

AD-A190 337

AUDITORY PATTERN MEMORY: MECHANISMS OF TONAL SEQUENCE
DISCRIMINATION BY H. (U) PURDUE UNIV LAFAYETTE IN DEPT
OF PSYCHOLOGICAL SCIENCES R D SORKIN 30 SEP 87
AFOSR-IR-87-1775 AFOSR-84-0302

1/1

UNCLASSIFIED

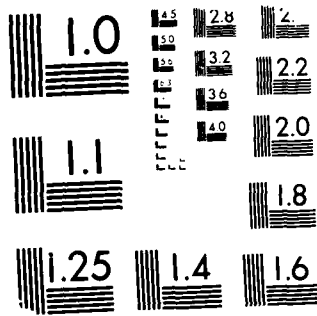
F/G 12/6

NL

END

DATE

48



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

2

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

AD-A190 337

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED		DTIC SELECTED		1b RESTRICTIVE MARKINGS	
2 SECURITY CLASSIFICATION AUTHORITY		SCHEDULED JAN 19 1988		3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.	
4 DECLASSIFICATION/DOWNGRADING SCHEDULE		D		5 MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR-87-1775	
6a NAME OF PERFORMING ORGANIZATION Purdue University		6b OFFICE SYMBOL (if applicable)		7a NAME OF MONITORING ORGANIZATION AFOSR/NL	
7 ADDRESS (City, State, and ZIP Code) West Lafayette, IN 47907				7b ADDRESS (City, State, and ZIP Code) Building 410 Bolling AFB DC 20332-6448	
8a NAME OF FUNDING/SPONSORING ORGANIZATION Office of Scientific Research		8b OFFICE SYMBOL (if applicable) NL		9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-84-0302	
10 ADDRESS (City, State, and ZIP Code) Bolling AFB, D.C. Building 410		10 SOURCE OF FUNDING NUMBERS			
		PROGRAM ELEMENT NO 61102F	PROJECT NO 2313	TASK NO A6	WORK UNIT ACCESSION NO
11 TITLE (Include Security Classification) Auditory Pattern Memory					
12 PERSONAL AUTHOR(S) Sorkin, Robert D.					
13a TYPE OF REPORT Final		13b TIME COVERED FROM 84-9-1 TO 87-8-31		14 DATE OF REPORT (Year, Month, Day) 87-9-30	
15 PAGE COUNT 38					
16 SUPPLEMENTARY NOTATION					
17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP	Auditory perception, auditory sequence discrimination auditory patterns, auditory memory, temporal uncertainty, models of auditory signal processing		
19 ABSTRACT (Continue on reverse if necessary and identify by block number) A two-process model of pattern discrimination was developed to describe how tonal sequences are processed, stored, and discriminated by the human auditory system. The model was tested in tasks in which subjects were required to discriminate between the frequency patterns encoded in two sequences of tones. The experimental results strongly supported the assumptions of a trace and context coding mechanism and indicated that the trace mechanism is relatively insensitive to temporal transformations made to the stimulus. An attempt to model the pattern discrimination mechanism with specific computational algorithms was less successful. A technique was developed to assess the manner in which information is accumulated from elements of an auditory or visual stimulus. Results indicate that the technique may be useful in the design of display systems.					
20 DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT <input type="checkbox"/> DTIC USERS			21 ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED		
22a NAME OF RESPONSIBLE INDIVIDUAL JOHN F. TANGNEY			22b TELEPHONE (Include Area Code) 202 767-5021		22c OFFICE SYMBOL NL

Form 1473, JUN 86

Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

87 12 9

AFOSR-TR- 87-1775

AUDITORY PATTERN MEMORY

Mechanisms of Tonal Sequence Discrimination by Human Observers

Robert D. Sorkin
Department of Psychological Sciences
Purdue University
West Lafayette, Indiana 47907

30 September 1987

Final Technical Report for Period 1 September 1984 - 31 August 1987

Prepared for

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Bolling Air Force Base, DC 20332-6448

SUMMARY

A two-process model of pattern discrimination was developed to describe how tonal sequences are processed, stored, and discriminated by the human auditory system. In this model, the comparison of auditory stimuli involves processing in two separate concurrent modes, with the mode contributing the least variance having the dominant effect on performance. In the 'sensory trace' mode, the subject compares a trace of the auditory sensation to the corresponding trace of a subsequent stimulus. The trace is assumed to decay with time, but be independent of the number of possible stimuli. In the 'context coding' mode, stimuli are discriminated by comparing encoded representations, each representation consisting of a comparison between the given stimulus and the overall context of other stimuli heard. The quality of the encoded representation is assumed to depend on the size of the auditory context.

This model was applied to tasks in which a subject compared two sequences of tones and judged whether the frequency patterns were the same or different. Three aspects of the sequences were manipulated: (a) the temporal variability of the sequences, (b) the correlation between the temporal envelopes of each pair of sequences presented on a trial, and (3) the time interval between the two sequences. In the 'correlated' condition, the two sequences on a trial had identical temporal patterns, while in the 'uncorrelated' condition, all sequence temporal patterns were generated independently. In order to study the properties of the trace mechanism, various temporal and spectral transformations were made to the second sequence of each pair.

Performance in the correlated condition was independent of temporal variability, but dropped with increases in the time interval between sequences. Performance in the uncorrelated condition decreased with increased temporal variability, and was independent of the length of the inter-sequence-interval. These results support the assumptions of the trace/context theory; good fits were obtained between data from human observers and the predictions of the two-process model. Further, performance in the correlated condition was remarkably insensitive to certain types of transformations made to the second sequence, suggesting that (a) sensory traces can be generated for separate aspects of a sensory stimulus, or (b) traces can be temporally transformed or "scaled" for comparison with other stimuli.

An analysis of the underlying pattern discrimination mechanism was done by evaluating two computational models of sequence comparison: a weighted average algorithm and an optimal sequence matching algorithm. The overall pattern of performance for both models was similar to that of the human subjects, but trial-by-trial correlations between the human and algorithm responses was low, suggesting that neither algorithm provided an accurate description of the human's discrimination mechanism.

Some general assumptions of the signal detection model were

✓
□
□
Codes



A-1

evaluated in experiments with auditory tonal sequences and visual displays generated by similar statistical rules. These assumptions concern the manner in which the observer processes information from individual elements of the auditory sequence or visual array. The experimental results were consistent with expectations: (1) In sequential auditory displays, the first arriving (primacy) and last arriving (recency) elements are emphasized; (2) with visual displays the fixated region is emphasized, the effective size of the region depending on the element coding and format.

TABLE OF CONTENTS

	Page
1. INTRODUCTION AND PURPOSE	1
1.1 Basic Experimental Paradigm	1
2. SEQUENCE DISCRIMINATION THEORY	3
3. SEQUENCE DISCRIMINATION EXPERIMENTS	5
3.1 General Method	6
3.2 Experiment 1	8
3.2.1 Procedure	8
3.2.2 Results	8
3.3 Experiment 2	9
3.3.1 Procedure	9
3.3.2 Results	9
3.4 Experiment 3	11
3.4.1 Procedure	11
3.4.2 Results	13
3.5 Estimation of model parameters	13
3.6 Discussion	15
4. SEQUENCE COMPARISON ALGORITHMS	16
4.1 Weighted Average Algorithm	18
4.2 Optimal Matching Algorithm	19
4.3 Analysis of Algorithm and Human Performance	20
5. LINEAR TRANSFORMATION OF SEQUENCE DURATION	23
5.1 Duration Scaling Experiment	24
5.2 Results and Discussion	24
6. PROCESSING INFORMATION FROM TONAL SEQUENCES AND VISUAL DISPLAYS	27
7. REFERENCES	34
8. PROJECT PERSONNEL	35
9. ADVANCED DEGREES	36
10. PUBLICATIONS	36
11. INTERACTIONS	37

LIST OF FIGURES

Figure		Page
1	Examples of the pairs of sequences present on trials of experiment 1. The first pair of sequences has different patterns of high and low frequency tones but the same pattern of intertone gaps. The second pair of sequences has the same frequency pattern but different temporal patterns.	2
2	Average performance (plotted points) as a function of the gap standard deviation and condition (C,U) in experiment 1 (8 tones, mean gap = 100 ms). Solid lines are theoretical functions (see text).	8
3	Average performance (plotted points) as a function of the gap standard deviation and condition (C,U) in experiment 2a (8 tones, mean gap = 50 ms). Solid lines are theoretical functions (see text).	10
4	Average performance (plotted points) as a function of the mean gap duration and condition (C,U) in experiment 2b (12 tones, gap standard deviation = 40 ms). Solid lines are theoretical functions (see text).	10
5	Average performance (plotted points) as a function of the number of tones and condition (C,U) in experiment 2c (mean gap = 50 ms, gap standard deviation = 40 ms). Solid lines are theoretical functions (see text).	10
6	Examples of the pairs of sequences present on trials of experiment 3 (see text).	11
7	Average performance (plotted points) as a function of the gap standard deviation and condition (C,S,U) in experiment 3a (8 tones, mean gap = 50 ms, inter-sequence-interval = 0.5 s). Solid lines are theoretical functions (see text).	12
8	Average performance (plotted points) as a function of the inter-sequence-interval (ISI) and condition (C,S,U) in experiment 3b (8 tones, mean gap = 50 ms, gap standard deviation = 30 ms). Solid lines are theoretical functions (see text).	12
9	Average performance (rating task) as a function of the gap standard deviation and condition (C,S,U) for a mean gap duration of 50 ms. for the human observers (plotted symbols) and for the two comparison algorithms (solid and dashed lines).	21

LIST OF FIGURES (continued)

Figure		Page
10	Receiver Operating Characteristic (ROC) curves for human observers RAT and AMC (plotted points) and best linear fits to the algorithms (solid and dashed lines).	21
11	Examples of pairs of tonal sequences in the expanded conditions of experiment 4.	24
12	Average performance as a function of inter-sequence-interval, gap standard deviation, and temporal transformation, for the correlated conditions of experiment 4.	26
13	Average performance as a function of inter-sequence-interval, gap standard deviation, and temporal transformation, for the uncorrelated conditions of experiment 4.	26
14	Average performance (d') of four observers to a frequency-coded sequential auditory stimulus as a function of the number of tones in the sequence. Solid line: square-root-of-n prediction. Dashed line: prediction of a partitioned variance model.	28
15	Average performance (d') of three observers to a visual display as a function of the number of display elements. Solid line: square-root-of-n prediction.	29
16	Probability of observer's "yes" response given the particular magnitudes (over trials) of the display element in a visual detection task (two different elements).	31
17	Slope of the Probability vs. Element Magnitude curves as a function of element sequential position (auditory display experiment). Curves are given for sequences composed of 6 tones and 10 tones.	31
18	Slope of the Probability vs. Element Magnitude curves as a function of element spatial position in the visual display-straight line array of digital elements.	32
19	Slope of the Probability vs. Element Magnitude curves as a function of element spatial position in the visual display-straight line array of analog elements.	32
20	Slope of the Probability vs. Element Magnitude curves as a function of element spatial position in the visual display-square array of digital elements.	33
21	Slope of the Probability vs. Element Magnitude curves as a function of element spatial position in the visual display-square array of analog elements.	33

LIST OF TABLES

Table		Page
1	Summary of variables and conditions for the sequence discrimination experiments. Each experiment (1,2,3) employed a different group of three subjects. (Frequencies in Hertz, durations in milliseconds.)	7
2	Parameter values for simplified model (Eqs. 4a and 4b).	14
3	Model parameter values (assume $c_2 = 0.001$).	14

1. INTRODUCTION AND PURPOSE

A "simple" sound, such as a brief pure tone, can be made into a "complex" sound by concatenating the "simple" sounds into sequences. Our goal is to understand how such tonal patterns are processed by the human auditory system. First, we extend an existing model of auditory signal discrimination to the complex sequence discrimination task. Second, we test some specific comparison algorithms which may form the basis for the sequence pattern comparison operation. Third, we extend the general discrimination model to the case where linear transformations in duration have been applied to the tonal patterns, and we clarify some special properties of auditory trace memory. Finally, we summarize an applications of the detection theory model to the analysis of multiple element auditory and visual displays.

1.1 The Basic Experimental Paradigm

Many investigators have noted that varying the parameters of a tonal sequence can have a large effect on performance in a discrimination or detection task. For example, Watson and his colleagues have reported a number of studies in which a subject had to detect small differences between tonal sequences whose parameters varied across the experimental trials. This type of trial-by-trial variation produced large changes in the subject's ability to detect small changes in the frequency, duration, or intensity of one component of the tonal sequence (Watson, Kelly and Wroton, 1976; Spiegel, and Watson, 1981; and Leek and Watson, 1984). Watson (1987) pointed out that these effects can be much larger than those obtained from single tone (2IFC) discrimination tasks in which the tone parameters vary over trials.

We reported similar effects in an experiment which tested the discrimination of temporal jitter in tonal sequences (Sorkin, Boggs and Brady, 1982). Subjects in that experiment were presented with a pair of tone sequences composed of high and low frequency tone bursts. The subject's task was to discriminate differences in the temporal envelopes of the two sequences and ignore differences in the frequency pattern of the sequences. We were interested in how temporal discrimination depended on the pattern of tone frequencies. When the frequency patterns were fixed over trials, performance was quite good, and was described well by a simple model of temporal discrimination. When the frequency pattern for each sequence was varied randomly over trials, but fixed for the pair of sequences within a trial, performance was poor. When the binary pattern was random both across and within trials, performance was even worse.

The present experiments continue our examination of the effects of spectral and temporal variation on sequence discrimination. In the current experiments, the subject must discriminate differences between the frequency patterns of two tonal sequences and ignore differences in the temporal envelopes of the sequences. We were interested in the effect of variations in the temporal envelope of the sequences on the subject's

ability to discriminate different patterns of tone frequency.

The subject's ability to discriminate between the sequence frequency patterns was evaluated with a same/different task. On each trial, subjects were presented with a pair of sequences and reported whether the frequency pattern in the two sequences was the same or different. On approximately half of the trials, the two sequences had different frequency patterns; on approximately half of the trials, they had the same frequency sequence. Subjects were instructed to ignore any variations in the temporal envelope of the sequences, such as jitter in the tone onsets, durations, or inter-tone gaps.

An important aspect of the experimental paradigm was the manner in which the temporal envelopes of the sequences were varied. In one condition, the correlated condition, the temporal envelopes were varied across trials but were identical within trials, e.g. the pair of sequences within a trial had identical temporal envelopes. In a second condition, the uncorrelated condition, the temporal envelopes were varied both across and within experimental trials. Thus, in the uncorrelated condition, the pair of tone sequences within a trial had different temporal patterns.

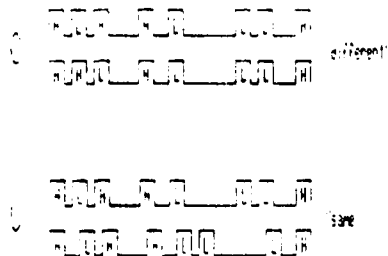


FIG. 1. Examples of the pairs of sequences present on trials of experiment 1. The first pair of sequences has different patterns of high- and low-frequency tones but the same pattern of intertone gaps. The second pair of sequences has the same frequency pattern but different temporal patterns.

These conditions are illustrated in Figure 1, which shows some of the tonal sequences possible on trials of a typical experiment. These tonal sequences were composed of tones of one of two frequencies: a high (H) frequency, or a low (L) frequency. In the first example the frequency sequences are different, therefore the correct response is "different". This example illustrates the correlated condition, in which the two sequences have identical envelopes. The second example illustrates a pair of sequences in the uncorrelated condition. The frequency pattern of these two sequences is the same, but the temporal envelopes are different; the correct response is now "same".

Suppose that the observer has available a memory trace of each input sequence, and that this trace is a close

representation of the acoustic input. In the correlated condition, the memory trace of the two sequences can be compared directly, since any difference between the traces of the two sequences will be due to differences in their respective frequency patterns. This assumes that the trace will not be degraded during the time interval it must be held prior to the comparison operation. In the uncorrelated condition the two input sequences do not have the same temporal envelope, and a direct comparison of the two memory traces is not appropriate. In the uncorrelated condition, a direct comparison of the traces of the two inputs may confound differences in the frequency patterns of the sequences with differences in their temporal envelopes. In the uncorrelated condition, it will be necessary for the observer to first transform or encode the input sequences, in order to remove those aspects of the sequences that are due to differences in the temporal envelopes.

2. SEQUENCE DISCRIMINATION THEORY

Tanner (1961) proposed a general framework for categorizing the memory requirements of a variety of single and two-interval psychophysical tasks; this framework has been extended and quantified by a number of workers including Sorkin (1962) in a study of the same/different task, and Macmillan, Kaplan, and Creelman (1977) in a study of categorical perception. Tanner's model included a short-term decaying memory for the acoustic input to the system plus various interference and long-term memory factors. The trace decay component of the Tanner model was extended by Kinchla and Smyzer (1967) and later incorporated as the trace component of the trace-context model developed by Durlach and his colleagues (e.g. Durlach and Braida, 1969; Berliner and Durlach, 1973). In Sorkin (1984), we proposed that the Durlach and Braida dual mode model could be extended to the sequence discrimination task.

According to the Durlach-Braida model, a subject can employ two different processing modes in a discrimination task: a trace mode and a context mode. In the trace mode, the comparison operation is performed on internal memory traces of the input signals. The trace is a direct representation of the acoustic input, e.g. a precategorical replica, and the trace decays or accumulates noise over time. Thus, performance in the trace mode will deteriorate as the time interval between the two inputs to be discriminated is increased. In the context mode, each input is first categorized into one of a defined set of codes. The comparison operation is then performed on these encoded representations. Once encoded, the internal data are not subject to degradation over time. However, there is a context noise present which is a function of the difficulty of the encoding process. Task variables such as the stimulus range are assumed to increase this context noise. Performance in a given task will usually involve trade offs in the internal noise associated with each mode of operation.

These considerations are incorporated into the Durlach and

Braida model in the following manner:

$$d' = c_1 / [\sigma_s^2 + (\sigma_c^2 + \sigma_t^2)^{-1/2}] \quad (1)$$

where c_1 is a constant associated with the mean distance between stimuli and,

σ_s^2 , σ_c^2 , and σ_t^2 are variance components associated with the

sensory, context, and trace aspects of the task, respectively. The sensory variance term refers to the noise from internal sources other than those associated with encoding and storing the sequence information. It can be seen from an examination of equation 1 that the trace and context variances will operate together in determining which process, or whether either process, will have an effect on performance. For example, if the variance associated with one process is much smaller than the variance associated with the other process, the process with the smaller variance will have the dominant effect on performance.

In the sequence discrimination tasks of the present study, we assume that the sensory variance is very small relative to the context and trace variances; therefore we shall ignore this term in the remainder of our analysis. The relative size of the trace and context variance components will depend on the magnitude and type of variability present in the sequence temporal envelopes. In general, we would expect that the magnitude of the trace variance would be high in the uncorrelated condition relative to the correlated condition, and therefore context mode processing would be dominant in the uncorrelated condition.

We assume that the variance of the trace component will increase with the presence of any differences between the two traces that are irrelevant to the frequency discrimination task. These differences may be produced by two main factors: (a) the time period during which the trace of the first sequence must be held in order to make a comparison with the trace of the second sequence, and (b) differences between the sequence traces, such as caused by temporal variability, that are irrelevant to the discrimination task. These variables are summarized in the following equation:

$$\sigma_t^2 = c_2 [ISI + nd + (n-1)g] + c_3 (1 - r) \quad (2)$$

In the equation, c_2 and c_3 are constants which apply to the storage time and temporal variability components, respectively. The first term specifies the mean duration of the interval from the beginning of one stimulus sequence to the beginning of the second sequence; ISI is the inter-sequence-interval, n is the

number of tones in each sequence, d is the average duration of each tone, and g is the average duration of the gaps between the tones. The magnitude of the second term depends on some measure of the correlation, r , between the temporal envelopes of the two sequences.

In the correlated conditions, r is approximately unity; in the uncorrelated conditions, r is a function of the variability of the temporal gap between tones, σ_g , the variability of the tone duration, σ_d , and the task parameters n , d , and g . Because we do not have a closed-form expression for the correlation of the sequence temporal envelopes as a function of σ_g , σ_d , g , n , and d , we use the expression $c_3(\sigma_g/g)^2$ as an approximation to the righthand term of equation 2, in the uncorrelated condition. This assumption yields the following expressions for performance in the sequence tasks:

Performance in the uncorrelated condition,

$$d'(\text{unc}) =$$

$$c_1 \left\{ \frac{\sigma_g^2}{c} + [c_2 (\text{ISI} + nd + (n-1)g) + c_3 \left(\frac{\sigma_g}{g} \right)^2] \right\}^{-1/2} \quad (3a)$$

and performance in the correlated condition,

$$d'(\text{cor}) = c_1 \left\{ \frac{\sigma_g^2}{c} + [c_2 (\text{ISI} + nd + (n-1)g)] \right\}^{-1/2} \quad (3b)$$

The context variance is assumed to depend on the complexity of the encoding operation required. Thus, the context variance is assumed to increase with the number of different frequency patterns that are possible during a block of trials of the sequence experiment. This is a slight departure from the approach taken by Durlach and Braida (1969) in which they assumed that the context variance is a function of the total range of the stimulus set. Task factors which do not increase the encoding requirement should not affect the context variance. Thus, we assume that the context variance will not be a function of variability in parameters related to the sequence temporal envelope, such as σ_g and σ_d . Of course one could argue that increasing the number of possible inputs, such as by adding temporal variability, will increase the difficulty of the encoding operation, and hence will increase context variance. Our preliminary assumption is that these factors will have a much greater influence on the trace comparison process than on the context encoding process.

3. SEQUENCE DISCRIMINATION EXPERIMENTS

3.1 General Method

We conducted experimental tests of several aspects of sequence discrimination performance described by the model. Table I summarizes the conditions and variables in our experiments. The major focus in these experiments was on the different pattern of performance in the correlated and uncorrelated conditions. In experiments 1, 2a, and 3a, the independent variable was the magnitude of the variability of the inter-tone gaps. These three experiments provide a picture of how performance depends on gap variability in the correlated and uncorrelated conditions. In experiment 2b, the independent variable was the mean gap between tones. This experiment allows performance in the correlated and uncorrelated conditions to be compared at different gap intervals and a constant gap variability. In experiments 2c and 3b, the independent variables were the number of tones in each sequence and the length of the inter-sequence interval. These experiments evaluate how performance in the correlated and uncorrelated conditions changes as the number of tones and inter-sequence-interval are increased.

These experiments employed different procedures for generating the set of frequency patterns and for generating the frequency patterns on "different" trials. The procedures were developed over a period of time using different groups of observers. The major procedural differences between experiments 1 and 2 and experiment 3 were (a) an increase in the complexity and number of possible frequency patterns and (b) the addition of variability to the tone durations. The motivation for increasing the complexity of the sequence patterns in experiment 3 was to make it more difficult for observers to learn the set of patterns used. In our early experiments we had observed that the dependence of performance on ISI (in the correlated condition) tended to be attenuated with extensive practice on the task. We believed that extensive practice with a small set of patterns could result in the context coding process becoming dominant in that condition. Variability in the tone durations was added in order to test the effect of synchronizing the tone onsets on the variability manipulation. These changes resulted in a large increase in the difficulty of the discrimination task. In order to bring observer performance into the same range as the other experiments, it was necessary to change the rule for generating patterns on "different" trials. The specific procedures relating to each experiment are described further below.

Three groups of subjects consisting of three subjects per group participated in the experiments. Thus, one group of subjects served in experiment 1; a second group of subjects served in experiments 2a, b, and c; and a third group served in experiments 3a and b. The subjects were undergraduate students paid an hourly rate plus a fractional incentive bonus for correct responses. All had normal hearing and extensive experience with psychoacoustic tasks. They performed the task for approximately two hours per day, three days per week. Subjects were seated in a double-walled Industrial Acoustics chamber. Stimuli were

presented to the subject's right ear over TDH-39 headphones. All signals were presented at 71 dB SPL.

TABLE I. Summary of variables and conditions for the sequence discrimination experiments. Each experiment (1, 2, 3) employed a different group of three subjects. (Frequencies are in Hertz, durations in milliseconds.)

Expt.	Difference rule (see text)	Frequencies	Number of patterns	Conditions	Number of tones (n)	Interval (ISI)	Tone duration (d)		Intertone gap (g)	
							mean	s. d.	mean	s. d.
1	two adjacent tones transposed	669,1621	54	correlated, uncorrelated	8	500	30.5	0	100	variable 20,40,60,80
2a	new pattern chosen from basic set of 16	669,1621	16	correlated, uncorrelated	8	500	30.5	0	50	variable 10,20,40
2b	new pattern chosen from basic set of 16	669,1621	96	correlated, uncorrelated	12	500	30.5	0	variable 50,100,150,200	40
2c	new pattern chosen from basic set of 16	669,1621	variable 16,96,576	correlated, uncorrelated	variable 8,12,16	500	30.5	0	50	40
3a	3 random tones changed	500,909,1667,2857	2520	correlated, uncorrelated, synchronized	8	500	40.0	12.25	50	variable 5,10,20,30,40
3b	3 random tones changed	500,909,1667,2857	2520	correlated, uncorrelated, synchronized	8	variable 500,1500,3000	40.0	12.25	50	30

The experimental task was a same/different task: subjects were presented with a pair of tonal sequences and had to determine if the frequency pattern in the two sequences was the same or different. Subjects answered via a set of two (same/different) pushbuttons. Subjects were instructed to ignore variations in the temporal structure of the sequences, such as jitter in the tone onsets, durations, or inter-tone gaps. Feedback about the correct response on each trial was provided to the subjects. All tone bursts were sinusoidal segments. Tones were synthesized via 12 bit (experiment 3) or 13 bit (experiment 1 and 2) digital to analog converters at a sampling rate of 20 KHz and low-pass filtered at 5KHz. A symmetric 2 ms linear envelope was used as a gating function. Temporal variability was implemented by a subroutine that added positive or negative time increments to the inter-tone gaps according to a Gaussian approximation. The maximum size of an increment or decrement was 2.5 times the gap standard deviation; negative gaps were not allowed.

Subjects were trained on each of the types of temporal matching conditions until consistent performance was obtained. Daily testing sessions consisted of from 5 to 7 blocks of 100

trials. The experimental conditions were fixed within each block; only one temporal matching condition or mean gap duration was run in a given session. The gap standard deviation was changed randomly from block to block within each session.

3.2 Experiment 1

3.2.1 Procedure

The eight tone binary sequences were constructed by sampling from tone frequencies of 669 and 1621 Hz. The first sequence on a trial was created by sampling from a set of 54 different binary patterns. These 54 sequences were formed by starting with the set of all sequences of 8 tones which contain equal numbers of high and low tones, and then arbitrarily deleting all sequences which had regular appearing patterns, such as HLHLHLHL. On same trials, the second sequence generated was the same pattern as the first. On different trials, two adjacent (different frequency) tones in the sequence were transposed in position. Neither the first pair or the last pair of tones in the sequence were subject to this manipulation. Figure 1 illustrates some sample sequence pairs on typical trials of experiment 1. Subjects were run on at least 300 trials per condition.

3.2.2 Results

The average performance of the three subjects as a function of the gap standard deviation is shown in figure 2. The data from individual subjects had essentially the same form as that shown in the figure. The standard error of data points from individual subjects was typically 0.15 or less. Performance in the correlated condition was independent of the jitter magnitude. Performance in the uncorrelated condition, however, dropped well over one d' unit as the gap standard deviation increased.

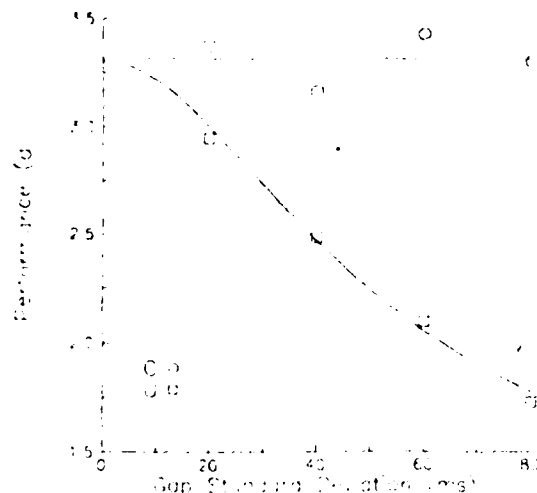


FIG. 2. Average performance (plotted points) as a function of the gap standard deviation and condition (C,U) in experiment 1 (8 tones, mean gap 100 ms). Solid lines are theoretical functions (see text).

3.3 Experiment 2

3.3.1 Procedure

Experiment 2 was similar to experiment 1, except in the manner of choosing the frequency patterns and in the employment of a new group of three subjects. The sequences were again binary tonal sequences of frequencies 669 and 1621 Hz, with an equal number of high (H) and low (L) tones. On each trial the first tonal sequence was selected randomly from a previously defined set of 16 different sequence patterns composed of 8 tones each. On same trials, the first sequence was presented twice during the trial. On different trials, a second sequence was sampled from the set. Tonal sequences composed of 12 or 16 tones were created by adding a four tone prefix or a four tone prefix plus a four tone suffix to the basic 8 tone set. These four tone additions were random sequences subject to the equal H and L constraint.

In experiment 2a, performance in the uncorrelated condition was evaluated at a mean gap duration of 50 ms and gap standard deviations of 10, 20, and 40 ms; performance in the correlated condition was assessed only at a gap standard deviation of 40 ms. In experiment 2b, performance was evaluated in the correlated and uncorrelated conditions at a fixed gap standard deviation of 40 ms, and at mean gap durations of 50, 100, 150, and 200 ms. In experiment 2c, performance was evaluated in the correlated and uncorrelated conditions at a mean gap duration of 50 ms and a gap standard deviation of 40 ms, for sequences composed of 8, 12, and 16 tones. Subjects ran 300 trials per condition of experiment 2a, and 200 trials in each condition of experiments 2b and 2c.

3.3.2 Results

Figure 3 illustrates the averaged results of experiment 2a. The pattern of results was consistent with that of experiment 1 shown in figure 2; the standard error of individual subject data points was less than 0.2. Figure 4 is a plot of the average performance of the three subjects in experiment 2b. Performance is plotted as a function of the mean gap duration, g , at a gap standard deviation of 40 ms. Performance of the individual subjects was similar to the averaged data shown. Performance in the correlated condition decreased with increases in mean gap duration. Performance in the uncorrelated condition was quite poor at short gap durations. As the mean gap duration increased, performance in the correlated and uncorrelated conditions converged. Figure 5 illustrates the results of experiment 2c; this shows the effect of the number of tones in each sequence on performance in the two temporal correlation conditions. Performance in both conditions dropped as n increased from 8 to 16; the largest drop occurred in the correlated condition.

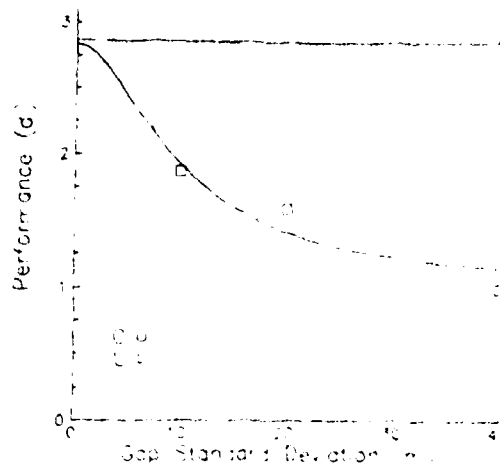


FIG. 3 Average performance (plotted points) as a function of the gap standard deviation and condition (C,U) in experiment 2a (8 tones, mean gap = 50 ms). Solid lines are theoretical functions (see text).

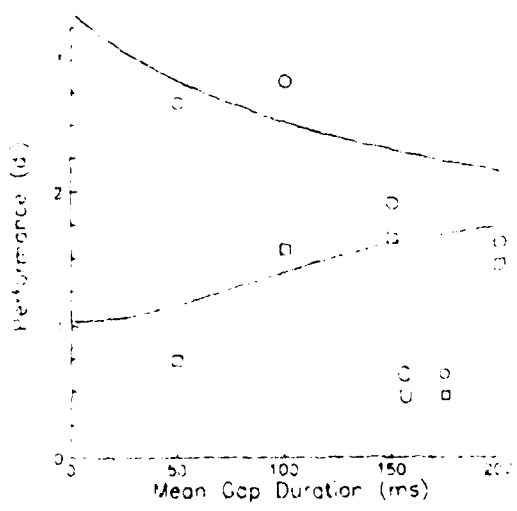


FIG. 4 Average performance (plotted points) as a function of the mean gap duration and condition (C,U) in experiment 2b (12 tones, gap standard deviation = 40 ms). Solid lines are theoretical functions (see text).

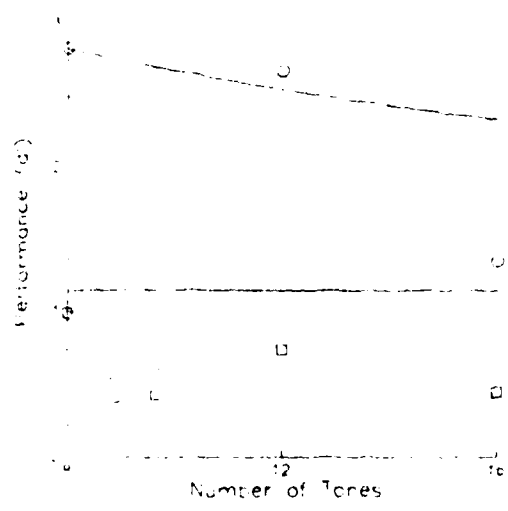


FIG. 5 Average performance (plotted points) as a function of the number of tones and condition (C,U) in experiment 2c (mean gap = 50 ms, gap standard deviation = 40 ms). Solid lines are theoretical functions (see text).

3.4 Experiment 3

3.4.1 Procedure

In experiment 3a an additional temporal condition, the synchronized condition, was added to the correlated and uncorrelated conditions of experiment 1 and 2. In order to make the context coding component of the task more demanding, the number of possible frequencies was increased from 2 to 4, and the number of different patterns (on same trials) from 16 to 2520. Figure 6 illustrates some tonal sequences possible on experimental trials from the three types of temporal variability conditions.

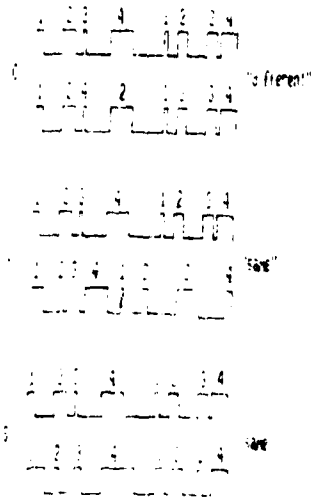


FIG. 6. Examples of the pairs of sequences present on trials of experiment 3 (see text).

In the correlated condition (C) shown in figure 6, the pattern of tone onsets, durations, and inter tone gaps is identical in the two sequences. Note that for this particular pair of sequences the frequency pattern is different in the two sequences, so that the correct response is "different". The second pair of sequences illustrates the uncorrelated condition (U). Here the tone onsets, durations, and gaps are uncorrelated between the two sequences. The frequency pattern is identical on the trial shown, hence the correct response is "same". The last pair of sequences illustrates the synchronized condition (S). In this condition the tone onsets are identical in the two sequences, but the durations and gaps are varied. In this particular pair of sequences the frequency pattern is identical, so the correct response is "same".

The tonal sequences were composed of eight tones; each tone

was selected from a set of four frequencies: 500, 909, 1667, and 2857 Hz. The frequency pattern of the first sequence of each trial was randomly determined but subject to the constraint that each frequency appeared twice. On different trials, three of the tone positions of the sequence were randomly selected to be changed from their original frequency to one of the three remaining frequencies. The first and last tone position were excluded from this selection.

Manipulation of the temporal characteristics of the two sequences was accomplished in a manner different from the previous experiments. The tone durations were randomly selected from a set of five durations (20, 30, 40, 50, and 60 ms) having a mean duration of 40 ms and a standard deviation of 12.25 ms. In Experiment 3a., same/different performance was evaluated in the same manner as in the previous experiments at a mean gap duration of 50 ms and at gap standard deviations of 5, 10, 20, 30, or 40 ms (omitting the 5 ms correlated condition). In experiment 3b., performance was assessed at a mean gap duration of 50 ms, at a gap standard deviation of 30 ms, and at inter-sequence-intervals (ISIs) of 500, 1500, and 3000 ms, using a rating response instead of the two-response same/different task.

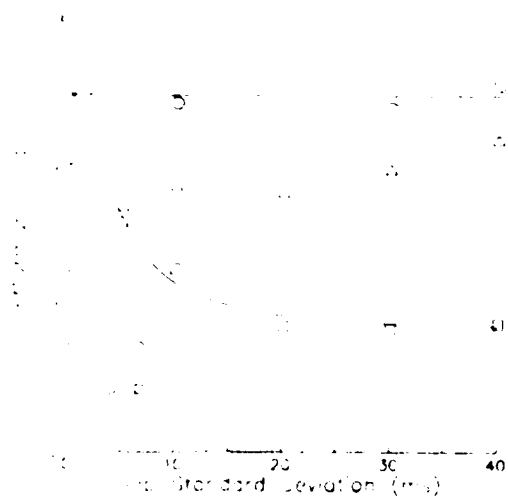


FIG 7. Average performance (plotted points) as a function of the gap standard deviation and condition (C,S,U) in experiment 3a (8 tones, mean gap = 50 ms, intersequence interval = 0.5 s). Solid lines are theoretical functions (see text).

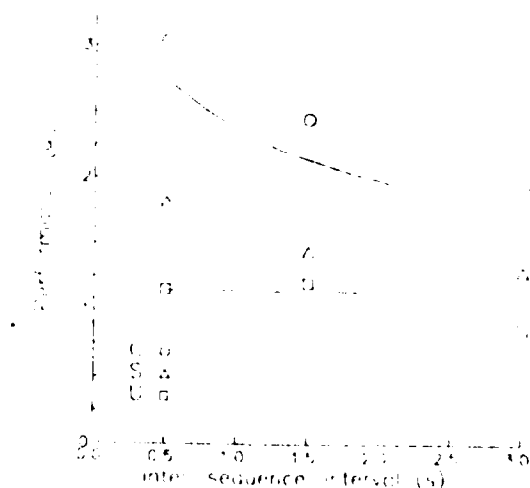


FIG 8. Average performance (plotted points) as a function of the inter-sequence interval (ISI) and condition (C,S,U) in experiment 3b (8 tones, mean gap = 50 ms, gap standard deviation = 30 ms). Solid lines are theoretical functions (see text).

3.4.2 Results

The average performance of three subjects in experiment 3a is illustrated in Figure 7, which shows the effect of gap variability on the discriminability of the stimulus sequences in the three conditions run, at a mean gap of 50 ms. Individual subject data was based on at least 200 trials per point. The standard error of individual data points was less than 0.2. Performance in the correlated condition was essentially independent of gap variability. Performance in the synchronized condition improved slightly at high gap variabilities. Performance in the uncorrelated condition dropped approximately 1 d' unit and then leveled off.

The results of Experiment 3b are shown in Figure 8, which is a plot of performance at a gap standard deviation of 30 ms, as a function of the inter-sequence-interval. Performance in the correlated and synchronized conditions decreased with increased ISI. Performance in the uncorrelated condition was independent of ISI.

3.5 Estimation of Model Parameters

In order to evaluate the relationship between the model predictions and the data, equations 3a and 3b were simplified as follows:

$$d'(\text{unc}) =$$

$$\{ A + [B(\text{ISI} + nd + (n-1)g) + C(\sigma_g/g)^2]^{-1} \}^{1/2} \quad (4a)$$

and

$$d'(\text{cor}) = \{ A + [B(\text{ISI} + nd + (n-1)g)]^{-1} \}^{1/2} \quad (4b)$$

where A, B, and C are constants which incorporate the model parameters:

$$A = \frac{c_1^2}{c_2^2}, \quad B = \frac{c_2}{c_1}, \quad \text{and} \quad C = \frac{c_3^2}{c_1^2}$$

Equation 4a was fit separately to the data of experiments 1, 2a, and 3a, using the Gauss-Newton method. We assumed that at a gap standard deviation of zero, performance in the uncorrelated condition was equal to average performance in the correlated condition. Thus, the three parameters were fit to 5, 4, and 6 data points in experiments 1, 2a and 3a, respectively. Table II summarizes the parameter values computed by this procedure.

The theoretical curves generated with these parameters are shown in figures 2, 3, and 7. The parameter values derived from the data of experiment 2a were then used to generate the

theoretical curves shown in figures 4 and 5, and in similar fashion the parameter values derived from experiment 3a were used to generate the theoretical curves shown in figure 8. The theoretical curve for the uncorrelated condition in figure 5 actually decreases slightly as a function of n, as does the theoretical curve for the uncorrelated condition shown in figure 8.

=====
 Table II. Parameter Values for Simplified Model (Equations 4a,b)

Experiment	Parameter		
	A	B	C
1	1.155	0.000071	0.618
2a	1.069	0.000131	5.930
3a	0.724	0.000155	34.05

=====
 Table III. Model Parameter Values (assume $c_2 = 0.001$)

Experiment	Parameter		
	c_1^2	c_1	c_3
1	12.19	3.75	8.70
2a	7.14	2.76	45.27
3a	8.91	2.54	219.69

Because the parameter c_2 describes a general property of the trace decay process, we expected it to be constant over the different experiments. Assuming a particular value for c_2 enables calculation of the remaining parameter values for purposes of comparison. Table III summarizes the model parameters which result from the assumption that c_2 is equal to 0.001 in experiments 1, 2a, and 3a. It is apparent that the computed value of the context coding variance, c_1^2 , was highest in experiment 1, although the values do not differ greatly over the three experiments. Because the size and nature of the pattern set varied across the three experiments (see Table I), c_1 should not have been constant across experiments 1, 2a, and 3a.

However, the largest value of c_c^L was obtained in experiment 1, contrary to our expectation that the size of the pattern set would be the main determinant of context coding variance. It may be easier for an observer to encode patterns from the larger, four-frequency, pattern set of experiment 3a than from the smaller, two-frequency, set of experiment 1. It is also evident from Table III that the values of parameter c_1 were approximately equal in experiments 1, 2a, and 3a.

Finally, it is clear from Table III that the value of c_3 varied widely across the three experiments. We used the expression $c_3[\sigma_g/g]^2$ as an estimate of the effect of the envelope correlation on the trace variance, and ignored the influence of other variables such as tone duration and number of tones. The large variation in the obtained values of c_3 indicates that the $c_3[\sigma_g/g]^2$ expression provides an inaccurate estimate of the effect of envelope variability.

In order to improve our estimate of this variance component, we examined the magnitude of the correlation between the temporal envelopes of sequences in our experimental conditions, using computations made on data files of the actual stimulus sequences. This correlation was equal to unity in the correlated conditions of all experiments. In the synchronized condition of experiment 3a, the envelope correlation was essentially constant at a value of 0.7, for all values of σ_g . In the uncorrelated conditions of experiments 1a and 2a, the envelope correlation was initially equal to unity (at $\sigma_g = 0$) and then dropped rapidly to a value of 0.1 at a gap variability of 20 ms. In the uncorrelated condition of experiment 3a, the presence of additional variability in the tone durations caused the initial correlation (at $\sigma_g = 0$) to be approximately 0.16. The correlation then fell to less than 0.1 at a gap variability of 5 ms. We attempted to fit equation 3a to the data of experiments 1, 2a, and 3a, using an approximation to the envelope correlation function in place of $[\sigma_g/g]^2$. This procedure resulted in relatively poor fits to the data because of the flatness of the function relating envelope correlation to gap variability.

3.6 DISCUSSION

The experimental results provided qualitative support for the main assumptions of the discrimination model. In particular, the results are consistent with the idea that trace mode processing is dominant in the correlated conditions. Performance was independent of gap variability in the correlated condition, but decreased rapidly with gap variability in the uncorrelated condition (Figures 2, 3, and 7). It follows from an assumption of trace mode processing that performance should decrease with increases in the time interval that the trace must be held in storage, such as the time between the onset of the first sequence and the onset of the second sequence. That prediction was confirmed in experiment 3b (figure 8) which showed that performance in the correlated condition decreased as a function

of the inter-sequence-interval. Performance in the uncorrelated condition dropped very little as the inter-sequence-interval increased. The prediction was also confirmed in experiment 2b by the drop in correlated condition performance when the average duration of the gap interval was increased. Increasing the gap duration from 50 ms to 200 ms caused an increase in the required storage period of 1650 ms (11 x 150) and resulted in a drop in performance of about 1 d' unit. Performance in the uncorrelated condition increased during this manipulation.

Performance was consistently lower in the uncorrelated conditions, and decreased with the addition of temporal variability to the sequence envelopes, as shown in figures 2, 3, 7 and 8. We attribute these decreases to increased noise in the trace process produced by the presence of variability in the sequence traces in the uncorrelated condition. As the magnitude of the sequence variability increases in the uncorrelated condition, the variance of the trace process increases rapidly, and eventually context mode processing dominates. At that point, task variables which affect the context variance, such as the number of possible different patterns, influence performance.

Performance in both the correlated and uncorrelated conditions of experiment 2c decreased with increases in the number of tonal elements as shown in figure 5. The largest drop was observed in the correlated condition. The small drop in performance in the uncorrelated condition may not be surprising since one would expect that context coding variance would increase relatively slowly, perhaps as a logarithmic function of the number of tonal patterns. Part of the drop in performance for the correlated condition is attributable to the increase in the total storage period that is produced by increasing n (adding one tone increases the total storage time by 80.5 ms, the duration of one tone plus one gap). The very large performance decrease in the 16 tone case may be due to other factors such as an upper limit on the length or duration of the trace.

An interesting result concerns performance in the synchronized condition of experiment 3. Performance in this condition resembles that of the correlated condition in several respects: it does not decrease with increasing temporal variability (figure 7) and it decreases with increases in the magnitude of the inter-sequence-interval (figure 8). The former result is consistent with our calculation of a constant correlation of approximately 0.7 between the temporal envelopes of the synchronized sequences of experiment 3. Lower performance in the synchronized condition than in the correlated condition is consistent with the idea that there is additional trace variance due to the temporal variability of the sequences.

4. SEQUENCE COMPARISON ALGORITHMS

The effects of gap variance and correlation described in the preceding sections of this report are entirely consistent with the extension of the two-process model. However, we were

interested in pursuing several additional questions about the processing mechanisms involved in performance of the sequence comparison task. The first question concerned the specific mechanism underlying the comparison operation. As a preliminary approach to this problem we evaluated some specific sequence comparison algorithms in the context of this experimental task. An important aspect of this approach is that it is not necessary to make assumptions about observer processing strategy or memory mode for particular conditions; a comparison algorithm is postulated and then applied without modification to all the experimental conditions. If successful, the postulated comparator would exhibit the observed sensitivity to the temporal coherence of the stimulus sequences, as well as to other aspects of the sequence context.

Consider the general problem of sequence comparison: A sequence, a , is a series of elements, a_1, a_2, \dots, a_n where the elements a_i are drawn from some defined alphabet. One wishes to compute a measure of the difference or similarity between sequence a and b . Traditional methods for comparing sequences include Euclidean distance metrics, the Hamming distance measure, and others.

These traditional measures sum the results of all comparisons between element pairs, a_i and b_i . In the case of tone burst sequences, this measure would be a function of the frequency difference between the tones. For the Euclidean distance measure the difference, d^{ab} , between sequence a and b is given by:

$$d^{ab} = \left[\sum_{i=1}^n (a_i - b_i)^2 \right]^{1/2}$$

For the Hamming distance measure,

$$d^{ab} = \sum_{i=1}^n d_i \quad \text{where } d_i = 0 \text{ if } a_i = b_i \\ d_i = 1 \text{ if } a_i \neq b_i$$

Both these measures sum the results of comparisons made on corresponding elements in each sequence. When the sequences are composed of different numbers of elements, or when there is no correspondence between the elements of each sequence, more sophisticated measures are needed. The requirement for comparing sequences in which there are non-corresponding elements is a problem common to many sequence comparison situations. For example, a deleted or added element may upset the normal correspondence between a_i and b_i at position i in the sequence. Distortions in the time periods between the sequence elements or in the duration of the elements also may occur. These are the sort of sequence transformations to be expected in natural stimuli such as speech. We would expect that the human pattern processing mechanism would be able to handle these types of distortions and transformations without undue difficulty. Many

applications of sequence comparison algorithms (including speech) are described in a book edited by Sankoff and Kruskal (1983).

4.1 Weighted Average Algorithm

An example of a more complex distance measure is one derived from the weighted average of comparisons made between all pairs of elements, a_i, b_j , in each sequence. Element pairs that are offset by a large difference in their respective position receive a smaller weight in the comparison measure. The weighted average measure which we employed is based on one used by Bradley and Bradley (1983) in a study of sequence comparison techniques applied to the study of bird songs. Bradley and Bradley compared two different types of sequence comparison algorithms: a weighted average algorithm and an optimal matching algorithm. We used modified versions of those algorithms as candidate comparator mechanisms for our tonal sequence task with human subjects.

The weighted average algorithm employed by Bradley and Bradley uses the average of all inter-element distances for all pairs, i, j , of elements from the two sequences. Measures on the distance between each pair of elements are weighted by the function, W_{ij} , which is a linear function of the distance between positions i and j . Their algorithm is defined as follows:

$$M = \frac{1}{2} \left[\sum_{i=1}^m \sum_{j=1}^n W_{ij} D_{ij} + \sum_{k=1}^m \sum_{l=1}^n W_{kl} D_{kl} \right]$$

where:

M is the distance from a to b

D_{ij} is the distance function on element i, j

m, n are the number of elements in a and b

W_{ij} is the weighting for element pair at position i, j

and
$$W_{ij} = k \left(1 - \frac{i-1}{m-1} - \frac{j-1}{n-1} \right)$$

with k chosen so
$$\sum_{i=1}^m \sum_{j=1}^n W_{ij} = 1.$$

The weighted average algorithm we employed was the same as the Bradley and Bradley procedure except that we allowed the computation of negative distances between sequences, and we employed a modified weighting function. If the span between element position i and j exceeded a constant (about 20% of the sequence duration), the weight was set equal to zero. If the span was less than or equal to that constant, a unit weight was applied. The distance between elements, D_{ij} , was defined as a function of the tone frequency. The lowest tone frequency was assigned the value 1, the next 2, etc. D_{ij} was then defined as two times the absolute difference between those values. If either, but not both, of the elements was a silent period, D_{ij} was set equal to unity.

4.2 Optimal Matching Algorithm

The second algorithm employed by Bradley and Bradley was an optimal matching procedure based on the concept of the Levenshtein distance metric (Kruskal, 1983). These metrics allow comparisons of sequences which have suffered insertion/deletion or time compression/expansion transformations. These metrics evaluate the summed total cost of the basic operations needed to transform one sequence into another. For example, the substitution of very dissimilar tones would add more cost than would the substitution of similar tones. The minimum possible cost for the entire set of operations then defines the distance between the sequences. Following Bradley and Bradley (1983), the algorithm for the specific optimal matching metric we used is recursively defined as follows:

$$S_{mn}^{ab} = \frac{ab}{\max(m, n)}$$

where

$$S_{ij}^{ab} = \min \left\{ S_{i-1, j-1}^{ab} + D_{ij}^{ab}, S_{i-1, j}^{ab} + C_i^a, S_{i, j-1}^{ab} + C_j^b \right\}$$

$$S_{00}^{ab} = 0, S_{i0}^{ab} = \sum_{k=1}^i C_k^a, S_{0j}^{ab} = \sum_{k=1}^j C_k^b$$

S_{ij}^{ab} is the distance from a to b

D_{ij}^{ab} is the element substitution cost

C_k^a is the insertion and deletion cost for element k of sequence a

and m, n are the lengths of sequence a and b .

A constant value of 1.9 was used for the insertion/deletion cost of a silent element, and a constant value of 3.6 was used for the cost of inserting or deleting any tone element. The cost of substituting an element of tone frequency i for an element of tone frequency j was set equal to 2.3 times the absolute difference between the tone numbers (lowest frequency equals tone number 1, next frequency equals 2, etc., as described for the weighted average algorithm). (A Pascal routine is available from the author which implements the optimal matching procedure using a dynamic-programming technique, rather than explicit recursion.)

4.3 Analysis of Algorithm and Human Performance

Each algorithm was tested on transformed versions of the stimulus sequences used in the human experiments. The computer representation of each stimulus sequence was sliced into a sequence of segments 10 ms in duration. These segments then became the elements of the sequences subjected to the comparison algorithms. Both of the algorithms incorporate parameters which will affect the resulting discrimination performance. No attempt was made to evaluate a large volume of the parameter performance space. Values of the parameters were tried until a reasonable range of performance was obtained for the stimulus sequences employed in the experiment. The goal was to determine whether some of the general features of the observed human behavior would be matched by the algorithm, rather than to compute a fit of the algorithm to the human data.

The performance of the two algorithms tested is shown as the solid and dashed lines in Figure 9; the plotted points were obtained from human subjects performing a rating task version of the conditions described in section 1.1. The standard deviation of the algorithm's responses was different on same and different trials, preventing use of a direct d' measure. The performance of the algorithms was evaluated by determining the area under the Receiver Operating Characteristic (ROC) curves (based on the algorithm's trial by trial responses) and then converting to equivalent yes/no d' values. The performance of both algorithms was perfect in the correlated condition, and is not shown. In the uncorrelated condition, the performance of both algorithms decreased as a function of gap variability.

Figure 10 shows representative ROC curves (plotted on normal probability axes) obtained for both algorithms and for two of the subjects. The slope of the normal probability ROC functions depended somewhat on the algorithm employed; slopes generally were less than 1.0 for the weighted average algorithm and exceeded 1.0 for the warp algorithm. Slopes for the human ROC curves were consistently near 1.0.

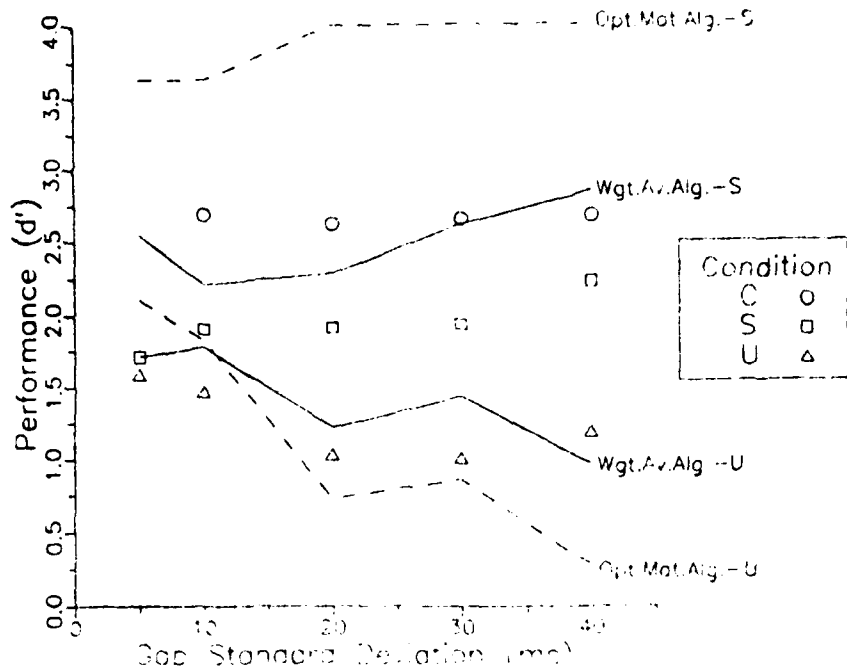


Figure 9. Average performance as a function of the gap standard deviation and condition (C,S,U) for the human subjects (plotted symbols) and for the two comparison algorithms (solid and dashed lines).

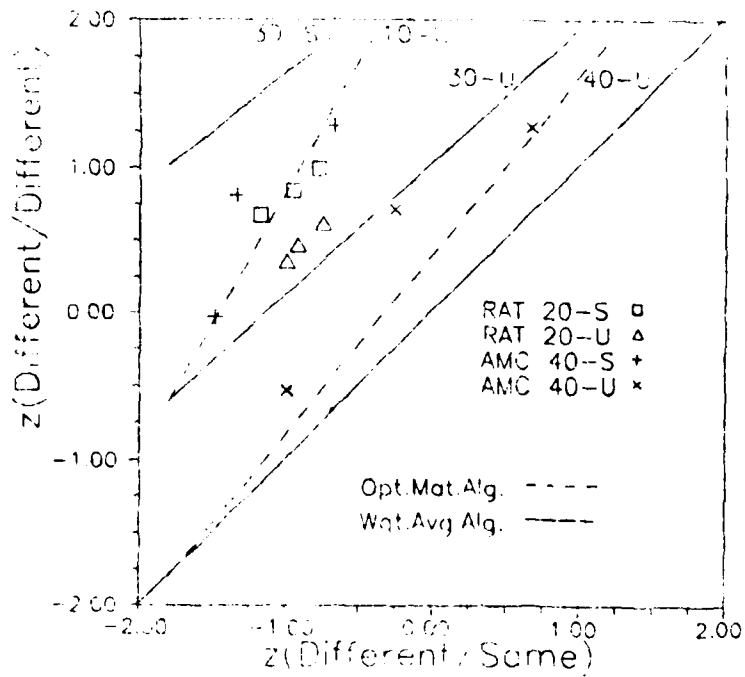


Figure 10. Receiver Operating Characteristic (ROC) curves for human observers RAT and AMC (plotted points) and best linear fits to the algorithms (solid and dashed lines).

The correlation between human and algorithm performance was computed on the trial by trial responses of the human subjects (ratings) and the algorithms (computed responses). Because the magnitude of such correlations depends on the level of performance, the most revealing correlations are on the conditions having the poorest performance. At the uncorrelated 30 ms or 40 ms conditions, the correlations obtained between individual subjects and either of the algorithms was generally less than 0.4. Neither algorithm correlated consistently higher than the other. Correlations between the algorithms or between pairs of subjects also were generally less than 0.4. Apparently, the subjects and the algorithms (for the particular parameter values chosen) order the similarity of sequence pairs in idiosyncratic ways.

Both of the algorithms tested produced performance which was similar to the profile of human performance over the different experimental conditions studied. However, note that these comparison algorithms are not models of human tone pattern discrimination. To construct such a model, we would need to postulate internal noise in order to obtain less than perfect model performance in the correlated condition. A reasonable assumption would be that this internal noise would take the form of uncorrelated temporal jitter added to the tone onsets. To estimate the magnitude of this internal jitter, we could observe the performance of the algorithm in low levels of uncorrelated jitter. When the performance of the algorithm in the uncorrelated jitter matched that of the human in the correlated jitter, an estimate of the internal jitter could be obtained.

This application of sequence comparison procedures was not intended to serve as a test of whether these specific algorithms will be good models of tonal sequence discrimination. Rather, our analysis demonstrates that such algorithms may be a useful starting point in modelling how humans process serially defined stimuli. A number of other computational procedures and variations of these algorithms could have been tested. For example, optimal matching procedures can incorporate the costs of changes in the durations of the tones or in the gaps between the tones. Algorithms of this type have been developed for the machine processing of speech. Such algorithms could be evaluated in tone pattern discrimination or in other tasks involving stimulus sequences. Other investigators have reported the effects of specific types of sequence context manipulations on the salience and discriminability of subsequences (see Leek, 1987). It would be interesting to see whether the performance of specific comparison algorithms would resemble the human observer's sensitivity (or lack of sensitivity) to such contextual manipulations and to the kinds of transformations discussed in the next section.

our current research plan is to evaluate different sequence comparison algorithms directly, by use of a rhythm discrimination paradigm. This procedure requires the observer to report whether or not two sequences of tones have the same pattern of durations

and inter-tone gaps, e.g. the same rhythm of tones. On "different" trials, the second sequence on each trial has a temporal jitter introduced into its inter-tone gaps. Performance will be evaluated as a function of the magnitude of this added inter-tone gap jitter; we expect that performance will increase as a function of this parameter. We will also modify the second sequence of each trial by uniformly compressing or expanding all sequence gaps and durations. This manipulation maintains the rhythm pattern but disturbs the temporal correlation between the envelopes of the two sequences. Our interest is in defining how performance, at a given jitter magnitude, depends on the amount of compression or expansion. Thus, we hope to use this experimental paradigm to discriminate between sequence comparison mechanisms which require correlation-like processes and those, such as the optimal matching algorithm, which do not.

5. LINEAR TRANSFORMATION OF SEQUENCE PARAMETERS: PROPERTIES OF THE SENSORY TRACE

An important question in sequence comparison is the sensitivity of the trace and context mechanisms to particular types of transformations in the sequences to be compared. The context mechanism is assumed to be insensitive to the length of the ISI and to the effects of certain events occurring during the ISI period. The trace mechanism, on the other hand, is assumed to be highly sensitive to the duration of the ISI and to certain other potentially interfering events, such as the occurrence of similar stimuli during the ISI period. The specific effects of other types of transformations of the sequences are of considerable theoretical interest. Linearly scaling the duration of the initial or final sequence, for example, should have a different effect on performance while in the two modes. Such a scaling may or may not impose additional processing demands on either mode.

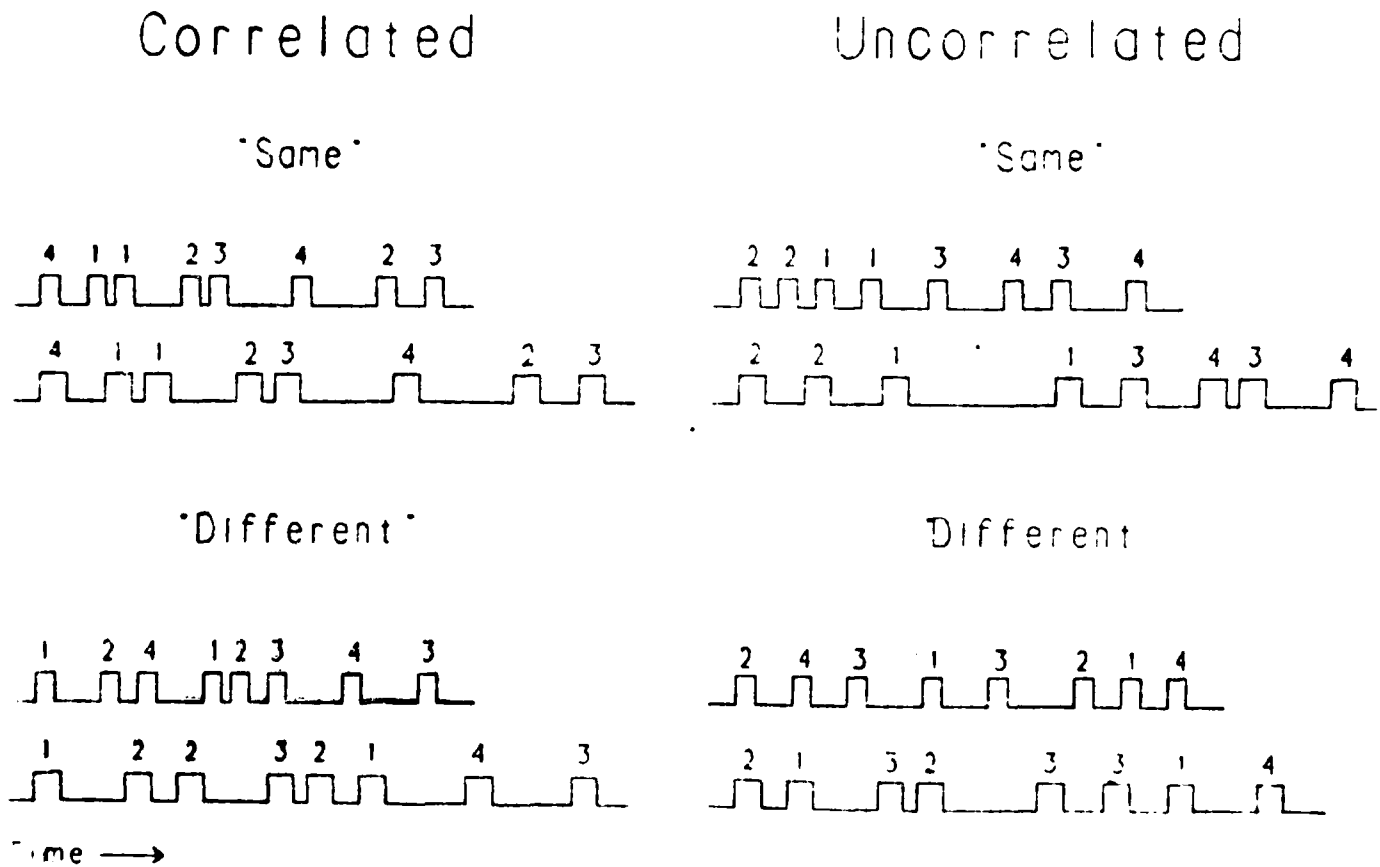
For example, one interesting question concerns whether duration compression or expansion will produce decrements in the trace mode which are a function of the ISI. That is, will duration scaling interact with the increase in trace variance over time? The answer depends on one's concept of the trace process. If the trace is a true raw representation of the input, then compressing or expanding the sequence should degrade performance in a fashion similar to the addition of temporal jitter to the sequence intervals, and an interaction will result. If, however, the trace is (effectively) only a partially encoded representation of the tone bursts and the temporal (rhythmic, etc.) relationships among the tone bursts, then such transformations may have a quite different effect. In that case, the effect of such a transformation would not produce interactions with ISI, but rather would resemble the effects of adding temporal jitter in the context mode: e.g. produce a fixed decrement, depending on the magnitude of the transformation. These predictions were subject to experimental test.

5.1 Duration Scaling Experiment

In experiment 4, we modified the manner in which the durations of the tones and gaps of the second sequence in a trial were generated. In the 'normal' duration transformation condition, the second sequence in a trial was generated without modification. Trials in this condition were of the same type as those shown earlier. In the 'expanded' duration transformation condition, the durations of all tones and gaps in the second sequence of the trial were multiplied by a factor of 1.4, producing a sequence 40% longer than it would have been without the transformation.

Duration transformation was manipulated factorially with temporal correlation, producing the 'correlated-expanded' condition, in which the two sequences in a trial had the same temporal pattern prior to expansion of the second sequence, and the 'uncorrelated-expanded' condition, in which the temporal patterns of the two sequences in a trial were different prior to expansion of the second sequence. Typical trials from each of these conditions are shown in figure 11.

Figure 11. Examples of pairs of tonal sequences in the expanded conditions of experiment 4.



We were interested in whether operation in the trace mode requires that the two sequences have correlated temporal envelopes. Another way to put the question is to ask whether subjects would treat expansion of the second sequence of a trial as a difference in the sequence temporal patterns- similar to that found in the uncorrelated condition- or if, instead, they would be able to scale the duration of the sensory-trace of one sequence to match that of the other. If such a scaling process were possible, then one would expect to see the same general pattern of results in the correlated-expanded condition as in the correlated-normal condition. If such a scaling process were not possible, then the pattern of performance in the correlated-expanded condition should closely resemble that found in the uncorrelated-normal condition.

The results can be seen in figure 12 and 13. Figure 12 shows performance averaged across subjects for the correlated conditions. Performance in both the normal and expanded cases was independent of gap standard deviation, but decreased with increasing ISI. Figure 13 shows performance in the uncorrelated conditions. In this condition, performance was independent of ISI in the conditions involving large differences in sequence temporal patterns within a trial (i.e. those in which the gap standard deviation was 40), but decreased with ISI in those conditions which did not involve large temporal pattern differences (i.e. those in which the gap standard deviation was 10). Note that performance in both the uncorrelated-normal and uncorrelated-expanded conditions decreased with increasing gap standard deviation. The results from the 'normal' duration transformation conditions replicate the results of experiments 1 through 3, while the results from the 'expanded' conditions support the duration-scaling hypothesis previously proposed. While performance in the expanded conditions is poorer than that in the normal conditions, perhaps due to trace-decay during duration scaling, the general pattern of results is the same, leading one to believe that S's were able to compensate for sequence expansion in these conditions, perhaps by "scaling" the sensory traces.

Another interesting transformation to be evaluated is that of frequency. One particular question relates to frequency transformations of an octave or less (either up or down). Because of the octave generalization phenomenon, we would expect that the trace (or "encoded" trace, if our duration transformation results are confirmed) would be insensitive to octave frequency transformations, but possibly not to other transformations. Depending on the outcome of these experiments, we will modify the concept of the trace mechanism in the two-process discrimination model to incorporate the observed results. Preliminary data from a frequency transformation experiment indicates that observers probably cannot scale frequency in the same way as duration; performance as a function of the intersequence interval flattens out as the magnitude of the frequency transformation is increased.

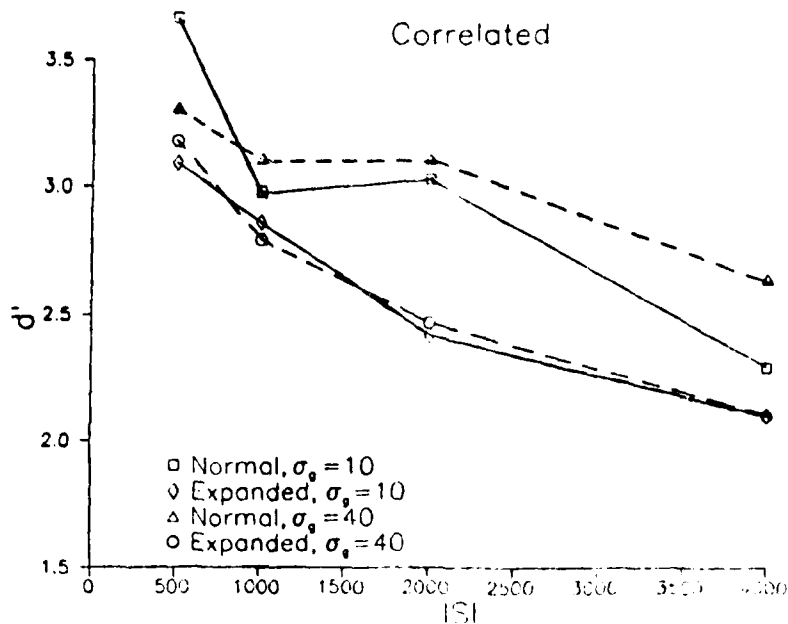


Figure 12. Average performance as a function of inter-sequence interval, gap standard deviation, and temporal transformation, for the correlated conditions of experiment 4.

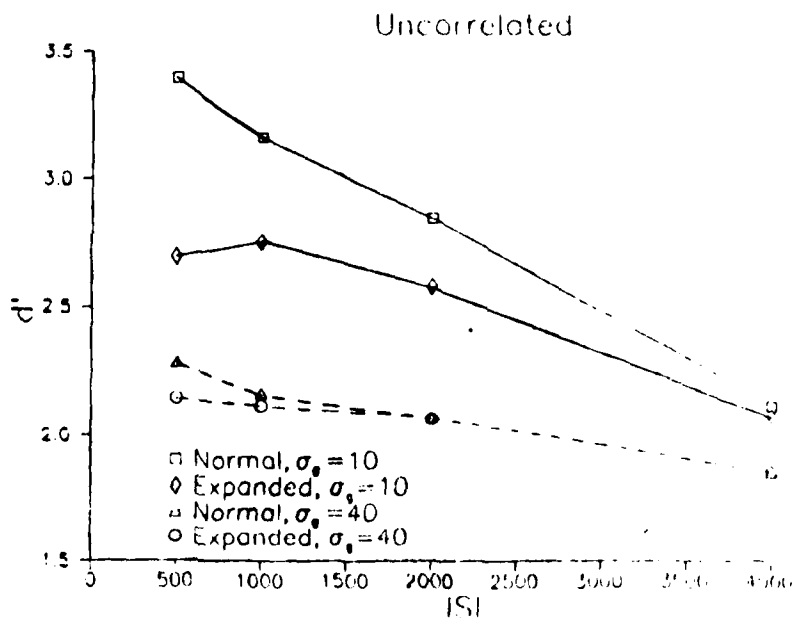


Figure 13. Average performance as a function of inter-sequence interval, gap standard deviation, and temporal transformation, for the uncorrelated conditions of experiment 4.

6. PROCESSING INFORMATION FROM TONAL SEQUENCES AND VISUAL DISPLAYS

Our interest in the perception of auditory tonal sequences and in signal detection models of human detection and discrimination led us to study some problems concerning the display of information via the auditory or visual modalities. This section discusses a signal detection theory based-method for evaluating visual and auditory displays. The method is used to study to how a subject acquires information from a multiple element display. The method allows assessment of the processing characteristics associated with individual elements of the display. Using the method, it is possible to compare the effectiveness of different display codes and of different arrangements of display elements. This summary concentrates on the mechanics of the method and summarizes data from some experiments with visual and auditory displays. This work was performed in collaboration with Donald E. Robinson and Bruce G. Berg of Indiana University.

The basic experimental paradigm is a signal detection task in which the subject must decide whether event A or event B led to the displayed data. Either event A and B occurs randomly on a trial. The display consists of n display elements; each element conveys information about the trial event which led to the displayed data. The displayed elements are generated by a statistical process whose parameters depend on which event (A or B) was selected on that trial. For example, consider a display consisting of the four, two-digit, numerical elements:

[1.3 3.1 0.2 2.4]

These two-digit elements are generated from one of two normal distributions: distribution "A", having a mean of 1.5 and a standard deviation of 1.2, and distribution "B", having a mean of 2.5 and a standard deviation of 1.2. On an "A" trial, each of the four display elements is independently generated from the A distribution. On a "B" trial, each is drawn from the B distribution. Given the subject's hit and false alarm rates, we can then compute a measure of his/her performance with the signal detection measure, d' . For a display containing only a single element, d' for an optimal, noiseless, observer would be equal to the difference between the means of the A and B distributions, divided by the common standard deviation, that is, $d' = 0.83$.

It should be evident that the displayed information (the normal deviates) could be presented to the subject in different ways; in analog form as in a straight line or square array of line gauges, or in an auditory display in which the information is coded by the frequency of tones in a sequence of tones. How would an ideal observer perform with a display composed of n independent display elements? It can be shown the performance of an ideal observer increases with the square root of the number of elements in the display. Hence, in the four element case described above, ideal performance is equal to a d' of 0.83×2

or 1.67.

Naturally, for certain display formats and display durations, we would expect that processing limitations would prevent the human subject's d' from increasing at the same rate as the ideal observer. And, in some cases, human performance may level off at some asymptotic level; passed some point, increasing the number of elements would not result in increased performance.

Figure 14 shows data obtained in auditory experiments using sequentially presented tones. It is clear that the subject fails to accumulate information at the \sqrt{n} rate. The auditory data was well fit by a model which incorporated two sources of internal noise: one associated with the observation of each element and one with the formation of the decision statistic.

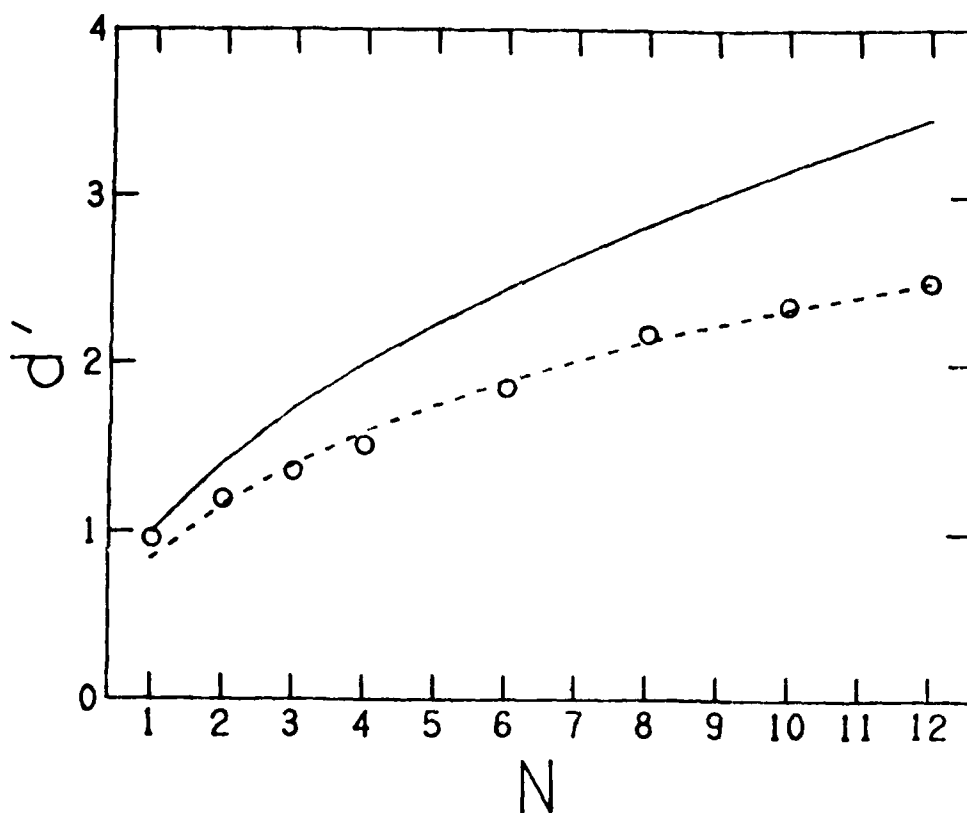


Figure 14. Average performance (d') of four observers to a frequency-coded sequential auditory stimulus as a function of the number of tones in the sequence. Solid line: square-root-of- n prediction. Dashed line: prediction of a partitioned variance model.

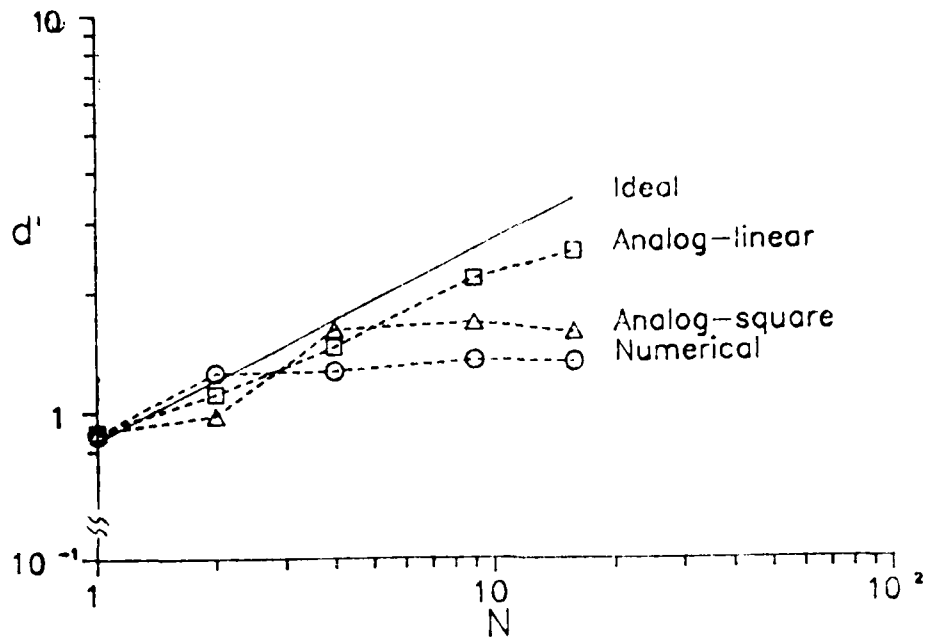


Figure 15. Average performance (d') of three observers to a visual display as a function of the number of elements. Solid line: square-root-of- n prediction.

Figure 15 shows performance averaged across three subjects in a visual display experiment using digital and analog (line gauge) display elements arranged in both straight line and square arrangements. The plotted data were obtained at a display duration of 255 msec; the solid line is ideal (\sqrt{n}) behavior. At low values of n , the human data are very close to the ideal. At higher values of n , performance in the square array analog and numerical conditions reaches asymptotic levels. Data obtained with the same displays at a duration of 1000 msec matched the ideal functions very closely at all values of n .

The asymptotic nature of the visual display performance, at high values of n , is consistent with the idea that the subject cannot obtain full information from all the display positions, at least in a brief stimulus exposure. An important question, then,

is whether the subject obtains full information from a small subset of display elements or partial information from every display element. If the subject uses partial information from all the elements to make a decision, one might ask whether information from some display elements weighs more heavily than from other elements in determining the decision. An analytical procedure, developed by D. E. Robinson and Bruce Berg of Indiana University, allows some resolution of these important questions about subject performance.

The amount of information that the subject obtains from different elements of a display will be reflected in the variability of the subject's response that may be attributed to each display element. Recall that on each trial, the subject responds "signal A" or "signal B". Thus, one can compute the probability that the subject responds " β " as a function of the particular values taken by the element at position i in the display array. Such a plot should generate the cumulative probability distribution for the " β " response given the value of element i , as shown in figure 16. Suppose that this function were flat- one would conclude that the subject's response was independent of the value of display element i ; similarly, if the function resembled a step function, we could perfectly predict the response by knowing the element value. The point is that, under certain assumptions, the variance of this distribution, e.g. the slope of this cumulative probability function, provides an estimate of the contribution of element i to the subject's decision. The lower the variability, the higher the contribution of that element to the subject's decision.

The next several figures illustrate the results of an analysis of this sort applied to the visual and auditory displays described earlier. Figure 17, from the experiment with a sequential auditory display, shows the slope of the element probability distribution as a function of temporal position. The parameter on the figure is the total sequence length, n . If each element in the sequence contributed equally to the final decision, the lines in Figure 17 would be horizontal. It is clear that the last tone in a sequence contributes more to the final decision than do tones in the middle, which contribute less than those near the beginning of the sequence. The result is similar to that observed in experiments on the recall of lists of verbal materials.

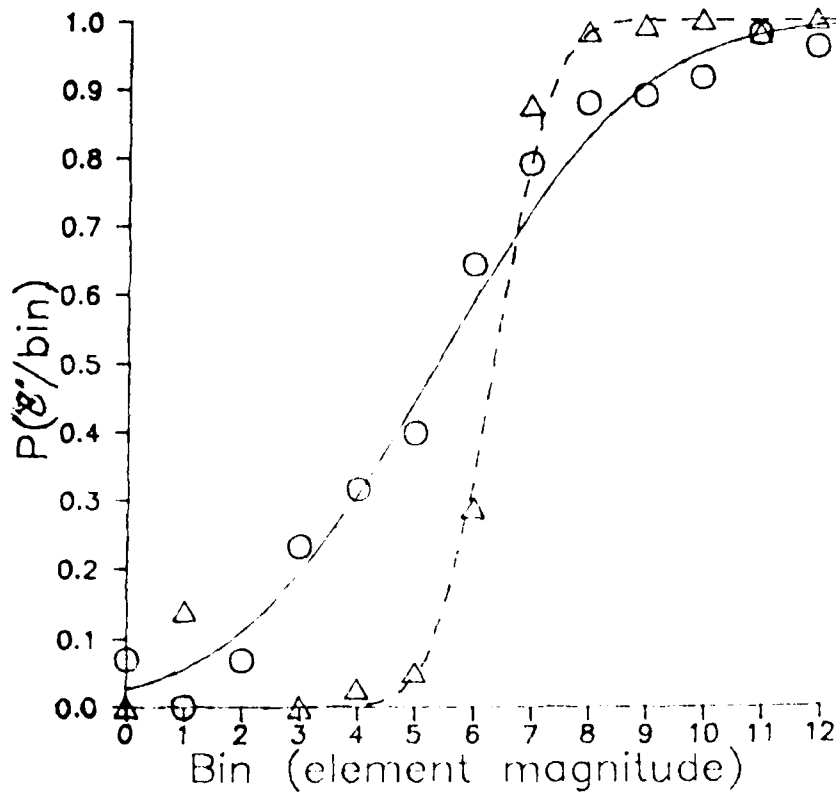


Figure 16. Probability of observer's "yes" response given the particular magnitudes (over trials) of the display element in a visual detection task (two different elements).

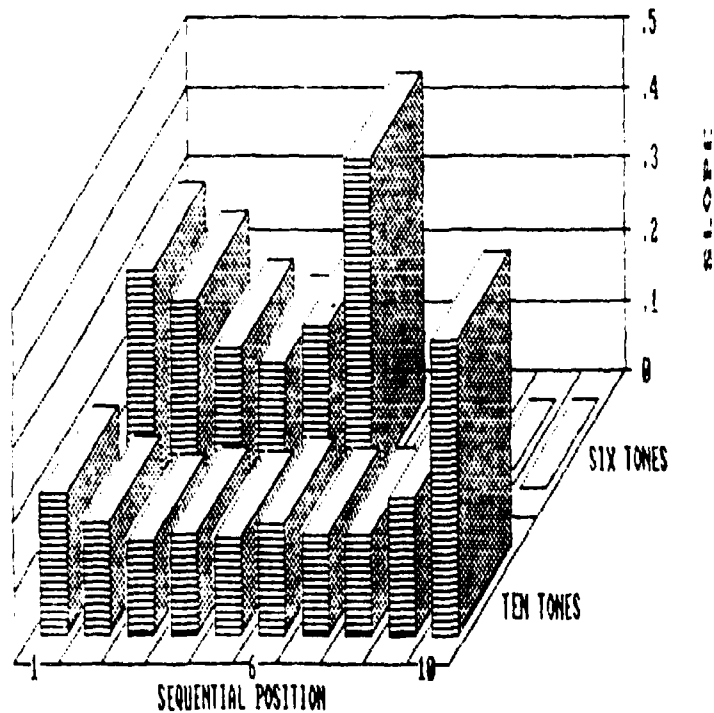


Figure 17. Slope of the Probability vs. Element Magnitude curves as a function of element sequential position (auditory display experiment). Curves are given for sequences composed of 6 tones and 10 tones.

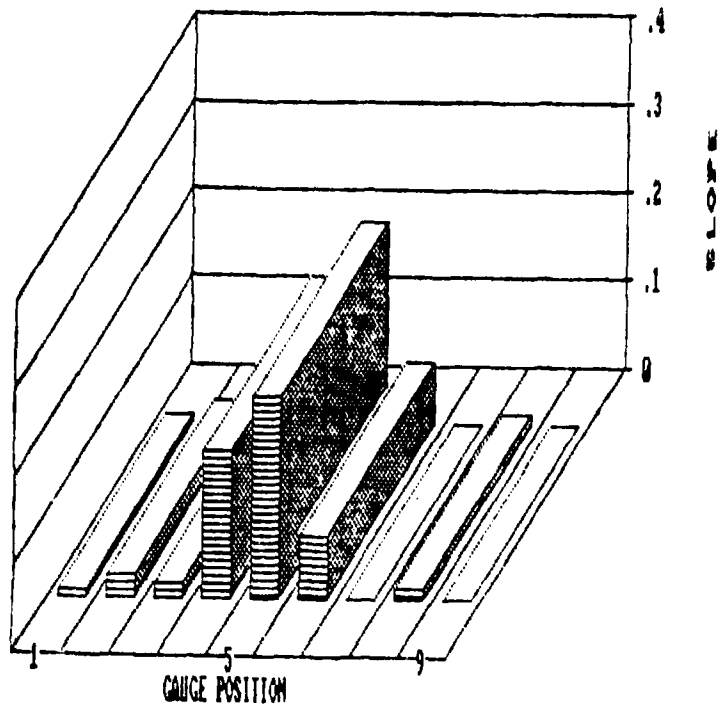


Figure 18. Slope of the Probability vs. Element Magnitude curves as a function of element spatial position in the visual display-straight line array of digital elements.

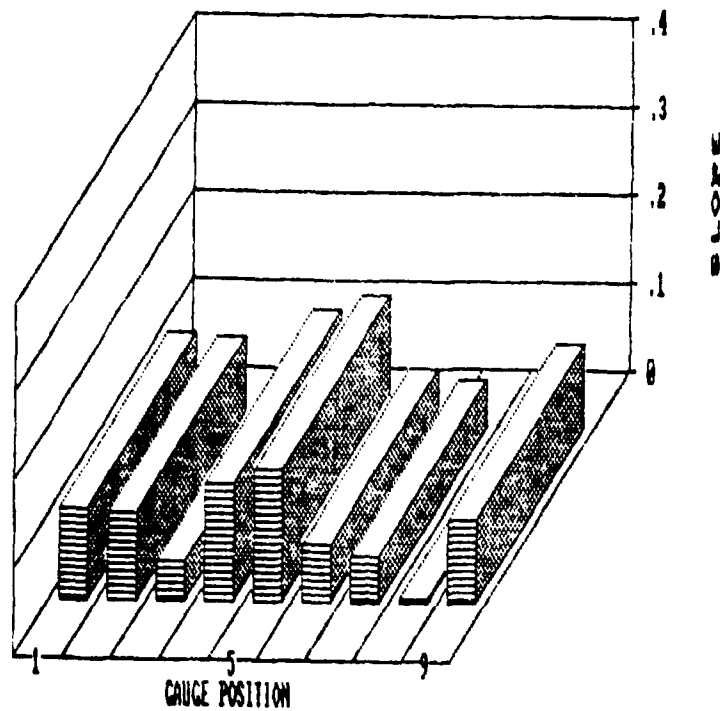


Figure 19. Slope of the Probability vs. Element Magnitude curves as a function of element spatial position in the visual display-straight line array of analog elements.

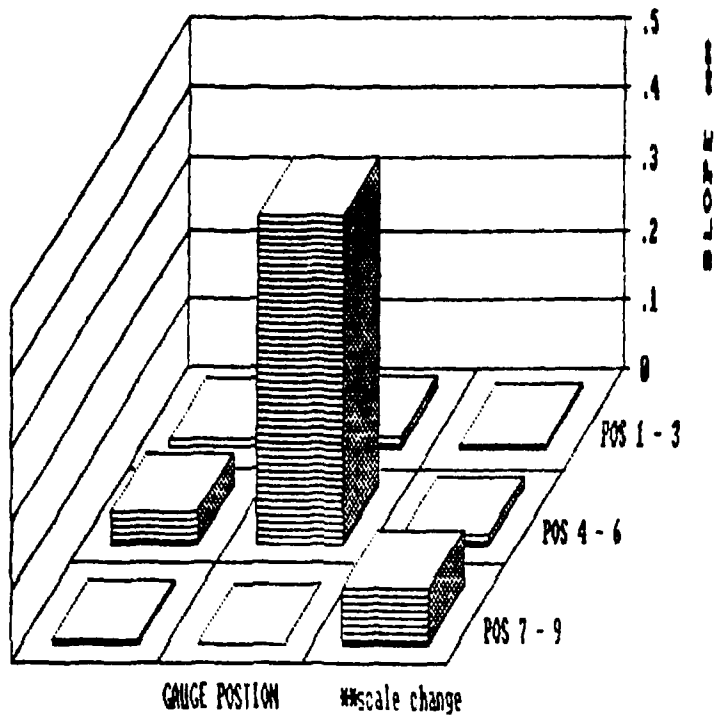


Figure 20. Slope of the Probability vs. Element Magnitude curves as a function of element spatial position in the visual display-square array of digital elements.

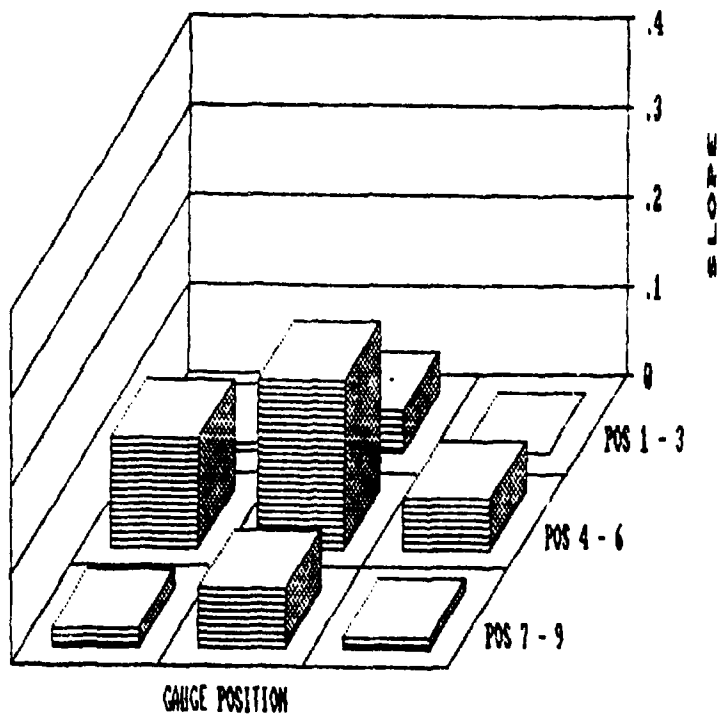


Figure 21. Slope of the Probability vs. Element Magnitude curves as a function of element spatial position in the visual display-square array of analog elements.

Figures 18, 19, 20, and 21 are similar plots of the slopes of the element probability functions for nine element visual displays. Figure 18 is average data from four subjects in a visual display experiment using nine digital elements arranged in a straight line. The major contribution to the subjects' response is from the three elements in the center of the display. Figure 19 shows the subjects' performance with a 9-element analog display arranged in a line. Now, it appears that information from many elements has an influence on the response. Figures 20 and 21 show similar plots of data for 9 element digital and analog displays arranged in a square spatial pattern. A similar picture emerges: information is accrued from most of the elements in the analog display, but from only a few, centrally located and spatially adjacent elements in a digital display. (The data from individual subjects resembled the averaged data very closely.)

This analytical technique may enable useful comparisons to be made between different types and arrangements of visual and auditory displays and may provide important information about how human observers process information from stimulus displays.

7. REFERENCES

Berliner, J. E., and Durlach, N. I. (1973). Intensity perception. IV. Resolution in roving-level discrimination. J. Acoust. Soc. Am., 53, 1270-1287.

Bradley, D. W., and Bradley, R. A. (1983). Application of Sequence Comparison to the Study of Bird Songs, in: Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison, D. Sankoff and J. B. Kruskal, eds., Addison-Wesley, Reading.

Deutsch, D. (1986). Recognition of durations embedded in temporal patterns. Perception and Psychophysics, 39(3), 179-186.

Durlach, N. I., and Braida, L. D. (1969). Intensity Perception. I. Preliminary theory of intensity resolution. J. Acoust. Soc. Am., 46, 372-383.

Kinchla, R. A., and Smyzer, F. A diffusion model of perceptual memory. Perception and Psychophysics, 1967, 2, 219-229.

Kruskal, J. B. (1983). An Overview of Sequence Comparison, in: Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison, D., Sankoff and J. B. Kruskal, eds., Addison-Wesley, Reading.

Leek, M. R. (1987). The role of attention in complex sound resolution, in: Complex Sound Perception, W. A. Yost and C. S. Watson, eds., Erlbaum, Hillsdale.

Leek, M. R. and Watson, C. S. (1984). Learning to detect auditory pattern components, J. Acoust. Soc. of Am., 76, 1037-1044.

Macmillan, N. A., Kaplan H. L., and Creelman, C. D. The psychophysics of categorical perception. Psychological Review, 1977, 84, 452-471.

Sankoff, D., and Kruskal, J. B. (1983). Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison, Addison-Wesley, Reading.

Sorkin, R. D. Extension of the theory of signal detectability to matching procedures in psychoacoustics. Journal of the Acoustical Society of America, 1962, 34, 1745-1751.

Sorkin, R. D. (1984). Discrimination of binary tone sequences. J. Acoust. Soc. Am., 75(Suppl. 1), S21.

Sorkin, R. D. (1987). Computational models of tonal sequence discrimination. In W. A. Yost and C. S. Watson (Eds.), Auditory Processing of Complex Sounds. Hillsdale, NJ: Erlbaum, 104-117.

Sorkin, R. D., Boggs, G. J., and Brady, S. L. (1982). Discrimination of temporal jitter in patterned sequences of tones. J. Exper. Psych.: Human Performance and Perception, 8, 46-57.

Spiegel, M. F., and Watson, C. S. (1981). Factors in the discrimination of tonal patterns. III. Frequency discrimination with components of well-learned patterns. J. Acoust. Soc. Am., 69(1), 223-230.

Tanner, W. P., Jr. Physiological implications of psychophysical data. Ann. N.Y. Acad. Sci., 1961, 89, 752-765.

Watson, C. S. (1987). Uncertainty, informational masking, and the capacity of immediate auditory memory. In W. A. Yost and C. S. Watson (Eds.), Auditory Processing of Complex Sounds. Hillsdale, NJ: Erlbaum, 267-277.

Watson, C. S., Kelly, W. J., and Wroton, H. W. (1976). Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty. J. Acoust. Soc. Am., 60, 1176-1186.

8. PROJECT PERSONNEL

Elvers, G., Graduate Student, Department of Psychological Sciences, Purdue University.

Mabry, T., Graduate Student, Department of Psychological Sciences, Purdue University.

Sorkin, R. D., Professor of Psychological Sciences, Purdue University.

Snow, M. P., Graduate Student, Department of Psychological Sciences, Purdue University.

9. ADVANCED DEGREES

Awarded:

Master of Science to M. P. Snow. Evidence for a Duration Scaling Process in Auditory Memory, Purdue University, August, 1987.

Pending:

Master of Science to T. R. Mabry. Detection Theory Analysis of Visual Displays, Purdue University, December, 1987.

10. PUBLICATIONS

Adkins, J. M. and Sorkin, R. D. Effect of channel separation on earphone presented tone, noise, and stereophonic material. Journal of Audio Engineering Society, 1985, 33, 234-239.

Sorkin, R. D. and Woods, D. D. Systems with human monitors: A signal detection analysis. Human-Computer Interaction, 1985, 1, 49-75.

Robinson, D. E. and Sorkin, R. D. A contingent criterion model of computer assisted detection. In R. Eberts and C. G. Eberts (Eds.) Trends in Ergonomics/Human Factors, Vol. II. Amsterdam: North Holland, 1985.

Sorkin, R. D. and Kantowitz, B. H. Speech Communication. In G. Salvendy (Ed.) Handbook of Human Factors/Ergonomics. New York: J. Wiley & Sons, 1987.

Kantowitz, R. D. and Sorkin, R. D. Allocation of Functions. In G. Salvendy (Ed.) Handbook of Human Factors/Ergonomics. New York: J. Wiley & Sons, 1987.

Sorkin, R. D. Auditory and Tactile Displays. In G. Salvendy (Ed.) Handbook of Human Factors/Ergonomics. New York: J. Wiley & Sons, 1987.

Sorkin, R. D. Computational models of tonal sequence discrimination. In W. A. Yost & C. S. Watson (Eds.), Auditory Processing of Complex Sounds. Hillsdale, NJ: Erlbaum, 1987.

Sorkin, R. D. Tonal Sequence Discrimination. Temporal factors in the discrimination of tonal sequences. J. Acoust. Soc. Am., 1987, 82, (in press).

Snow, M. P. Evidence for a Duration Scaling Process in Auditory Memory, M.S. Thesis, Purdue University, August 1987.

Sorkin, R. D., Kantowitz, B. H., and Kantowitz, S. C. Likelihood alarm displays. Human Factors, (accepted for publication pending

revision).

Sorkin, R. D., Mabry, T. R., and Weldon, M. A. Detection theory analysis of performance with visual displays. (in preparation)

Snow, M. P. and Sorkin, R.D. Duration and frequency transformations in tonal sequence discrimination. Perception and Psychophysics, (in preparation).

Sorkin, R.D. and Snow, M.T. Discrimination of time-expanded tonal sequences. J. Acoust. Soc. Am., 1987, 82, (abstract, in press).

Sorkin, R.D., Robinson, D.E., and Berg, B.G. A detection theory method for the analysis of visual and auditory displays. Proceedings of the Human Factors Society, 1987, (in press).

Sorkin, R.D. Review of Loeb, M., Noise and Human Efficiency, in American Journal of Psychology, (in press).

11. INTERACTIONS

Sorkin, R. D. Attended Acoustical Society Meetings, October 1984, Minneapolis, Minn. and April 1985, Austin, Texas; discussed research with workers in complex pattern perception area.

Sorkin, R. D. Lecture and briefing on research at AFOSR/NL, Bolling AFB, Washington, D.C., July 1985.

Sorkin, R. D. Attended Acoustical Society Meetings, November 1985, Nashville, Tenn., and May 1986, Cleveland, Ohio; discussed research with workers in complex pattern perception area.

Sorkin, R. D. Convened symposium on recent developments in hearing, Midwestern Psychological Association, Chicago, Ill., May 1986.

Sorkin, R. D. Panelist: National Science Foundation Epscor Program, May, 1986

Sorkin, R. D. "Models of Tone Sequence Comparison", Air Force Office of Scientific Research Workshop on Complex Sound Perception, Sarasota, Fla., April 1986.

Sorkin, R. D. "Computational Models of Tonal Sequence Discrimination". Psychonomics Society Meeting November 1986, New Orleans, La.

Sorkin, R. D. Invited lecture: Computer-assisted signal detection. Wright State University, January 16, 1987.

Sorkin, R. D. Invited lecture: Likelihood displays. Indiana University, April 2, 1987

Sorkin, R. D., Attended Acoustical Society Meeting, May 1987, Indianapolis, Indiana; discussed research with workers in complex pattern perception area.

Sorkin, R. D. Member of National Research Council Panel on The Classification of Complex, Non-speech Sounds, 1987-1989.

Sorkin, R. D. Member of National Research Council Panel on the Removal of Noise from a Speech/Noise Signal, 1986-1988.

Contract with Naval Weapons Center, China Lake, California: "Analysis of Automated Decision Systems", March 1987-December 1987.

Pending Interactions:

Sorkin, R. D., Robinson, D. E., and Berg, B. G. "A Detection theory method for the analysis of visual and auditory displays", Human Factors Society meeting, October 1987, New York, N.Y.

Sorkin, R. D. and Snow, M. P. "Discrimination of time-expanded tonal sequences", Acoustical Society meeting, Miami, Fl. November 1987.

Sorkin, R. D. "Two-process model of tonal sequence discrimination", AFOSR conference, Chicago, December 1987.

~~FIL~~MED
48