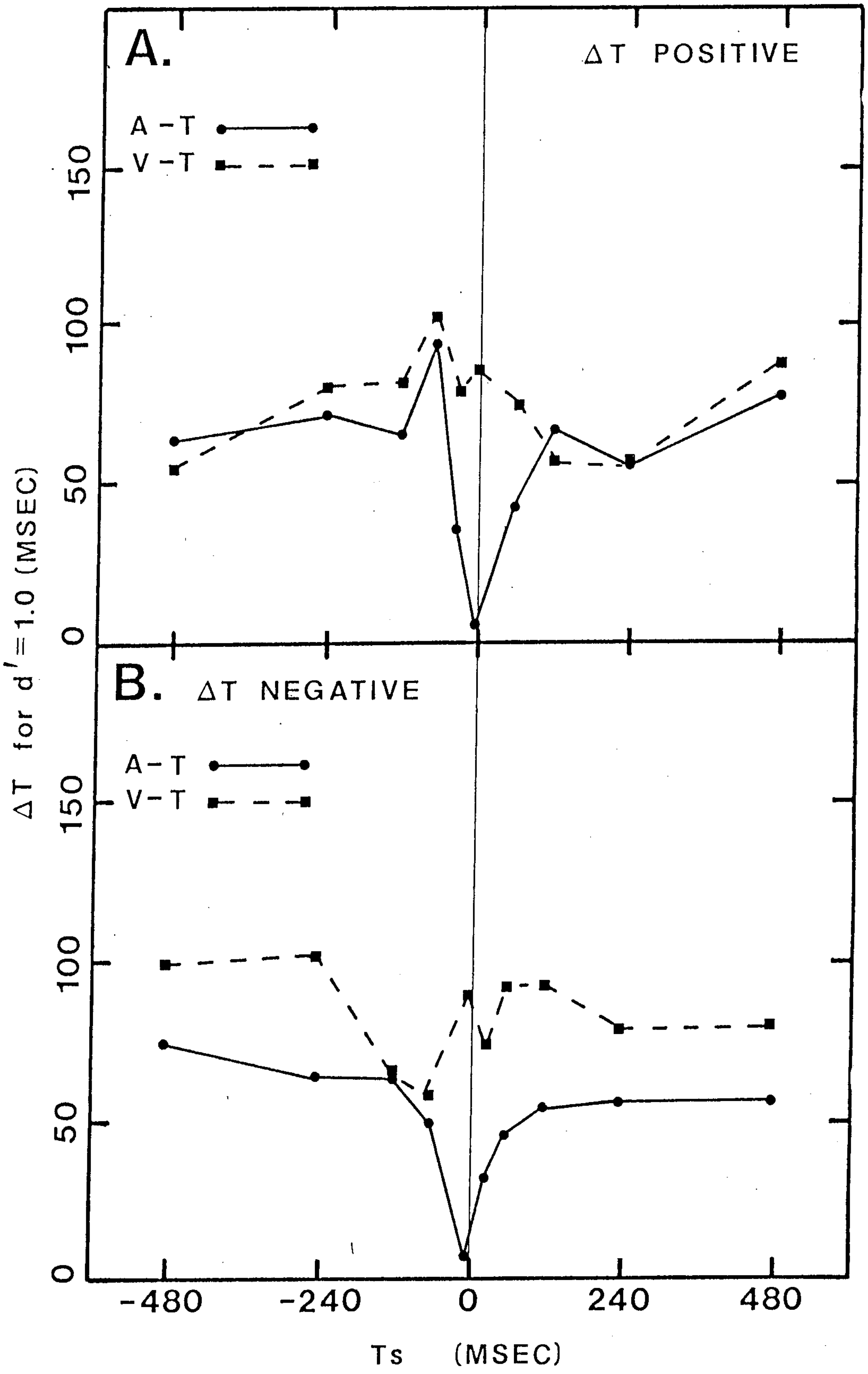

Figure 3.10. Comparison of auditory-tactile and visual-tactile median discrimination functions. The data shown in Figs. 3.6 and 3.9 are shown together. For Ts>120 msec, the ΔT for d'=1.0 is similar in each cross-modality case. The similarity is most apparent in panel a. For Ts<120 msec, auditory-tactile temporal resolution is superior.

V-T performance does not show a maximum for objective simultaneity. It shows a more gradual slope, generally negative in panel a and positive in panel b. For V-T stimuli, the best performance is seen when Ts is approximately 120-240 msec ($\Delta$T positive) and -60 msec ($\Delta$T negative).

For Ts greater than 120 msec, the A-T and V-T results are more similar than were the results for smaller values of Ts. For positive $\Delta$T, no difference between the $\Delta$T for d'=1.0 can be observed (although it is the case that the data for one observer are not represented in the data for V-T stimuli). In panel b, $\Delta$T negative, the dashed line is shifted upwards from the solid line by 20-30 msec. However, a qualitative difference between cross-modality combinations is not apparent, as it was for the conditions in which Ts was less than 120 msec. For Ts greater than 120 msec, the just-discriminable change is nearly constant, for each cross-modality combination. There is no indication from these results that a change representing a constant proportion of Ts is resolvable at a constant performance level.

The results from these two cross-modality experiments may also be compared to the V-A cross-modality experiment discussed in Chapter Two. In that experiment, performance was more similar to the performance observed in Experiment

2, using V-T stimulus pairs. A more complete consideration of the similarities is presented below.

## III. Experiment 3. Same-different judgments using within-modality stimulus pairs.

In Sections I and II, the data from discrimination experiments using two different cross-modality stimulus pairs to define Ts were discussed. In some conditions of those experiments, no difference between cross-modality combinations was observed. For other conditions, especially the smaller values of Ts, V-T discrimination was poor, relative to A-T discrimination. One means of evaluating the absolute level of cross-modality temporal resolution is to compare it to the discrimination of stimulus pairs presented to single modalities. This control experiment was conducted, using the same-different psychophysical procedure. These results may also be compared to existing reports of temporal resolving power for single modalities, to evaluate the procedures used in Experiments 1-3.

### A. Methods.

The same-different psychophysical procedure was used. The details of the procedure were described in section I.A. Minor changes were made. For the control experiments, fewer conditions were tested. For the within-modality cases, Ts was marked by a pair of identical pulses, delivered to a single transducer: the earphone, the vibrator, or the LED. For stimulus pairs marked in this way, negative values of Ts have no significance, since the elements of the pairs are indistinguishable. In the tactile (T-T) experiment, data were obtained for Ts between 30 and 480 msec, with both positive and negative $\Delta T$. The same conditions were included in the auditory (A-A) experiment; in addition Ts=15 msec was tested with positive $\Delta T$. For the visual (V-V) case, data were obtained with Ts between 30 and 480 msec and $\Delta T$ positive, and with Ts between 120 and 480 msec and $\Delta T$ negative.

Each condition was tested on one occasion only. Estimates of $\Delta T$ for d'=1.0 were taken from fitted psychometric functions based on 288 trials, as before. The A-A conditions were completed at the beginning of the series of same-different experiments. The T-T and V-V data were collected at the conclusion of the experimental series, approximately five months later. The same group of subjects participated in the cross-modality and the within-modality experiments.

### B. Results.

Since only one psychometric function per subject per condition was obtained, individual results in this experiment are less reliable than the cross-modality results, which were selected from three separate estimates.

The data shown as medians of individual performance are more systematic. The median $\triangle T$ for d'=1.0 by condition for each modality is listed in Table 3.3.

The medians are plotted as functions of Ts in Fig. 3.11. Although these results are based on fewer trials than the cross-modality results, they are generally simpler to summarize than were those data. In panel a, those conditions for which $\triangle T$ is positive, the greatest deviations from linearity are seen. The A-A and T-T results seem not to obey Weber's Law; the just-discriminable change for Ts from 60 to 480 msec is approximately constant. In all three modalities, however, the best performance is observed for the smallest values of Ts. In panel b, conditions with $\triangle T$ negative, the discrimination functions are well fit by straight lines. The slope of the fitted line is equivalent to the Weber fraction. For lines fit by eye, the slopes are 0.13 for A-A and T-T discrimination, and 0.25 for V-V discrimination. Each of these is within the range of values reported previously (Woodrow, 1951), although few visual interval discrimination results are available. The A-A case has been studied more often. The present data are similar to those presented by Abel (1972a). For this direction of change as well, the best performance occurs when Ts is smallest.

TABLE 3.3 RESULTS OF WITHIN-MODALITY SAME-DIFFERENT JUDGMENTS

| Cond: | | Median $\Delta T$ for $d'=1.0$ (msec): | | |
|---|---|---|---|---|
| Ts | $\Delta T$ | A-A | T-T | V-V |
| +15 | + | 7 | - | - |
| +30 | + | 21 | 10 | 21 |
| +60 | + | 41 | 28 | 10 |
| +120 | + | 55 | 39 | 57 |
| +240 | + | 56 | 44 | 80 |
| +480 | + | 46 | 49 | 122 |
| +30 | - | 15 | 14 | - |
| +60 | - | 15 | 22 | - |
| +120 | - | 35 | 24 | 24 |
| +240 | - | 49 | 46 | 49 |
| +480 | - | 70 | 69 | 120 |

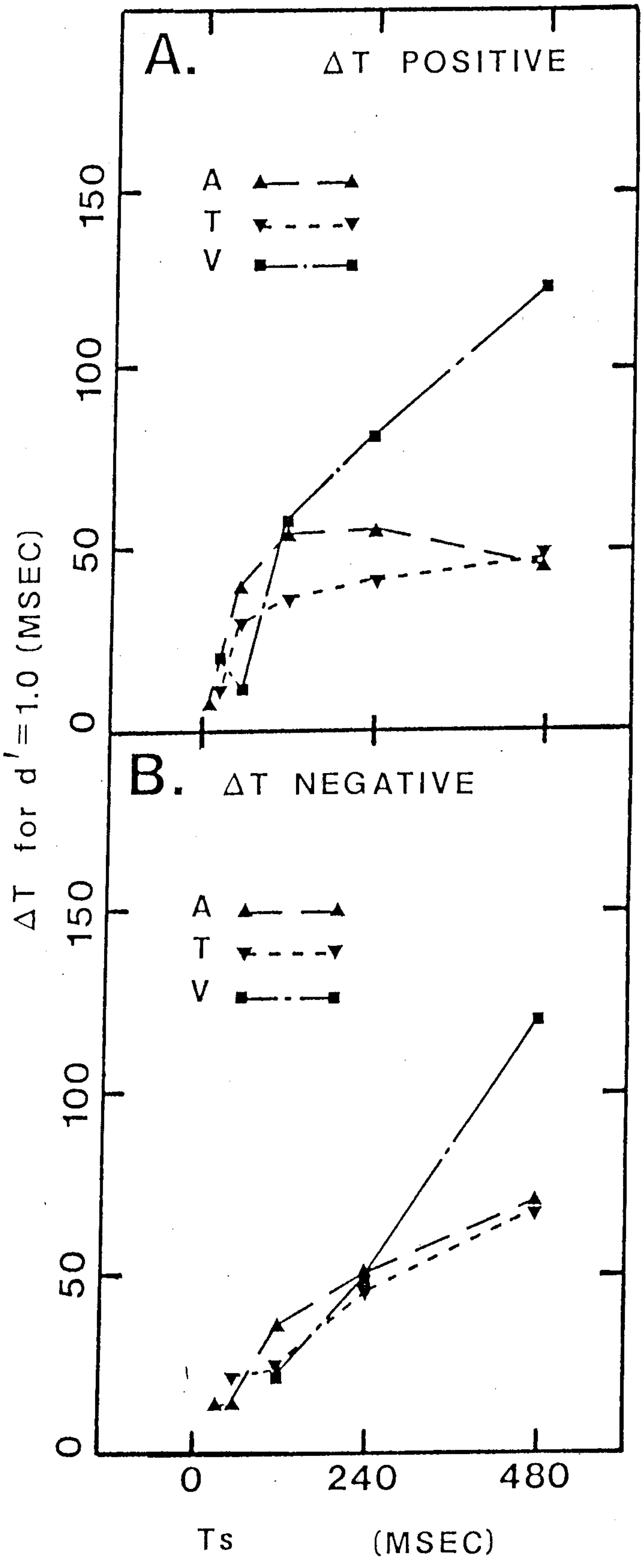

Figure 3.11. Within-modality median discrimination functions. Only positive values of Ts were used. Differences between modalities are greatest at Ts=480 msec. For negative ΔT, a peak in the ΔT for d'=1.0 is not seen.

## C. Discussion.

For small Ts, the median $\Delta T$ for d'=1.0 in within-modality temporal discrimination is slightly lower than that observed in the best cross-modality cases, those using A-T pairs. It is therefore much smaller than the $\Delta T$ for d'=1.0 for V-T and V-A temporal resolution. The $\Delta T$ for d'=1.0 is, with minor exceptions, a monotonic function of Ts, and it always drops sharply as Ts approaches zero. In this respect, A-T temporal resolution is similar to within-modality temporal resolution, and V-T and V-A temporal resolution are clearly different. At larger Ts, the control results are unfortunately too unreliable to permit strong conclusions about relative performance. The $\Delta T$ for d'=1.0 within modalities appears to be slightly larger than the $\Delta T$ for d'=1.0 for single modalities, but it is likely that no real difference exists. Within-modality temporal resolution, measured with the same-different procedure, is probably slightly more precise than the A-T cross-modality case. It also appears to be unconfounded by the tendency to categorize stimuli that has been hypothesized to interfere with the cross-modality cases. It should be noted that the values of Ts and Tc used in the control experiments were large enough to be experienced as successive events. No errors arising from order-labeling could occur, since the elements were identical. Errors resulting from the use of the label, simultaneous, might have occurred if smaller values of Ts had been presented.

Evidence that the within-modality results constitute a valid baseline for evaluating the cross-modality results can be found in the Abel (1972a) experiment, which measured the resolution of temporal gaps marked by noise bursts. Abel used standard intervals over approximately the same range as the present study. The results of the two experiments are very similar. However, estimates of auditory temporal resolution from this and Abel's experiments are not in agreement with those of Getty (1976) and Divenyi and Danner (1977). The latter two experiments found smaller changes to be discriminable. As noted, the present A-A experiment was conducted first in the series, when the subjects had only a few hours of practice. The difference in results thus may reflect to some extent the experience of the listeners.

## IV. Conclusion

At this point, a brief summary of all experiments using the same-different procedure is given. Six cases have been studied: Ts has been defined by A-T, V-T, and V-A cross-modality stimulus pairs, and by A-A, T-T, and V-V within-modality pairs. Of these, the V-A experiment was conducted at CID. The other data were collected at The Boys Town Institute, using a single group of subjects. More replications of each condition were obtained from the subjects in the CID group; otherwise, there are only insignificant differences in the way the experiments were conducted at the two laboratories.

Two forms of the median psychophysical function are distinguished by the magnitude of the $\Delta T$ for d'=1.0 at Ts=0 msec. For within-modality and A-T discriminations, the $\Delta T$ for d'=1.0 decreases rapidly as Ts approaches 0 msec. These cases are shown in Fig. 3.12. The functions in this figure were shown previously in the chapter. For the longer Ts, visual temporal resolution seems poorer, but it is not possible to make inferences about relative levels of performance in the A-T and within-modality cases. In panel b of Fig. 3.12, a clear superiority of within-modality temporal resolution is seen when Ts equals +30 to +120 msec. Within this range, local minima in cross-modality temporal resolution are often observed, as was discussed above. No such non-monotonicities occur in the psychophysical functions for the within-modality conditions.

Figure 3.13 presents the other form of the discrimination function. In these cases, V-T and V-A, the smallest $\Delta T$ for d'=1.0 never occurs at Ts=0 msec, and characteristic peaks and valleys in the functions are seen. For most other stimulus dimensions, the just-resolvable change is a linear, monotonic increasing function of the magnitude of the standard stimulus, as predicted by Weber's

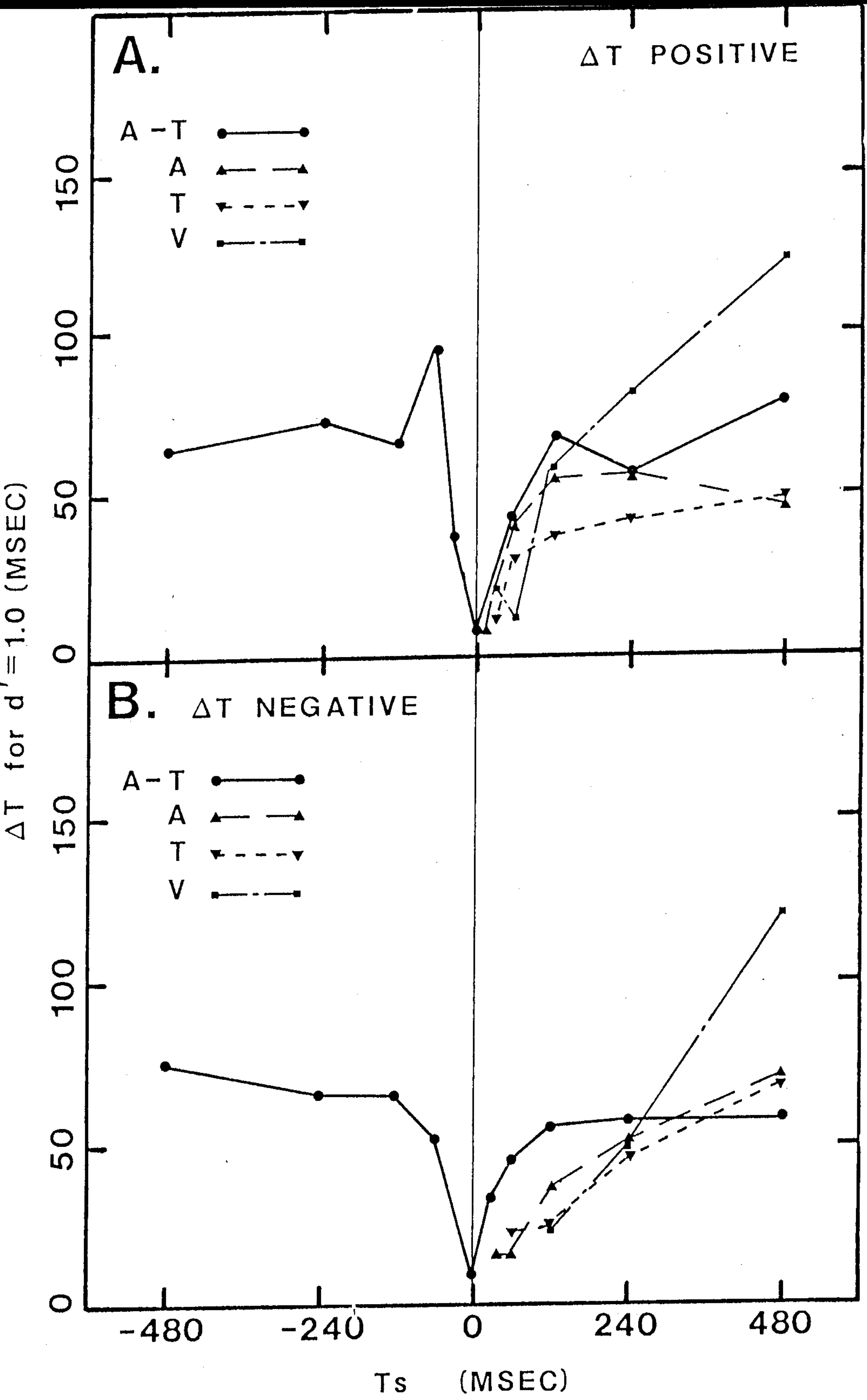

Figure 3.12. Comparison of auditory-tactile and within-modality discrimination functions. The medians of ΔT for d'=1.0 are plotted. These modality combinations are characterized by the discriminability of small changes at Ts=0 msec and by the simple relation between the ΔT for d'=1.0 and Ts.

Figure 3.13. Comparison of visual-tactile and visual-auditory discrimination functions. Median $\Delta T$ for $d'=1.0$ are plotted. The visual-auditory data are from the experiment conducted at CID. Minima in these data never occur at Ts=0 msec. Features common to the two sets of functions include the presence of discrimination minima at Ts=+60 or +120 msec (panel a) or -60 or -120 msec (panel b), and maxima at Ts=-60 msec (panel a) or +60 msec (panel b).

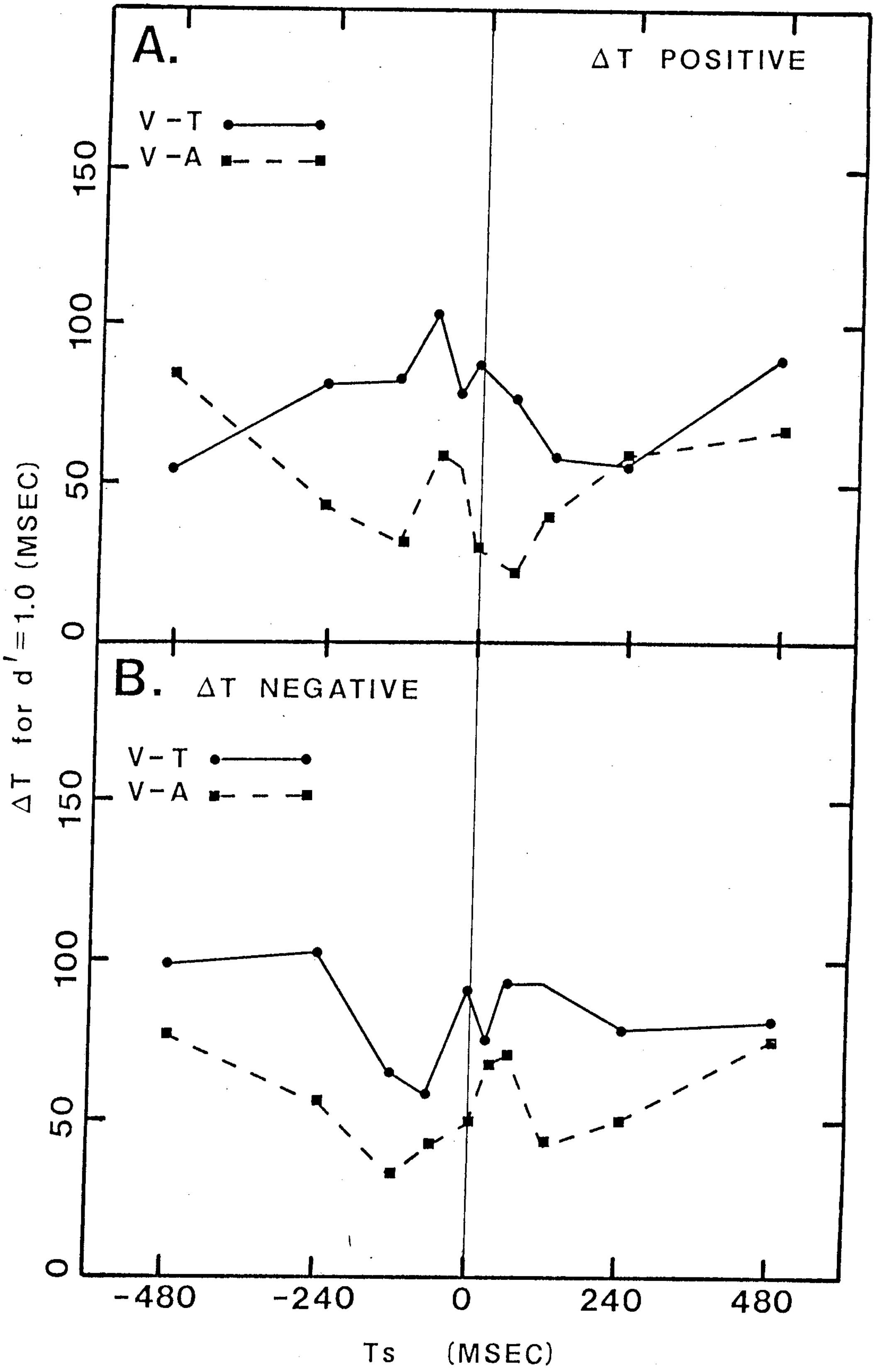
A.
ΔT POSITIVE
V-T
V-A
B.
ΔT NEGATIVE
V-T
V-A
ΔT for d′=1.0 (MSEC)
0
50
100
150
-480
-240
0
240
480
Ts (MSEC)

Figure 3.13

Law. In these cases, however, it has been hypothesized that the internal representation of Ts may be ambiguous with respect to the exact extent of the standard, or memory for its extent may be poor. In any case, the data are similar to those seen in instances of categorical perception (Liberman, Harris, Kinney, and Lane, 1961; Miller et al., 1976), in which the just-resolvable change is larger or smaller than predicted by Weber's Law, depending on the location of the stimuli to be discriminated with respect to the stimulus categories. Pairs that elicit the same label may thus become difficult to discriminate, while pairs normally given different labels may be easily resolved. In the figure, peaks occur at Ts=-60 and $\Delta T$ positive, and at Ts between 0 and +120 when $\Delta T$ is negative. The smallest $\Delta T$ for d'=1.0 occurs in these functions when Ts is +60 or +120 msec, $\Delta T$ positive, and at Ts equal to -60 or -120 msec and $\Delta T$ also negative.

The interpretation offered for these results has relied heavily on assumptions about the observer's strategy. If performance is strongly affected by strategy, it is obviously possible that the observer is capable of making more precise discriminations under some other psychophysical procedure. The same-different procedure may not be the best for unbiased measurements of temporal resolution. Additional experiments were designed to investigate that possibility.

# CHAPTER FOUR. Cross-modality temporal resolution measured with a forced-choice procedure.

The poor temporal resolution measured on certain conditions of Experiments 1 and 2 was interpreted as resulting from the use of a labeling strategy when the stimuli presented as standard and comparison could not be identified unambiguously. It was suggested at the conclusion of Chapter Three that with other psychophysical procedures, estimated temporal resolution might be greater in those difficult conditions. In Experiment 4, a three-alternative forced-choice adaptive procedure was used in an attempt to neutralize the effects of criterion and to focus the subjects' attention on the relevant aspect of the stimulus pair. The same-different procedure requires the subject to keep a constant criterion for responding "different". In the review of labeling experiments, it was often noted that judgments of perceived temporal relations are quite variable and are subject to influences of attention, as reflected in the prior entry effect. It was also seen that a single presentation of a stimulus pair cannot be labeled as precisely as can a pair presented repeatedly (Gengel and Hirsh, 1970). The experiments reported in Chapters Two and Three do not explain effects such as prior entry, but they do seem to indicate that the criterion for sameness may be unstable over the same range of Ts for which the other effects are observed (Stone, 1926).

To reduce the criterion problem in psychophysics, it is common to use a forced-choice procedure, with a signal presented on each trial. The two-alternative forced choice (2AFC) procedure is common; it was used in several of the temporal discrimination experiments reported in Chapter One. However, in 2AFC, the subject must be able to describe the aspect of the signal that differentiates it from the non-signal alternative of the trial. For example, he would have to know that Tc was a longer interval than Ts. It seemed that in experiments in the current group, the 2AFC procedure might encourage subjects to label the stimulus pairs. Therefore, the three-alternative forced choice (3AFC) procedure was used, under the traditional assumption that the odd alternative, the one containing the Tc, could be identified without being labeled and with minimal demand on the subject to maintain a constant criterion for responding. The procedure was used in conjunction with a Levitt (1971) stepping rule to enable the subjects to work at the stimulus level that was just-resolvable for them. The rule chosen was one that causes $\Delta T$ to converge on the value that can be identified on 71% of the trials. For three alternatives, this corresponds to a d' of 1.2 (Elliott, 1959). This is a slightly higher performance level than was followed in Experiments 1 and 2. However, the $\Delta T$ for the same level of performance in those experiments can be interpolated from the psychometric functions.

## Experiment 4. Temporal resolution for auditory-tactile and visual-tactile stimulus pairs measured in a forced-choice procedure.

### A. Methods.

A three-alternative forced-choice (3AFC) procedure was used. The symbol conventions, apparatus, and stimulus elements were unchanged from the previous experiments. A positive value of Ts or Tc indicates that A or V leads T. A positive $\Delta T$ indicates that the tactile element occurs later in time in Tc than in Ts.

The 3AFC trial sequence is diagrammed in Fig. 4.1. In the figure, Tc occurs in alternative 2; it could occur as each alternative with probability of 0.33. The interpair interval was randomized as in Experiments 1 and 2. The base interval was 600 msec. An additional interval, I, selected with uniform probability from 30 to 200 msec, was added between alternatives 1 and 2. Between alternatives 2 and 3, 300-I msec was added. The response interval began 10 msec after the second element of the third pair had been presented and continued until a button was pressed. Feedback lights identified the correct alternative after a response had been recorded.

The data were collected in 96-trial blocks. The starting $\Delta T$ for the second and later blocks in a session

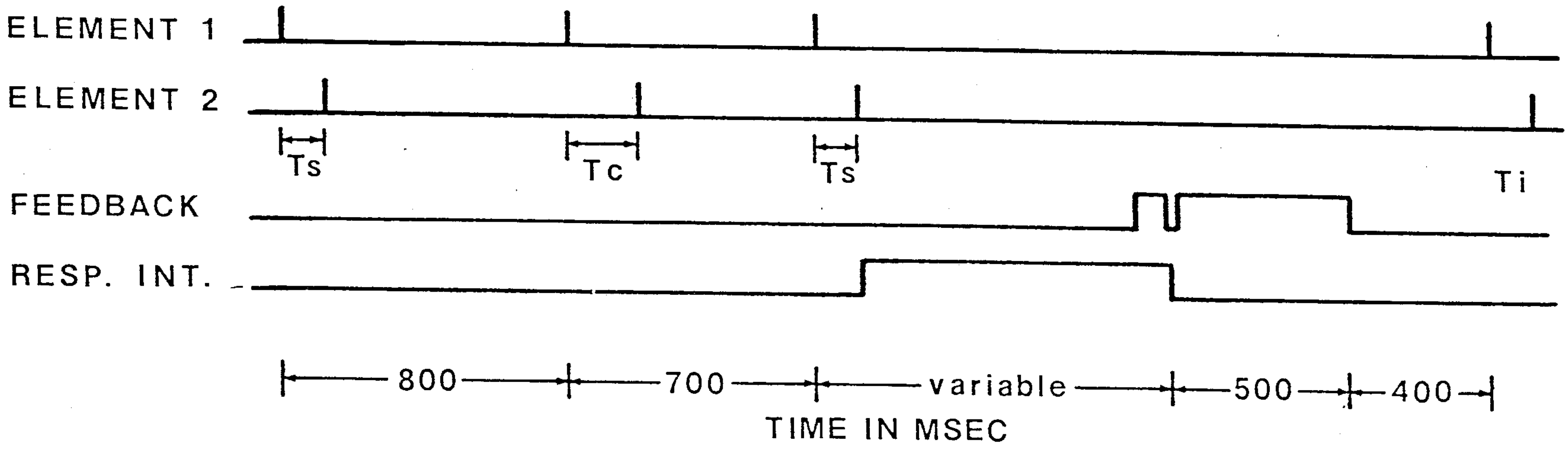

Figure 4.1. Three-alternative forced-choice trial sequence. In the example shown, the correct alternative is two. Ts, 120 msec, occurs in alternatives one and three; Tc, 200 msec, is in alternative two. The interpair intervals were randomized, with the constraint that the total elapsed time from the onset of the first pair to the onset of the third pair was always 1500 msec. The response interval began 10 msec after the trailing element of the third pair and continued until a button was pressed. Feedback lights associated with the response box buttons were used. When a button was pressed, a brief (100 msec average) light identified the subject's response. A longer light, 500 msec, identified the correct response. The next trial was preceded by 400 msec of delay. In the figure, the first pair from the next trial is labeled Ti to indicate that it may be either Ts or Tc.

was the same value at which the previous block had ended. Three to four blocks of trials were obtained per 30-minute session. A condition was tested for at least two consecutive sessions. The track was continued in the later sessions; that is, the starting level was always the same $\Delta T$ at which the previous run had stopped. Step size could be changed by the experimenter; based on the form of the track, changes were made at the beginning of the later sessions.

The initial level of the comparison stimulus, Tc, was chosen to be easily discriminable. Tc was then changed according to a Levitt (1971) stepping rule. A rule that converges on $p(c)=0.71$ was used. To determine the $\Delta T$ that is resolvable at this level of performance, the absolute value of $\Delta T$ was reduced by a fixed step size after two consecutive correct responses. The absolute value of $\Delta T$ was increased after each incorrect response. The step size was arbitrarily chosen by the experimenter, based on the absolute level of $\Delta T$ and whether the track was approaching its asymptotic level. The most common starting step size was 10 msec; smaller steps were used when $\Delta T$ was small and had stabilized. In general, the standard error of the estimate of the asymptotic level was proportional to the step size.

The group of subjects who participated in Experiments 1-3 also participated in Experiment 4. Following at least 1500 training trials, data were collected in two A-T conditions and one V-T condition for each subject. All subjects were tested with Ts=0 msec and negative $\Delta T$ in the A-T case. Performance in this condition had been uniformly good in Experiment 1; it was included for two reasons. First, it allowed the procedures to be compared in a condition for which little difficulty had previously been observed. If different estimates of the just-discriminable change were obtained in the second condition but not this one, it could be concluded that the improvement was not just a constant difference between procedures. Second, since this condition was highly resolvable, its inclusion in this experiment was seen as an opportunity to trace improvement, if any, with additional training. Each subject also was tested in a condition from each cross-modality case (A-T and V-T) for which the $\Delta T$ for $d'=1.2$ in Experiment 1 or 2 had been relatively large and for which it was suspected that the labeling strategy had been used.

### B. Results.

For this experiment, the data from individual subjects are first considered separately, since the relevant comparisons are between the $\Delta T$ for $d'=1.2$ from two psychophysical procedures. The data are presented in the form of stimulus tracks, which display $\Delta T$ as a function of time. In a typical stimulus track, the starting level of the stimulus will be high enough to be easily discriminable. The stimulus magnitude decreases toward an asymptote, at which performance is relatively stable and is close to the target level specified by the stepping rule. The $\Delta T$ for $d'=1.2$ then is given by a summary value that reflects the asymptotic stimulus level. For the rule used here, that is the level required for 71% correct responses, or $d'=1.2$ (Elliott, 1959). The mean of the values of $\Delta T$ at which successive reversals (one in each direction) of the slope of the track occur is an estimate of the target value of $\Delta T$. The grand mean of a number of such midpoints is used as the estimate of the $\Delta T$ for $d'=1.2$. For example, if a track begins with $\Delta T=200$ msec, decreases to 50 msec, increases to 100 msec, decreases to 65 msec, and finally increases to 125 msec, the mean of midpoints is obtained as follows. The first two reversals occur when $\Delta T=50$ and 100 msec; the midpoint is 75 msec. The next two reversals occur at $\Delta T=65$ and 125 msec; the midpoint is 95 msec. The mean of these two midpoints is then 85 msec. The midpoints of increasing segments only are used to ensure that the estimates are independent. Since the midpoints are successive estimates of the stimulus level required for the criterion level of performance, the standard deviation of the midpoints is the standard error of the estimate.

### a. Results for S1.

Data were obtained for S1 in three conditions. For the A-T case, the conditions were Ts=-60 msec, $\Delta T$ positive, and Ts=0 msec, $\Delta T$ negative. In the V-T case, Ts=-120 msec, $\Delta T$ positive was tested.

For Ts=-60 msec in the A-T experiment, 1056 trials were run. The track was started with $\Delta T$=200 msec and the fixed step size was 10 msec. For the final session (n=288 trials), the step was reduced to 5 msec. The track for all 1056 trials is shown in Fig. 4.2a. In this figure and in the other stimulus tracks presented in this chapter, the ordinate is the absolute value of $\Delta T$, the stimulus level selected by the adaptive procedure, and the abscissa is the number of the trial. Points have been plotted for those trials on which the slope of the track changes sign. The plotted values of $\Delta T$ are the values that are used to calculate the mean stimulus level required for d'=1.2. The points to the right of the track are the target level of the track and the mean $\Delta T$ for d'=1.2 from experiment 1 or 2, each with its associated standard error.

Two aspects of this track may be noted. First, no long term decrease in $\Delta T$ is observed. The final level is actually higher than the initial level. In cases such as this, no asymptote can be estimated. The mean $\Delta T$ and standard error are taken from the portion of the track

Figure 4.2. Stimulus tracks from the adaptive procedure for S1. For this and the three following figures, the abscissa is trial number and the ordinate is the stimulus value, $\angle T$, presented on that trial. The target level of $\Delta T$ is that which is discriminable at a $d'$ of 1.2, or 71% correct. The vertical lines indicate the breaks between sessions. The points to the right of each track are the target level ± one standard error reached in the final adaptive session (circle) and the mean ± one standard error of 1-3 estimates of the $\Delta T$ for $d'=1.2$ obtained in Experiment 1 or 2 (square). Error bars are not shown if they are smaller than the symbol. An arrow indicates that only one previous estimate was available. Panel a: A-T, Ts= -60 msec. $\Delta T$ positive. N=1056 trials. Panel b: A-T, Ts= 0 msec. $\Delta T$ negative. N=576 trials. Note that the ordinate of panel b is different from panels a and c. Panel c: V-T, Ts= -120 msec. $\Delta T$ positive. N=960 trials.

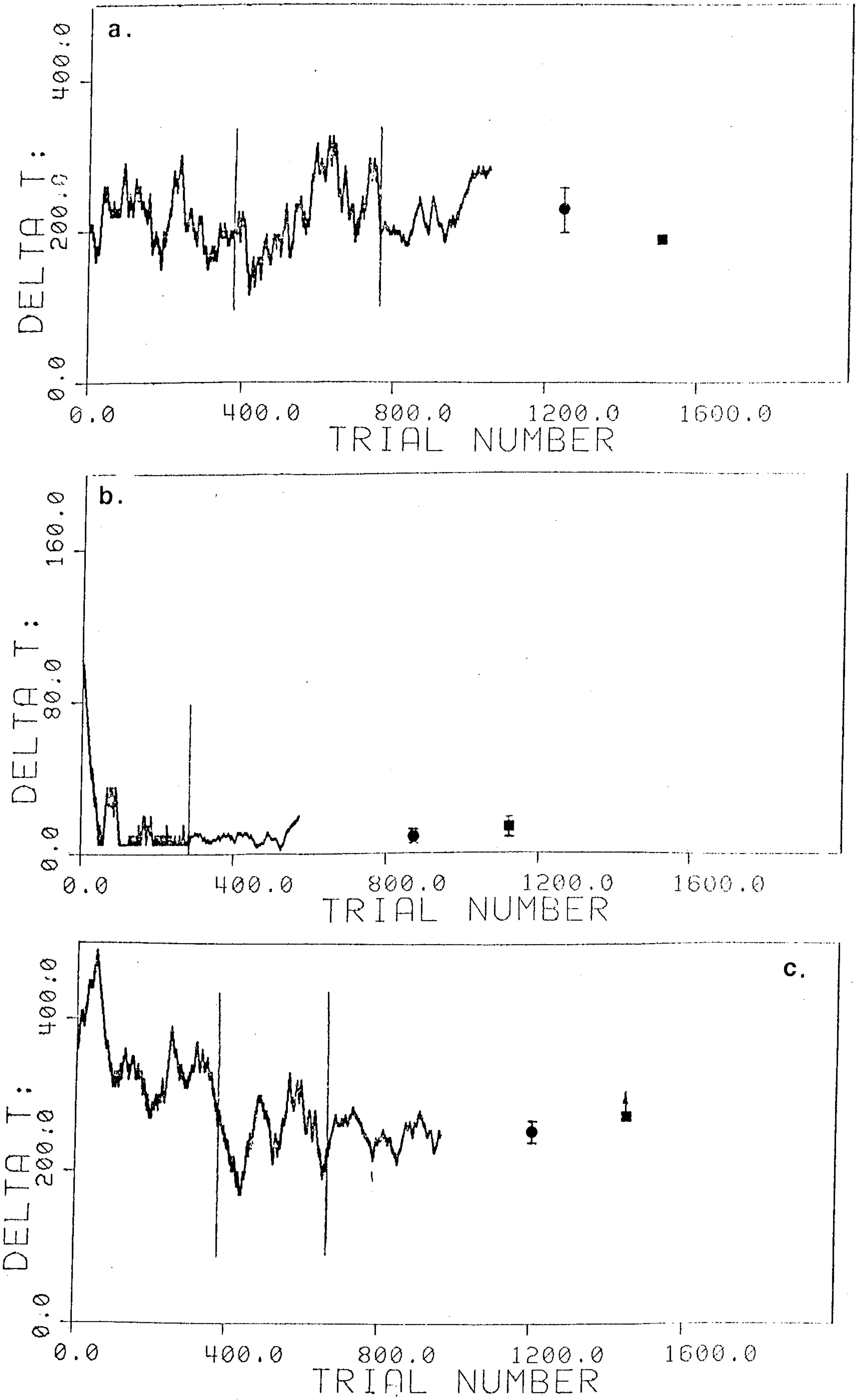
a.
DELTA T:
0.0
200.0
400.0
0.0
400.0
800.0
1200.0
1600.0
TRIAL NUMBER
b.
DELTA T:
0.0
80.0
160.0
0.0
400.0
800.0
1200.0
1600.0
TRIAL NUMBER
c.
DELTA T:
0.0
200.0
400.0
0.0
400.0
800.0
1200.0
1600.0
TRIAL NUMBER

Figure 4.2

obtained in the final session. For this condition, the mean $\Delta T$ for $d'=1.2$ is 227 ± 34 msec. That value does seem to summarize the track. The mean $\Delta T$ from this procedure may be compared to the value interpolated from the psychometric functions fitted in Experiments 1-2. In this case, the $\Delta T$ for $d'=1.2$ was 191 msec, with a standard error of 0.9 msec. Given the large standard error in the stimulus track, it cannot be concluded that a real difference between the two estimates exists. It is certainly clear that the hoped-for decrease in the just-discriminable change in the adaptive procedure was not obtained.

The second feature of the track is its instability. The track does not seem to have negative slope, as it would if the asymptote had not been reached, but it does not appear to have stablilized at one level either. At trial number 400, $\Delta T$ had been reduced to 110 msec. However, the subject could not keep the stimulus at that level, and over the next 200 trials, it steadily increased to over 300 msec. In Appendix One, a method for predicting the excursions in $\Delta T$ from the subjects' psychometric functions obtained in Experiments 1-2 is described. Statements about the variability of $\Delta T$ under the adaptive procedure are based upon comparisons to tracks simulated using that procedure.

The track for the other A-T condition is easier to interpret. The track is shown in Fig. 4.2b. In this condition, Ts=0 msec and negative $\Delta T$, $\Delta T$ drops rapidly from its initial value of 100 msec to 5 msec, the smallest $\Delta T$ possible in the first 288 trials, since the experimental program did not allow $\Delta T$ of 0 msec to be presented. The flat portions of the track in Fig. 4.2b are for those trials on which a level decrease was appropriate under the Levitt rule but not under the constraint described. For the final 288 trials, the step size was reduced to 1 msec. The appearance of the track changes at that point, as can be seen in the figure. Values of $\Delta T$ between 5 and 1 msec were presented as required in the final session. Over those trials, the track was stable; the mean $\Delta T$ for d'=1.2 was 9 ± 3 msec. For comparison, the mean $\Delta T$ for d'=1.2 from Experiment 1 was quite similar, 13 ± 5 msec.

The track for the single V-T condition is shown in Fig. 4.2c. The condition was Ts=-120 msec, with $\Delta T$ positive. The later portions of this track are more stable than the track in Fig. 4.2a, but there is still a rather prominant decrease in $\Delta T$ to a level that cannot be maintained. The mean $\Delta T$ for d'=1.2 here was 253 ± 15 msec. Only one estimate of the just-discriminable change from Experiment 2 is available for this subject; from that psychometric function, d' was 1.2 when $\Delta T$ was 272 msec. The variability of the estimate is large but cannot be quantified. There is again good agreement between estimates of performance in the two procedures.

b. Results for S2.

Data were obtained for S2 in the same A-T conditions, and for Ts=0 msec, $\Delta T$ negative in the V-T case. For S2, estimates of the $\Delta T$ for d'=1.2 in Experiments 1 and 2 had been consistently lower than those obtained for the other subjects. In particular, S2 had none of difficulty with nearly-simultaneous Ts and Tc that the other subjects did. For this reason, other criteria were used to select the V-T condition for study in Experiment 4. The simultaneous standards from both cross-modality pairs were used, to demonstrate further that a difference in capability exists across the two cross-modality cases.

The data for 672 trials with A-T, Ts=-60 msec, $\Delta T$ positive, are shown in Fig. 4.3a. The mean $\Delta T$ for d'=1.2 from Session 2 is 31 ± 9 msec. For S2, the mean from Experiment 1 was 20 ± 3 msec.

For Ts=0 msec and $\Delta T$ negative in the A-T combination, the track is shown in Fig. 4.3b. A total of 1056 trials was obtained in this condition. Fluctuations in the level were reduced in the final session when the step size was reduced. For that session, the mean $\Delta T$ is 13 ± 3 msec. For comparison, the $\Delta T$ for d'=1.2 estimated from Experiment 1 was 13 ± 0.6 msec. The agreement between procedures here is excellent.

Figure 4.3. Stimulus tracks for S2. Symbols are the same as in Fig. 4.2. Panel a: A-T, Ts= -60 msec. ΔT positive. N=672 trials. Panel b: A-T, Ts= 0 msec. ΔT negative. N=1056 trials. Panel c: V-T, Ts=0 msec. ΔT negative. N=576 trials.

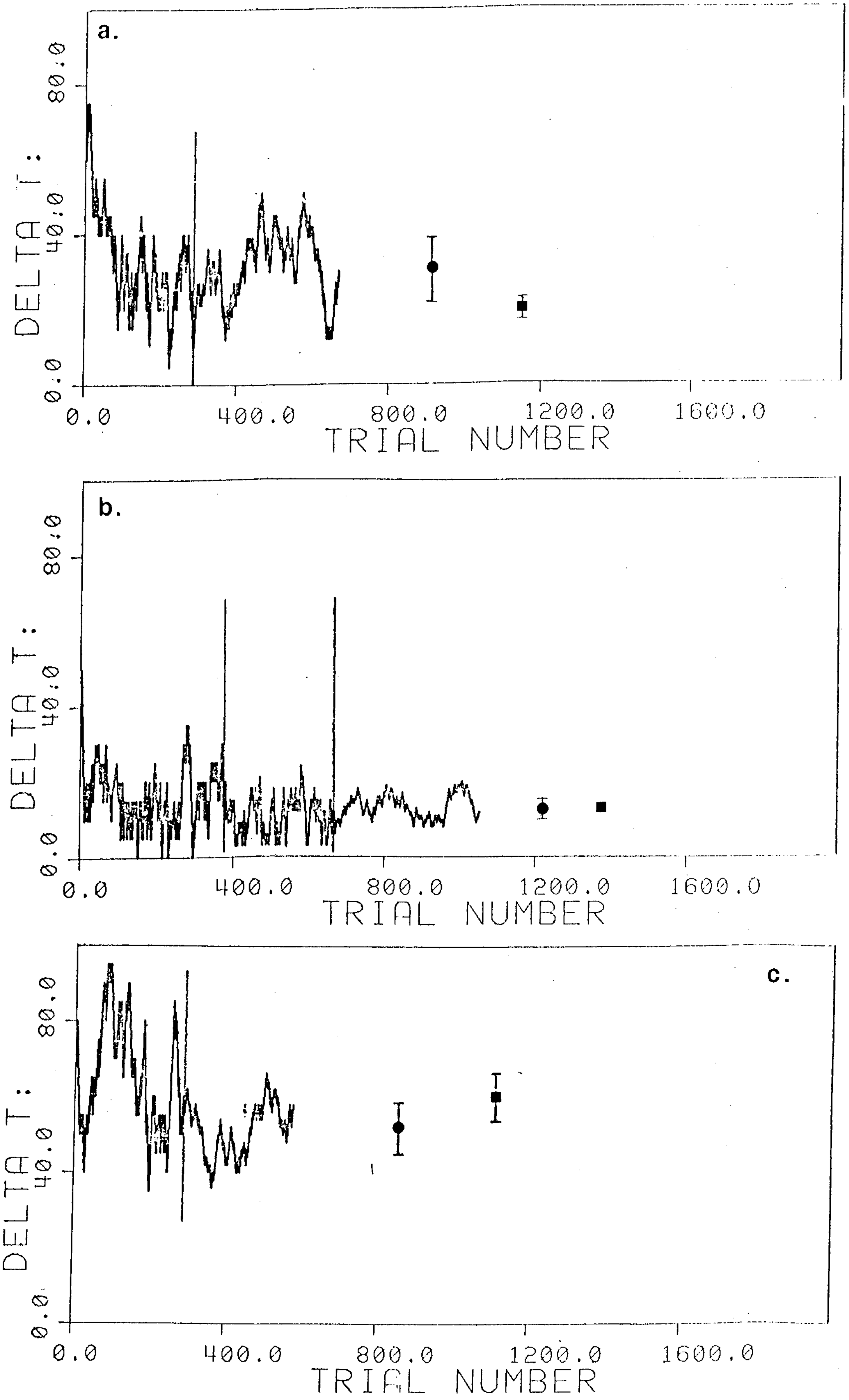
a.
DELTA T:
0.0
40.0
80.0
0.0
400.0
800.0
1200.0
1600.0
TRIAL NUMBER
b.
DELTA T:
0.0
40.0
80.0
0.0
400.0
800.0
1200.0
1600.0
TRIAL NUMBER
c.
DELTA T:
0.0
40.0
80.0
0.0
400.0
800.0
1200.0
1600.0
TRIAL NUMBER

Figure 4.3

The data from the V-T condition are plotted in Fig. 4.3c. Ts in this condition is 0 msec, as noted, and $\Delta T$ is negative. The final level of $\Delta T$ was 51 ± 7 msec. The corresponding $\Delta T$ for d'=1.2 estimated from Experiment 2 was 60 ± 6 msec. Again, quite similar results were found. The just-discriminable change for this condition is larger than the just-discriminable change for the corresponding A-T condition by 38 msec, nearly a factor of four.

c. Results for S3.

The same A-T conditions were also presented to S3. In the V-T case, Ts=-60 msec was used, with $\Delta T$ positive. The stimulus track for the first A-T condition, Ts=-60 msec and positive $\Delta T$, is shown in Fig. 4.4a. A total of 672 trials is shown. The form of the track here is similar to the form seen for S1; there is no consistent decrease in the $\Delta T$ presented. The track has segments of steep slope, both positive and negative. Once again, attention seems to fluctuatate over time; the level of $\Delta T$ can be driven down but cannot be kept at the low level. The mean $\Delta T$ for d'=1.2 from the second session is 186 msec, but data from the second session suggest poorer temporal resolution than do those from the first. The final few trials return to the level that characterized the first session. Accordingly, the mean level from the first session has been used as the summary statistic for this track. The mean $\Delta T$ was 136 ± 26

Figure 4.4. Stimulus tracks for S3. Symbols are the same as in Figure 4.2. Panel a: A-T, Ts= -60 msec. $\Delta T$ positive. N=672 trials. Panel b: A-T, Ts= 0 msec. $\Delta T$ negative. N=1248 trials. Note that the ordinate of panel b is different from panels a and c. Panel c: V-T, Ts= -60 msec. $\Delta T$ positive. N=960 trials.

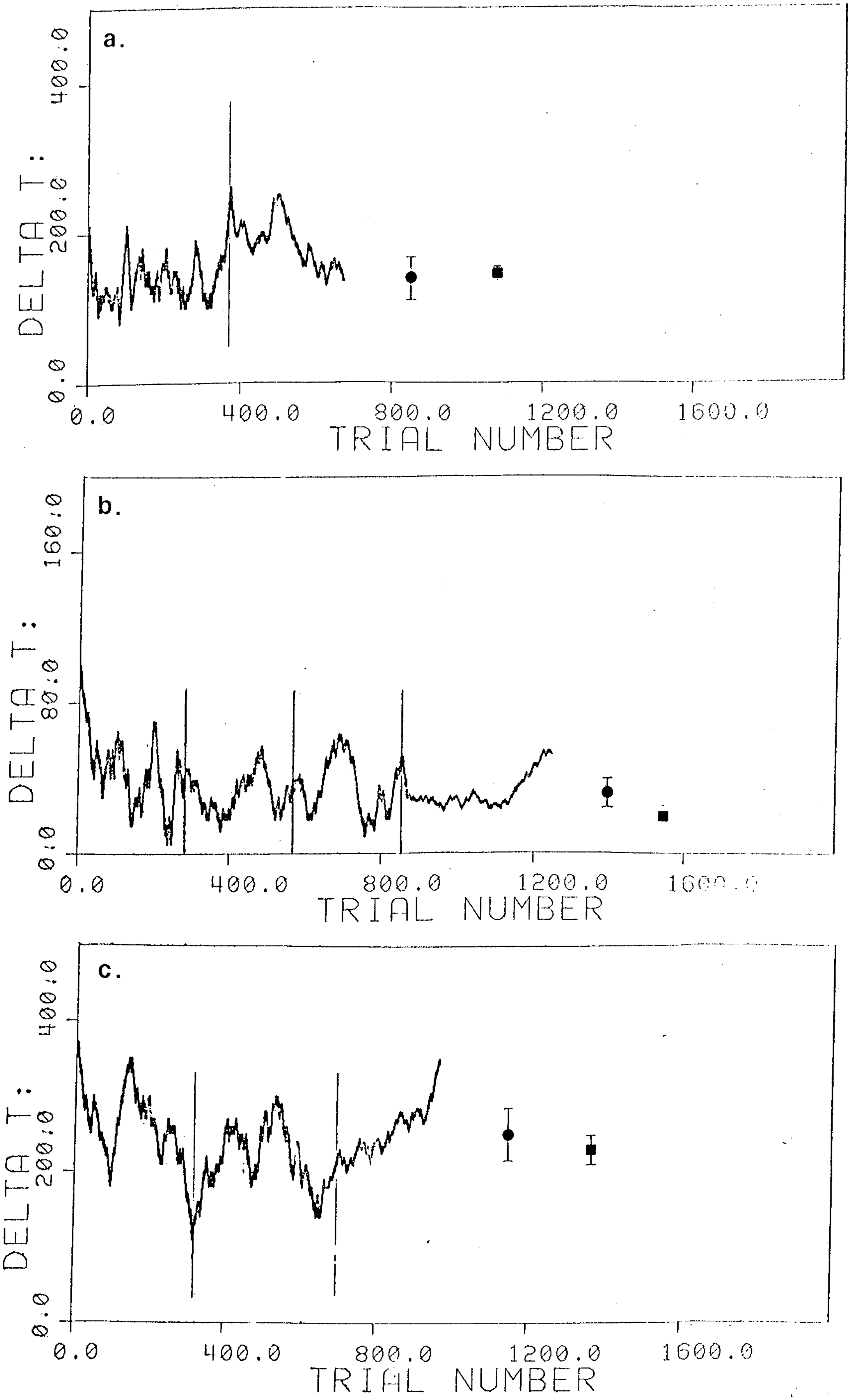
a.
DELTA T:
0.0
200.0
400.0
0.0
400.0
800.0
1200.0
1600.0
TRIAL NUMBER
b.
DELTA T:
0.0
80.0
160.0
0.0
400.0
800.0
1200.0
1600.0
TRIAL NUMBER
c.
DELTA T:
0.0
200.0
400.0
0.0
400.0
800.0
1200.0
1600.0
TRIAL NUMBER

Figure 4.4

msec in that session. From Experiment 1, the mean just-discriminable change was 143 ± 10 msec.

At Ts=0 msec and negative ΔT, the same subject performs more consistently. The track, seen in Fig. 4.4b, generally decreases although increases are seen, including one at the end of the track. The mean ΔT for d'=1.2 over the final 384 trials is 31 ± 8 msec. This is slightly higher than the mean ΔT for d'=1.2 of 19 ± 3 msec obtained from Experiment 1.

The data from 960 trials with V-T stimulus pairs are shown in Fig. 4.4c. The standard is -60 msec and changes are in the positive direction. The pattern of this track is, again, erratic, as the level of ΔT can be as small as 100 msec or as large as 320 msec. The change in ΔT does not appear to be due to learning, since it reaches its minimum value early in the run and approaches a maximum after nearly 1000 trials. The mean ΔT for d'=1.2 for these data is 248 ± 33 msec. The mean ΔT for d'=1.2 from Experiment 2 was also large and variable, 232 ± 22 msec.

d. Results for S4.

Subject 4 had encountered difficulty with a different condition of Experiment 1 than had the other subjects. For this subject, the A-T conditions tested were Ts=+60 msec and Ts=0 msec, both with ΔT negative. For the V-T case, Ts was

-60 msec and ΔT was positive. The track of ΔT over 576 trials of the first condition is shown in Fig. 4.5a. The track is fairly stable over the final few hundred trials. The mean level of ΔT is 185 ± 16 msec. Only one acceptable psychometric function had been estimated in Experiment 1 for S4. The interpolated value of ΔT for d'=1.2 was 202 msec in that case. The arrow indicates that other estimates would have been higher.

The track obtained for Ts=0 msec is seen in Fig. 4.5b. This track is highly systematic, dropping from 100 msec to an eventual asymptote at ΔT of 24 ± 3 msec in approximately 400 trials. The final level is comparable to the value determined in experiment 1, 27 ± 4 msec.

The track for V-T Ts=-60 msec and ΔT positive is also quite stable at its final level, as shown in Fig. 4.5c. In this case, a sharp decrease had characterized the early part of the run. Again, it is observed that over a brief period, a relatively small change is resolvable. However, the subjects tend to be unable to keep ΔT at such levels. The track stabilizes at 80 ± 5 msec in the final session. The estimates of the ΔT for d'=1.2 from Experiment 2 had a mean of 81 ± 21 msec. The just-discriminable changes are the same in each procedure.

## C. Discussion.

Figure 4.5. Stimulus tracks for S4. Symbols are the same as in Fig. 4.2. Panel a: A-T, Ts= +60 msec. △T negative. N=576. Panel b: A-T, Ts=0 msec. △T negative. N=1152. Note that the ordinate of panel b is different from panels a and c. Panel c: V-T, Ts=-60 msec. △T positive. N=960.

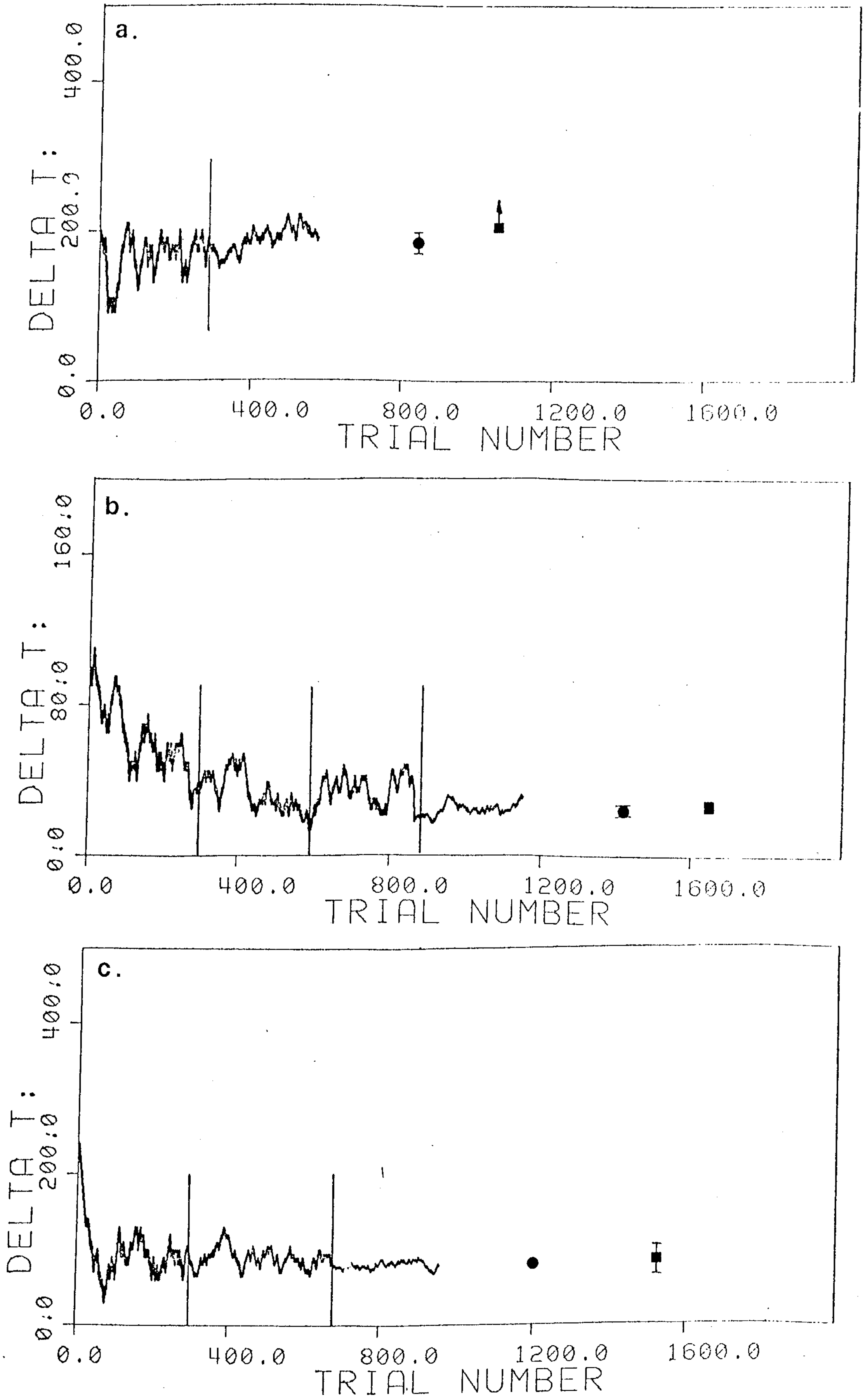
a.
DELTA T:
0.0
200.0
400.0
0.0
400.0
800.0
1200.0
1600.0
TRIAL NUMBER
b.
DELTA T:
0.0
80.0
160.0
0.0
400.0
800.0
1200.0
1600.0
TRIAL NUMBER
c.
DELTA T:
0.0
200.0
400.0
0.0
400.0
800.0
1200.0
1600.0
TRIAL NUMBER

Figure 4.5

It had been hypothesized that the just-discriminable change in conditions such as Ts=-60 msec, positive $\Delta T$, would be smaller under an adaptive, criterion-free psychophysical procedure. However, the hypothesis was not supported. In the condition for which each had required a relatively large change for d'=1.2 in Experiments 1 and 2, the four subjects required nearly the same $\Delta T$ in the adaptive procedure. For Ts=0 msec, using A-T pairs, the same result was found. That condition required only small changes in each experiment. The extent to which the estimates of the mean $\Delta T$ for d'=1.2 agree is evident in Fig. 4.6, a scatterplot of the twelve pairs of means. Double logarithmic coordinates are used to accomodate the large range of values. The line has a slope of one. The Pearson product moment correlation coefficient for these data is 0.987, which with 10 degrees of freedom occurs with $p < 0.001$. A least-squares fitted line to the points has the equation,

$$y = 0.991x + 2 \text{ msec} \qquad \text{(eq. 4.1)}$$

where x refers to the mean $\Delta T$ for d'=1.2 from the same-different procedure and y refers to the mean $\Delta T$ for d'=1.2 from the adaptive procedure.

Two conclusions may be drawn from these results. First, because the mean values of the $\Delta T$ for d'=1.2 obtained with this procedure are essentially identical to the estimates obtained in the corresponding conditions of

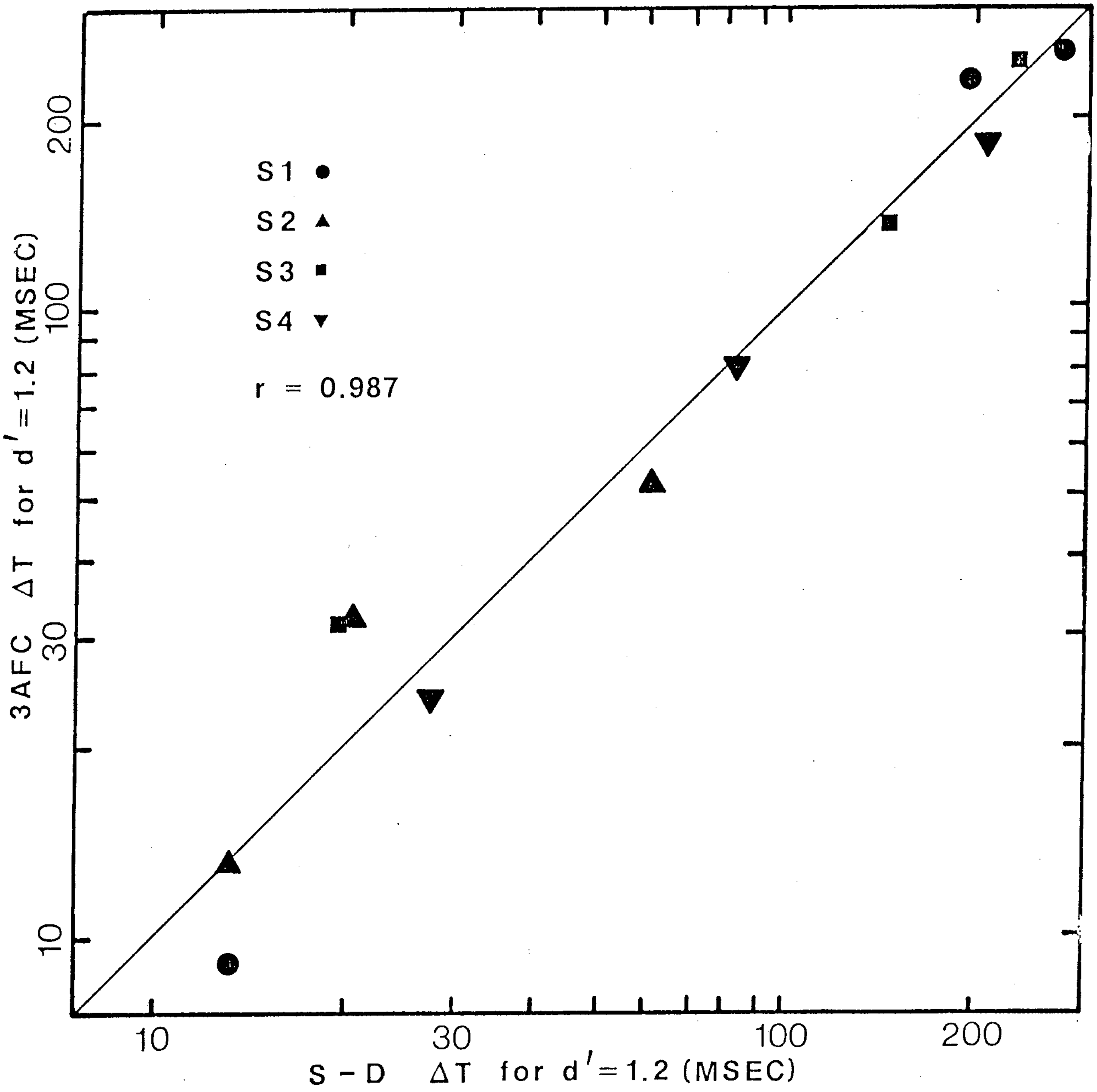

Figure 4.6. Scatter diagram of just-discriminable ΔT from two experimental procedures. The pairs of points are estimates of the ΔT for d'=1.2 obtained in Experiment 1 or 2 and in Experiment 4. The values along the abscissa are the means of successive estimates taken from fitted psychometric functions. The values of the ordinate are taken from the final session of the 3AFC/adaptive procedure. The correletion coefficient for the 12 pairs of points is 0.987.

Experiments 1 and 2, the data support the validity of the results of those experiments, which included a much broader range of conditions than could have been studied with the adaptive procedure. The first two experiments had been characterized by variability across subjects, an occasional inability to generate a reliable psychometric function, and for certain conditions, surprisingly large just-discriminable changes. It could have been argued that all of these findings were due to the experimental design, which called for the testing of a large number of conditions at the expense of obtaining a large number of trials in each. In this experiment, a few of the conditions for which performance was relatively poor were retested over a larger number of trials; with the trials concentrated at the threshold level of $\Delta T$; and with a different psychophysical task. The results replicate those of Experiments 1 and 2.

In some of the tracks shown in this chapter, performance may be stable over a long run but be characterized by large excursions in each direction. Examples are seen in Figs. 4.2a; 4.3c; 4.4a; and 4.5a. In these figures, the tracks have regions over which the level of $\Delta T$ can be driven down to a relatively small value but cannot be kept at that level. There are also regions over which the subject seems to lose the ability to make the discrimination, and the level increases steeply. These features are most prominant in the tracks for the non-zero

Ts. Compared to these, the tracks for Ts= 0 msec are compact; that is, excursions in each direction are less pronounced. Since these subjects had served in Experiments 1 and 2, fitted psychometric functions were available for each subject for each of the conditions tested in Experiment 4. These functions were used to simulate adaptive runs under the circumstances of the actual experiment (same starting values of $\Delta T$, same step sizes, and same number of trials). The simulation procedure is discussed in Appendix One. The expected variability of the stimulus track depends on the slope of the underlying psychometric function and the step size. One way to obtain the statistics necessary for the evaluation of a track is by simulating the run with identical stimulus conditions. Thus by using the psychometric function from a previous experiment and duplicating the step size from the present experiment, an estimate of variability can be obtained. The simulation results indicate that the variability of the actual runs is significantly greater than the variability of the simulations. This result supports the conclusion that an invariant probability function (given by the psychometric function) does not govern the subject's responses. On some proportion of the trials, a different process must operate. For those conditions for which the just-discriminable change is unusually large, the difference in variability is more apparent than is the difference when the target level is small. When the conditions for this experiment were selected, it was hypothesized that performance when Ts=0 msec was qualitatively different from performance when, for example, Ts=-60 msec and $\Delta T$ is positive. In the former condition, the target level of $\Delta T$ is so small that statistical tests of the difference between the actual data and the simulated data are inconclusive. If it could be shown that the difference between simulated data and real data characterizes one set of conditions but not another, support for the argument that there are qualitative differences in the way the discrimination task is performed would be gained. The data do not permit such a conclusion at the present time. The data do support the hypothesis that the underlying probability function is not fixed; this finding is consistent with the suggestion that the subject's attention shifts or that his strategy is changed during the run. Greater variability would arise if for parts of a run, performance were better or worse than that predicted from a single psychometric function.

It had been argued in Chapter Two that subjects find it difficult to compare the sensory representations of the temporal relation in successive pairs of cross-modality stimuli. If that is the case, the labeling strategy will be used on some proportion of the trials. By the labeling strategy, it was meant that the discrimination response was based on the labels assigned to the pairs when coding or memory limitations prevented a more direct comparison. In the present results, the form of the stimulus track may provide some insight into the time course of attention to the temporal relation between the stimulus events. In the tracks shown in the figures listed above, intervals of 40-80 trials during which $\Delta T$ changes monotonically may be seen. For intervals in which the level decreases, it may be hypothesized that the subject is highly focussed on the basic sensory representation of the stimuli. Within such intervals, the track may be driven to values that are well below the eventual "asymptotic" level (see Fig. 4.5a, trials 20-40 for an example). It was noted in Chapter Two that the precision of simultaneity judgments over a five-minute period is also greater than their long-term precision. The time course of improved performance is similar to that seen in the tracks. For segments over which the track increases sharply, it is likely that the subject has lost the focus and has resorted to the labeling strategy. In this situation, the subject possibly directs his attention toward guessing the temporal relation in the stimulus pairs; he is attending to a higher representation of the stimulus events. Over these intervals, the level of $\Delta T$ will increase to a level at which the labeling strategy is effective.

## D. Summary of Experiments 1-4.

At this point, the facts and inferences about cross-modality temporal resolving power that have emerged from the first four experiments can be restated and summarized. The change in Ts that is resolvable at a fixed level of performance such as d'=1.0 or 1.2 is not predictable from Weber's Law. Some generalizations about the nature of performance can be made. The just-discriminable change can be predicted from the way in which stimulus pairs might be labeled. For a pair that elicits a reliable and unambiguous percept, the just-discriminable change is consistent across subjects and across successive measurements in the same subject. The clearest and possibly the only cross-modality example of this is observed when A-T stimuli are presented as a simultaneous pair (Ts=0 msec). For these cases, 10-15 msec changes will be discriminable. For the condition tested adaptively, the track of the stimulus levels presented decreases to an asymptotic level that is relatively stable and consistent with predictions based on the psychometric function obtained in Experiment 1.

For values of Ts for which labels are not consistently applied, ie, small non-zero Ts, resolution is difficult if Tc is likely to be given the same label as Ts. In the adaptive procedure, the level of $\Delta T$ selected according to performance is unstable, as noted above. The subjects' performance appears to reflect periods of enhanced discrimination or attentional focus, and periods of confusion or loss of focus. During these periods, the tracks show a steady decrease or increase in the level of $\Delta T$.

E. Conclusion.

The labeling of temporal relations and temporal resolution are correlated to a considerable extent but each has some independent variance. Under the appropriate test conditions, it appears that each may limit the other. To find the factor that limits both aspects of temporal processing, it is necessary to consider the encoding of the stimulus elements. Some pairs may be ambiguously encoded; as a result they would not be labeled consistently or resolved well. However, the tracks from Experiment 4 suggest that for brief periods, on the order of five minutes, discrimination can be significantly better than the long term average. That implies that temporal information is available to the subject that is sufficiently precise to allow better performance than has been observed. The data obtained in Experiments 1-4 implicate the range of Ts between ±120 msec as being most subject to attentional bias. The evidence from the first three experiments is seen in the form of the discrimination function when Ts and Tc are both within that range. In the adaptive procedure, it is reflected in the variability of the stimulus tracks. In previous labeling experiments, attentional influences were noted in such phenomena as the prior entry effect (Stone, 1926) and the instability of the PSS (Gengel and Hirsh, 1970). These effects also occur over the nearly-simultaneous range of Ts.

If labeling and temporal resolution are limited by the same underlying mechanism, then procedural options that improve one type of performance might also be applied to the other. Order labeling is made more precise if the subject is allowed to observe the stimulus pair repeatedly prior to labeling it (Gengel and Hirsh, 1970). This variation in procedure succeeds in lowering the DL where another, the use of feedback, does not. The discrimination procedures used in the present series of experiments have been single-presentation procedures that employ feedback in an effort to specify the signal for the subject. As has been described, these procedures may not reveal the limits of cross-modality temporal resolution. In the final experiment, the range of Ts between ±120 msec was studied using the method of adjustment. In that experiment, trial-by-trial feedback was not used, but the stimulus pair was recycled until the subject was satisfied with his judgment, as in the repeated-presentation method for order labeling used by Hirsh (1959; Hirsh and Sherrick, 1961; Gengel and Hirsh, 1970). With this procedure, estimates of the just-discriminable change in Ts were obtained that are believed to specify the lower limits of cross-modality temporal resolution.

# CHAPTER FIVE. Cross-modality tenporal resolution measured using the method of adjustment.

The experiment described in this chapter was intended to summarize cross-modality temporal resolving power for values of Ts within 120 msec of physical simultaneity. This region of Ts has proved difficult to study with other psychophysical procedures. For this experiment, the method of adjustment was used. The order labeling literature has indicated that the precision with which a stimulus pair can be identified increases by a factor of about three when unlimited rather than single presentations of the stimulus pair can be observed (Gengel and Hirsh, 1970). In the method of adjustment, the same increase in the precision of temporal resolution might be seen. Most importantly, the adjustment procedure was tried in an effort to avoid the effects seen in other procedures that have been attributed to failures of attention. The results of this experiment have been presented as an abstract (Sinex, 1978).

## A. Methods.

The stimulus elements used in this experiment were the same ones presented in Experiments 1-4, with one exception. The tactile stimulus in this experiment was delivered from a piezoelectric ceramic bender (Linden Corp.). This device differs from the electromechanical device used in Experiments 1-4 in that its bandwidth is broader, and as a result it delivers a sharper tap with less ringing. However, the deflections that could be produced with these devices are of small amplitude. The bender was mounted in the same enclosure with the LED. The characteristics of the device are summarized in Fig. 5.1. The tactile stimulus was presented at 10-15 dB SL, the highest level available. At the conclusion of Experiment 5, some conditions were replicated with the electromagnetic vibrator. The results of those comparisons, which indicated no difference in resolving power as measured with each device, are summarized in Appendix Two. For this experiment, a ten-turn potentiometer for adjusting Tc replaced the standard response box. The voltage across the potentiometer was read by an analog-to-digital (A-D) converter in the computer interface. The other apparatus was identical to the previous experiments.

Data were obtained using the method of adjustment. Each trial consisted of repeated presentations of a sequence that included a standard stimulus pair, Ts, and a comparison pair, Tc, in which the temporal relation between the elements was under the control of the subject. The trial structure is diagrammed in figure 5.2. Subjects always knew the temporal relation, Ts, in the standard, and they were instructed to make the comparison pair match the standard. The trial was terminated by a button press when a satisfactory match had been made. Trials commonly included twenty to forty presentations; the number was unrelated to any of the experimental variables.

Figure 5.1. Description of the piezoelectric vibrator. Panel a: Schematic diagram of the device. The bender is clamped in a cantilever mounting position, just below the surface of the enclosure. When energized, the bender deflects in the direction indicated by the arrow. The shaded area is the wooden base of the enclosure. Panel b: Vibrator frequency response. Each point is the bender output given a sinusoidal input at 20 V RMS. The ordinate is force, expressed in dB re 1 dyne. During the measurements, the vibrator was loaded with a Bruel and Kjaer Artificial Mastoid, model 4930, used in conjunction with a model 2209 Sound Level Meter. Panel c: Vibrator response to a 5.0 msec pulse. Measurement apparatus is the same as in panel b.

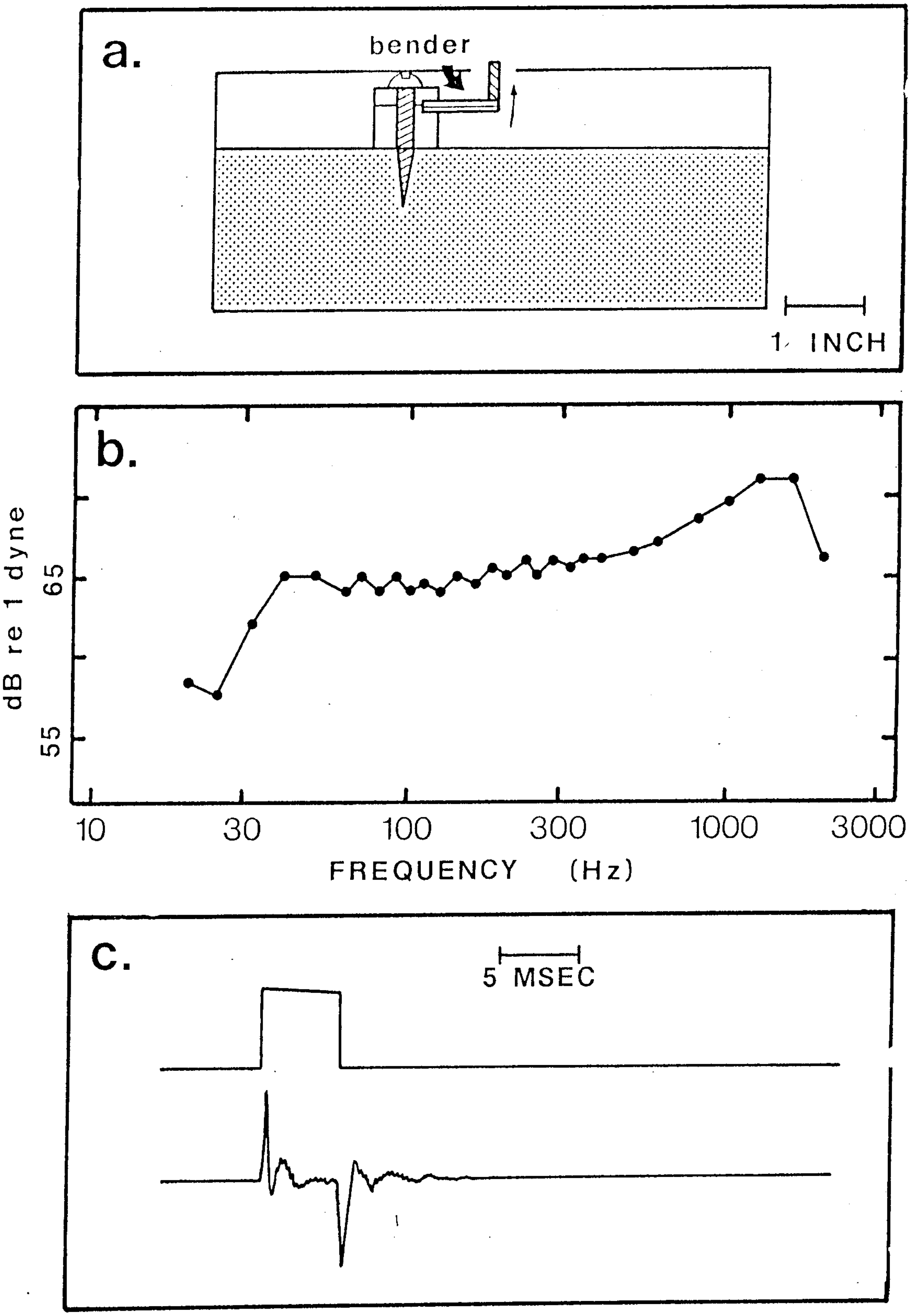
a.
bender
1 INCH
b.
dB re 1 dyne
65
55
10
30
100
300
1000
3000
FREQUENCY (Hz)
c.
5 MSEC

Figure 5.1

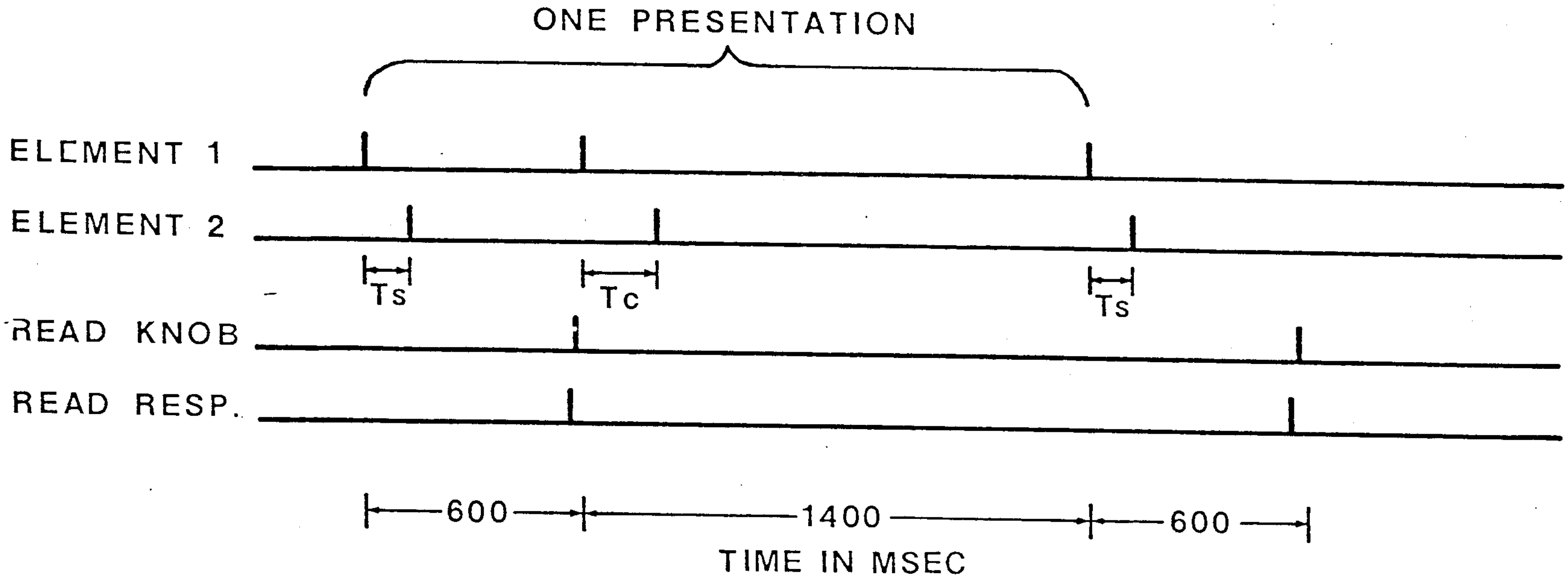

Figure 5.2. Matching-to-standard trial sequence. In the example shown, Ts is 120 msec and the current value of Tc is 200 msec. Ts and Tc are separated by a fixed 600-msec interval. The next presentation begins 1400 msec after the onset of Tc. Just prior to Tc, a check of the response button is made. If no response has been made, the knob position is read and the value of Tc is updated for the next presentation.

Two control procedures were used to minimize cues related to the position of the adjustment knob. First, the value of Tc associated with a given knob position varied from trial to trial over a 100-msec range. Second, the next trial would not begin until the knob had been turned a distance sufficient to start the next trial with a value of Tc differing from the previous judgment by 100 msec or more.

a. Cross-modality matches.

Three cross-modality stimulus combinations were tested: visual-auditory; visual-tactile; and auditory-tactile. Within each, seven values of Ts were used. These included a simultaneous condition (Ts = 0 msec) and three conditions with each element leading (Ts = 30, 60, or 120 msec). In this experiment, Ts and Tc are given as unsigned numbers with a descriptive phrase such as "auditory leading" whenever possible. When signed numbers must be used to establish order, an explanation is given. Signs are omitted since for each cross-modality pair, conditions in which Ts has a different order are discussed separately. Three subjects participated in the experiment; each was tested with three cross-modality stimulus pairs. These subjects had not participated in any of the previous experiments. The seven values of Ts were tested in a fixed order that alternated the magnitude of Ts and the order of elements; the cycle of values was tested once, then repeated within each cross-modality case. Each subject encountered the three cross-modality combinations in a different order, so that each was presented first to one subject, second to another, and last to the third. Thus, when data are presented as averages across subjects, effects due to learning may be expected to cancel. Improvement in performance due to learning was negligable after the experiment had begun.

In each one-hour experimental session, only one standard was presented. Approximately forty judgments per session were obtained. Subjects were instructed to match the standard interval and to minimize the variance in their judgments. Following each block of five trials, they were given feedback based on the standard deviation of the judgments. They were told that small constant errors (CE) could be disregarded. The constant error is defined as

$$CE = |Tc| - |Ts| \qquad \text{(eq. 5.1).}$$

Thus a positive CE means that the interval in Tc, the adjusted pair, was larger than the interval in Ts, the standard.

Each observer was given practice with two values of Ts marked by the cross-modality stimulus pair that would be tested first. Approximately five hours of training were spent with Ts=0 msec and approximately five hours with Ts=60 msec (visual leading tactile for S5; visual leading auditory for S6; and tactile leading auditory for S7). There is no indication that increased familiarity with these standards led to better performance when they were retested as part of the actual experiment.

### b. Within-modality control matches.

At the conclusion of the experiment, a brief series of within-modality matches was obtained from the same group of subjects. Intervals were defined by delivering two successive pulses to the same transducer. In the auditory-auditory experiment, one click and one tone pip (1 kHz, 9 msec duration) were used. In addition, visual-visual conditions using two LEDs at different locations were also used. Simultaneous standards were omitted or replaced by Ts = 15 msec when the stimulus elements were generated by a single transducer.

Within-modality judgments were obtained in half-sessions and the standards were tested only once. The number of trials per standard is only about 1/4 of the number of cross-modality matches per standard. Otherwise, the testing situation did not differ.

## B. Results.

Following practice sessions, the three observers made approximately 1700 judgments in each of the three cross-modality conditions. Estimates of temporal resolving power are taken from the distributions of judgments. Figure 5.3 summarizes the data on which the conclusions are based. In this figure, the mean of Tc is plotted as a function of Ts for each individual. Each panel summarizes a different cross-modality combination. The line in each panel has a slope of one; perfect matches would fall on the line. Most of this section will focus on the standard deviation of judgments, averaged across observers. Following the terminology of Sternberg and Knoll (1973), inferences about the precision of temporal judgments are based on the standard deviation of the matches; this is considered to be the major component of temporal resolving power for reasons that are discussed later. The data also indicate the accuracy of temporal judgment. The constant error is the measure of accuracy.

Estimates of capability for individual subjects were obtained in the following way. The summary estimate for each condition is the mean of the standard deviations of matches from two sessions. The mean standard deviation was used, rather than the standard deviation of the pooled matches, because occasionally, a significant shift in the CE occurred upon retest of a condition. The CE could change with no loss in precision, as indicated in Fig. 5.4. The two frequency polygons present the distributions of matches for S5 in the visual-auditory Ts=0 msec condition. The first forty matches are described by the solid line. Those

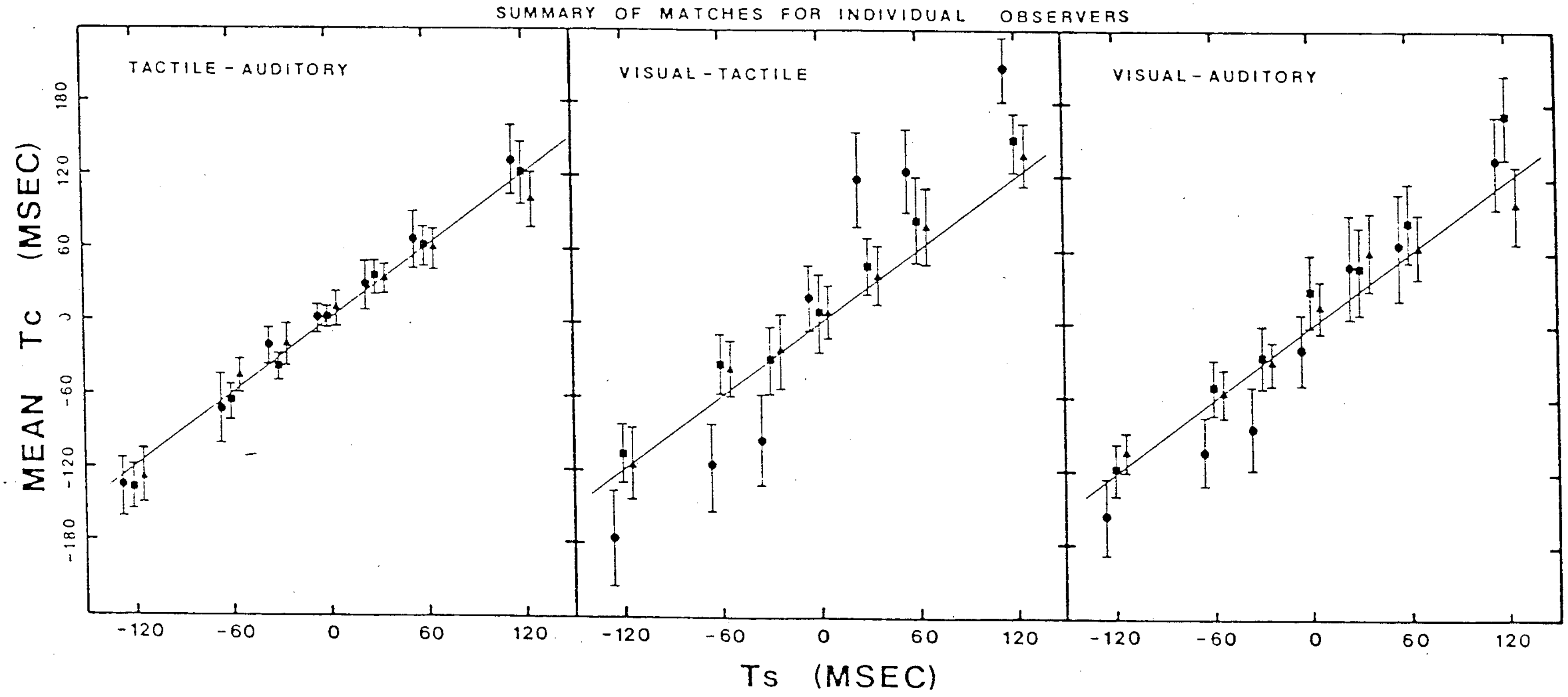

Figure 5.3. Summary of matches for individual observers. The abscissa is Ts; negative values indicate that auditory leads tactile; tactile leads visual; and auditory leads visual. The ordinate is the mean of Tc for each observer (● = S5; ■ = S6; ▲ = S7). Each point is based on approximately 80 matches, as explained in the text. The vertical bars indicate ± 1 session standard deviation. The line has slope = 1 and intercept = 0. Note that tactile-auditory matches have smaller standard deviations and lie closer to the line than do matches to other cross-modality stimuli.

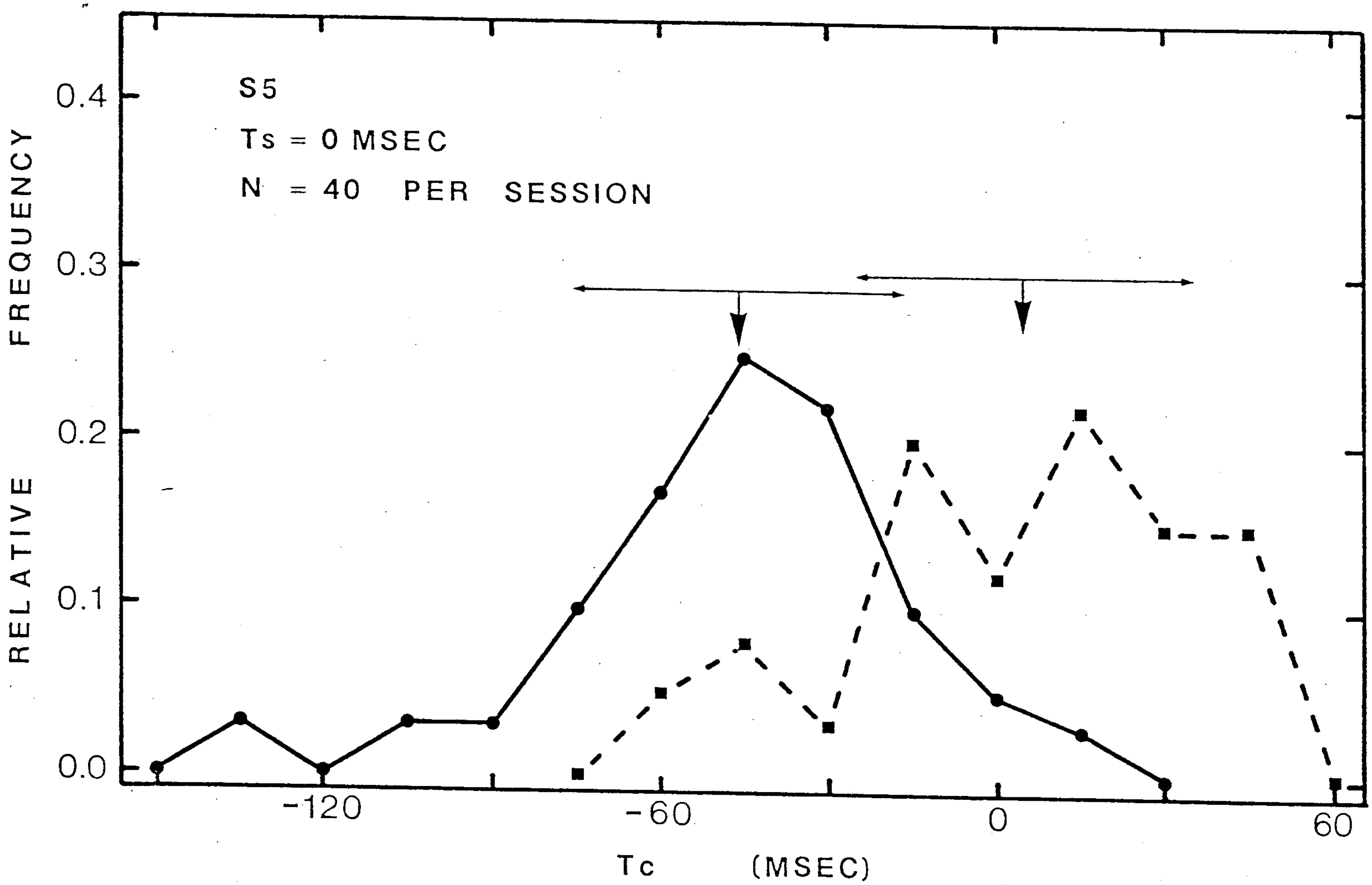

Figure 5.4. Changes in the constant error across sessions. The data were obtained from S5 with visual-auditory stimuli and Ts = 0 msec. The abscissa is Tc, and the ordinate is the relative frequency of Tc in 15-msec bins. The standard deviation of each distribution is approximately 29 msec, but the means differ by 51 msec. Similar shifts were observed in one or two conditions for each subject.

matches had a mean of -46 msec (negative means tactile leading) and a standard deviation of 29.0 msec. The second forty judgments, obtained approximately two weeks later, are described by the dashed line. This distribution has a mean of +5 msec (visual leading), but the standard deviation is still 29.5 msec. The standard deviation of all eighty matches is 39 msec, which does not reflect the precision of the matches made with the same criterion. A similar criterion shift was observed for each subject in at least one condition. Large criterion shifts, however, were not observed in the auditory-tactile experiment.

Figure 5.5 presents the group mean standard deviations. The data are also summarized in Table 5.1. Two conditions with a common leading element are plotted in panels a-c to facilitate comparisons between them. Panel d shows the data for the within-modality conditions. Within each modality, all conditions for which the absolute value of Ts was the same have been pooled to obtain the value plotted. As indicated by the error bars, differences in performance between stimulus configurations were not large. In this as well as in the previous figures, two conclusions are suggested. First, cross-modality temporal resolution is consistently poorer, i.e., is less precise, than within-modality temporal resolution. Second, among cross-modality conditions, the auditory-tactile condition once again produces significantly smaller standard

Figure 5.5. Group mean standard deviations of matches. Each point is the session standard deviation, averaged across three subjects and two sessions per subject. The abscissa is Ts. Results for cross-modality stimuli are shown in panels a-c; the lines in each panel represent stimuli with a common leading element. All within-modality combinations are shown in panel d. The vertical bars indicate ± 1 standard error of the mean. Only auditory-tactile and tactile-auditory pairs resemble within-modality combinations.

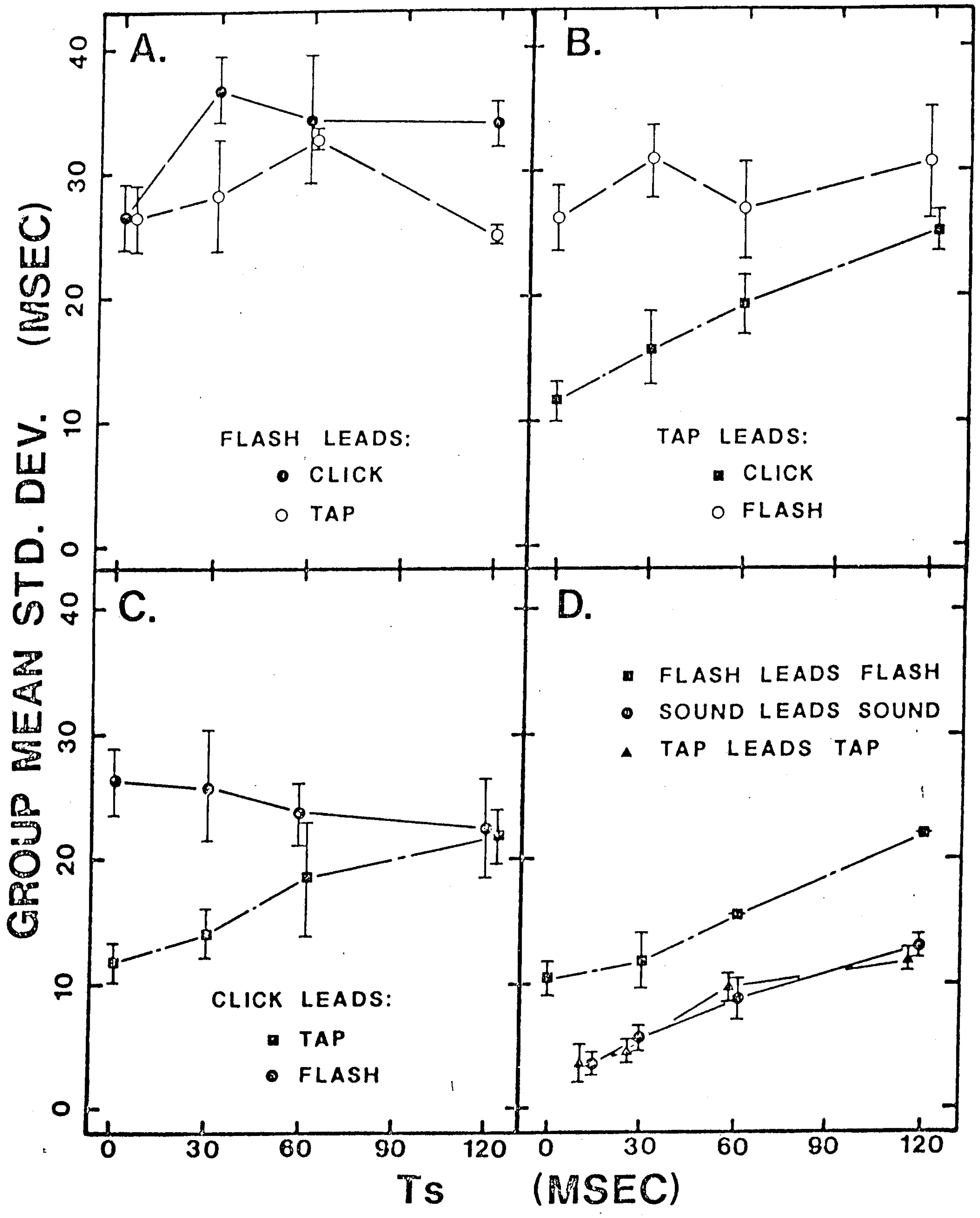
A.
B.
C.
D.
GROUP MEAN STD. DEV. (MSEC)
Ts (MSEC)
0
10
20
30
40
0
30
60
90
120
FLASH LEADS:
● CLICK
○ TAP
TAP LEADS:
■ CLICK
○ FLASH
CLICK LEADS:
■ TAP
● FLASH
■ FLASH LEADS FLASH
● SOUND LEADS SOUND
▲ TAP LEADS TAP

Figure 5.5

## TABLE 5.1 RESULTS OF CROSS-MODALITY MATCHES TO STANDARD

### a. Auditory-tactile standard deviations

| | Session std. dev. (msec) | | | Across subjects: | |
|---|---|---|---|---|---|
| Ts: | S5: | S6: | S7: | Mean: | Std. err: |
| -120 | 28.5 | 25.0 | 22.1 | 25.2 | 1.8 |
| -60 | 23.9 | 16.7 | 17.1 | 19.2 | 2.4 |
| -30 | 21.6 | 13.6 | 12.0 | 15.7 | 3.0 |
| 0 | 11.2 | 8.8 | 14.7 | 11.6 | 1.7 |
| 30 | 14.7 | 10.1 | 16.8 | 13.9 | 2.0 |
| 60 | 27.9 | 13.6 | 13.6 | 18.4 | 4.8 |
| 120 | 25.4 | 17.8 | 22.4 | 21.9 | 2.2 |

### b. Visual-tactile standard deviations

| | Session std. dev. (msec) | | | Across subjects: | |
|---|---|---|---|---|---|
| Ts: | S5: | S6: | S7: | Mean: | Std. err: |
| -120 | 39.1 | 23.6 | 29.2 | 30.6 | 4.5 |
| -60 | 34.7 | 23.6 | 22.4 | 26.9 | 3.9 |
| -30 | 36.7 | 27.1 | 28.5 | 30.8 | 3.0 |
| 0 | 27.0 | 30.4 | 21.0 | 26.1 | 2.7 |
| 30 | 37.6 | 22.8 | 24.3 | 28.2 | 4.7 |
| 60 | 33.5 | 33.3 | 30.5 | 32.4 | 1.0 |
| 120 | 25.5 | 23.0 | 25.9 | 24.8 | 0.9 |

### c. Visual-auditory standard deviations

| | Session std. dev. (msec) | | | Across subjects: | |
|---|---|---|---|---|---|
| Ts: | S5: | S6: | S7: | Mean: | Std. err: |
| -120 | 29.9 | 21.1 | 15.9 | 22.3 | 4.1 |
| -60 | 27.9 | 23.4 | 19.3 | 23.5 | 2.5 |
| -30 | 34.0 | 25.2 | 18.4 | 25.9 | 4.5 |
| 0 | 29.3 | 29.5 | 20.7 | 26.5 | 2.9 |
| 30 | 42.1 | 34.6 | 32.8 | 36.5 | 2.9 |
| 60 | 44.0 | 32.1 | 26.3 | 34.1 | 5.2 |
| 120 | 37.1 | 32.6 | 30.7 | 33.4 | 1.9 |

deviations. Over the range of Ts examined in this experiment, it is seen that in cross-modality conditions involving the visual modality, the just-resolvable change, expressed as the standard deviation of judgments, is nearly constant. In the auditory-tactile case, however, the standard deviation systematically changes in proportion to Ts. The same result is also seen for all within-modality control conditions. Such a proportional change is, of course, consistent with Weber's Law. In these conditions, the data would be well-fit by lines with slopes of 0.08-0.10, which is in agreement with the value often given as the Weber fraction for duration (Woodrow, 1951). The lines for visual-visual and auditory-tactile matches would have larger intercepts. The results of an analysis of variance confirm that auditory-tactile temporal resolving power is significantly greater than either visual-tactile or visual-auditory temporal resolution ($F=32.1$; 2,4 df.; $p<0.01$). The two-way analysis examined the effects of the three cross-modality stimulus combinations and of the magnitude of Ts on the standard deviation of matches. The results of the analysis are shown in Table 5.2a. It was found that the group standard deviation is significantly smaller for the auditory-tactile case. A significant relation between the standard deviation and Ts was also found. This relation is entirely due to the auditory-tactile data, as indicated by the significance of the interaction between cross-modality stimulus combination and Ts.

Although it is considerably more precise than either of the cross-modality stimulus combinations that include the visual system, auditory-tactile temporal resolution is not as good as within-modality temporal resolution, especially if the single modality is the auditory or the tactile. A separate analysis of variance was conducted using the within-modality data. The summary is shown in Table 5.2b. For these data, there is no significant difference between modalities, nor is there an interaction between modality and Ts as was observed in the cross-modality experiments. The dependence of the standard deviation on Ts noted in Fig. 5.5 is statistically reliable, however ($F=19.0$; 5,10 df.; $p<0.01$).

The conditions order themselves in the same manner with regard to the accuracy of temporal matches although the differences are less prominant than were the differences in standard deviations. The group mean CE is plotted as a function of Ts in figure 5.6. The data are also presented in Table 5.3. Visual-tactile and visual-auditory judgments are least accurate, followed by auditory-tactile and within-modality conditions. Two additional analyses of variance were conducted using the constant errors. Although the trends in the data are as indicated, the differences in

## TABLE 5.2 SUMMARY OF ANALYSIS OF VARIANCE

### a. Cross-modality standard deviation

| Source: | F: | df: | p: |
|---|---|---|---|
| Ts | 4.5 | 6,12 | $p<0.05$ |
| Modality | 32.1 | 2,4 | $p<0.01$ |
| Ts x Modality | 5.7 | 12,24 | $p<0.01$ |

### b. Within-modality standard deviation

| Source: | F: | df: | p: |
|---|---|---|---|
| Ts | 19.0 | 5,10 | $p<0.01$ |
| Modality | 1.4 | 2,4 | ns, $p>0.05$ |
| Ts x Modality | 0.3 | 10,20 | ns, $p>0.05$ |

### c. Cross-modality constant error

| Source: | F: | df: | p: |
|---|---|---|---|
| Ts | 1.5 | 6,12 | ns, $p>0.05$ |
| Modality | 1.6 | 2,4 | ns, $p>0.05$ |
| Ts x Modality | 0.4 | 12,24 | ns, $p>0.05$ |

### d. Within-modality constant error

| Source: | F: | df: | p: |
|---|---|---|---|
| Ts | 1.6 | 5,10 | ns, $p>0.05$ |
| Modality | 0.8 | 2,4 | ns, $p>0.05$ |
| Ts x Modality | 1.3 | 10,20 | ns, $p>0.05$ |

Figure 5.6. Group mean constant error. Each point is the mean constant error (CE), averaged across three subjects and two sessions per subject. In each panel, the selection of stimulus combinations is identical to Fig. 5.5. The summary statistic E reflects the deviation from zero error. Its magnitude is smallest within modalities and for auditory-tactile and tactile-auditory pairs.

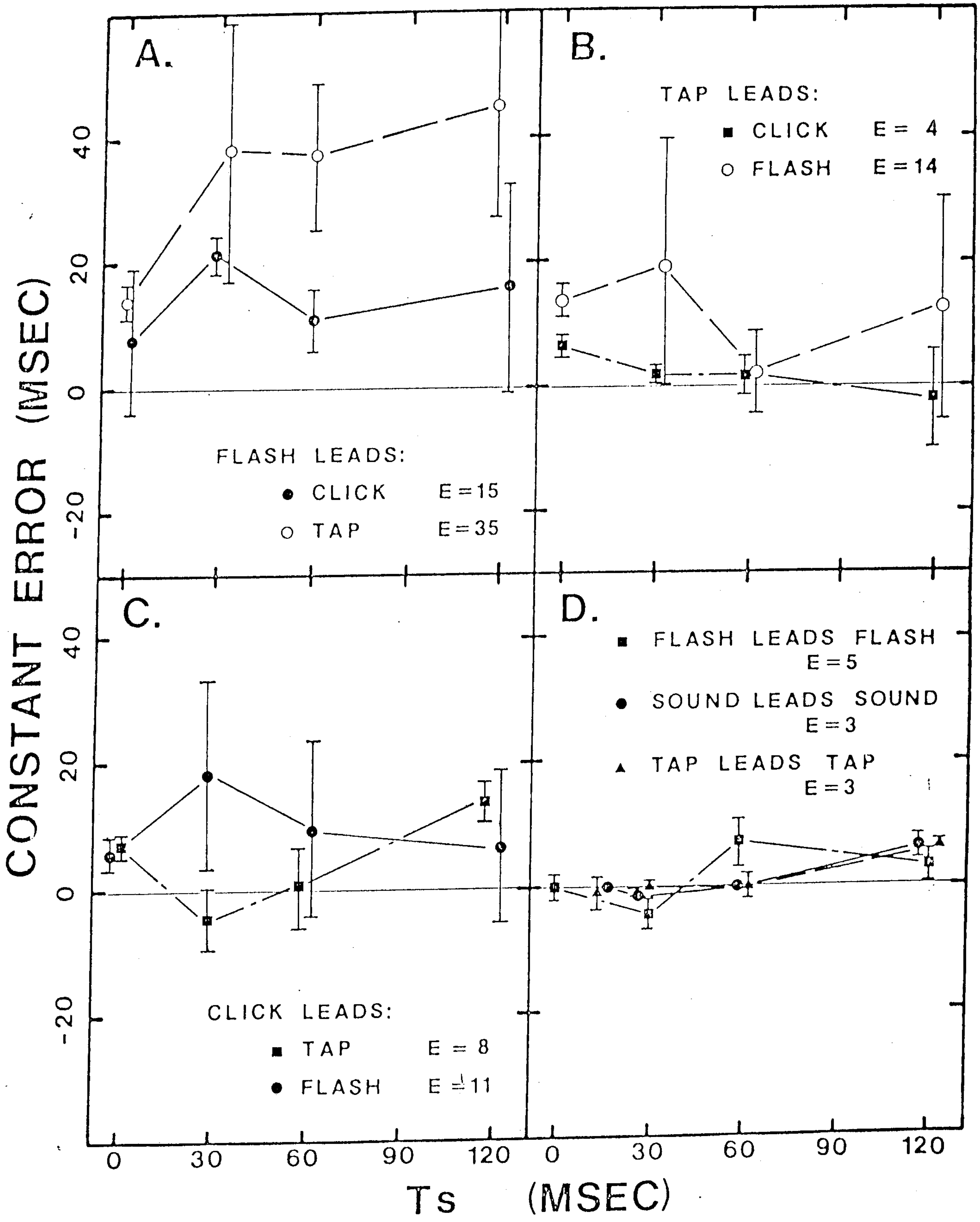
A.
FLASH LEADS:
● CLICK E=15
○ TAP E=35
B.
TAP LEADS:
■ CLICK E= 4
○ FLASH E=14
C.
CLICK LEADS:
■ TAP E = 8
● FLASH E = 11
D.
■ FLASH LEADS FLASH E=5
● SOUND LEADS SOUND E=3
▲ TAP LEADS TAP E=3
CONSTANT ERROR (MSEC)
40
20
0
-20
Ts (MSEC)
0
30
60
90
120

Figure 5.6

# TABLE 5.3 RESULTS OF CROSS-MODALITY MATCHES TO STANDARD

## a. Auditory-tactile mean

| | Mean Tc: (msec) | | | Across subjects: | | |
|---|---|---|---|---|---|---|
| Ts: | S5: | S6: | S7: | Mean: | Std. err: | CE: |
| -120 | -132 | -123 | -99 | -118 | 7.9 | -2 |
| -60 | -68 | -61 | -56 | -62 | 2.9 | 2 |
| -30 | -30 | -35 | -31 | -32 | 1.2 | 2 |
| 0 | -5 | -4 | -10 | -6 | 1.7 | 6 |
| +30 | 21 | 40 | 19 | 26 | 5.4 | -4 |
| +60 | 72 | 65 | 46 | 61 | 6.6 | 1 |
| +120 | 139 | 136 | 126 | 134 | 3.3 | 14 |

## b. Visual-tactile mean

| | Mean Tc: (msec) | | | Across subjects: | | |
|---|---|---|---|---|---|---|
| Ts: | S5: | S6: | S7: | Mean: | Std. err: | CE: |
| -120 | -178 | -107 | -113 | -133 | 18.5 | 13 |
| -60 | -117 | -34 | -37 | -63 | 22.1 | 3 |
| -30 | -98 | -30 | -22 | -50 | 19.7 | 20 |
| 0 | 21 | 10 | 11 | 14 | 2.7 | 14 |
| +30 | 118 | 47 | 39 | 68 | 20.5 | 38 |
| +60 | 125 | 85 | 80 | 97 | 11.6 | 37 |
| +120 | 208 | 149 | 137 | 165 | 17.8 | 35 |

## c. Visual-auditory mean

| | Mean Tc: (msec) | | | Across subjects: | | |
|---|---|---|---|---|---|---|
| Ts: | S5: | S6: | S7: | Mean: | Std. err: | CE: |
| -120 | -156 | -118 | -105 | -126 | 12.6 | 6 |
| -60 | -104 | -49 | -55 | -69 | 14.1 | 9 |
| -30 | -86 | -26 | -32 | -48 | 15.6 | 18 |
| 0 | -21 | 28 | 16 | 8 | 12.1 | 8 |
| +30 | 47 | 47 | 60 | 51 | 3.5 | 21 |
| +60 | 65 | 84 | 64 | 71 | 5.3 | 11 |
| +120 | 135 | 172 | 101 | 136 | 16.8 | 16 |

CE do not achieve statistical significance for either the cross-modality data or the within-modality data (Table 5.2c and 5.2d). Accuracy has been quantified by the index E, given by

$$E = \sqrt{\sum (CE)^2 / N} \qquad \text{(eq. 5.2)}$$

where CE is the mean constant error in each of N conditions. This index is the RMS deviation from the abscissa, which represents perfect accuracy.

C. Discussion.

The data obtained with the matching-to-standard procedure provide systematic estimates of cross-modality temporal resolution over the range of Ts within 120 msec of physical simultaneity. From these results, three conclusions may be stated.

First, as suggested in Experiments 1-3, all cross-modality temporal resolution is less precise than within-modality temporal resolution. The estimates of within-modality temporal resolution obtained in the control conditions of this experiment are in good agreement with those from previous studies. The general finding was that the function relating the standard deviation of within-modality matches to Ts is linear and monotonically increasing, with slope approximately equal to 0.1. If the intercept of the function is zero, the slope is identical to the Weber fraction. For the auditory and tactile modalities, the intercept of a line fitted to the standard deviations was approximately 2 msec. In the data from the visual modality, the intercept was 10 msec rather than zero, but the slope did not differ. For these data, the intercepts of lines fitted to the standard deviations should approximate the values reported as two-pulse thresholds in Chapter One. In two modalities, Ts=0 msec was not examined, and in the third, the visual, it was studied with Ts marked by spatially-separate light sources. For these reasons, the comparison may be questionable. Nevertheless, in audition, the intercept of 2.5 msec compares well with the two-pulse threshold reported by Gescheider (1966). However, the intercept of 2.7 msec for tactile discrimination is much smaller than Gescheider's estimate of 10 msec. As noted, Gescheider's figure is the result of measurement, while the value mentioned here is an extrapolation, assuming that the discrimination function is linear. For vision, the intercept of 10 msec is most comparable to the data of Sweet (1953), who found that stimuli from adjacent sources, one viewed foveally and one peripherally, were reported to be non-simultaneous when separated by 6-12 msec. The indication is that modality differences in the just-discriminable change become less prominant as Ts increases. For example, regression equations fitted to the discrimination functions may be solved to predict the standard deviation and the Weber fraction for Ts=1000 msec.

The predicted Weber fraction is 0.090 for auditory stimuli, 0.083 for tactile stimuli, and 0.111 for visual stimuli. Goodfellow (1934) obtained the data, using the method of adjustment and Ts=1000 msec. He reported Weber fractions to be 0.800, 0.100, and 0.107 for the three modalities. The agreement is surprising, considering the magnitude of extrapolation from the present data.

The second finding had also been anticipated from Experiment 2. Among the cross-modality stimulus combinations, temporal resolution is significantly better for intervals defined by auditory-tactile stimulus pairs. Temporal resolution for these stimuli is easily summarized; the standard deviation of matches increases as a linear function of Ts. The slope of the discrimination function is again 0.1. The difference between the discrimination functions for auditory-tactile intervals and interval defined within modalities is in the intercept; at a given Ts, the auditory-tactile standard deviation is approximately 10 msec larger than the auditory or tactile standard deviation. Again, this relation suggests that the relative difference between cross-modality and within-modality performance will decrease as larger values of Ts are used. In Experiment 2, it was observed that, within measurement error, the larger values of Ts, 240 and 480 msec, were resolved equally well in the cross-modality and within-modality cases.

Finally, temporal resolution for cross-modality stimulus pairs that include a visual element is worse than temporal resolution for auditory-tactile intervals, for small Ts. The standard deviation for visual-auditory and visual-tactile pairs is nearly constant, at approximately 30 msec, for all values of Ts tested. When Ts=120 msec, the difference between visual-tactile or visual-auditory and auditory-tactile temporal resolution is relatively small.

It was stated that the standard deviation of adjustments is believed to be the most important indicator of temporal resolving power. Accuracy of reproduction is limited by temporal resolving power, but it seems to be subject to other influences as well. As Fig. 5.4 shows, the central tendency of the matches can change significantly from day to day. In this and the other instances of criterion shift, the session standard deviation was more constant than the CE. Criterion fluctuation has been observed in other cross-modality data, such as simultaneity judgment (Chapter Two). The shift was less surprising in that experiment, since no standard was present. In this procedure, the criterion is for similarity only, not for a more complex quality of the stimulus. If similarity is difficult to judge, the difficulty may account for some of the variation in performance in Experiment 2, which used the same-different procedure. Criterion shift has been a recurring theme in the consideration of the perception and discrimination of cross-modality stimulus pairs. Once a criterion has been established, it seems that it can be maintained for a brief period. But the exact criterion that will be adopted is sometimes difficult to predict. It also seems that a stable criterion takes some time to develop. Single observations of a pair or of successive pairs may be difficult to compare to a criterion for this reason.

Cross-modality temporal resolution is not easily predictable from the within-modality results. For example, the time constant of a linear system may be used to predict various aspects of its response to two inputs (Green, 1971). Such models may be applied to temporal resolution, if assumptions are made about the ability to resolve differences in the response to sucessive pairs of events. Given the results of the current within-modality experiments, it would be concluded that the auditory and tactile modalities can transmit equally precise temporal information. In terms of the model, the encoded events would be identical with respect to their usefulness as time markers, and auditory-tactile temporal resolution would be expected to be as good as either modality alone, if a common comparator mediated the discriminations. The data show that it is not. The argument based on time constants fails for pairs involving vision as well. The data show that the visual time constant is longer than either the auditory or the tactile. Thus if a visual element were replaced with an auditory or tactile element, the fact that the non-visual event is more precisely encoded should improve the resolution of the cross-modality pair over that seen for a visual within-modality pair. Again, the results, especially of Experiment 5, are quite different from the prediction. The failure of this line of reasoning to describe cross-modality temporal resolution suggests that the mechanisms that underlie within-modality temporal resolution are different from and are more peripheral than those that mediate cross-modality temporal resolution. This is hardly surprising, but it is in contrast to the conclusion often drawn from the order-labeling result, that temporal organization is imposed upon stimulus pairs in all modalities by a single neural center (Hirsh and Sherrick, 1961; Efron, 1963a). The data reported here support the proposition that order labeling and temporal discrimination are different aspects of temporal information processing. In the final chapter, the model is reconsidered in terms of the physiological data reviewed in Chapter One. It is found that the differences between within- and cross-modality temporal resolution can be reconciled if further assumptions are made about the places at which the two tasks might be mediated.

# CHAPTER SIX. Summary, conclusions, and a possible explanation for the results.

In this chapter, the results of the five experiments are summarized and the conclusions restated. Finally, some speculations about physiological constraints on temporal resolving power and about the significance of these findings to other areas of research are offered. Some general principles have emerged from this work that were not previously available and that could not have been predicted from the published data. The observed differences between within-modality and cross-modality temporal resolution do not support the hypothesis that temporal discrimination is mediated by a single neural center. The results of labeling experiments have often been interpreted in terms of such a center. The fact that discrimination experiments measure a different aspect of temporal processing than do labeling experiments has been emphasized here as well as by others (Hirsh and Sherrick, 1961); thus, the possibility remains that a single processing center mediates other psychophysical judgments. The summary of cross-modality temporal resolving power provided here might prove useful in the development of schemes for recoding information for presentation to alternate sensory modalities. The recoding of speech for hearing-impaired listeners is a significant potential application of this information.

In the current investigations, an interaction between the adequacy of a given psychophysical procedure and the range of Ts under study was found. For nearly-simultaneous standards, single-presentation procedures do not reliably estimate the limits of temporal resolution. It was proposed that values of Ts in this range, within ±120 msec of physical simultaneity, are ambiguously perceived or are subject to memory limitations that interfere with performance in such procedures. The reason for the interference is unclear, as is the reason why it should be more prominant over a narrow range of stimulus values. One intriguing possibility was mentioned in Chapter One. It was noted there that when naturally-occurring events are observed, the delays between the onset time of the distal stimulus and the development of a central neural representation of that stimulus that occur in each modality will rarely be identical. In the auditory modality, the variability in the delay arises as a result of the propagation velocity of acoustic energy. Three milliseconds elapse for each additional meter of distance between source and receiver. In the tactile modality, the source of variability is the difference in neural conduction delays, according to the site of stimulation. Visual information, though transmitted more slowly in the nervous system, is not subject to these kinds of variation, since for objects near enough to be touched or heard, propagation delay may be neglected. To maximize the information that may be received from paired sources, it is possible that mechanisms to compensate for arrival asynchronies exist. Some evidence that they do has been offered by König (1965) and Dodd (1977), who find that speech intelligibility is enhanced by cross-modality presentation, even when a physical asynchrony exists between channels. Such a mechanism would probably operate over the range of Ts for which criterion or memory-related interference with discrimination performance has been observed. If such a mechanism existed, it would have to be selective for information in two modalities that can be associated with a single source. One basis for the decision as to whether a common source produces two signals could be the magnitude of the delay between them. Delays greater than 120-240 msec will occur in natural situations (for example, between lightning and thunder), but most often such delays are indicative of independent sources. Such a mechanism might have a greater effect when only a single presentation of the stimulus is available, either for labeling or discrimination. This is also the range over which the prior entry effect was observed (Stone, 1926).

The effects of criterion were most apparent in the data obtained using the same-different procedure. In each of the three cross-modality stimulus conditions tested, selected combinations of Ts and a direction of change resulted in a condition for which discrimination was difficult, relative to that observed with values of Ts that were only slightly different. The common feature to the difficult conditions was that the values of Ts and Tc were both taken from the range over which percepts are labeled with the greatest uncertainty. That is, when Ts, the standard, was approximately 60-120 msec, and Tc, the comparison, was approximately 0-60 msec, some subjects were unable to make the discrimination. An explanation analogous to the phenomenon of categorical perception (Liberman, Cooper, Shankweiler, and Studdert-Kennedy, 1967) was proposed to account for this finding. Briefly, it was suggested that under circumstances in which the stimulus pairs are perceived to have an uncertain temporal relation, the subject may tend to label them as being "simultaneous". If he uses that label when the standard pair is 60-120 msec, he is very likely to apply the same label to the comparison pair if its physical temporal relation is closer to objective simultaneity. The result is a decrease in performance, relative to values of the standard for which the use of labels can be avoided. This pattern of performance, which was designated the labeling strategy, was observed for corresponding conditions in the experiments that used visual-auditory, visual-tactile, and auditory-tactile stimulus pairs. It did not occur, however, in the within-modality control conditions using the same-different procedure.

Data obtained with the method of adjustment were relatively free from confounding and generally provide the lowest estimates of the just-resolvable $\Delta T$. Under the assumption that the true resolving power of the system is at least as great as the most precise reliable measurements available, conclusions with respect to temporal resolution in individual modalities and in cross-modalitiy combinations are based on these estimates.

Within-modality temporal resolution has been studied extensively by others and was measured in the present study to establish a baseline for evaluating cross-modality performance. The results of the control experiments were in agreement with previous reports (Goodfellow, 1934; Woodrow, 1951; Abel, 1972b; Getty, 1975; Divenyi and Danner, 1977). With one exception, the earlier studies were restricted to auditory stimuli. The discrimination functions are linear with Ts, and the slopes, which approximate Weber fractions, are 0.08 for auditory and tactile stimuli, and 0.11 for visual. The greatest difference between modalities is observed as Ts approaches 0 msec. For auditory and tactile temporal resolution, the just-discriminable change approaches 2 msec. For visual stimuli, it approaches 10 msec. These values are in rough agreement with the previously-reported values of the two-pulse threshold (Gescheider, 1966, 1967) or the successiveness threshold (Sweet, 1953).

Among the cross-modality stimulus combinations, temporal resolution for A-T pairs most resembles within-modality performance. The just-resolvable change reaches a minimum at Ts=0 msec, at which the value is approximately 12 msec. The just-discriminable change increases systematically with increasing Ts.

The remaining cross-modality data are obtained from stimulus combinations involving vision, for which temporal resolution must be described differently. For these conditions, the just-discriminable change reaches a plateau for Ts between ±120 msec. There is no minimum in the just-discriminable change at Ts=0 msec, as there is for within-modality and A-T data. The minimum just-discriminable change for V-A and V-T stimuli is approximately 30 msec, and it occurs over a range of Ts, generally at 30-100 msec above or below Ts=0 msec. With the method of adjustment, in which bias due to memory or criterion uncertainty is minimized, the just-discriminable change is approximately a constant over the range of Ts between ±120 msec. With the method of adjustment, greater improvement in these two cross-modality combinations is seen than was the case for the A-T combination. For the larger values of Ts, the differences between cross-modality stimuli involving vision and the A-T and within-modality cases are less pronounced. The greatest differences between cross-modality pairs are found when Ts=0 msec, as had been the case with within-modality stimuli.

### Tentative model of temporal resolution based on

physiological data.

The relation between cross-modality temporal resolution and within-modality temporal resolution was considered in Chapter Five. It was suggested there that the level of temporal resolution seen in each class of stimulus pairs was inconsistent with the idea that a single processor operating at a localized neural center mediates both within- and cross-modality temporal discrimination. It is stressed that such tasks as order labeling may be subject to different rules, in which case the idea of a single center to organize events in time may still be viable (Hirsh and Sherrick, 1961). It is shown here that the differences between within-modality and cross-modality temporal resolution can be partially accounted for by assuming that the tasks are mediated at successively higher neural centers and that the variance in neural arrival times provides the first constraint on resolving power. This tentative model is based on estimates of the standard deviation of single neuron latencies at various levels of the primary sensory pathways.

In Chapter One, the preservation of temporal information at the level of single neurons was summarized. The index of the precision with which an onset time can be encoded is the variance in the neuron's latency to a pulsed stimulus. Often, the variance was inferred from other information, such as the highest frequency at which the neuron could discharge in phase with a modulated stimulus. For such data, the following conservative approximation of the standard deviation of arrival times is used. When phase-locked firing is observed, discharges occur predominantly in one half-cycle of the stimulus. At the highest frequency for which phase-locking is observed, an interval equal to one-half of the period of the excitatory half-cycle is taken as the standard deviation. Thus if phase-locking fails at 100 Hz, the estimated standard deviation is 2.5 msec.

For audition, little temporal information has been lost at the cortical level. deRibaupierre, Goldstein, and Yeni-Komshian (1972) reported that a subpopulation of cortical neurons, called "lockers", had latencies that did not vary by more than 2 msec to repeated click stimuli. For the model, a standard deviation of 1 msec will be assumed. The same value will be used for all lower centers, since precision cannot be worse at those levels.

In the tactile system, Mountcastle, Talbot, Sakata, and Hyvarinen (1969) described a class of units in the somatosensory cortex, designated "thin-spike" units, that could be entrained by sinusoidal stimulation at frequencies up to 100 Hz. The response to single pulses was not described; a standard deviation of 2.5 msec will be used for prediction. Mountcastle et al. (1969) report that all thalamocortical projection fibers will follow the stimulus waveform at the same frequency; the same estimate of precision is used to characterize that level.

In the visual system alone, a monotonic decrease in the precision of temporal representation as reported as higher-ordered neurons are examined (van de Grind, Grusser, and Lunkenheimer, 1973). In the lateral geniculate nucleus (LGN), Spekreisje, van Norren, and van den Berg (1971) reported that individual neurons responded to flickering stimuli with many of the characteristic properties of human flicker perception. For their data, the highest effective frequency was approximately 30 Hz. From this, a standard deviation of 10 msec is estimated. These authors noted that further stages of visual processing could only degrade the precision of temporal information. In summarizing the data for visual cortical neurons, van de Grind et al. (1973) set an upper limit of approximately 5 Hz. If phase-locked responding cannot be maintained at that frequency, it is assumed that the standard deviation of arrival times can be no smaller than 50 msec.

Following the lead of Mountcastle et al. (1969) and Levick (1973), these standard deviations are used to predict the upper bounds of temporal resolution for two cases for which data were collected in the present experimental series. The Theory of Signal Detectability has shown that the distribution of events along a stimulus axis can be used to predict discrimination performance (Green and Swets, 1966). The detectability index, d', is the separation of two such distributions in standard deviation units. Thus, assuming that the distributions have equal variance, two intervals should be discriminable with d'=1.0 when the means of the two distributions are separated by one standard deviation.

The model can be tested by predicting the two-pulse threshold for each of the three modalities. In keeping with the findings of Spekreisje et al. (1971), it is assumed that the discrimination can be mediated at the thalamic level. The model predicts that the neural response to two elements will be discriminably different, with d'=1.0, from the response to one when Tc equals one standard deviation. This value is 1 msec for the auditory modality; 2.5 msec for the tactile modality; and 10 msec for the visual modality. These values slightly underestimate the published psychophysical data: 1.6 msec for auditory and 10 msec for tactile stimuli (Gescheider, 1966), and 20 msec for visual stimuli (Keitzman, 1967). However, the model is intended to predict the best possible performance. Later influences that could degrade the resolution estimated from peripheral encoding are not considered. In addition, the present data suggest that the figure given for tactile stimuli may be slightly high. In any case, the predictions agree with the data in ordering the thresholds across modalities.

The resolution of a cross-modality stimulus pair can be predicted assuming similar limitations in neural encoding. It is assumed that three independent sources of variance contribute to limit temporal resolution. Two of these are associated with the encoding of the onset time of the marker stimulus elements, as noted above. The third is associated with the neural timing of the interval between markers. This source of variance has often been modeled by a Poisson process (ie, Creelman, 1962). Here, the only assumption that is made regarding the timing variance is that it increases in proportion to the standard interval. Since these sources of variance are independent, they combine according to the relation,

$$\sigma_{ttl}^2 = \sigma_a^2 + \sigma_b^2 + \sigma_{tim}^2 \qquad \text{(eq. 6.1)}$$

where $\sigma_a^2$ and $\sigma_b^2$ are the variances of the marker latencies; $\sigma_{tim}^2$ is the variance associated with timing the interval; and $\sigma_{ttl}^2$ is the total variance of the encoded interval. The values of $\sigma_a^2$ and $\sigma_b^2$ depend upon the neural level at which access to the temporal information may be gained. For cross-modality cases, it is assumed that sub-cortical information is not available. Therefore, the time of occurance of the visual element can be specified only to within 50 msec. Thus the standard deviations of cross-modality intervals when Ts=0 msec are $\sqrt{2500 + 1}$ , or

50 msec for V-A stimuli; $\sqrt{2500 + 6}$, or 50 msec for V-T stimuli; and $\sqrt{1 + 6}$, or 3 msec for A-T stimuli. The model predicts that pairs of intervals that differ by these standard deviations will be discriminable with $d'=1.0$. Here, the disagreement with the data is greater. From Experiment 5, the obtained values of the standard deviation of adjustments were 26, 26, and 12 msec for V-T, V-A, and A-T stimulus combinations. Quantitatively, these predictions are poor. They are better approximations to the results of Experiments 1 and 2, and the V-A discrimination experiment conducted at CID. However, several aspects of the model are encouraging. It correctly predicts the ordering of performance for the cross-modality stimulus pairs. It predicts that the largest variance associated with the representation of the stimulus elements or with the timing of the interval will be the limiting factor in temporal resolution for those cross-modality combinations. Thus for cross-modality stimulus combinations, the involvement of the visual modality will account for the constancy of the just-discriminable change when Ts is less than 120-240 msec. However, as Ts increases, the timing variance will eventually become the dominant source. With A-T stimuli, the timing variance is the dominant source for small Ts as well, since the marker stimuli are precisely encoded. It was observed in all the experiments that modality differences are less pronounced at large values of

Ts, as the model suggests they should be. The model also provides an explanation for the discrepancy between within-modality and cross-modality performance that is intuitively acceptable. Finally, it provides some insight into the finding from Experiment 5 that the just-discriminable change for V-A and V-T stimulus pairs is essentially a constant for the 120 msec range tested. Under the model, the major source of variance over that range is that associated with the encoding of the visual stimulus, not that associated with timing the interval itself. Thus, the constancy of the just-discriminable change is expected, since the marker stimulus is encoded independent of its temporal relation to the stimulus in the other modality. The inability of the model to predict the absolute level of discrimination performance is not surprising, given the coarseness of the estimates of the physiological parameters and the fact that it estimates a lower limit for the just-resolvable change in Ts. The model does not at present provide a quantitative account of the findings of the experiments. It does provide a means for categorizing the results, and it seems to provide a first-order explanation for the major effects reported: the differences between within-modality and cross-modality temporal resolution; the form of the psychophysical function for cross-modality temporal resolution, including the constancy of the just-discriminable change in Ts when the stimulus includes a visual element; and the finding that whatever differences in capability characterize the stimulus conditions tested will tend to decrease as the standard interval increases.

The range of conditions investigated in these experiments was quite broad. No recent study of temporal resolving power has compared three sensory modalities. No study has compared cross-modality temporal resolution for three stimulus combinations or for a range of standard intervals. It is a characteristic of cross-modality temporal processing that varibility within and across subjects is quite high. However, an attempt was made to repeat the measurements under a variety of procedures until a stable pattern of performance had been identified, and the results are presented with confidence in their validity.

Those who study the feasibility of recoding speech, a complex time-varying information source, for presentation in other sensory modalities or combinations of modalities might find these results applicable to their problem. While no cross-modality combination approaches the temporal resolution of the auditory system, it seems likely that significant amounts of information can be transmitted by means of tactile or auditory-tactile devices. In terms of temporal capabilities, auditory-tactile presentation is superior to auditory-visual. Electrotactile and vibrotactile aids for the deaf are, of course, being developed. The present results merely confirm that the somatosensory system is the closest rival to audition in the time domain.

# LIST OF REFERENCES

Abel, S.M. (1972a). "Duration discrimination of noise and tone bursts," J. Acoust. Soc. Am., 51, 1219-1223.

Abel, S.M. (1972b). "Discrimination of temporal gaps," J. Acoust. Soc. Am., 52, 519-524.

Allan, L.G. (1975). "The relationship between judgments of successiveness and judgments of order," Percept. and Psychophys., 18, 29-36.

Allan, L.G. (1976). "Is there a constant minimum perceptual duration?" Quart. J. Exp. Psychol., 28, 71-76.

Allan, L.G. and Kristofferson, A.B. (1974). "Successiveness discrimination: two models," Percept. and Psychophys., 15, 37-46.

Babkoff, H. (1975). "Dichotic temporal interactions: fusion and temporal order," Percept. and Psychophys., 18, 267-272.

Babkoff, H. and Sutton, S. (1963). "Perception of temporal order and loudness judgments for dichotic clicks," J. Acoust. Soc. Am., 35, 574-577.

Barlow, H.B. (1961). "Comment on neural quanta" in Sensory Communication, edited by W. Rosenblith (Wiley, New York), pp. 786-790.

Bartley, S.H. (1958). Principles of perception (Harper and Brothers, New York), 482 pp.

Boring, E.G. (1921). "The stimulus error," Am. J. Psychol., 32, 449-471.

Boring, E.G. (1933). The physical dimensions of consciousness (Century, New York), 251 pp.

Boring, E.G. (1942). Sensation and Perception in the History of Experimental Psychology (Appleton, New York), 644 pp.

Boring, E.G. (1959). A History of Experimental Psychology, 2nd edition (Appleton-Century-Crofts, New York), 777 pp.

Boynton, R.M. (1972). "Discrimination of homogeneous double pulses of light," in Handbook of Sensory Physiology Vol. VII/4: Visual Psychophysics, edited by Hurvich, L. and Jameson, D. (Springer-Verlag, Berlin), pp. 202-232.

Briggs, R. and Perrott, D.R. (1972). "Auditory apparent movement under dichotic listening conditions," J. Exp. Psychol., 92, 83-91.

Ciganek, L. (1965). "A comparative study of visual and auditory EEG responses in man," Electroenceph. and Clin. Neurophysiol., 18, 625-635.

Cobb, W.A. and Dawson, G.D. (1960). "The latency and form in man of the occipital potentials evoked by bright flashes," J. Physiol., 152, 108-121.

Collyer, C.E. (1974). "The detection of a temporal gap between two disparate stimuli," Percept. and Psychophys., 16, 96-100.

Corwin, T.R. and Boynton, R.M. (1968). "Transitivity of visual judgments of simultaneity," J. Exp. Psychol., 78, 560-568.

Creelman, C.D. (1962). "Human discrimination of auditory duration," J. Acoust. Soc. Am., 34, 582-593.

Davis, H. (1973). "Electric response audiometry, with special reference to the vertex potentials," in Handbook of Sensory Physiology Vol V/3: Clinical and Special Topics, edited by Keidel, W.D. and Neff, W.D. (Springer-Verlag, Berlin), pp 85-103.

de Lange, H. (1954). "Relationship between CFF and a set of low-frequency characteristics of the eye," J. Opt. Soc. Am., 44, 380-389.

deRibaupierre, F.; Goldstein, M.H. Jr.; and Yeni-Komshian, G. (1972). "Cortical coding of repetitive acoustic pulses," Br. Res., 48, 205-225.

Divenyi, P.L. and Danner, W.F. (1977). "Discrimination of time intervals marked by brief acoustic pulses of various intensities and spectra," Percept. and Psychophys., 21, 125-142.

Dodd, B. (1977). "The role of vision in the perception of speech," Percept., 7, 31-40.

Efron, R. (1963a). "The effect of handedness on the perception of simultaneity and temporal order," Brain, 86, 261-284.

Efron, R. (1963b). "The effect of stimulus intensity on the perception of simultaneity in right- and left-handed subjects," Brain, 86, 285-294.

Efron, R. (1963c). "Temporal order perception, aphasia, and de ja vu," Brain, 86, 403-425.

Efron, R. (1970). "Effect of stimulus duration on perceptual onset and offset latencies," Percept. and Psychophys., 8, 231-234.

Efron, R. (1973a). "Conservation of information by perceptual systems," Percept. and Psychophys., 14, 518-530.

Efron, R. (1973b). "An invariant characteristic of perceptual systems in the time domain," in Attention and Performance IV, edited by Kornblum, S. (Academic, New York), pp. 713-736.

Egan, J.P. (1975). Signal detection theory and ROC analysis (Academic, New York), 277 pp.

Egan, J.P.; Lindner, W.A.; and McFadden, D. (1969). "Masking-level differences and the form of the psychometric function," Percept. and Psychophys., 6, 209-215.

Elliott, P.B. (1959). "Tables of d'," Technical Report No. 97, Electronic Defense Group, University of Michigan, 40 pp.

Exner, S. (1875). "Experimentelle untersuchnung der einfachsten psychischen Processe," Pflug. Arch. ges. Physiol., 11, 403-432.

Fraisse, P. (1963). The Psychology of Time (Harper and Row, New York), 343 pp.

Fukada, Y.; Motokawa, K.; Norton, A.C.; and Tasaki, K. (1966). "Functional significance of conduction velocity in the transfer of flicker information in the optic nerve of the cat," J. Neurophysiol., 29, 698-714.

Fukada, Y. and Saito, H. (1971). "The relationship between response characteristics to flicker stimulation and receptive field organization in the cat's optic nerve," Vis. Res., 11, 227-240.

Galambos, R.; Schwartzkopf, J.; and Rupert, A. (1959). "Microelectrode study of superior olivary nuclei," Am. J. Physiol., 197, 527-536.

Geldard, F.A. (1970). "Vision, audition, and beyond," in Contributions to Sensory Physiology, edited by Neff, W.D. (Academic, New York), pp. 1-17.

Gengel, R.W. and Hirsh, I.J. (1970). "Temporal order: effects of single versus repeated presentations, practice, and verbal feedback," Percept. and Psychophys., 7, 209-211.

Gescheider, G.A. (1966). "The resolving of successive clicks by the ears and skin," J. Exp. Psychol., 71, 378-381.

Gescheider, G.A. (1967). "Auditory and cutaneous temporal resolution of successive brief stimuli," J. Exp. Psychol., 75, 570-572.

Getty, D.J. (1975). "Discrimination of short temporal intervals: a comparison of two models," Percept. and Psychophys., 18, 1-8.

Gibbon, J. and Rutschmann, R. (1969). "Temporal order judgments and reaction time," Science, 165, 413-415.

Goodfellow, L. (1934). "An empirical comparison of audition, vision, and touch in the discrimination of short intervals of time," Am. J. Psychol., 46, 243-258.

Green, D.M. (1971). "Temporal auditory acuity," Psychol. Rev., 78, 540-551.

Haber, R.N. and Standing, L.G. (1970). "Direct estimates of the apparent duration of a flash," Canad. J. Psychol., 24, 216-229.

Green, D.M. and Swets, J.A. (1966). Signal detection theory and psychophysics (Wiley, New York; reprinted by Krieger, Huntington, N.Y.), 479 pp.

Halliday, A.M. and Mingay, R. (1964). "On the resolution of small time intervals and the effect of conduction delays on the judgment of simultaneity," Quart. J. Exp. Psychol., 16, 35-46.

Harter, M.R. (1967). "Excitability cycles and cortical scanning: a review of two hypotheses of central intermittancy in perception," Psychol. Bull., 68, 47-58.

Hirsh, I.J. (1959). "Auditory perception of temporal order," J. Acoust. Soc. Am., 31, 759-767.

Hirsh, I.J. and Fraisse, P. (1964). "Simultaneite et succession de stimuli heterogenes," L'Annee Psychologique, 64, 1-19.

Hirsh, I.J. and Sherrick, C.E. (1961). "Perceived order in different sense modalities," J. Exp. Psychol., 62, 423-432.

Ikeda, H. and Wright, M.J. (1975a). "Spatial and temporal properties of sustained and transient neurones in area 17 of the cat's visual cortex," Exp. Br. Res., 22, 363-383.

Ikeda, H. and Wright, M.J. (1975b). "Retinotopic distribution, visual latency, and orientation tuning of 'sustained' and 'transient' cortical neurones," Exp. Br. Res., 22, 385-398.

James, W. (1890). Principles of Psychology (Henry Holt, New York).

Jeffress, L.A. (1948). "A place theory of sound localization," J. Comp. Physiol. Psychol., 41, 35-39.

Jeffress, L.A. (1972). "Binaural signal detection: vector theory," in Foundations of Modern Auditory Theory, Vol II, edited by Tobias, J. (Academic, New York), pp. 349-368.

Julesz, B. and Hirsh, I.J. (1972). "Visual and auditory perception: an essay of comparison," in Human Communication: A Unified View, edited by David, E.E. Jr. and Denes, P.B. (McGraw-Hill, New York), pp. 282-340.

Keitzman, M.L. (1967). "Two-pulse measures of temporal resolution as a function of stimulus energy," J. Opt. Soc. Am., 57, 809-814.

Keitzman, M.L. and Sutton, S. (1968). "The interpretation of two-pulse measures of temporal resolution in vision," Vis. Res., 8, 287-302.

Keitzman, M.L. and Gillam, B.J. (1972). "Visual temporal integration and simple reaction time," Percept. and Psychophys., 11, 333-340.

Kelly, D.H. (1961). "Visual responses to time-dependent stimuli. I. Amplitude sensitivity measurements," J. Opt. Soc. Am., 51, 422-429.

König, E. (1965). Contribution to "Discussion to the third round table: Modifications of Speech Audiometry," Int. Audiol., 4(2), 72-74.

Kristofferson, A.B. (1967). "Attention and psychophysical time," Acta Psychol., 27, 93-100.

Kristofferson, A.B. and Allan, L.G. (1973). "Successiveness and duration discrimination," in Attention and Performance IV, edited by Kornblum, S. (Academic, New York), pp. 737-749.

Lane, H. (1965). "The motor theory of speech perception: a critical review," Psychol. Rev., 72, 275-309.

Lele, P.P.; Sinclair, D.C.; and Weddell, G. (1954). "The reaction time to touch," J.Physiol., 123, 187-203.

Leshowitz, B. (1971). "Measurement of the two-click threshold," J. Acoust. Soc. Am., 49, 462-466.

Levick, W.R. (1973). "Variation in the response latency of cat retinal ganglion cells," Vis. Res., 13, 837-853.

Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am., 49, 467-477.

Liberman, A.M.; Cooper, F.S.; Shankweiler, D.P.; and Studdert-Kennedy, M. (1967). "Perception of the speech code," Psychol. Rev., 74, 431-461.

Liberman, A.M.; Harris, K.S.; Kinney, J.A.; and Lane, H. (1961). "The discrimination of relative onset time of the components of certain speech and non-speech patterns," J. Exp. Psychol., 61, 379-388.

Maffei, L.; Cervetto, L.; and Fiorentini, A. (1970). "Transfer characteristics of excitation and inhibition in cat retinal ganglion cells," J. Neurophysiol., 33, 276-284.

Matteson, H.H. (1970). "Effects of surround luminance on perceptual latency," J. Opt. Soc. Am., 60, 1125-1131.

McKay, D.M. and Jeffreys, D.A. (1973). "Visual evoked potentials and visual perception in man," in Handbook of Sensory Physiology, Vol. VII/3B: Central Visual Information, edited by Jung, R. (Springer-Verlag, Berlin), pp 647-678.

Miller, J.D.; Wier, C.C.; Pastore, R.E.; Kelly, W.J.; and Dooling, R.J. (1976). "Discrimination and labeling of noise-buzz sequences with varying noise lead times: an example of categorical perception," J. Acoust. Soc. Am., 60, 410-417.

Mountcastle, V.B.; Talbot, W.H.; Sakata, H.; and Hyvarinen, J. (1969). "Cortical neuronal mechanisms in flutter-vibration studied in unanesthetized monkeys," J. Neurophysiol., 32, 452-484.

Neuhaus, W. (1930). "Experimentelle Untersuchnung der Scheinbewegnung," Arch. die ges. Psychol., 75, 315-458.

Nilsson, T.H. (1974). "Two-pulse interval vision thresholds," J. Opt. Soc. Am., 59, 753-757.

Oatley, K.; Robertson, A.; and Scanlon, P.M. (1969). "Judging the order of visual stimuli," Quart. J. Exp. Psychol., 21, 172-179.

Pagni, C.A. (1967). "Somatosensory evoked potentials in thalamus and cortex of man," in The Evoked Potentials, edited by Cobb, W. and Morocutti, C. (Elsevier, Amsterdam), pp. 147-155.

Parzen, E. (1962). Stochastic Processes (Holden-Day, San Francisco), 324 pp.

Pastore, R.E. (1976). "Categorical perception: a critical re-evaluation," in Hearing and Davis, edited by Hirsh, S.K.; Eldredge, D.H.; Hirsh, I.J.; and Silverman, S.R. (Washington University Press, St. Louis), pp. 253-264.

Patterson, J.H. and Green, D.M. (1970). "Discrimination of transient signals having identical energy spectra," J. Acoust. Soc. Am., 48, 894-905.

Penner, M.J. (1975). "The perception of offset: a problem of decision criteria," Percept. and Psychophys., 17, 587-590.

Perrott, D.R. (1974). "Auditory apparent motion," J. Aud. Res., 14, 163-169.

Pieron, H. (1952). The Sensations, translated by Pirenne, M.H. and Abbott, B.C. (Yale University Press, New Haven), 468 pp.

Pisoni, D.B. (1977). "Identification and discrimination of the relative onset time of two component tones: implications for voicing perception of stops," J. Acoust. Soc. Am., 61, 1352-1361.

Raab, D. and Fehrer, E. (1962). "Supplementary report: the effect of stimulus duration and luminance on visual reaction time," J. Exp. Psychol., 64, 326-327.

Raab, D.; Fehrer, E.; and Hershenson, M. (1961). "Visual reaction time and the Broca-Sulzer phenomenon," J. Exp. Psychol., 61, 193-199.

Robinson, D.N. (1967). "Visual discrimination of temporal order," Science, 156, 1263-1264.

Roufs, J.A.J. (1963). "Perception lag as a function of stimulus luminance," Vis. Res., 3, 81-91.

Rousseau, R. and Kristofferson, A.B. (1973). "The discrimination of bimodal temporal gaps," Bull. Psychon. Soc., 1, 115-116.

Rutschmann, J. and Link, R. (1964). "Perception of temporal order of stimuli differing in sense mode and reaction time," Perc. and Mot. Skills, 18, 345-352.

Rutschmann, R. (1966). "Perception of temporal order and relative visual latency," Science, 152, 1099-1101.

Rutschmann, R. (1973). "Visual perception of temporal order," in Attention and Performance IV, edited by Kornblum, S. (Academic, New York). pp. 687-701.

Schmidt, M.W. and Kristofferson, A.B. (1963). "Discrimination of successiveness: a test of a model of attention," Science, 139, 112-113.

Sherrick, C.E. and Rogers, R. (1966). "Apparent haptic movement," Percept. and Psychophys., 1, 175-180.

Siegel, S. (1956). Nonparametric statistics for the behavioral sciences (McGraw-Hill, New York), 312 pp.

Simmons, J.A. (1973). "The resolution of target range by echolocating bats," J. Acoust. Soc. Am., 54, 157-173.

Sinex, D.G. (1978). "Cross-modality temporal resolution for auditory, vibrotactile, and visual stimuli," J. Acoust. Soc. Am., 63, S52 (Abstract).

Smith, W.F. (1933). "The relative quickness of visual and auditory perception," J. Exp. Psychol., 16, 239-257.

Spekreisje, H.; van Norren, D.; and van den Berg, T.J.T.P. (1971). "Flicker responses in monkey lateral geniculate nucleus and human perception of flicker," Proc. Nat. Acad. Sci., U.S.A., 68, 2802-2805.

Sternberg, S. and Knoll, R.L. (1973). "The perception of temporal order: fundamental issues and a general model," in Attention and Performance IV, edited by Kornblum, S. (Academic, New York), pp. 629-685.

Stevens, S.S. (1961) "Is there a quantal threshold?" in Sensory Communication, edited by Rosenblith, W. (Wiley, New York), pp. 806-814.

Stone, J. and Hoffman, K.-P. (1971). "Conduction velocity as a parameter in the organization of the afferent relay in the cat's lateral geniculate nucleus," Br. Res., 32, 454-460.

Stone, S.A. (1926). "Prior entry in the auditory-tactual complication," Am. J. Psychol., 37, 284-287.

Stroud, J.M. (1955). "The fine structure of psychological time," in Information Theory in Psychology, edited by Quastler, H. (Free Press, Glencoe, Ill.), pp. 174-205.

Studdert-Kennedy, M.; Liberman, A.M.; Harris, K.S.; and Cooper, F.S. (1970). "Motor theory of speech perception: a reply to Lane's critical review," Psychol. Rev., 77, 234-249.

Sweet, A.L. (1953). "Temporal discrimination by the human eye," Am. J. Psychol., 66, 185-198.

Swets, J.A.; Tanner, W.P. Jr.; and Birdsall, T.G. (1961). "Decision processes in perception," Psychol. Rev., 68, 301-340.

Swisher, L. and Hirsh, I.J. (1970). "Brain damage and the ordering of two temporally successive stimuli," Neuropsychol., 10, 137-152.

Taylor, M.M. and Clarke, D.P.J. (1971). "Monaural detection with a contralateral cue (MDCC). II. Interaural delay of cue and signal," J. Acoust. Soc. Am., 49, 1243-1253.

Thor, D.H. (1967). "Dichoptic viewing and temporal discrimination: an attempted replication," Science, 158, 1704-1705.

Titchener, E.B. (1908). Lectures on the Elementary Psychology of Feeling and Attention (MacMillan, New York), 404 pp.

van de Grind, W.A.; Grusser, O.-J.; and Lunkenheimer, H.-U. (1973). "Temporal transfer properties of the afferent visual system," in Handbook of Sensory Physiology, Vol. VII/3A: Central Visual Information, edited by Jung, R. (Springer-Verlag, Berlin), pp. 431-573.

Vaughan, H.G. and Hull, R.C. (1965). "Functional relation between stimulus intensity and photically-evoked cerebral responses," Nature, 206, 720-722.

Verillo, R. (1963). "Investigation of some parameters of the cutaneous threshold for vibration," J. Acoust. Soc. Am., 34, 1768-1773.

Watson, C.S. (1973). "Psychophysics," in Handbook of General Psychology, edited by Wolman, B.B. (Prentice-Hall), pp. 275-306.

Watson, C.S.; Sinex, D.G.; and Kelly, W.J. (1976). "Relations between judged auditory-visual simultaneity and bimodal temporal resolution," Bull. Psychonom. Soc., 8, 247 (Abstract).

Westheimer, G. and McKee, S.P. (1977). "Perception of temporal order in adjacent visual stimuli," Vis. Res., 17, 887-892.

Whipple, G.M. (1899). "On nearly-simultaneous clicks and flashes," Am. J. Psychol., 10, 280-286.

Wier, C.C. and Green, D.M. (1975). "Temporal acuity as a function of frequency difference," J. Acoust. Soc. Am., 57, 1512-1515.

Williams, K.N. and Perrott, D.R. (1972). "Temporal resolution of tonal pulses," J. Acoust. Soc. Am., 51, 644-647.

Woodrow, H. (1951). "Time perception," in Handbook of Experimental Psychology, edited by Stevens, S.S. (Wiley, New York), pp. 1224-1236.

Woodworth, R.S. (1938). Experimental Psychology (Holt, New York), 889 pp.

Woodworth, R.S. and Schlosberg, H. (1954). Experimental Psychology (Holt, Rinehart, and Winston, New York), 948 pp.

Yund, E.W. and Efron, R. (1974). "Dichoptic and dichotic micropattern discrimination," Percept. and Psychophys., 15, 383-390.

# APPENDIX ONE

This appendix describes a Monte Carlo method for estimating the expected value of the variability of the stimulus track from the three-alternative forced-choice adaptive psychophysical procedure employed in Experiment 4. The estimation procedure takes advantage of the availability of fitted psychometric functions from Experiments 1 and 2 for the subjects and stimulus conditions used in Experiment 4.

For the simulation of an adaptive run, it is assumed that the response on each trial is governed by a stationary process. On each trial, the probability of a correct response can be obtained by evaluating the equation for the fitted psychometric function using the current stimulus level. This yields a value of d', which may be converted to the probability of a correct response according to the approximation formula given by Elliott (1959). The probability may be used in conjunction with a random number generator to simulate the occurence of correct responses on the appropriate proportion of the trials. Once a response has been generated by this procedure, the Levitt (1971) stepping rule is applied exactly as it is in the collection of real data.

A simulation of each session for each subject was carried out, with the parameters of the simulated run chosen to match those of the actual run. The starting stimulus level, the step size, and the number of trials were always equated. For each subject, the fitted psychometric function from Experiment 1 or 2 that had yielded the lowest estimate of the $\Delta T$ for $d'=1.0$ was used to generate the responses. A total of 33 sessions were simulated.

The variability of the stimulus track is given by the standard error of the estimate of the target stimulus level. The method for calculating the standard error is described in Chapter Four. A scatterplot af the actual standard errors and the standard errors of the corresponding simulations is shown in Fig. A1.1. The simulated standard error is consistently smaller than the actual value. The difference may be assessed by comparing the two sampling distributions, one made up of the actual standard errors and one consisting of the values obtained by simulation. The results of a Wilcoxin matched-pairs signed-ranks test (Siegel, 1956) indicate that the difference is highly significant ($z=4.31$; 33 cases; $p<0.001$ for a 2-tailed test). The mean levels of $\Delta T$ reached by the actual data and by the simulated runs also differ ($z=2.17$; 33 cases; $p<0.05$ for a 2-tailed test); the simulated runs end at slightly lower levels. The difference is not large, relative to the standard errors, however. It is concluded that the process governing the subjects' responses on successive trials is nonstationary.

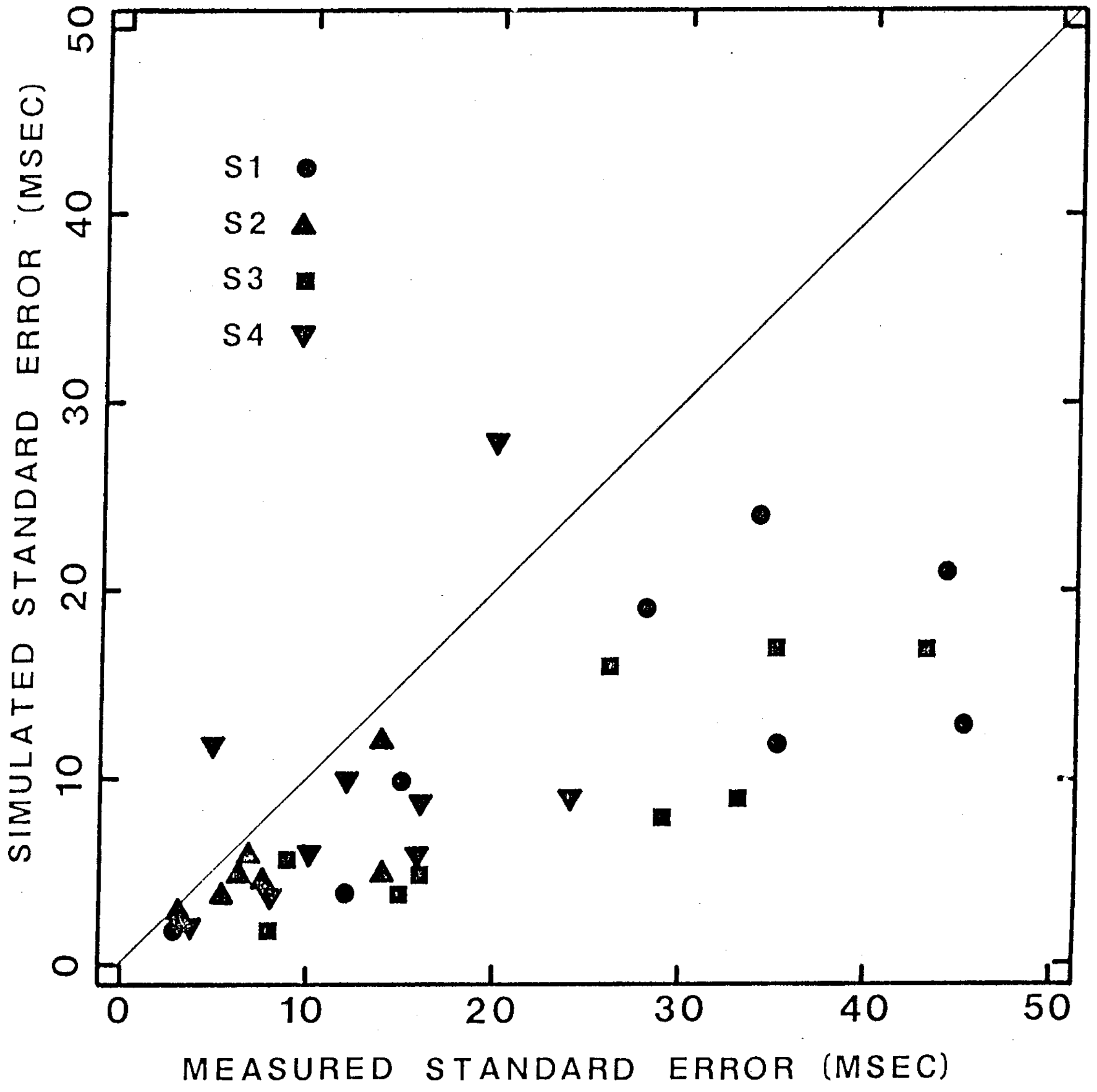

Figure A1.1. Comparison of the variability of actual and simulated stimulus tracks. Scatter diagram comparing the variability of the stimulus tracks obtained in Experiment 4 to the variability of tracks obtained from Monte Carlo simulations of the experimental sessions. Abscissa: standard error of the target level of the actual runs; data from four subjects. Ordinate: standard error of the target level of the simulated runs. Thirty-three sessions are represented. The variability of simulated data is significantly lower than that of the actual data.

# APPENDIX TWO

Two devices were used to generate tactile stimuli in the experiments. In Experiment 5, piezoelectric benders were used. In all other experiments the electromagnetic vibrator delivered the stimulus. Benders have been used by others to deliver pulse stimuli because of their frequency characteristics. They are less resonant than other vibrators, such as the Goodmans V-47. Unfortunately, the benders selected for these experiments proved to be unsatisfactory for single-presentation procedures because deflections of sufficient amplitude could not be generated. Therefore, the more intense electromagnetic device was used.

There is no reason to suspect that the devices used would effect the form of the results. However, as a precaution, a brief series of control conditions that compare the two stimuli were run as part of Experiment 5. Subjects 5-7 participated. Tactile within-modality temporal resolution was measured using each device, and in addition, each subject was tested with a cross-modality stimulus combination using the electromagnetic device.

No differences in performance were observed. Fig. A2.1a plots the group mean standard deviation of tactile within-modality matches as a function of Ts; the parameter is the source of the stimulus. In Fig. A2.1b, the group mean constant error (CE) is shown, again for each device.

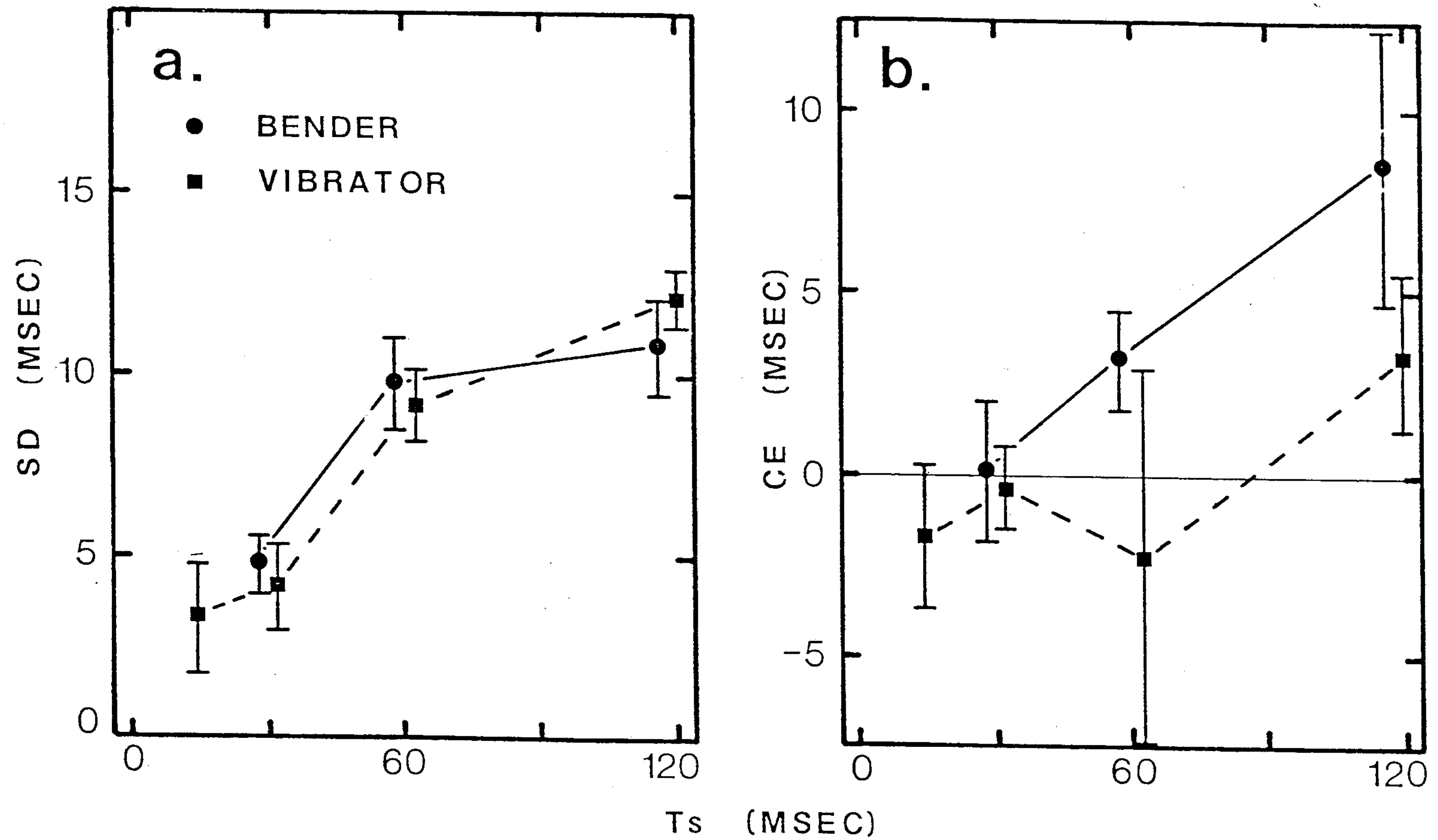

Figure A2.1. Comparison of two vibrators. Temporal resolution for tactile-tactile pairs defined by elements generated by two devices is compared. Panel a compares session standard deviations. The solid line summarizes data obtained using the piezoelectric vibrator, and the dashed line summarizes data obtained using the electromagnetic vibrator. No difference in results is apparent. Panel b summarizes the constant error in matches with each device. Small differences are found, but fluctuations in the CE have been suggested to be independent of temporal resolution.

The standard devations are essentially identical. Individual data also reveal no differences. There is a difference in the CE measured with each device. However, the CE is within a standard deviation of Ts in all cases. Slight fluctuations in the CE are expected.

In addition to the within-modality control, a cross-modality replication was conducted with each subject, using the electromagnetic vibrator. For S5, the stimuli were V-T; for S6 and S7 the stimuli were A-T. For each subject, the results of the replication were within the range of performance measured on the original tests using the piezoelectric vibrator. It is concluded that the results of the experiments were not influenced by the type of tactile stimulus delivered and that comparisons across procedures are not confounded by the stimulus change.