73-21,284

CHEN, Wen-Yen, 1940-
TWO-DIMENSIONAL STIMULUS GENERALIZATION FOLLOWING REDUNDANT DISCRIMINATION TRAINING.

The City University of New York, Ph.D., 1973
Psychology, general

University Microfilms, A XEROX Company, Ann Arbor, Michigan

TWO-DIMENSIONAL STIMULUS GENERALIZATION FOLLOWING REDUNDANT DISCRIMINATION TRAINING

by

WEN-YEN CHEN

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1973

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

March 24, 1973 — Eric G. Heinemann

date — Chairman of Examining Committee

May 3, 1973 — Leonard S. Kogan

date — Executive Officer

Dr. Solomon Weinstock

Dr. Neil Macmillan

Supervisory Committee

The City University of New York

Abstract

TWO-DIMENSIONAL STIMULUS GENERALIZATION FOLLOWING REDUNDANT DISCRIMINATION TRAINING

by

WEN-YEN CHEN

Adviser: Professor Eric G. Heinemann

Five Ss trained to discriminate between two compound stimuli consisting of combinations of white noise and light. The presentation ratio of the two compound training stimuli and their discriminability were varied. Following discrimination training generalization testing was conducted. It was found that the generalization surface had the highest proportion of response associated with combination of the weakest sound and weakest light, not with the positive training stimulus. S. trained under a condition of saml1 differences in sound level showed gradients along the sound dimension which were almost flat, indicating that the dimension of sound exerted very little control over responding. Similarly, small differences in light intensity during discrimination training produced little control by the light dimension. However, different presentation ratio failed to produce effect during the generalization testing.

The measure of choice reaction time (CRT) showed that the median CRT and response gradients are inversely related: high response probability is associated with short CRT; and vice versa.

The generalization gradients were analyzed in terms of a signal recognition model.

Acknowledgement

The author is deeply grateful to Dr. Eric G. Heinemann and Dr. Solomon Weinstock for their guidance, encouragement and assistance in the conduct of this research and preparation of this manuscript.

Thanks are also due to Dr. David Raab for his assistance in calibrating the noise intensity and in designing the auditory circuit.

Acknowledgement is further made to Dr. Neil MacMillan for his constructive criticism on the early draft of this manuscript.

The subjects who participated in this experiment were partly paid by the Department of Psychology, Brooklyn College of the City University of New York and partly through the grant (United States Public Health Service Grant MH 13955) to Eric G. Heinemann.

Finally, the author would like to express his appreciation to his wife, Chih-mei, for her patience and understanding throughout these years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

The term "stimulus generalization" refers to the empirical phenomenon that an organism conditioned to a CS also responds, to a lesser degree, to stimuli to which it has not been exposed. This phenomenon was first demonstrated by Pavlov who also offered a theoretical account. The ensuing theoretical expositions (Hull, 1943; Lashley and Wade, 1946; Spence, 1956) and experimental clarification (Jenkins and Harrison, 1960; Heinemann and Rudolph, 1963) greatly increase the understanding of the phenomenon.

Currently, in describing generalization, "stimulus control" is often used to refer to the empirical fact that variation in some aspect of stimulus properties brings about changes in behavior. The functional relationship between stimuli and behavior is called stimulus generalization gradient. The slope of the gradient is commonly used to indicate the degree of stimulus control. The sharper the slope of the gradient the greater the degree of stimulus control along the dimension.

Two-dimensional stimulus generalization

In many experimental situations stimuli that consist of several dimensions have been employed. The problem of how stimulus generalization takes place along more than one dimension is the subject of present paper. Pavlov (1927) conditioned a dog to a compound stimulus consisting of a light and a tone and later tested with each component of the stimulus. He found that when light was presented alone during the test, it failed to elicit a CR. Lashley (1938) also found that in a form discrimination experiment, some of the rats responded to discriminanda only on the basis of size; while others on the basis of brightness. A study done by Reynolds (1961) using an operant conditioning situation demonstrated that the two components of a complex S+ did not gain equal control over responding. In Reynolds' study the S+ was a white triangle on red background and the S- was a white circle on a green background. During the test, each component of the S+ and S- was presented alone, and it was found that one pigeon responded exclusively to the white triangle and the another responded only to the red background. Several recent investigators have found some control by both dimensions of a compound stimulus, for example, Butter (1963) and Johnson (1970).

The finding that the components of a compound stimulus may exert differing degree of control over responding has contributed to the revival of interest in the concept of "attention". The term "attention" has been used in many different senses. In discussions of animal discrimination and generalization it is often used as a synonym for "stimulus control" (Skinner, 1953) or of selective stimulus control (Terrace, 1966). But "attention" is sometimes also used to refer to a central process which, in turn, is used to explain the phenomenon of selective stimulus control (Mackintosh, 1956).

Undoubtedly there are many variables which affect the relative degree of control by the components of a compound stimulus. One of these variables is the specific differential training employed. A dimension subjected to specific differential training gains greater control than one which is not. For example, Newman (reported in Baron, 1965) showed that a pigeon trained to discriminate between a white vertical line on a green dircular surround (S+) and a green circular surround alone (S-) later showed a sharp gradient along the dimension of angularity of the line. On the other hand, when the bird was trained to discriminate between a white vertical line on green circular surround (S+) and a white vertical line on red circular surround (S-), and almost flat gradient was obtained along the angularity dimension. Similar results were reported in Purtle and Newman's study (1969) which employed the dimensions of angularity and texture. Purtle and Newman found that college students who received training to discriminate between values on the angularity dimension later showed a steeper gradient along the dimension of angularity than those who were trained to discriminate between compound stimuli.

The previous studies also suggest that the relative degree of control was affected by the temporal order of differential training with stimuli from each dimension (e.g., Chase, 1968; Kamin, 1969). The dimension having prior differential training tends to obtain prominent control and to block the development of stimulus control by a second dimension which is then added.

A third variable known to affect the relative degree of control is the relative discriminability between two training stimulus values from each dimension. At least with pigeons it has been found that a dimension having two stimulus values which are more discriminable will gain more control than the other dimension (Chase and Heinemann, 1972).

Theoretical analysis of two-dimensional gradients has focused on prediction of response strength to the compound stimuli from a knowledge of response strength associated with the separate components or, upon the combination rules. Three major models have been advanced on the basis of different theoretical orientations. A detailed discussion of these models and their properties can be found in Jones (1962), Butter (1963), Shepard (1964) and Cross(1965). Briefly, the first model, called the excitation model, is derived from the Pavlov-Hull-Spence tradition of theory in which generalization is conceived as a spread of excitation from the training CS to other similar stimuli. The generalization surface, according to this model, can be generated by rotating a uni-dimensional gradient around its peak, resulting in a surface of conical shape and circular isosimilarity contour. The rule of combination is multiplicative.

The second model, called the discrimination model, was advocated by Lashley and Wade (1946) and later by Guttman (1956). This model states that the generalization decrement for any two-dimensional stimulus is the sum of the decrement to each stimulus changing in one dimension alone. The isosimilarity contour of this model is of diamond shape and the combination rule is additive.

The third model is often referred to as the better criterion model. It is derived basically from Lashley's "principle of dominance organization" (Lashley, 1942). It asserts that stimulus control is selective. If one dimension gains control over responding the other dimension becomes ineffective. The generalization surface is a pyramid and produces square isosimilarity contours.

Numerous studies have been directed toward the evaluation of these three models (e.g., Jones, 1962; Butter, 1963; Shepard, 1964; Johnson, 1970).

This conventional approach to selecting combination rules may be worth pursuing, but it does not provide an analytic understanding of how the organism combines stimulus properties. In addition, the combination rule is usually found to be situation-dependent and very much influenced by the attention of the Ss (Shepard, 1964). It seems desirable to derive combination rules from a model of the underlying processes.

The signal recognition model of stimulus generalization

The development of detection theory in psychophysics has led to another approach to the problem of discrimination and generalization. The detection model has been used quite successfully in accounting for data obtained with a two-alternative choice method (Boneau and Cole, 1967; Heinemann, Avin, Sullivan and Chase, 1969; Chase and Heinemann, 1972).

A. The single dimension case: The Basic detection problem in psychophysics is to decide, on the basis of the events occurring in a fixed interval of time, whether the interval contains only background noise or signal plus noise. Within the structure of statistical decision theory, detection theory asserts that repeated presentation of a constant stimulus does not result in a constant sensory effect, but in a distribution of effects. In the detection situation there are only two possible stimulus presentations -- noise and signal plus noise. Let $f(x \mid N)$ denote the distribution density generated by noise alone and let $f(x \mid SN)$ be the distribution density generated by signal plus noise, where $\underline{x}$ might be regarded as the sensory effect of the stimulus. It is assumed that the observer is capable of computing the likelihood ratio $L(x) = f(x \mid SN) / f(x \mid N)$. For mathematical simplicity, it is further assumed that a monotonic transformation $\underline{y}$ of $L(x)$ exists such that the density function $f(y \mid N)$ and $f(y \mid SN)$ are normal. The observer places a fixed cut-off point or criterion X_c along the dimension $\underline{x}$. On each trial an observation $\underline{x}_i$ is made and compared to the criterion X_c. if $\underline{x}_i \leqslant X_c$, observer will respond 'No' and when $\underline{x}_i > X_c$, respond 'Yes'.

Usually, the variances σ_N^2 and σ_{SN}^2 of the two distributions are assumed to be equal: $\sigma_N^2 = \sigma_{SN}^2 = \sigma^2$. If the mean of the noise distribution is set equal to zero and if the mean of signal plus noise distribution is d, the "normalized" displacement of the distribution of signal plus noise, $d' = d/\sigma$, is customarily used as an index of the "sensitivity" of the sensory system, and is presumed to be independent of non-sensory factors. On the other hand, the placement of the criterion X_c is determined by such non-sensory variables as response contingent payoffs and the <u>a priori</u> probability of the signal.

Given the assumed density function $f(x \mid N)$, $f(x \mid SN)$ and criterion X_c, the probability of a "hit" $p(\text{Yes} \mid SN) = \int_{X_c}^{\infty} f(x \mid SN)\,dx$; and the probability of a "false alarm" $p(\text{Yes} \mid N) = \int_{X_c}^{\infty} f(x \mid N)dx$. These two probabilities are typically plotted against each other to form an ROC, or isosensitivity curve.

The detection model has been extended to account for discrimination and generalization in the pigeon (Heinemann, <u>et al</u>., 1969). Although the detection model does not provide any specific account on the acquisition process of discrimination learning, it does predict the asymptotic response probabilities to all stimuli along the dimension.

If X_j is a stimulus on the same dimension as two training stimuli X_1 and X_2 ($X_1 < X_2$), the probability of a response A_1 given X_j is equal to $\int_{X_c}^{\infty} f(x \mid X_j)\,dx$. If $p(A_1 \mid X_j)$ is plotted against X_j, a distribution function will be obtained. It is clear from the integral that when $X_j \longrightarrow \infty$, $p(A_1 \mid X_j) \longrightarrow 1.0$ and when $X_j \longrightarrow -\infty$, $p(A_1 \mid X_j) \longrightarrow 0$ (Luce, 1963). If the parameters, i.e., mean and variances of the assumed underlying normal distribution can be recovered independently and the stimuli can be properly located, discrimination and generalization can be described by the placement of the criterion X_c.

One of the obvious predicted consequences of the model is that the shape of the gradient is that of a normal ogive, not the conventional peaked gradients. Heinemann, et al.(1969), Pierrel and Sherrman (1960) and others (LaBerge, 1961; Cross and Lane, 1962; Pickett, 1967; Weinstock, Robbins and Chen, 1972) have obtained data supporting this prediction of the model.

B. The two-dimensional redundant case: The detection model can be readily extended to account for the generalization surface following training on a two-dimensional redundant discrimination: e.g., dim-light and soft-sound versus bright-light and loud-sound. The two dimensions, light and sound, are redundant in the sense that the stimulus from either dimension is sufficient to predict which choice will lead to reinforcement.

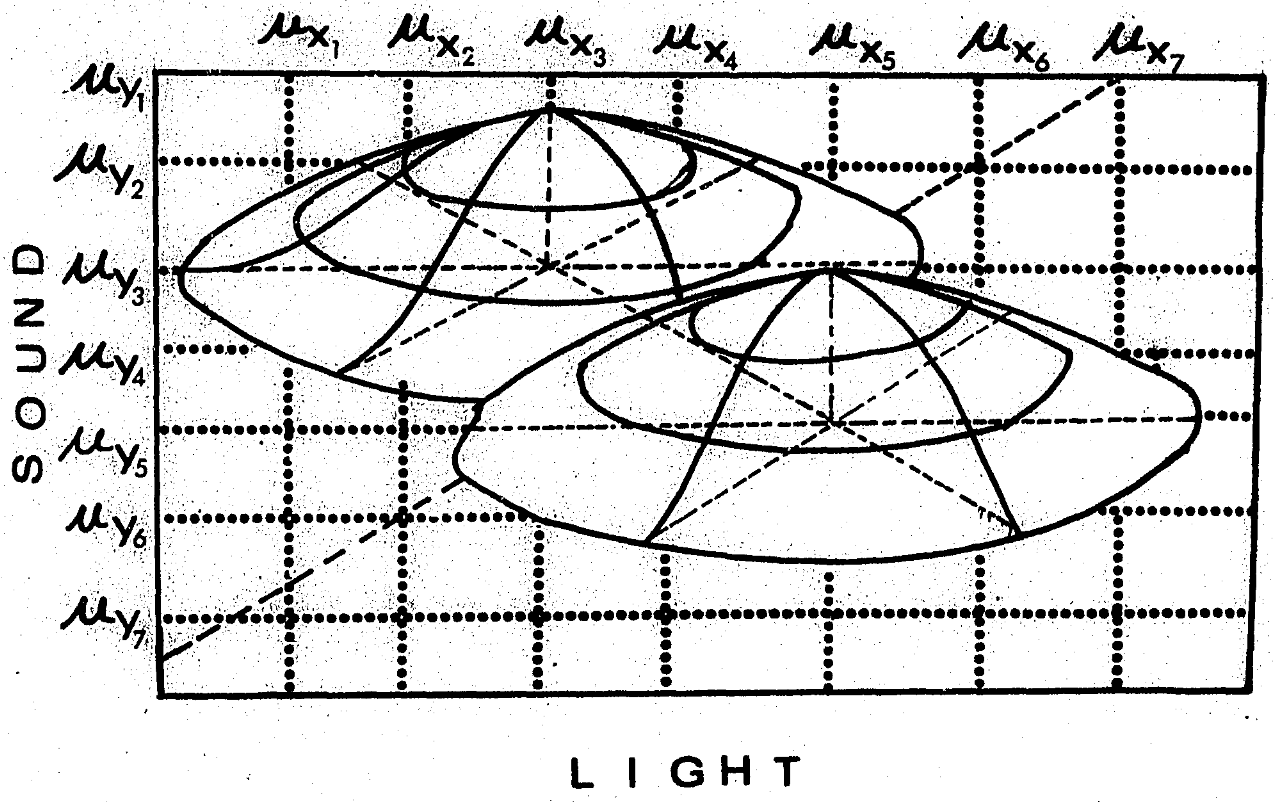
Consider the three-dimensional space shown in Figure 1. Let the random variables x and y represent the sensory effect of stimulation by light and sound respectively, and assume that the sensory effects of light and sound are orthogonal. Let the third dimension p (not labelled in the figure) represent distribution density of the random variables x and y which, for mathematical simplicity, are assumed to be normally distributed. Let X_1, X_2, ... X_7 represent seven different light intensities from dim to bright; Y_1, Y_2, ... Y_7 represent seven sound intensities, from soft to loud. If the mean and variance of the distribution of X_i is μ_{X_i} and $\sigma_{X_i}^2$ respectively, and of Y_j μ_{Y_j} and $\sigma_{Y_j}^2$, the joint density function is

$$f(X_i,Y_j) = [1 / (2\pi\sigma_{X_i}\sigma_{Y_j})] \exp [(X_i - \mu_{X_i})^2/\sigma_{X_i}^2 + (Y_j - \mu_{Y_j})^2/\sigma_{Y_j}^2]$$

Let the compound stimuli (X_3, Y_3) and (X_5, Y_5) serve as discriminative stimuli, indicated in Figure 1 as two bell-shaped volumes. The S

Figure 1. Three-dimensional representation of signal recognition model of two-dimensional discrimination and generalization.

Figure 1. Three-dimensional representation of signal recognition model of two-dimensional discrimination and generalization.



is trained to respond R when compound stimulus (X_3, Y_3) is presented and L when (X_5, Y_5) is presented. The task that the S faces is to find a decision function which will partition the x y plane in such a way that his decision will achieve a decision goal assumed to be maximalizing the expected values.

It can be shown (Green and Swets, 1966) that the optimal decision rule is one that is based upon a likelihood ratio λ where

$$\lambda = \frac{\exp - \frac{1}{2}[(X_i - \mu_{X_3})^2/\sigma_X^2 + (Y_j - \mu_{Y_3})^2/\sigma_Y^2]}{\exp - \frac{1}{2}[(X_i - \mu_{X_5})^2/\sigma_X^2 + (Y_j - \mu_{Y_5})^2/\sigma_Y^2]} \quad \ldots (1)$$

Assume that the distribution of xs have the common variance σ_X^2 and ys have the common variance σ_Y^2. If the stimulus presentation ratio of (X_3, Y_3) to (X_5, Y_5) is equal to 1.0, the optimal decision function will be a contour along which $\lambda = 1.0$. For $\lambda = 1.0$ the solution of (1) can be shown to be a linear function of the form

$$y = mx + b \quad \ldots (2)$$

where the slope is $m = -[(\mu_{X_5} - \mu_{X_3})\sigma_Y^2]/[(\mu_{Y_5} - \mu_{Y_3})\sigma_X^2]$ and the intercept $b = [\sigma_Y^2(\mu_{X_5}^2 - \mu_{X_3}^2)/2\sigma_X^2(\mu_{Y_5} - \mu_{Y_3})] + (\mu_{Y_5} - \mu_{Y_3})/2$.

If the presentation ratio of (X_3, Y_3) to (X_5, Y_5) is 1:3, the slope of the decision line will remain the same but intercept will move upward by a factor of $\log_e 3[\sigma_Y^2/(\mu_{Y_5} - \mu_{Y_3})]$ that is,

$$b = [\sigma_Y^2(\mu_{X_5}^2 - \mu_{X_3}^2)/2\sigma_X^2(\mu_{Y_5} - \mu_{Y_3})] + [(\mu_{Y_5} - \mu_{Y_3})/2] + [\log_e 3 \cdot \sigma_Y^2/(\mu_{Y_5} - \mu_{Y_3})]$$

On the other hand, if the presentation ratio is 3:1, the intercept of the decision line will be moved downward by the same factor, i.e.,

$$b = [\sigma_Y^2(\mu_{X_5}^2 - \mu_{X_3}^2)/2\sigma_X^2(\mu_{Y_5} - \mu_{Y_3})] + [(\mu_{Y_5} - \mu_{Y_3})/2] - [\log_e 3 \cdot \sigma_Y^2/(\mu_{Y_5} - \mu_{Y_3})].$$

It is obvious that this decision function is determined by the distance between X_5 and X_3; between Y_5 and Y_3; and variances σ_X^2 and σ_Y^2, all of which remained to be estimated.

Given this decision function, for any bivariate normal distribution on the decision plane with mean equal to (μ_X^* , μ_Y^*), the response probability to the left of decision line p(R) is given by

$$p(R) = \int_{-\infty}^{\infty}\int_{-\infty}^{Y=mx+b} [\, 1/(2\pi\sigma_X\sigma_Y)] \exp -\frac{1}{2} [(X - \mu_X^*)^2/\sigma_X^2 + (Y - \mu_Y^*)^2/\sigma_Y^2] \, dx \, dy \, dy$$

which can be shown to be equal to

$$p(R) = 1/2 + (1/2\sqrt{2\pi}\, \sigma_X^2) \left[\int_{-\infty}^{\infty} \exp - \frac{1}{2} (X - \mu_X^*)^2/\sigma_X^2 \cdot \mathrm{erf}(mx + b - \mu_Y^*)/\sqrt{2}\sigma_Y dx \right]$$

The solution of $A = \int_{-\infty}^{\infty} \exp - \frac{1}{2} (X - \mu_X^*)^2/\sigma_X^2 \cdot \mathrm{erf}(mx + b - \mu_Y^*)/\sqrt{2}\, \sigma_Y \, dx$

is very difficult. However digital computer program can provide an approximation to A with any desired percentage error. A method developed by S. Weinstock makes it possible to obtain the response probability for any stimulus combination (X_i, Y_j) by rotating the coordinate system so that one axis becomes superimposed on the decision line after normalization of the distribution. The response probability of any combination (X_i, Y_j) then, can be directly read from the available unit normal table (for details of the procedure see Sppendix 1).

The nature of the theoretical gradients generated by the model, either in the single or the two-dimensional case not only depends upon the placement of the decision criterion but upon the location of each stimulus also. How the physical intensity of the environment stimulus is transduced into psychological magnitude is the traditional problem of scaling. The mounting evidence collected by Stevens suggests that "... apparent, or subjective, magnitude grows as a power function of stimulus intensity " (Stevens, 1961, p.2), in opposition to Fechner's logarithmic law. Stevens' power function can be stated as follows:

$$\psi = k\phi^n$$

where ψ is the psychological magnitude and ϕ the physical intensity. The exponent n depends upon the stimulus continuum and k is a constant depending on one's choice of units. Both the traditional logarithmic transformation and Stevens' power function will be employed in placing the stimuli in connection with the theoretical analysis in latter section.

The signal recognition model of choice reaction time

Several choice reaction time models have been proposed in the framework of statistical decision theory. The earliest attempt made by Stone (1960) was an application of the classical theory of sequential sampling and optimal stopping developed by Wald (1947). The model states that upon the onset of a stimulus, a stream of information about the signal flows into the decision making mechanism of the S. The time taken for information to reach the decision mechanism, called input time T_i, is assumed to be independent of the signal and decision time T_d. The information which is arriving is regarded as a random variable X sampled by the S at short time intervals Δt. Consider a simple two-stimulus (X_1 and X_2), two-alternative (R_1 and R_2) discrimination situation and let $p(x \mid X_1)$ and $p(x \mid X_2)$ be the values of the random variable X when the stimuli presented are X_1 and X_2 respectively. The model assumes that at each discrete moment, S samples one point and computes a quantity $C(X_j)$, based on incoming information, and stores it in an adder. The S will continue to sample until $C(X_j)$ exceeds a preselected constant log B or falls below log A ($A < B$), and makes the appropriate response. According to Wald's sequential probability test, it can be shown that the quantity $C(X_j)$ which is based on likelihood of $\log p[(x \mid X_2)/p(x \mid X_1)]$ is optimal in the sense that the sample size needed to make the proper decision is minimal. Suppose the S takes n samples to reach the decision, the total response latency T will be equal to $T_i + n \cdot \Delta t + T_m$, where T_m indicates the time to execute the response, which is assumed to be independent of T_d and the stimulus.

Stone also shows (1960, Appendix 1) that if one is willing to assume a symmetric distribution of $p(x \mid X_1)/p(x \mid X_2)$ and $p(x \mid X_2)/p(x \mid X_1)$, then the distribution of sample size n's and therefore T_d's leading to a decision for X_1 is the same on correct and incorrect decision. In other words, if N_{ij} is the sample size for a decision in favor of X_i when X_j is presented, under the afromentioned restriction, $N_{11} = N_{12}$, i.e., $L(R_1 \mid X_1) = L(R_1 \mid X_2)$ and $N_{21} = N_{22}$, i.e., $L(R_2 \mid X_2) = L(R_2 \mid X_1)$. The prediction is testable in a simple detection situation.

A second model proposed by Stone assumes that S adopts a less efficient strategy, namely, that he uses a fixed sample size n for all trials choosing a value n which will give a certain accepted error rate. One of the consequences of the fixed sample size model is that, since T_d's are constant within the experimental settings, all variations in latencies must be attributed to either T_i or T_m.

Some other latency models (e.g., Edwards, 1965; Audley and Mercer, 1968; Thomas, 1970; Smith, 1968; LaBerge, 1962 Audley and Pike, 1968) will be discussed in a later section in relation to the findings of the present experiment.

The purpose of the study

The purpose of the present research is to examine the empirical generalization surface obtained after two-dimensional redundant discrimination training by manipulating: (a) the presentation ratio of two compound training stimuli (X_3, Y_3 vs X_5, Y_5); (b) the discriminability between two values of the training stimuli along each dimension, i.e., the discriminability between X_3 and X_5 and between Y_3 and Y_5. Comparisons will be made between empirical and theoretical gradients generated by the signal recognition model. The merit of employing human Ss is that it makes an independent estimation of parameters and scaling of stimuli along each dimension possible.

The latency gradient of stimulus generalization and its relation to the response probability gradient will also be examined. Most studies in the area of psychophysics and stimulus generalization have used measures of the relative frequencies of response as the dependent variable, and relatively few studies have concentrated on the measure of choice times. An analysis of choice times, in addition to the measure of response frequencies, might yield valuable information about the nature of decision process.

Method

Subjects: The Ss were five college students who were paid for their service. The Ss had no known defects in vision and hearing and were randomly assigned to one of five experimental conditions.

Apparatus: The apparatus consisted of a response panel and associated equipment for automatic presentation of stimuli and automatic recording of responses and latencies. On the response panel, there were three buttons: one located at the lower center, marked as the Center Key; and two response keys located at the upper right and left, equidistant from the lower Center Key, marked as R and L. At the upper center, there was a red indicator light which signaled the beginning of a trial. Directly above each response button, there was a green reinforcement light which could be lighted for 3 seconds after S pressed a response button. After response, the light located above the key which was correct for that trial was lighted, regardless of which key was pressed by the S.

A trial started with the onset of the red indicator light. Pressing the center button shut off the red indicator light and simultaneously started the presentation of both visual and auditory stimuli for 600 msec. The duration of the stimuli was controlled by a Hunter Interval Timer. After the stimuli were presented S responded either R or L, and the appropriate reinforcement light followed immediately. The termination of the reinforcement light concluded a trial. After a five seconds intertrial interval, the next trial began.

The auditory stimuli were bursts of white noise generated by a Grason-Stadler noise generator, Model 901B. A series of potentiometers was used to produce seven different intensity levels of white noise ranging from

-1.0 dB to 2.0 dB re 86.6 dB SPL.

Before each session of the experiment the noise intensity levels were measured with a Hewlett Packard 427A Voltmeter. The white noise was delivered to S through a set of earphones (Permoflux, PDR 8).

The visual stimuli were circular patches of light, 4 in. in diameter that were projected on a white wall. Wratten neutral density filters were used to vary the luminances of these circular patches over a range of -0.26 log ft.l. to 0.37 log ft. l., as measured by a Spectra Prichard Photometer. The stimuli described were surrounded by a circular black area 12 in. in diameter. This area, in turn, was surrounded by an evenly illuminated rectangular area 18 x 20 sq.in. The luminance of the "black" circular area was -1.03 log ft.l. and the luminance of the background was -0.58 log ft.l.

Responses and their latencies were recorded. The response latency was defined as the time between the onset of the stimulus and the occurrence of a response, as measured by a Hewlett Packard 5223L Electronic Counter with 1 msec. accuracy.

Seven levels of light intensity designated as X_1, X_2, X_3 ... X_7 ($X_1 < X_2 < X_3 \ldots < X_7$) and seven levels of noise intensity designated as Y_1, Y_2, Y_3, ... Y_7 ($Y_1 < Y_2 < Y_3 \ldots < Y_7$) were used in the experiment, constituting of a total of 7 x 7 = 49 combinations of light and sound that could be presented.

All equipment except the response panel was housed outside the experimental cubicle and a remote air conditioner was running throughout the experiment, providing a faint masking noise.

Discrimination training: The S was seated in front of the response panel and was given following instructions:

> " On the panel in front of you, there are three buttons marked 'L' and 'R' and 'Center Key'. Each trial of this experiment will begin when the light marked 'Signal' goes on. When you are ready, push the 'Center Key', this will shut off the signal light and a circular patch of light will be presented on the black circle on the wall. At the same time you will hear a burst of noise through the earphones. Each combination of light and noise you are to judge whether it is 'R' or 'L' and press the appropriate button. If you are right, a light will go on over the button you push, if you are wrong, the light over the other button will go on. In the beginning you have no way of knowing which combination of light and noise is 'R' or 'L' and you will have to guess. Try to respond correctly and rapidly throughout experiment.
>
> On each trial, before you push the 'Center Key' keep your eyes focused on the illuminated part of the rectangle on the wall. When you push the 'Center Key', then, shift your focus to the black circle. Any question ?"

The discrimination training conditions for five Ss are shown in Table 1. The entries R and L indicated the response that were reinforced in the presence of the combination (X_i, Y_j) of sound and light. The presentation percentage of a particular combination of stimuli is indicated in the parentheses beneath each response entry in Table 1. It should be noted in Table 1 that, for S AK, the two values along the light dimension were easily discriminable; while for S OP, they were difficult to discriminate.

After five min. of dark adaptation, the trial started with the onset of the red light. Each trial consisted of a simultaneous presentation of light and sound for 600 msec. The order of presentation of the stimulus combination was randomly determined in a block of 80 trials. After 80 trials, the stimuli were presented again in reverse order. Each session lasted from 20 to 30 min. Discrimination training continued until the performance reached asymptote, judged by inspecting the daily learning curves. The number of trials to reach the asymptote

Talbe 1. Discrimination Training Conditions.

S: JZ		
	Y_3: 0.0	Y_5: 1.0
X_3: -0.02	R (50 %)	
X_5: 0.05		L (50%)

S: AG		
	Y_3: 0.25	Y_4: 1.0
X_3: 0.03	R (25 %)	
X_5: 0.05		L (75 %)

S: RW		
	Y_3: 0.25	Y_5: 1.0
X_3: 0.03	R (75 %)	
X_5: 0.05		L (25 %)

S: AK		
	Y_3: 0.0	Y_4: 1.0
X_3: -0.02	R (50 %)	
X_5: 0.37		L (50 %)

S: OP		
	Y_3: 0.0	Y_5: 1.0
X_3: -0.02	R (50 %)	
X_4: 0.03		L (50 %)

ranged from 1200 to 1600.

Generalization testing: After the performances reached an asymptotic level, discrimination training continued for three more sessions. Then, generalization testing was started by inserting testing stimuli consisting of all possible combinations of sound and light intensity (a total of 47 not including training stimuli) among the training stimuli. The S was instructed that the experiment would proceed in the usual way except that on many trials the reinforcement light would not be given. However, on trials on which the training stimuli were presented feedback information was still provided. Within a block of 80 trials, 32 were training and 48 were testing trails. Testing stimuli were randomly mixed with training stimuli with the restriction that within a block of 80 trials each test stimulus was always presented exactly once. Over the course of all the generalization testing sessions, a total of 65 observation were obtained for each combination.

Unidimensional generalization gradients: In order to obtain gradients for the light and sound dimensions individually, the above discrimination and generalization procedure were repeated except that stimuli from one dimension were presented. This part of the experiment was conducted after the first part of experiment was concluded.

Magnitude estimation of the stimuli: The last stage of the experiment was to ask S to estimate directly the psychological magnitude of each stimulus of a dimension, using the magnitude estimation method developed by Stevens (1957). Again, the procedure was administered with respect to a single dimension only. The stimuli used were those employed in generalization testing. The S was given the following standard magnitude estimation instruction:

" I am going to give you a series of lights (sounds) with different intensities presented one at a time with a duration of 2 seconds. Your task is to tell me how bright (loud) they look (sound) by assigning a number to each of them. The first light (sound) you see (hear) will be called the standard and be given the number 10. In other words, if the standard is called 10 what would you call each of the rest of stimuli. Use whatever number seems to you appropriate -- fraction, decimal, or whole number. For example, if you feel the second stimulus is 7 times as bright (loud) as the standard, say 70. If it looks (sound) one fifth as bright (loud), say 2; if one twentieth say 0.5, etc.

Try to make the ratio between the number you assign to the different stimuli correspond to the ratio between the brightnesses (loudnesses) of the lights(sounds). In other words, try to make the number proportional to the brightness (loudness) as you see (hear) it.

The red light in front of you serves as a ready signal. When you are ready, press the "Center Key" the red light will go off and stimulus will be on for 2 seconds."

The standard stimuli were X_4 and Y_4. Each stimulus was estimated 20 times in a randomly assigned order.

Results

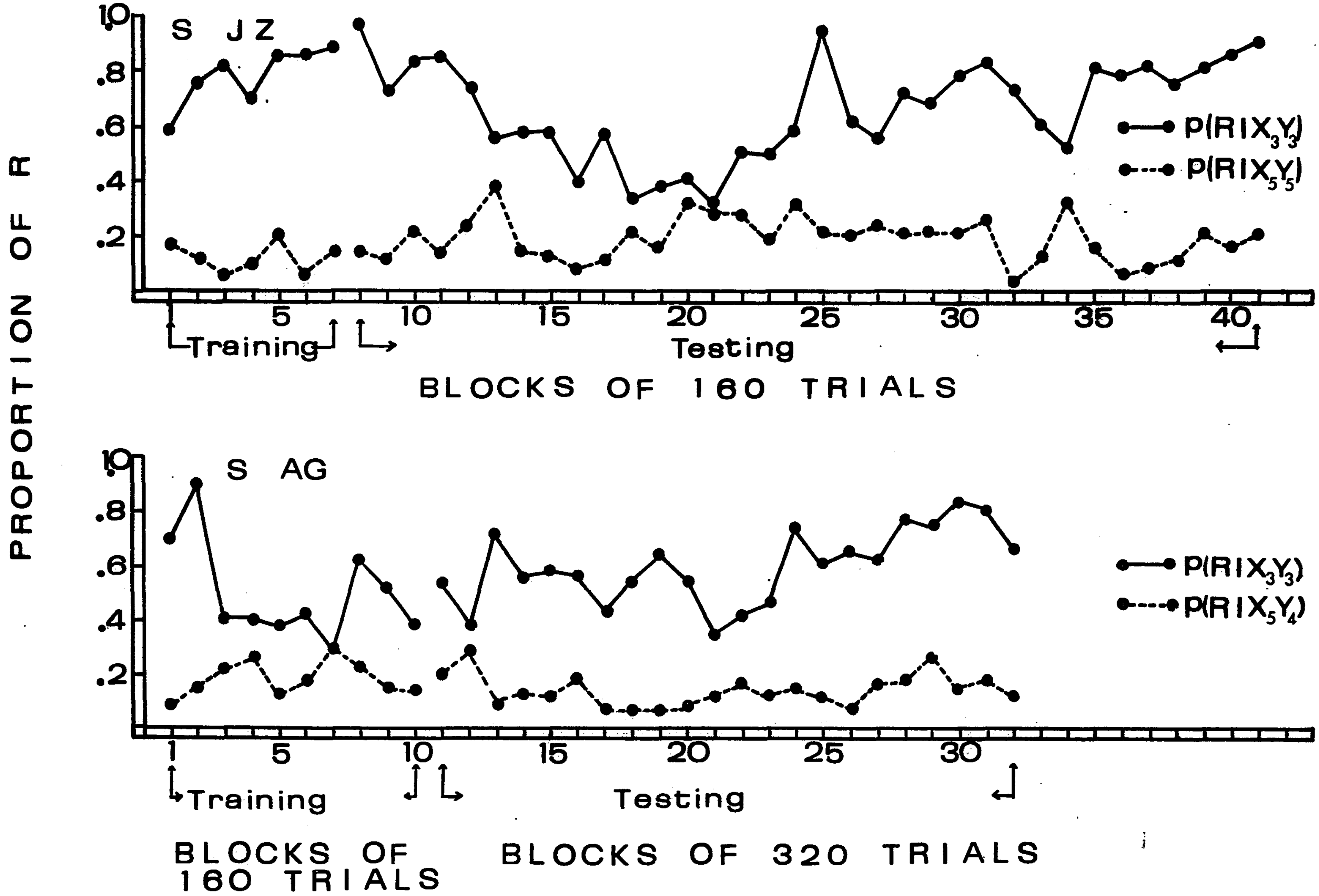
<u>Discrimination</u>: The performance during the training and the testing phase is depicted in Figure 2. Each graph shows the proportion of response R as a function of blocks of 160 or 320 trials for an individual <u>S</u>. Each point on the upper curve in each graph represents the estimate of $p(R \mid X_3,Y_3)$ and the lower curve represents the estimate of $p(R \mid X_5,Y_5)$. In a perfect discrimination, $p(R \mid X_3,Y_3)$ will reach 1.0 and $p(R \mid X_5Y_5)$ will be 0. It is clear from the graph that all five <u>S</u>s discriminated between the two training stimuli.

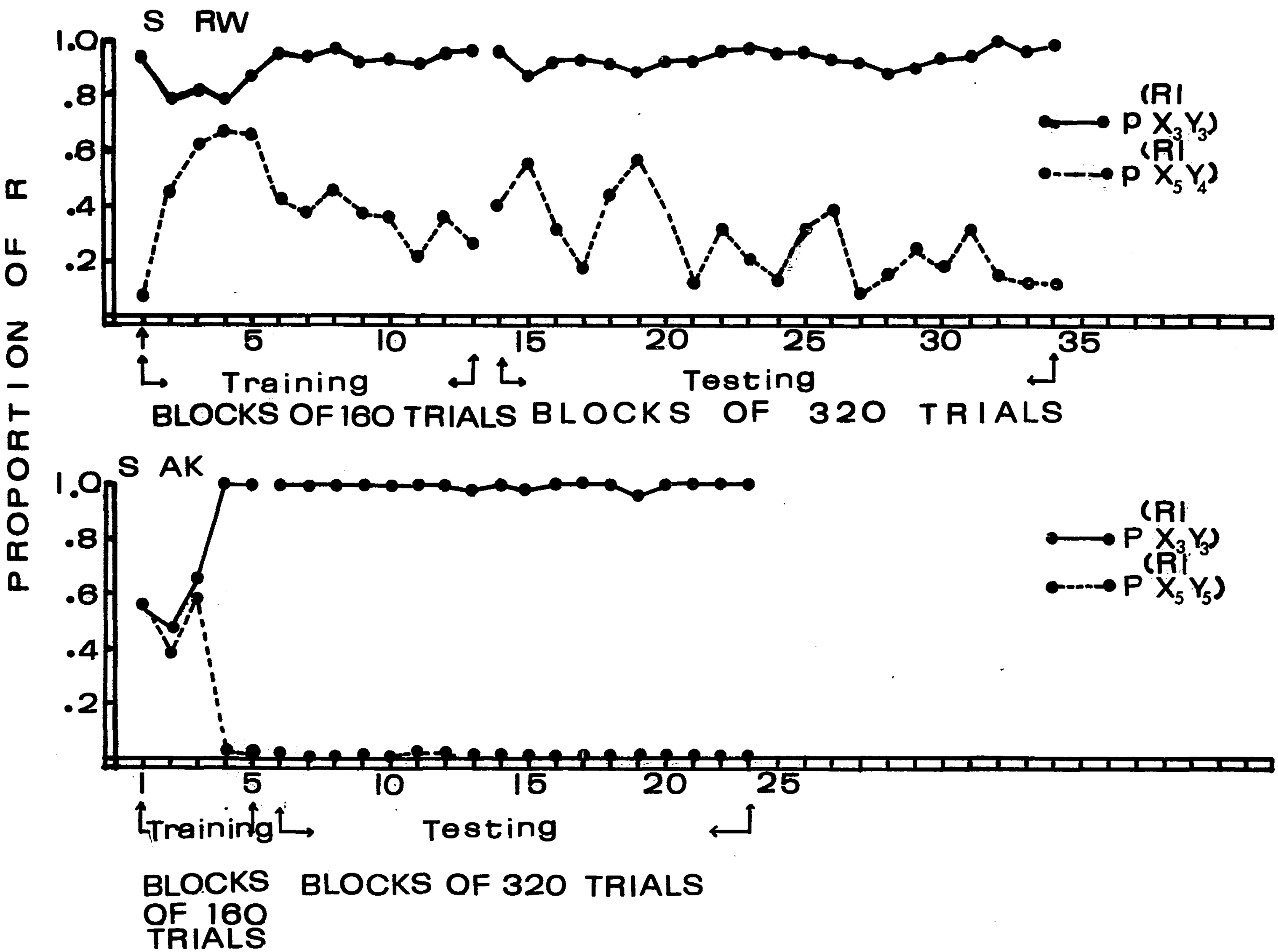
On the basis of each pair of conditional probabilities, a value of d' was computed. The results in d' during the training and testing phases are displayed in Figure 3 for each <u>S</u>. The value of d' in psychophysics is a measure of sensory sensitivity and is assumed to be independent of non-sensory variables. The greater d' the greater is the sensitivity of the sensory system. It can be seen from Figure 3 that the values of d' for all <u>S</u>s except AK fluctuate between 0 and 3. <u>S</u> AK, starting from the last block of training, shows an almost perfect discrimination throughout the rest of the experimental sessions.

It can also be seen from the figure that the introduction of testing stimuli alters the discriminative performance, although little consistent pattern was found among <u>S</u>s. The d's computed from the training data (excluding the first block) and the testing data are presented in Table 2.

<u>Generalization</u>: The results of generalization testing for each <u>S</u>

Figure 2. Discriminative performance during training and testing measured in terms of the proportion of $p(R \mid X_3, Y_3)$ and $p(R \mid X_5, Y_5)$ as a function of trial blocks. Each graph displays the performance of an individual S. The upper curve in each graph represents the estimate of $p(R \mid X_3, Y_3)$ and lower curve represents the estimate of $p(R \mid X_5, Y_5)$


S JZ
1.0
.8
.6
.4
.2
PROPORTION OF R
1
5
10
15
20
25
30
35
40
Training
Testing
BLOCKS OF 160 TRIALS
P(R I X3Y3)
P(R I X5Y5)
S AG
1.0
.8
.6
.4
.2
1
5
10
15
20
25
30
Training
Testing
BLOCKS OF 160 TRIALS
BLOCKS OF 320 TRIALS
P(R I X3Y3)
P(R I X5Y4)



S RW
PROPORTION OF R
1.0
.8
.6
.4
.2
(RI
p X3 Y3)
(RI
p X5 Y4)
5
10
15
20
25
30
35
Training
Testing
BLOCKS OF 160 TRIALS
BLOCKS OF 320 TRIALS
S AK
1.0
.8
.6
.4
.2
(RI
P X3 Y3)
(RI
P X5 Y5)
1
5
10
15
20
25
Training
Testing
BLOCKS OF 160 TRIALS
BLOCKS OF 320 TRIALS

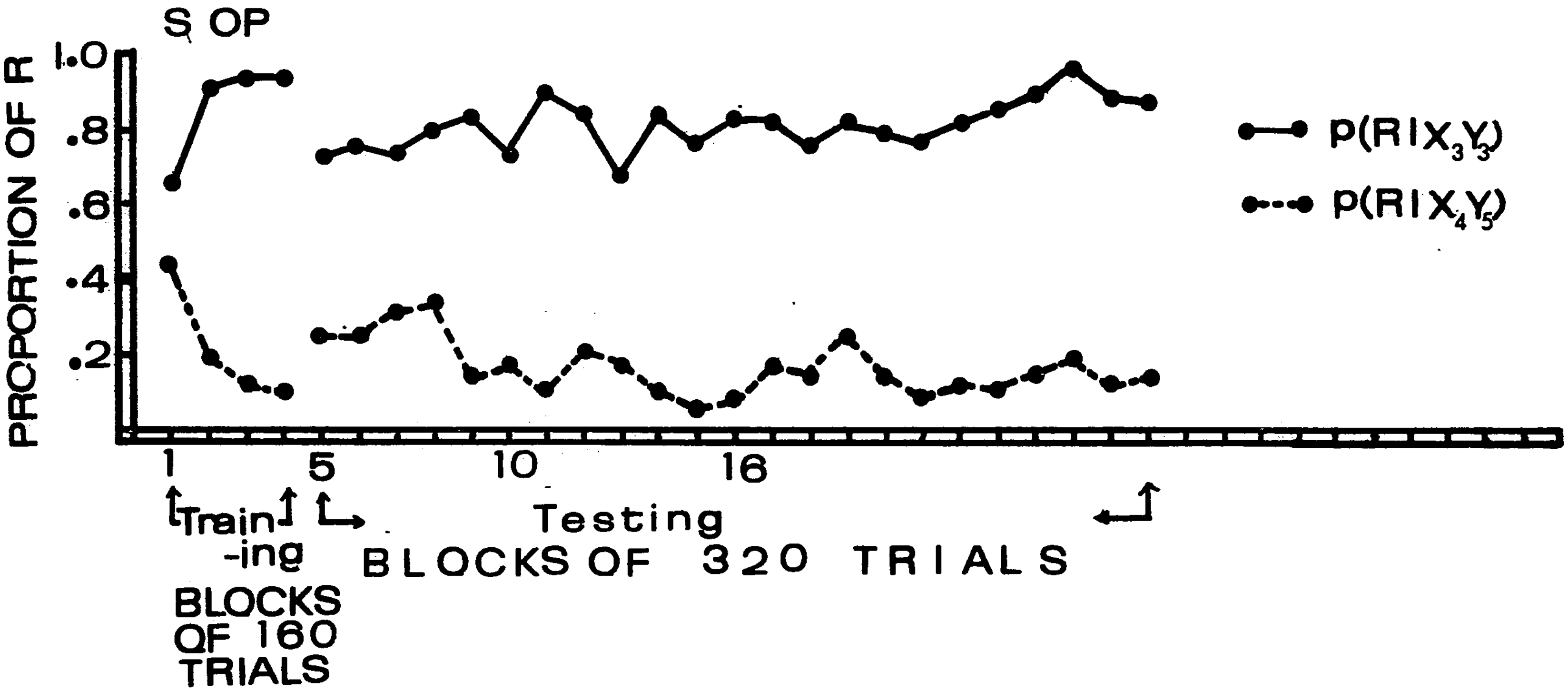
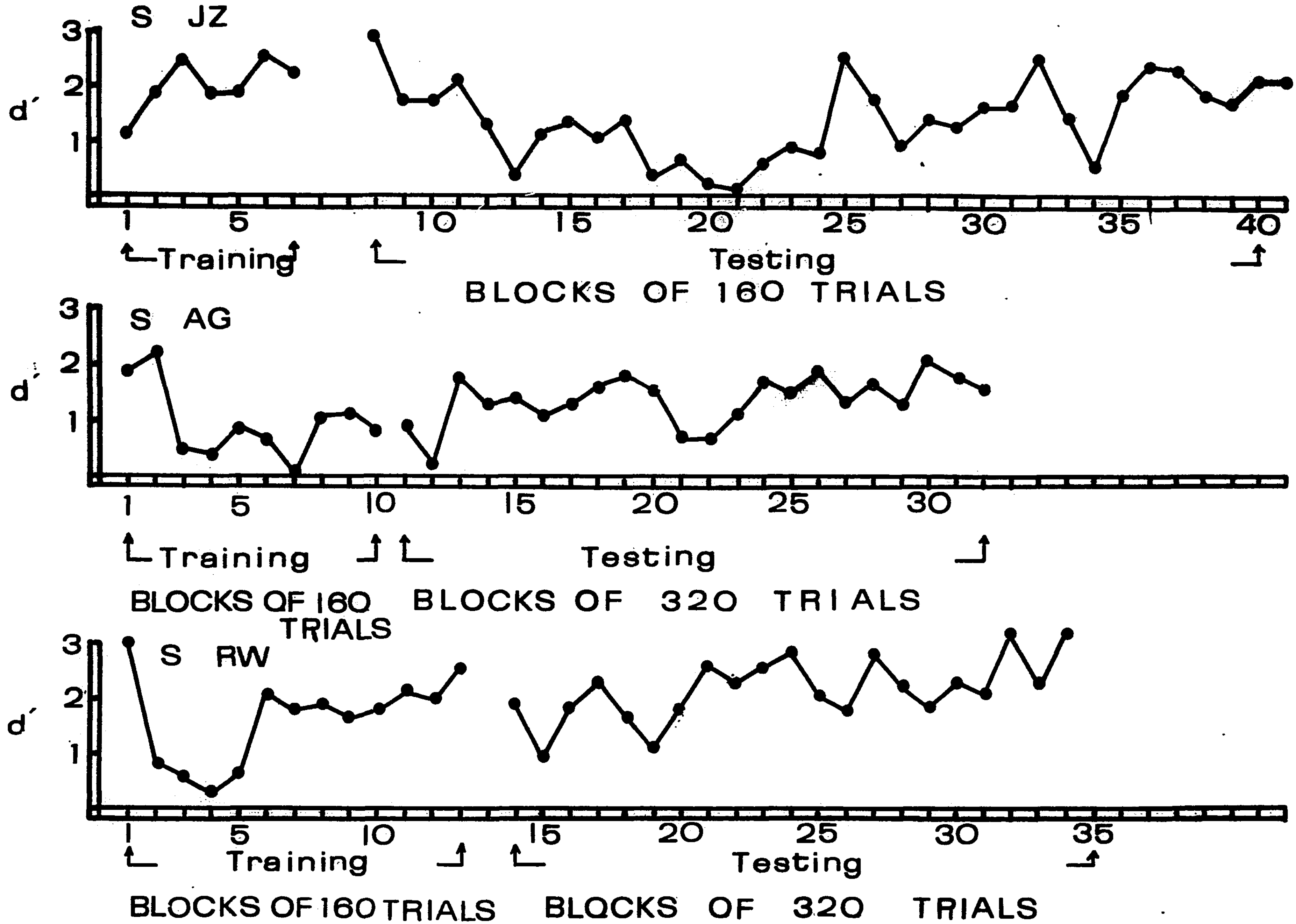

S OP
PROPORTION OF R
1.0
.8
.6
.4
.2
1
5
10
16
Training
Testing
BLOCKS OF 320 TRIALS
BLOCKS OF 160 TRIALS
p(R|X3Y3)
p(R|X4Y5)

Figure 3. Discriminative performance during training and testing measured in terms of d', as a function of trial blocks. Each graph in the figure represents the performance of an individual S.


S JZ
d′
3
2
1
1
5
10
15
20
25
30
35
40
Training
Testing
BLOCKS OF 160 TRIALS
S AG
d′
3
2
1
1
5
10
15
20
25
30
Training
Testing
BLOCKS OF 160 TRIALS
BLOCKS OF 320 TRIALS
S RW
d′
3
2
1
1
5
10
15
20
25
30
35
Training
Testing
BLOCKS OF 160 TRIALS
BLOCKS OF 320 TRIALS

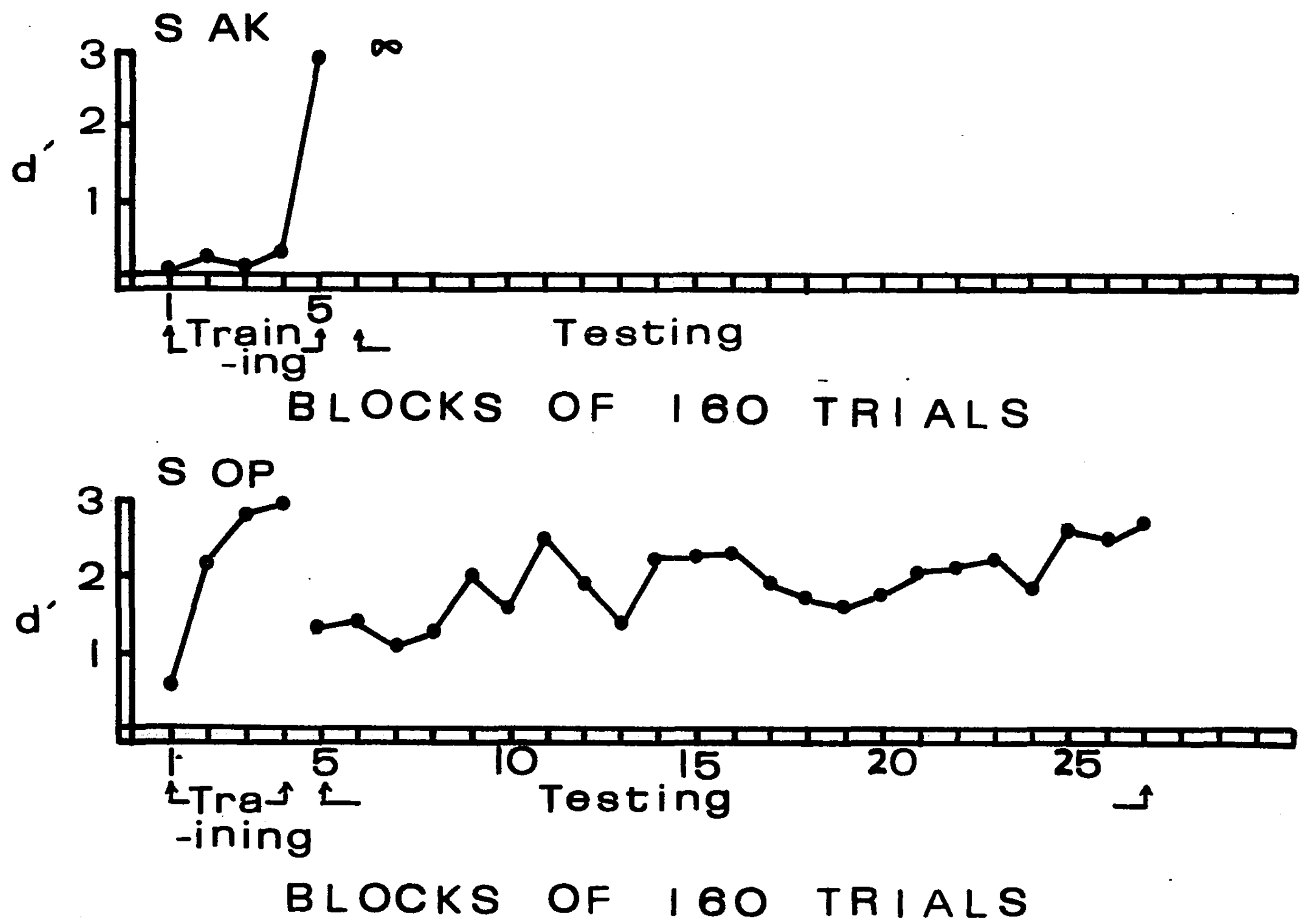

S AK
3
2
1
d´
∞
1
5
Training
Testing
BLOCKS OF 160 TRIALS
S OP
3
2
1
d´
1
5
10
15
20
25
Training
Testing
BLOCKS OF 160 TRIALS

Table 2. The average d's during training (excluding the first block) and generalization testing.

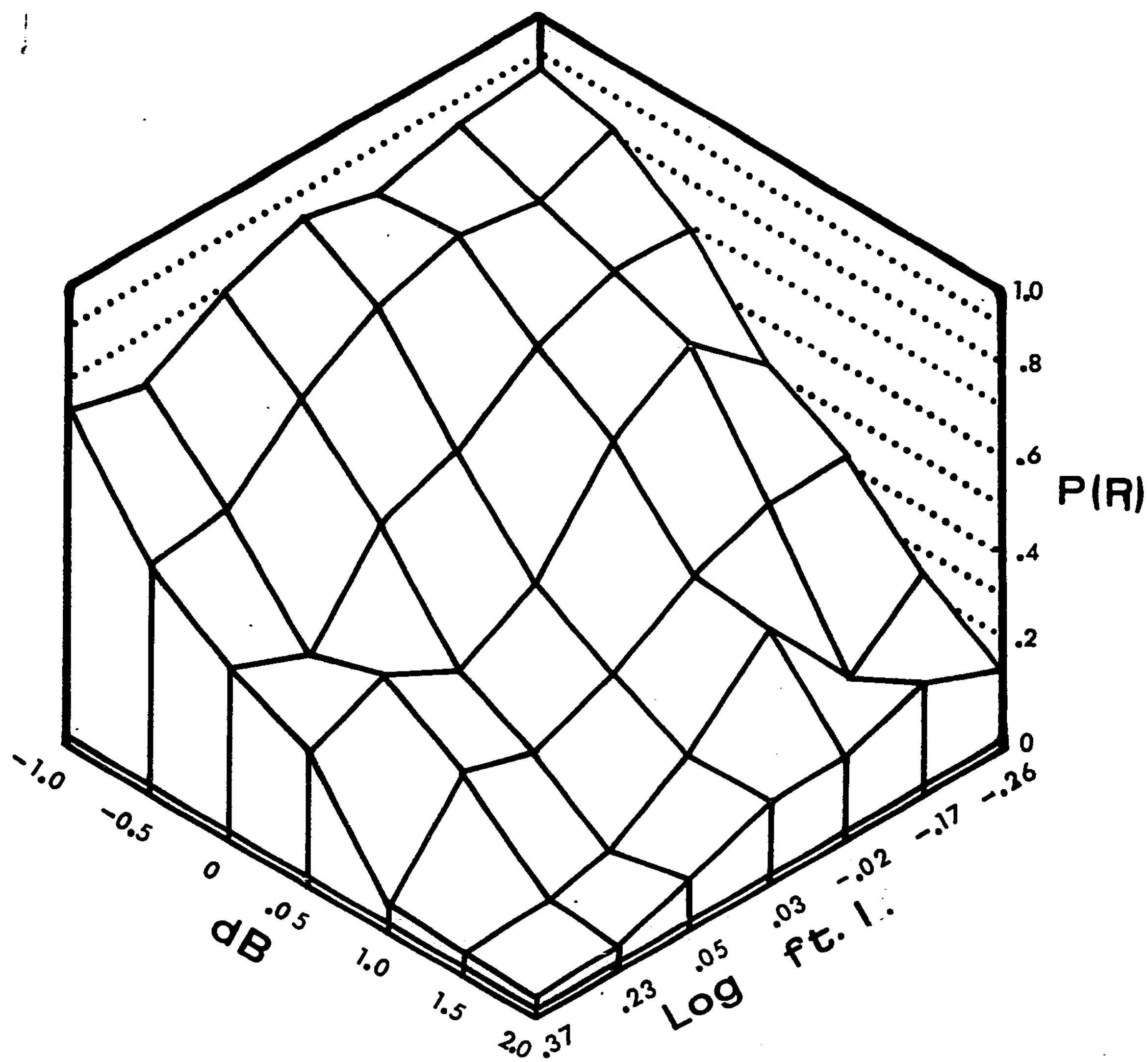
Subject	Training	Testing
JZ	2.16	1.41
AG	0.67	1.38
RW	1.54	2.21
AK	0.71	∞
OP	2.63	1.89

are shown in Figure 4. The graphs in Figure 4 are the three-dimensional representations of a generalization surface, showing how the response proportion of the R response, p(R), varies with the intensities of both sound and light, where p(R) for each testing stimulus is based on 65 observations. The total number of observations on training stimuli varies among Ss. For Ss JZ, AK and OP, each training stimulus was presented 1082 times. For S AG, stimulus (X_5,Y_4) was presented three times as often as (X_3,Y_3), that is, (X_5,Y_4) was presented 1668 times and (X_3,Y_3) 561 times. On the other hand, for S RW, the presentation frequencies were 561 and 1668, respectively, for the stimuli (X_5,Y_4) and (X_3,Y_3).

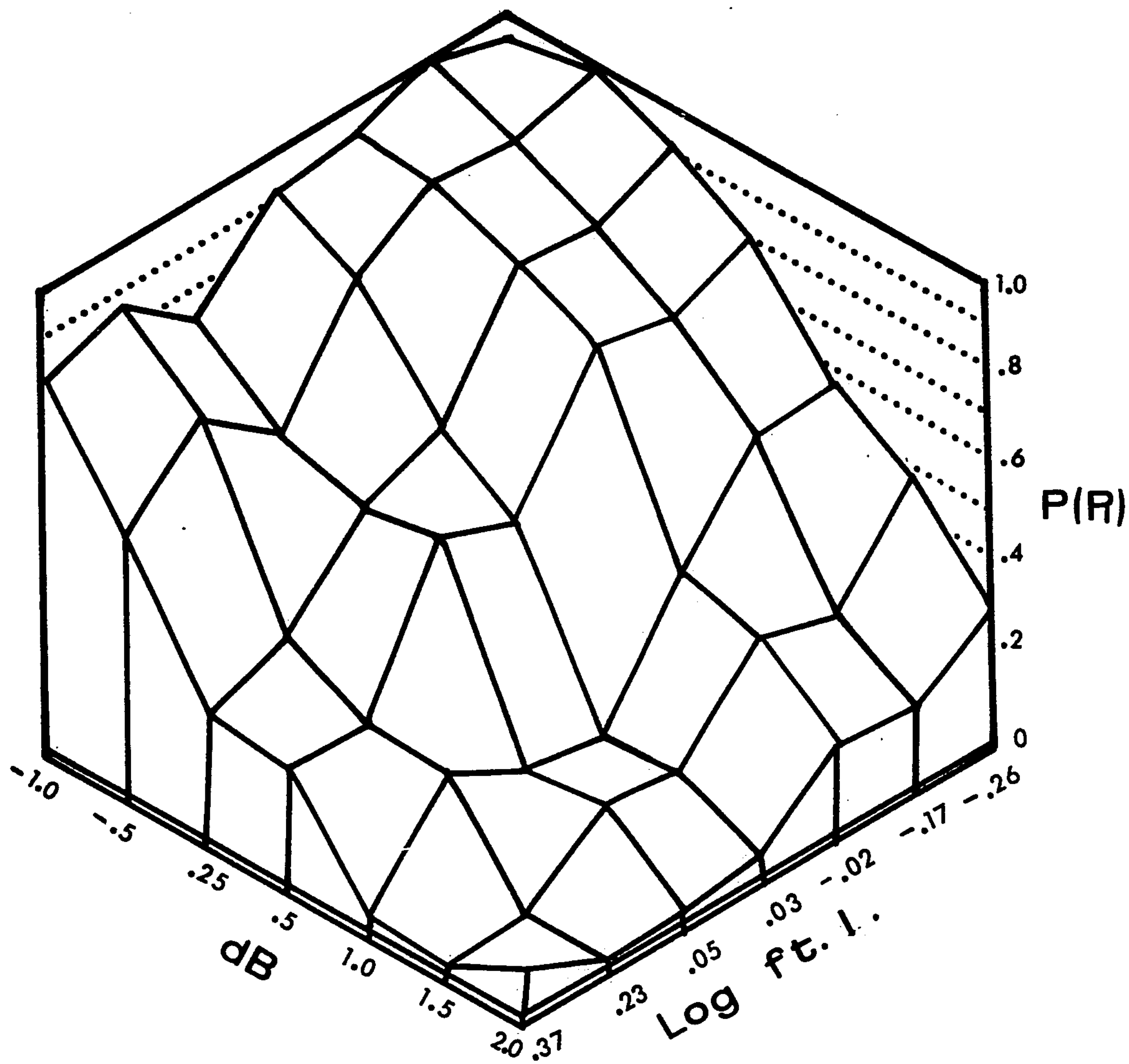
It can be seen from the figure that the generalization surfaces for the first three Ss generally display similar characteristics -- as the intensity levels of either sound or light increase p(R) decreases. The highest value of p(R) is associated with the combination of the weakest sound and weakest light, not with the positive training stimulus, while the lowest is associated with the combination of the strongest sound and the strongest light. Inspection of these three graphs suggests that the slopes of the gradients along the sound dimension are different from those along the light dimension. The generalization surfaces of the last two Ss display a slightly different form. S AK, it may be recalled, was trained under the condition of 'smaller sound difference' while S OP was under 'smaller light difference'. The results were that S AK who was subjected to 'smaller sound difference' discrimination training showed gradients along the sound dimension that are almost flat, indicating that the dimension of sound exerted very little control

Figure 4. Generalization surfaces. The horizontal axes indicate the various intensities of light and sound. The vertical axis indicates the proportion of response R. Each graph represents an individual S.

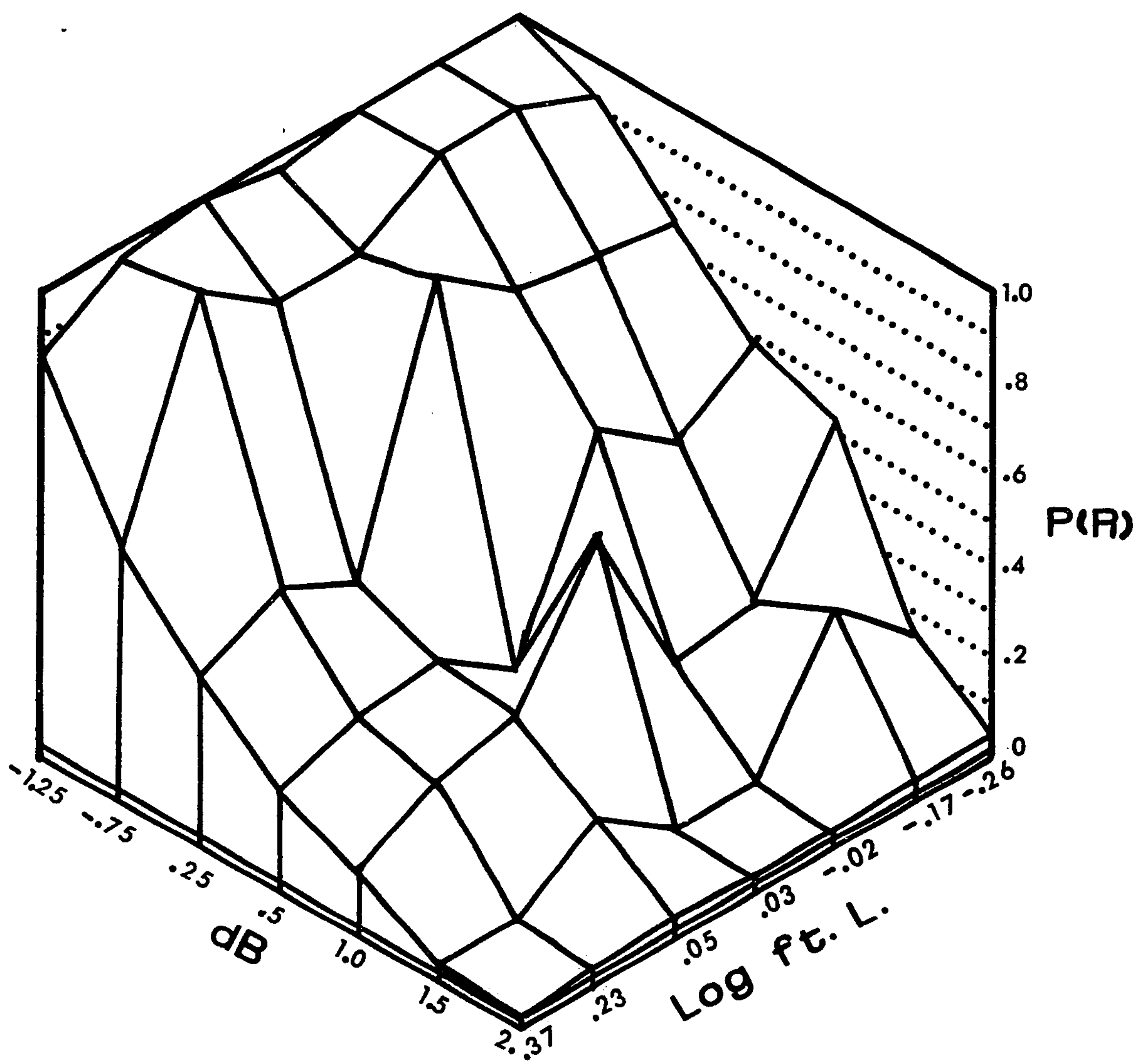
S JZ

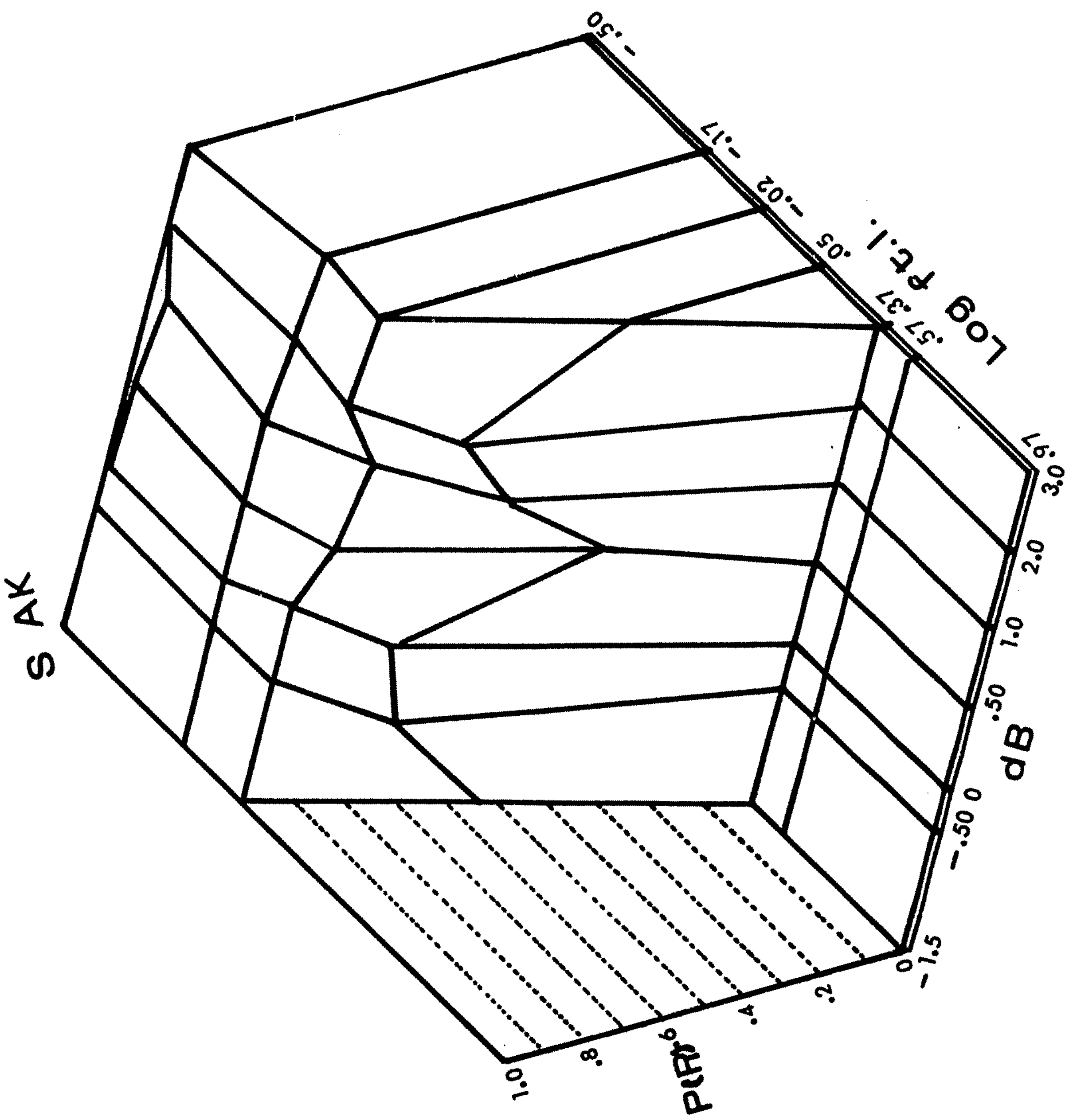


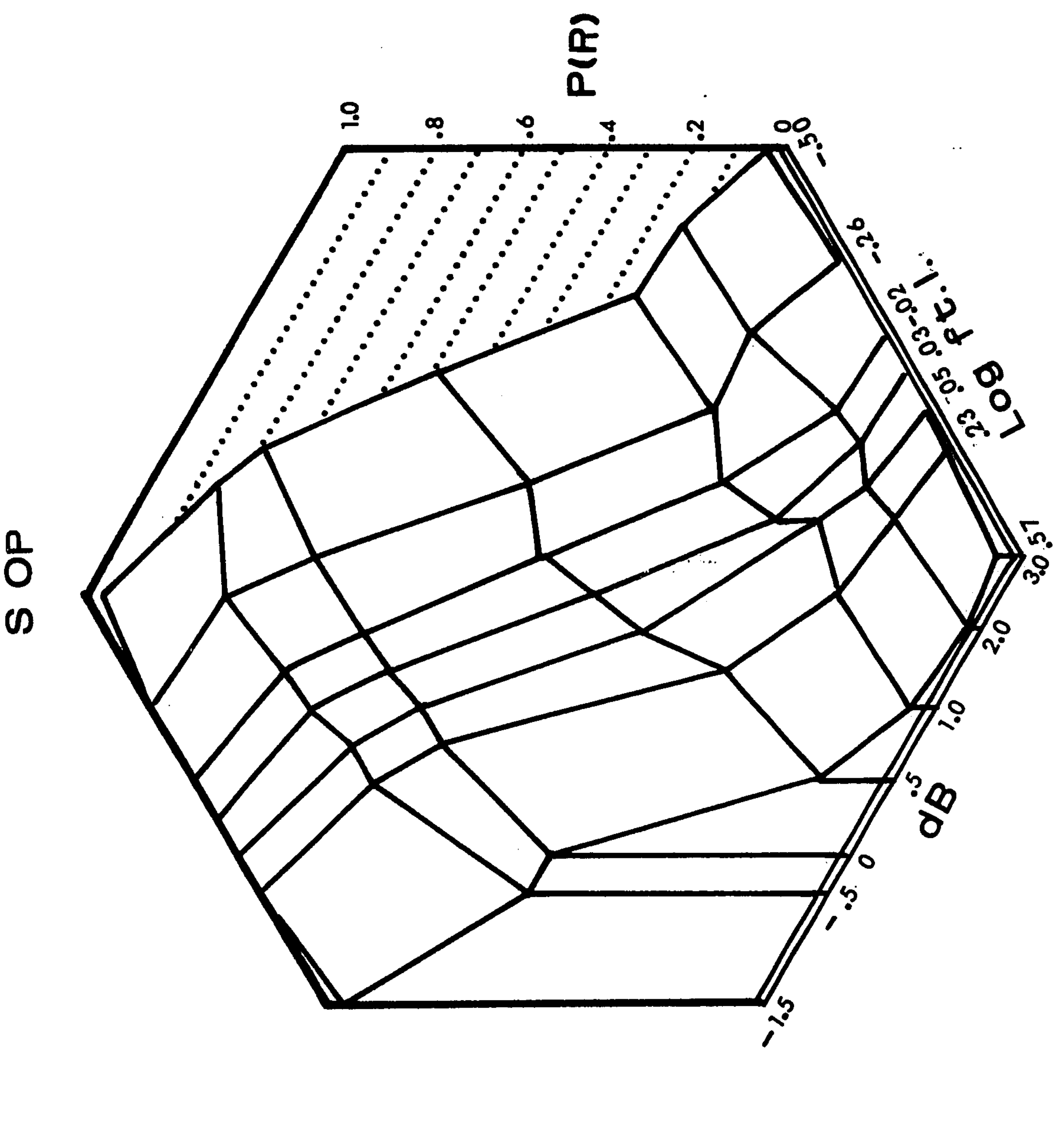
S AG

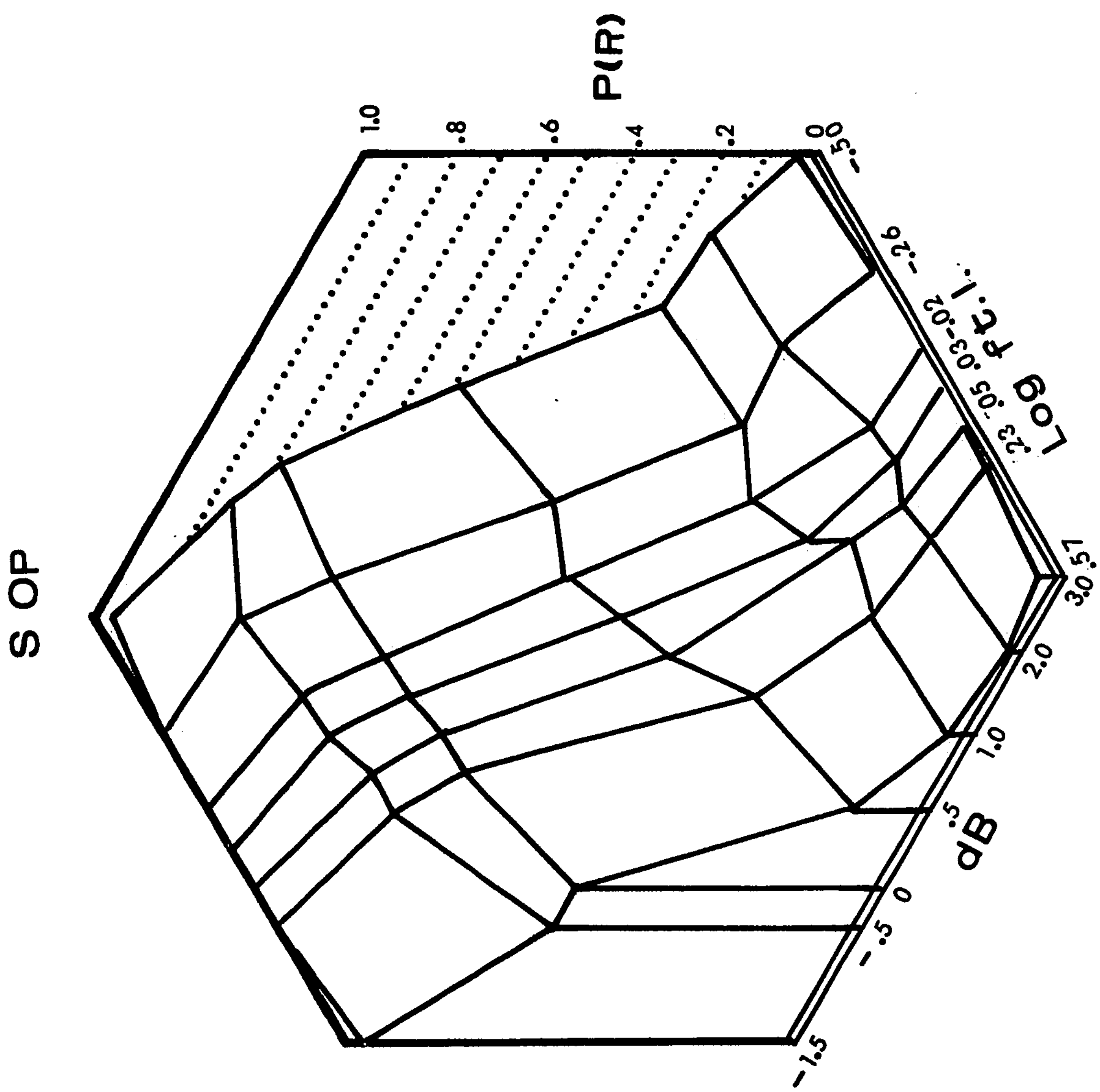


S RW




S AK
Log ft.l.
-.50
-.17
-.02
.05
.37
.57
.97
dB
3.0
2.0
1.0
.50
0
-.50
-1.5
P(R)
1.0
.8
.6
.4
.2
0


S OP
P(R)
1.0
.8
.6
.4
.2
0
-.50
-.26
-.02
.03
.05
.23
.57
Log ft. L.
-1.5
-.5
0
.5
1.0
2.0
3.0
dB


S OP
P(R)
1.0
.8
.6
.4
.2
0
-.50
-.26
-.02
.03
.05
.23
.51
Log ft-L
3.0
2.0
1.0
.5
0
-.5
-1.5
dB

over responding. On the other hand, 'smaller light difference' discrimination training produced little control over responding by the light dimension.

In order to assess the relative degree of control exercised by the sound and light dimensions the slopes of the gradients for given sound or light intensity levels based on logarithmic scales were calculated by the method of probit analysis (Finney, 1964): Each p(R) was first transformed into a probit by the table given by Bliss (1935), excluding those probability values greater than 0.95 and less than 0.05. Then for a given sound or light intensity, a linear probit regression line was fitted by the least squares method against seven light or sound intensity levels in logarithmic units. The obtained slope of the regression line was used as an approximation to the slope of the gradient. These estimated values of slope are summarized in Table 3.

It is clear from Table 3 that the average slope of the gradients along the light dimension was steeper than that along the sound dimension for Ss JZ, AG, and RW, indicating that the light dimension gained more control than the sound dimension at least on logarithmic scale. However, it is still not conclusive unless the proper scales for both dimensions are known or their respective intensities matched experimentally. For S AK, who was trained under the condition of 'smaller sound difference', the average slope of the gradients along the sound dimension was almost zero, while very steep slopes were obtained for gradients along the light dimension. The results of S AK suggest that responding was exclusively under the control of

Table 3. The estimated slope (probit/log) of gradient with respect to response R for given intensity of sound and light.

S JZ

Given Light Intensity Level(log ft.l.)	-0.26	-0.17	-0.02	0.03	0.05	0.23	0.37	Mean
Slope of Sound Gradient	-0.75	-0.73	-0.60	-0.68	-0.77	-0.69	-0.87	-0.73
Given Sound Intensity Level(dB)	-1.0	-0.5	0	0.5	1.0	1.5	2.0	Mean
Slope of Light Gradient	-1.25	-2.14	-2.11	-1.64	-0.83	-1.94	-0.85	-1.54

S AG

Given Light Intensity Level(log ft.l.)	-0.26	-0.17	-0.02	0.03	0.05	0.23	0.37	Mean
Slope of Sound Gradient	-0.81	-0.95	-0.91	-1.00	-0.69	-0.87	-0.75	-0.85
Given Sound Intensity Level(dB)	-1.0	-0.5	0.25	0.5	1.0	1.5	2.0	Mean
Slope of Light Gradient	-1.23	-2.66	-3.59	-3.03	-3.11	-2.69	-1.12	-2.49

S RW

Given Light Intensity Level(log ft.l.)	-0.26	-0.17	-0.02	0.03	0.05	0.23	0.37	Mean
Slope of Sound Gradient	-1.20	-1.00	-1.88	-1.39	-1.01	-0187	-0.95	-1.18
Given Sound Intensity Level(dB)	-1.0	-0.5	0.25	0.5	1.0	1.5	2.0	Mean
Slope of Light Gradient	-7.14	-4.26	-2.16	-1.90	-1.45	-1.50	0.0	-2.77

Table 3. The estimated slope (probit/log) of gradient with respect to response R for given intensity of sound and light (continued).

S: AK

Given Light Intensity Level(log ft.l.)	-0.51	-0.17	-0.02	0.05	0.37	0.57	0.97	Mean
Slope of Sound Gradient	0	0	0	0.07	0	0	0	0
Given Sound Intensity Level(dB)	-1.50	-0.50	0	0.50	1.00	2.00	3.00	Mean
Slope of Light Gradient	-3.86	-11.00	-5.07	-15.7	-36.00	-16.80	-51.10	-19.93

S: OP

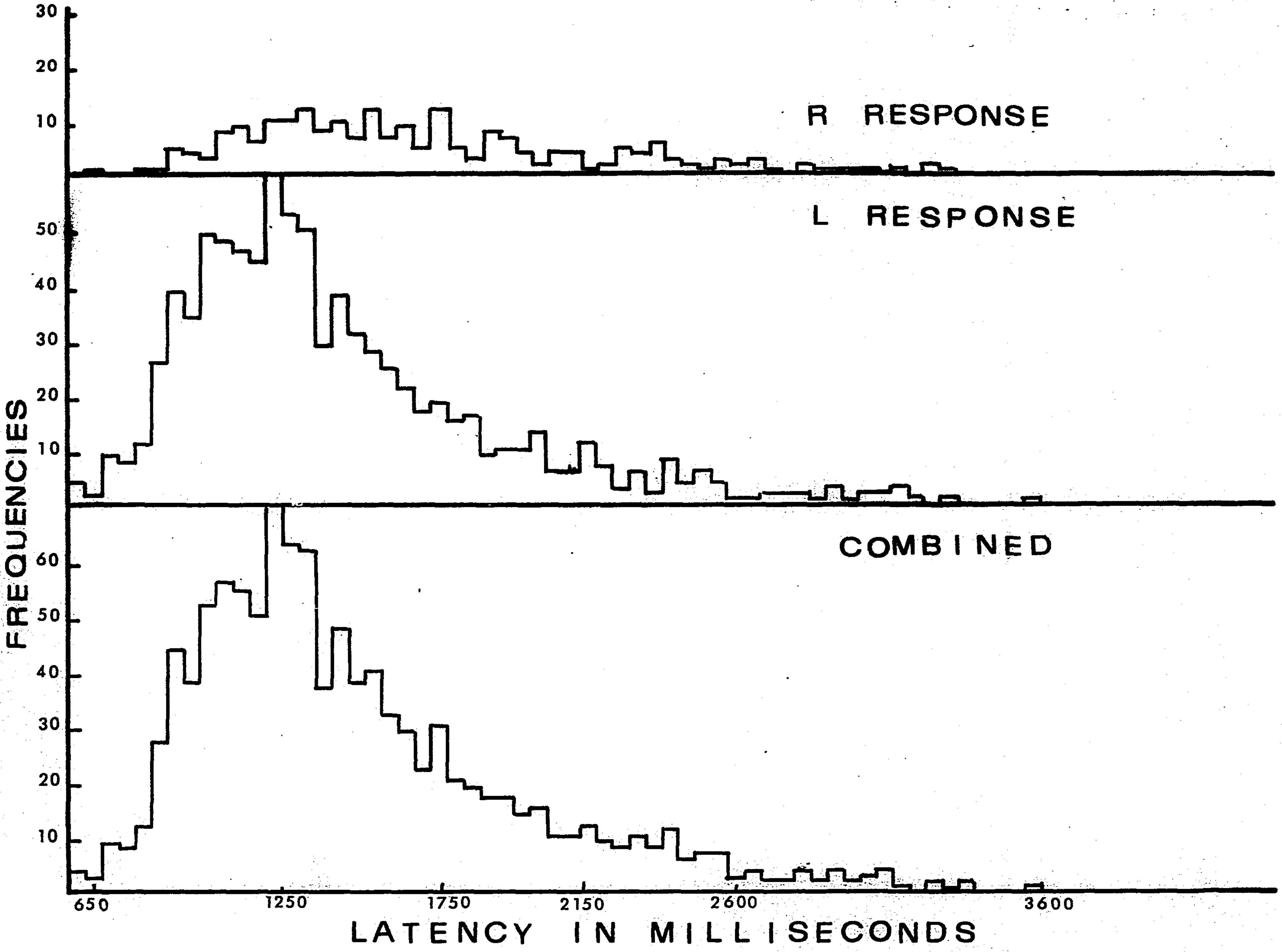
Given Light Intensity Level(log ft.l.)	-0.51	-0.26	-0.02	0.03	0.05	0.23	0.57	Mean
Slope of Sound Gradient	-1.13	-1.16	-1.63	-1.88	-1.57	- 2.06	-1.47	-1.76
Given Sound Intensity Level(dB)	-1.50	-0.50	0	1.0	1.50	2.0	3.0	Mean
Slope of Light Gradient	0.05	-0.70	-0.34	-0.96	-0.34	-1.11	-0.43	

the light dimension. On the other hand, S OP who was trained under the condition of 'smaller light difference' revealed the opposite trend, i.e., responding was mainly controlled by the sound dimension.

Choice reaction time (CRT): For each stimulus, the frequency distribution of CRT, like most response times, has a 'high tail'. A typical frequency distribution of CRTs for the stimulus (X_5, Y_5), S JZ, is shown in Figure 5. Median CRT associated with response R, L and with both responses combined were computed and are presented in Figure 6 for each of the five Ss (see also Appendix 2). The changes in p(R) are also shown in the same figure. In the left panel of each figure, gradients are plotted with sound intensity as the variable for each value of light intensity, while in the right panel, the same data are plotted with light intensity varying. Incomplete CRT gradients on these graphs indicate that the number of observations for these stimuli was too small ($N < 5$) to compute reliable median CRTs.

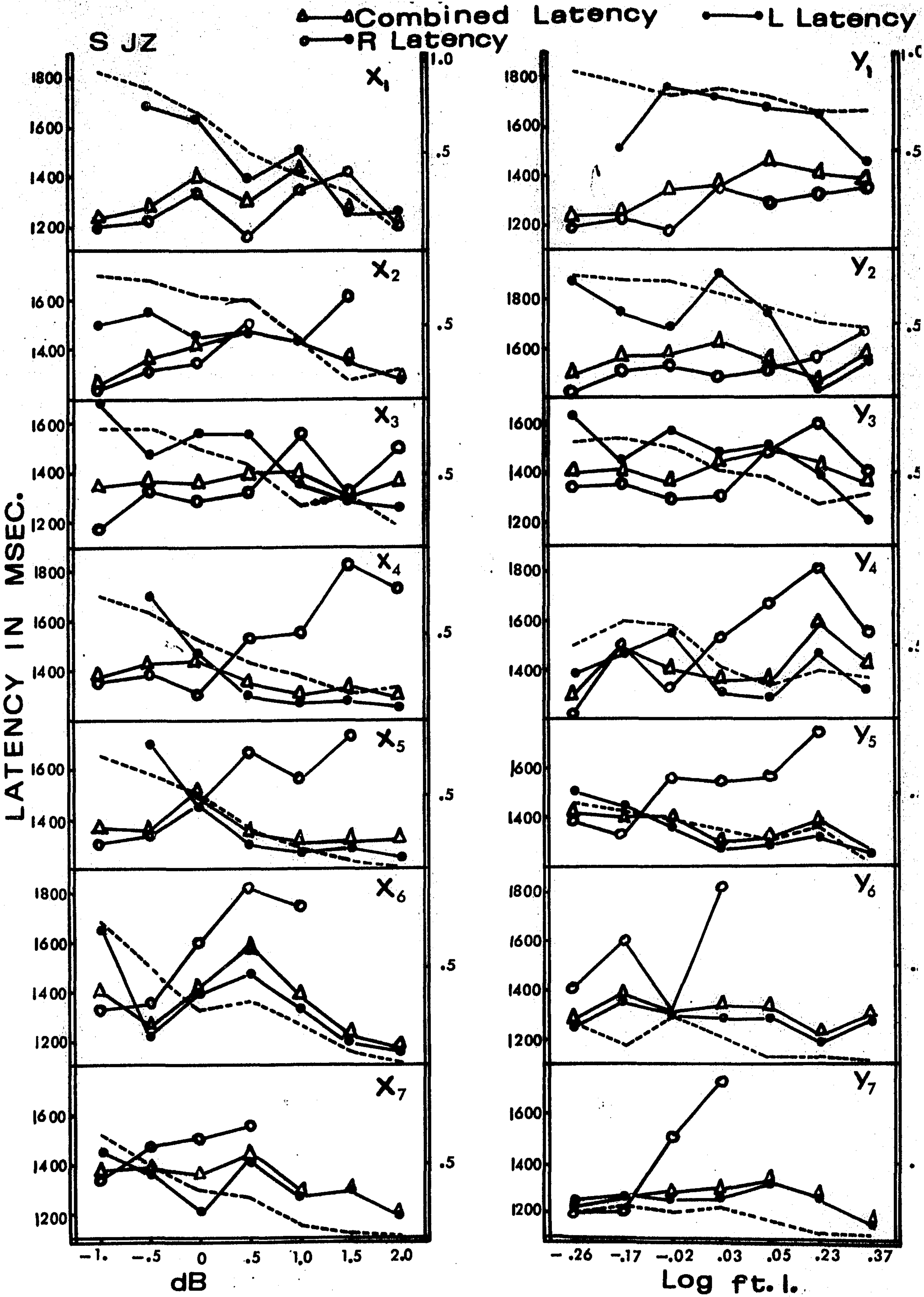
Inspection of these CRT gradients suggests that, for response R, the CRT increases with stimulus intensity, and for response L the trend is in the opposite direction. However, for Ss OP and AK, CRT gradients along those dimensions which do not exert any control over responding are almost flat. It is also interesting to note that the combined CRT gradients of Ss JZ, AG and RW have their peaks gradually shifted to the left as the intensity of either sound or light increased. For SsAK and OP whose responding was mostly controlled by a single dimension, the peaks were located approximately in the middle of the stimulus continua, and unaffected by variation

Figure 5. The frequency distribution of CRT for the stimulus (X_5, Y_5), S JZ. The distribution for response R is shown on the top panel and for response L in the middle and the combined frequency distribution on the bottom. The distribution was depicted to illustrate the 'high tail' characteristics of the distribution.

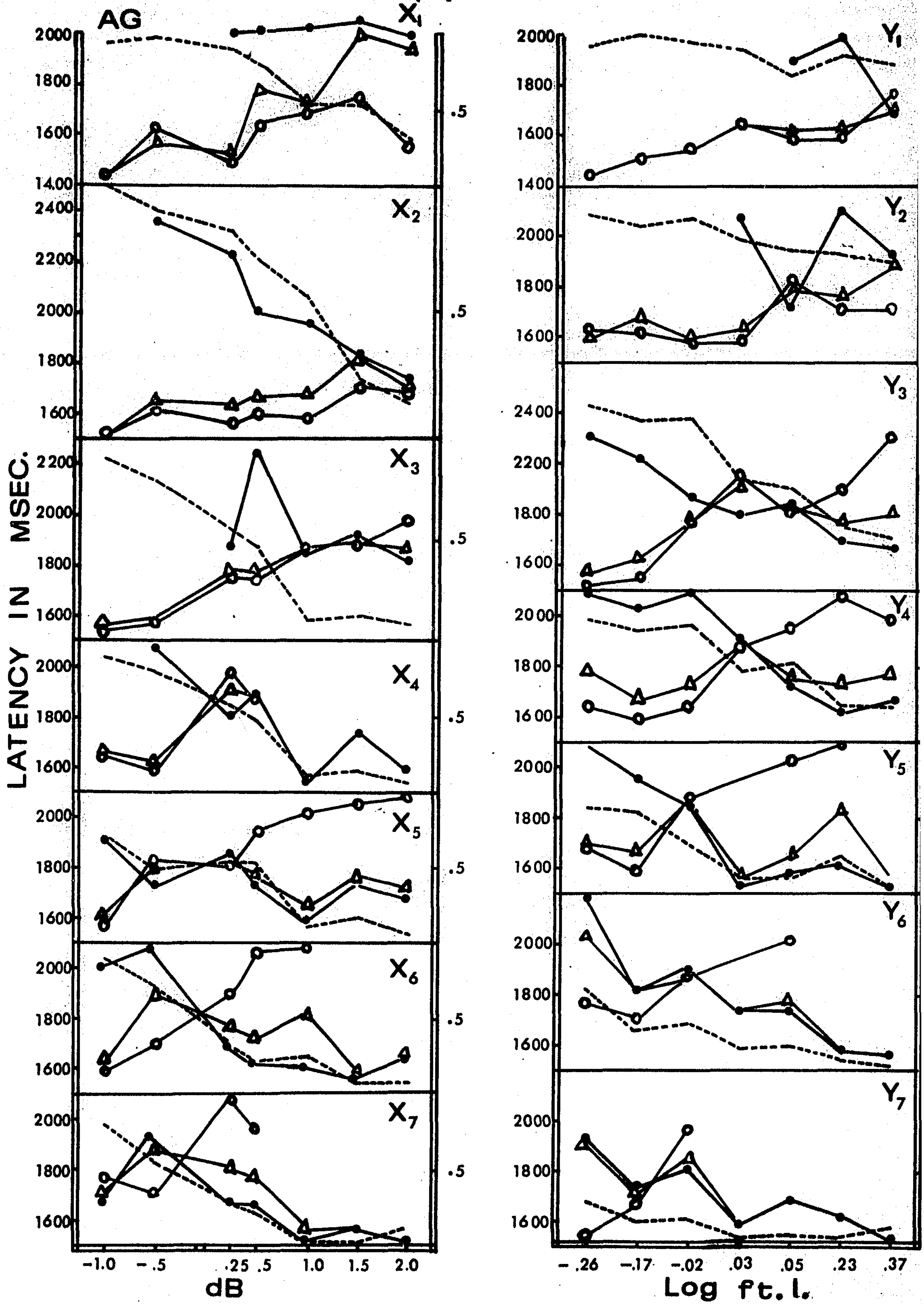


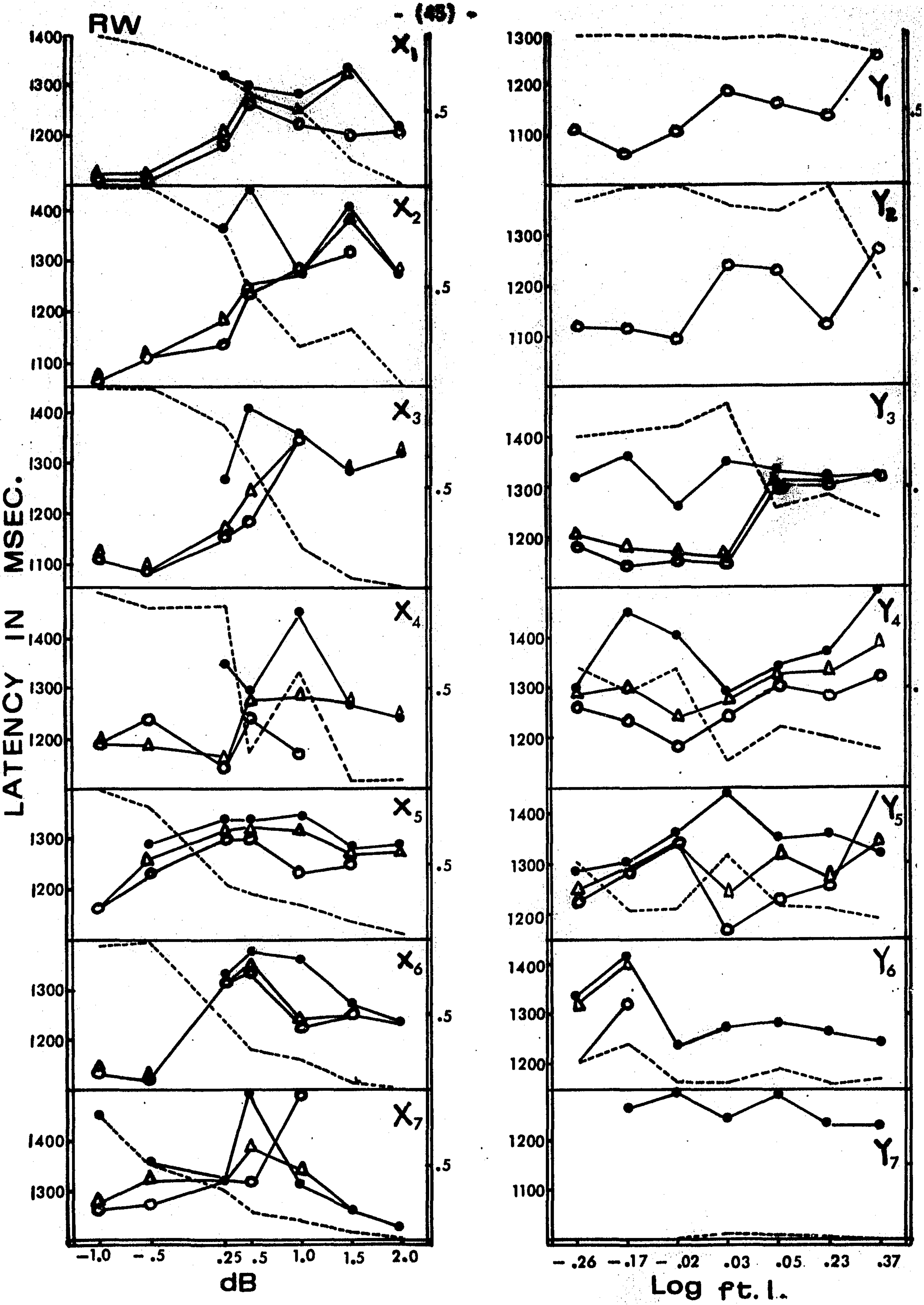
30
20
10
R RESPONSE
L RESPONSE
50
40
30
20
10
FREQUENCIES
COMBINED
60
50
40
30
20
10
650
1250
1750
2150
2600
3600
LATENCY IN MILLISECONDS

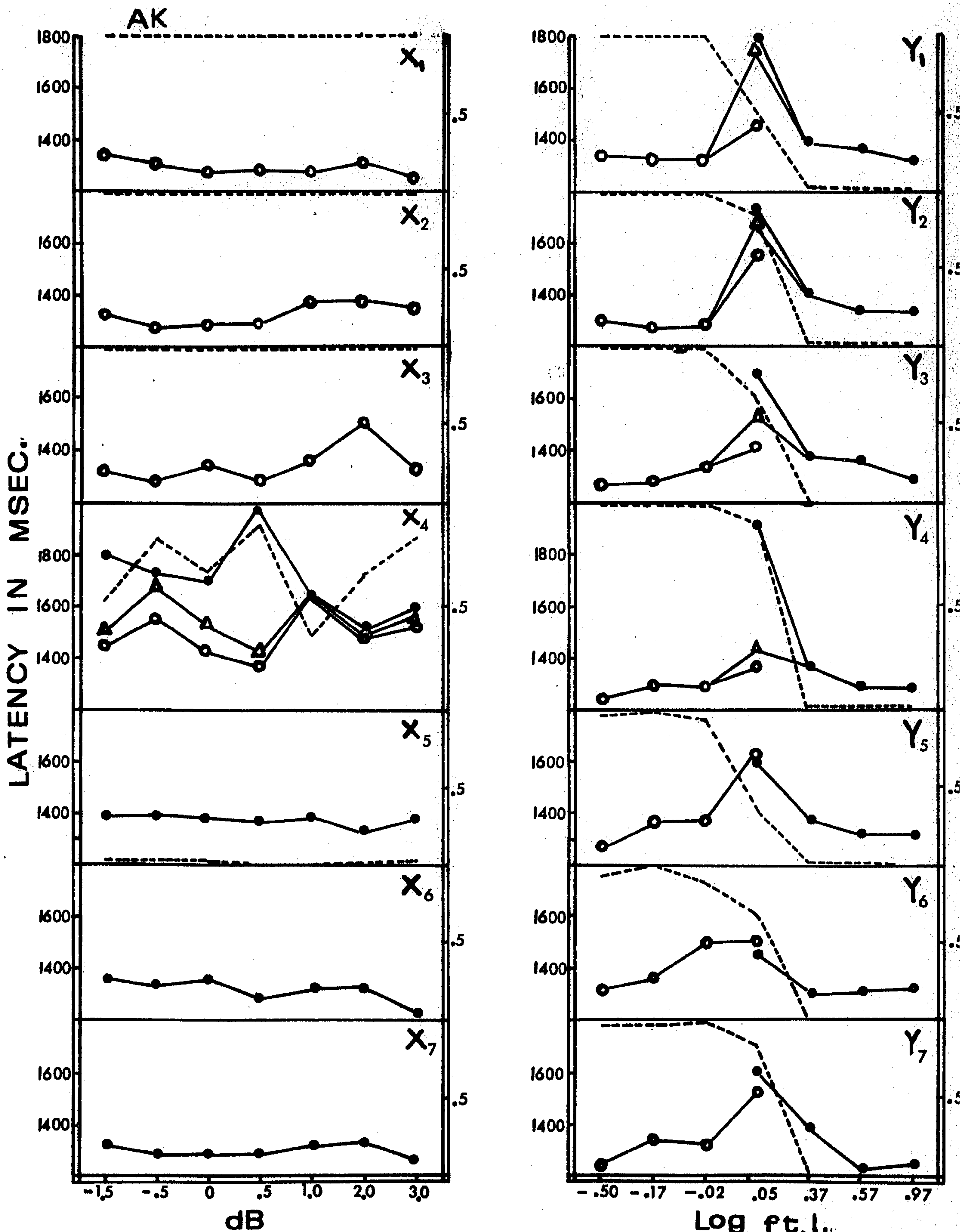
Figure 6. The gradients of median choice reaction time at different intensity levels of light and sound for all *Ss*. The graphs on the left and right panels are based on same data. The curve in the figure show the gradiens associated with responses R, L and both combined. The gradients of response probability T (dash line) are also reproduced here.

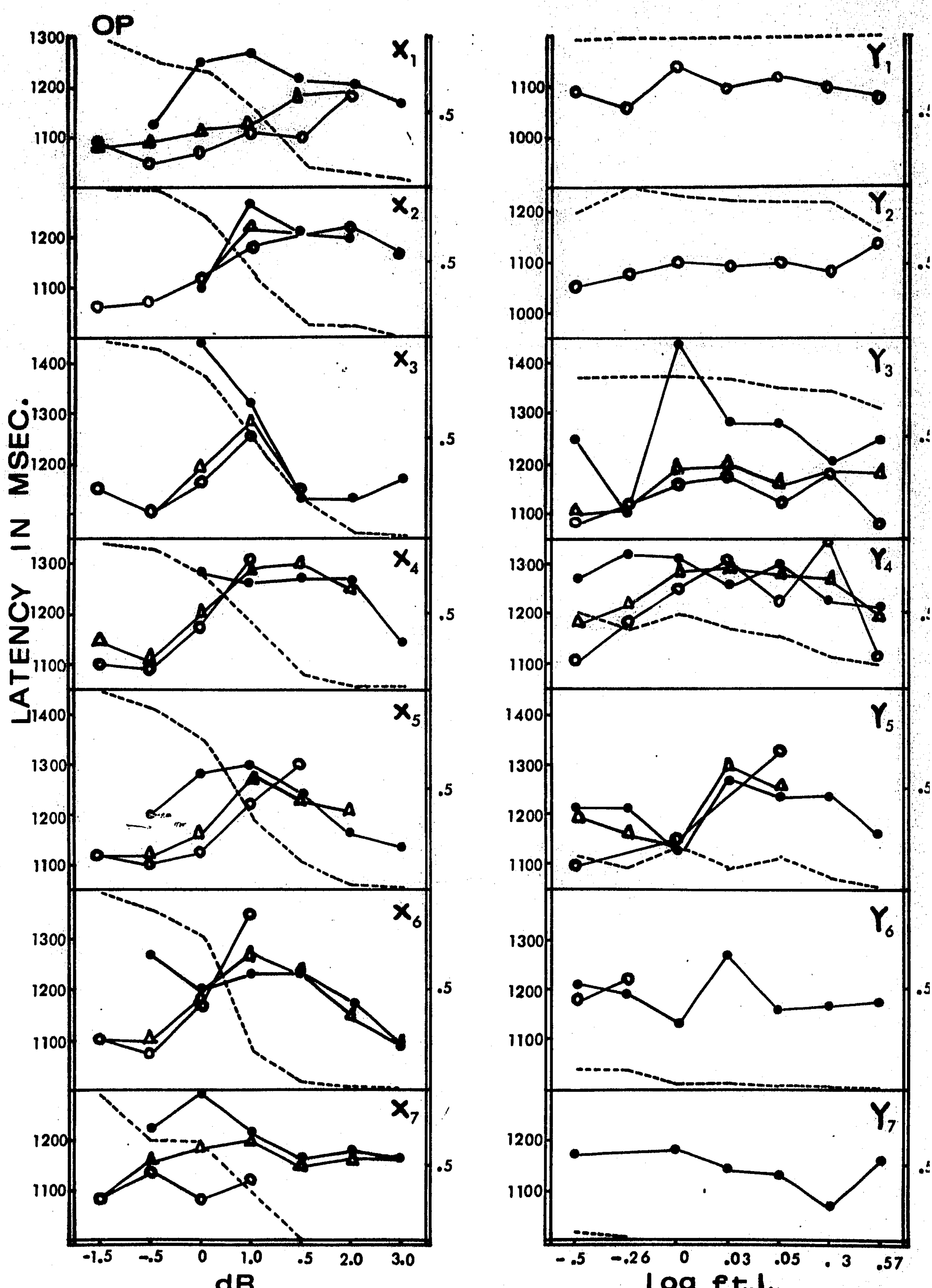


Combined Latency
L Latency
R Latency
S JZ
LATENCY IN MSEC.
X1
X2
X3
X4
X5
X6
X7
Y1
Y2
Y3
Y4
Y5
Y6
Y7
1800
1600
1400
1200
1.0
.5
-1.
-.5
0
.5
1.0
1.5
2.0
dB
-.26
-.17
-.02
.03
.05
.23
.37
Log ft. l.


AG
X_1
X_2
X_3
X_4
X_5
X_6
X_7
Y_1
Y_2
Y_3
Y_4
Y_5
Y_6
Y_7
LATENCY IN MSEC.
dB
Log ft. l.
-1.0
-.5
.25
.5
1.0
1.5
2.0
-.26
-.17
-.02
.03
.05
.23
.37


RW
LATENCY IN MSEC.
X1
X2
X3
X4
X5
X6
X7
Y1
Y2
Y3
Y4
Y5
Y6
Y7
1400
1300
1200
1100
.5
−1.0
−.5
.25
.5
1.0
1.5
2.0
dB
−.26
−.17
−.02
.03
.05
.23
.37
Log ft. l.


AK
LATENCY IN MSEC.
1800
1600
1400
.5
X1
X2
X3
X4
X5
X6
X7
Y1
Y2
Y3
Y4
Y5
Y6
Y7
-1.5
-.5
0
.5
1.0
2.0
3.0
dB
-.50
-.17
-.02
.05
.37
.57
.97
Log ft.L.


OP
LATENCY IN MSEC.
X1
X2
X3
X4
X5
X6
X7
Y1
Y2
Y3
Y4
Y5
Y6
Y7
-1.5
-.5
0
1.0
.5
2.0
3.0
dB
-.5
-.26
0
.03
.05
.3
.57
log ft.L.

of stimulus intensity.

For comparison purposes, response gradients of R are also reproduced in Figure 6, in dashed line. It can be seen that the median CRT and response gradient are inversely related: high response probability is associated with short CRT and vice versa.

This inverse relationship is more clearly shown by plotting the median CRT to each stimulus against its response probability. Such probability-CRT function for all Ss are presented in Figure 7. The solid lines in the figure were obtained by fitting second order polynomials to the data using the least squares method. It can be seen in Figure 7 that the slopes of all the probability-CRT function are negative, and that the functions closely approximate linear ones. The values of curvilinear correlation for five Ss are as follows: S JZ: -0.550; S AG: -0.716; S RW: -0.732; S AK: -0.763; S OP: -0.631.

Theoretical analysis

A. Prediction of the two-dimensional gradients from the unidimensional ones:

It was shown previously that the decision contour partitioning the (x,y) plane is a linear function with slope equal to

$$[(\mu_{X_5} - \mu_{X_3})\sigma_Y^2]/[(\mu_{Y_5} - \mu_{Y_3})\sigma_X^2]$$

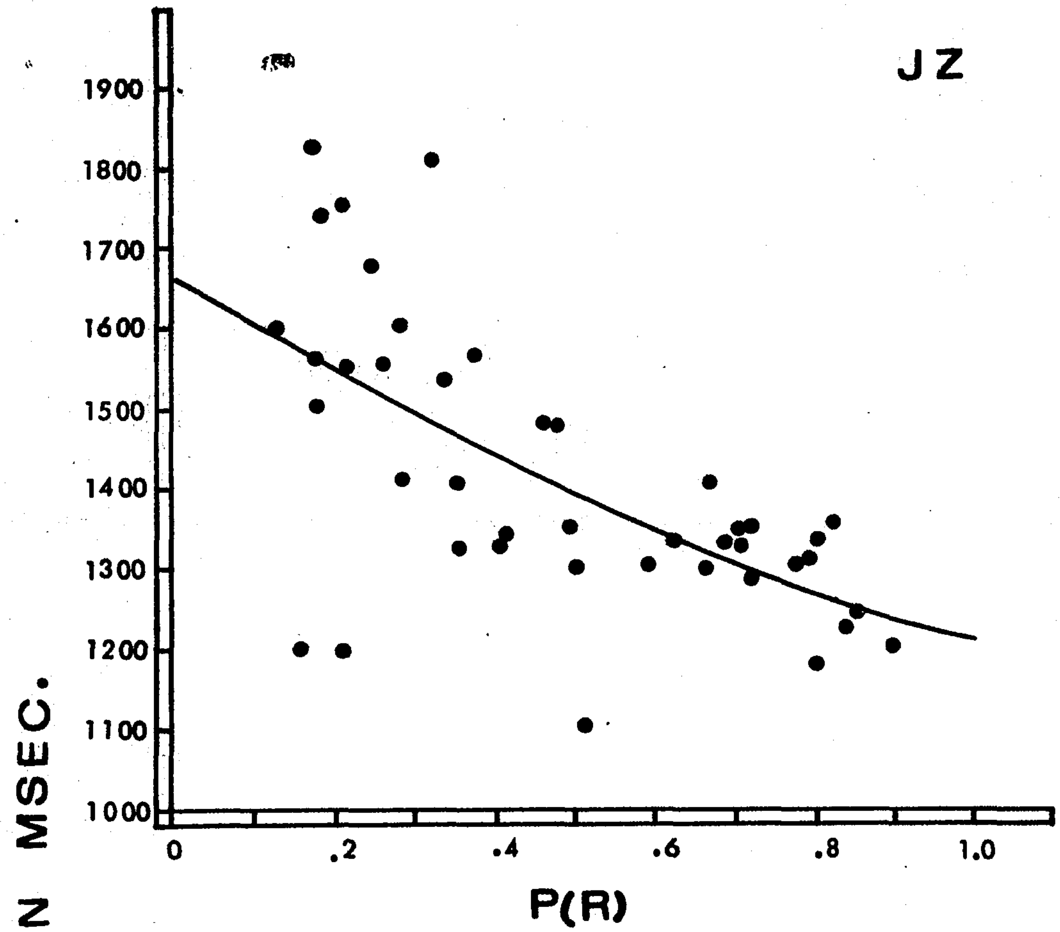
and intercept equal to

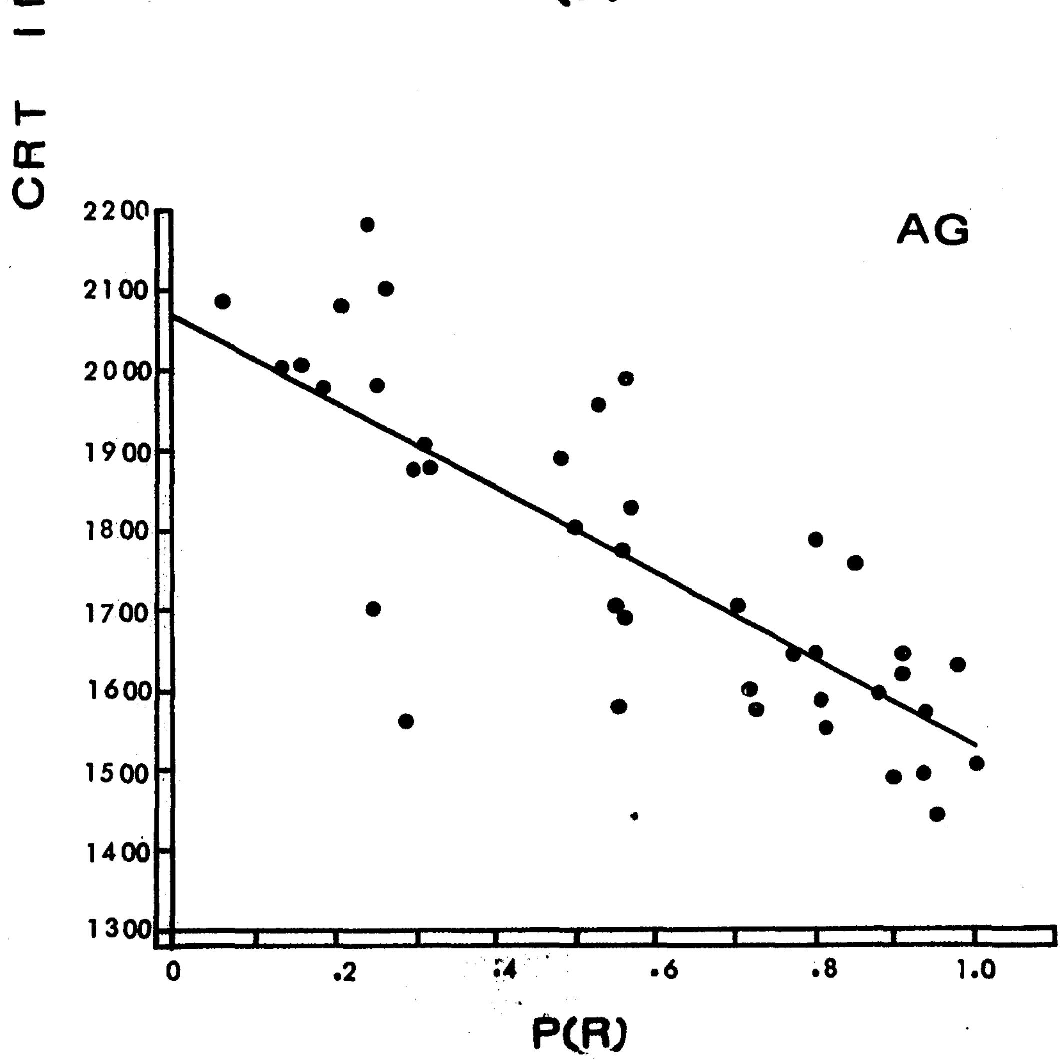
$$\frac{\sigma_Y^2(\mu_{X_5} - \mu_{X_3})}{2\sigma_X^2(\mu_{Y_5} - \mu_{Y_3})} + \frac{(\mu_{Y_5} - \mu_{Y_3})}{2} + \log K \cdot \frac{\sigma_Y^2}{(\mu_{Y_5} - \mu_{Y_3})}$$

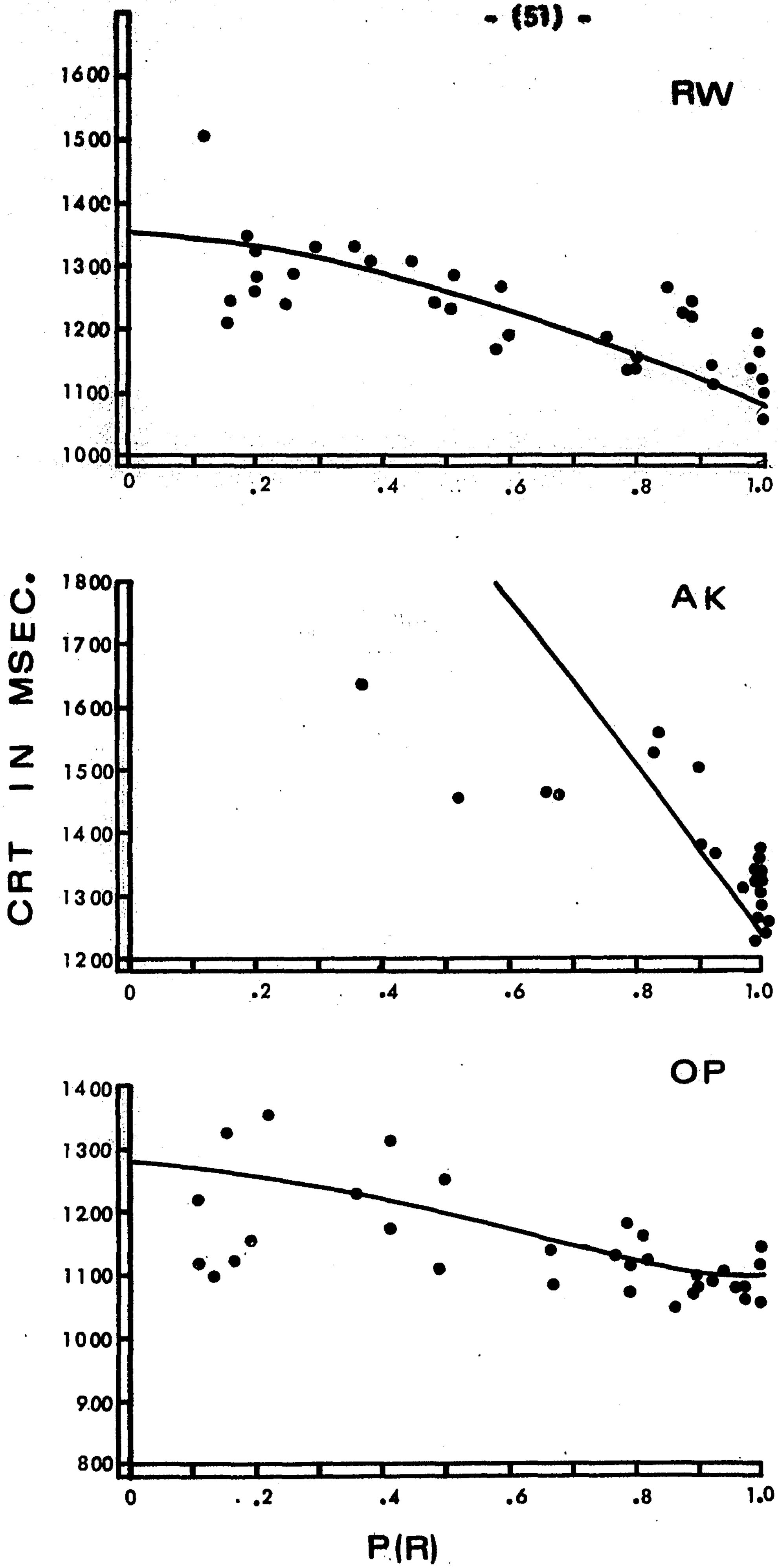
where K is the presentation ratio of the training stimuli (X_3,Y_3) to (X_5,Y_5).

The equation of the decision function when K=1 can be simplified by letting the origin of the coordinate system of decision axes fall on and bisect the line connecting the two means of the bivariate

Figure 7. Probability-median CRT function. Each point on the figure represents a median CRT and response probability of R value for a stimulus. The solid lines are the fitted polynomials of the second order.


JZ
1900
1800
1700
1600
1500
1400
1300
1200
1100
1000
0
.2
.4
.6
.8
1.0
P(R)
CRT IN MSEC.


AG
2200
2100
2000
1900
1800
1700
1600
1500
1400
1300
0
.2
.4
.6
.8
1.0
P(R)


RW
AK
OP
CRT IN MSEC.
1600
1500
1400
1300
1200
1100
1000
1800
1700
900
800
0
.2
.4
.6
.8
1.0
P(R)

distributions for the training stimuli. It can be shown that the origin of the new coordinate system falls at $\left(\frac{\mu_{X_5} - \mu_{X_3}}{2}, \frac{\mu_{Y_5} - \mu_{Y_3}}{2}\right)$ (see Appendix 3). The decision function can be rewritten as

$$Y' = \left[- \frac{(\mu_{X_5} - \mu_{X_3})\,\sigma_Y^2}{(\mu_{Y_5} - \mu_{Y_3})\,\sigma_X^2} \right] X' \qquad \ldots (3)$$

where X' and Y' indicate the decision variables in the new coordinate system. Using σ_X and σ_Y as the units of measurement, equation (3) can be further simplified to

$$Y'' = - [\, d'_x / d'_y \,]\, X''$$

where $d'_x = \frac{\mu_{X_5} - \mu_{X_3}}{\sigma_X}$, $d'_y = \frac{\mu_{Y_5} - \mu_{Y_3}}{\sigma_Y}$

and X'', Y'' represent the transformed decision variables in standard normal units. When $K \neq 1$, it can be shown that the decision function becomes

$$Y'' = - [\, d_x' / d_y' \,]\, X'' + (\log K/d_y') \qquad \ldots (4)$$

From equation (4), it can be seen that the slope of the decision line depends upon the values of d'_x and d'_y which, in turn, are function of σ_X and σ_Y, respectively, for a fixed set of training stimuli. In addition, the slope of the decision line also depends upon how stimuli are scaled. Since the exact relationship between the stimulus and the decision variable is not known, the following is based on estimates of σ_X and σ_Y derived from normal ogive that were fitted to the data obtained in the unidimensional experiment. The stimuli were spaced either according to logarithmic or power function.

<u>A. 1. Deriving theoretical decision lines on the basis of a logarithmic scale</u>

Sound intensities were expressed in units of dB and light intensities in log ft.l. The procedure used to estimate σ_X and σ_Y is that described by Finney (1964), which has been discussed in the previous section. A probit regression line was fitted to data and the reciprocal value of the slope of the function was taken as the maximum likelihood estimate of the standard deviation of the underlying distribution. The estimated values of σ_X and σ_Y are summarized in Table 4 (for unidimensional gradients, see Appendix 4).

To derive the theoretical decision line, these estimates of σ_X and σ_Y and the 'distance' between training stimuli of light and the 'distance' between training stimuli of sound were substituted into equation (4). The distance was simply defined as the difference between two distribution means.

The theoretical decision lines generated for each S together with the obtained 50% response points are shown in Figure 8. The solid lines in the figure represent theoretical decision lines derived from the model. Since there is no unique solution for calculating empirical 50% response points (Finney, 1964), they were represented by two straight lines: The dashed line (L line) represents the fitted 50% response points (open circles) which were obtained along the light dimension with sound intensity as parameters while the broken line (S line) was the fitted 50% response points (filled points) along the sound dimension with light intensity as parameters. The 'best' line, probably falls somewhere between these two lines.

In the case of Ss AG and RW who were trained with unequal presentation ratio, the solid lines with non-zero intercept represent

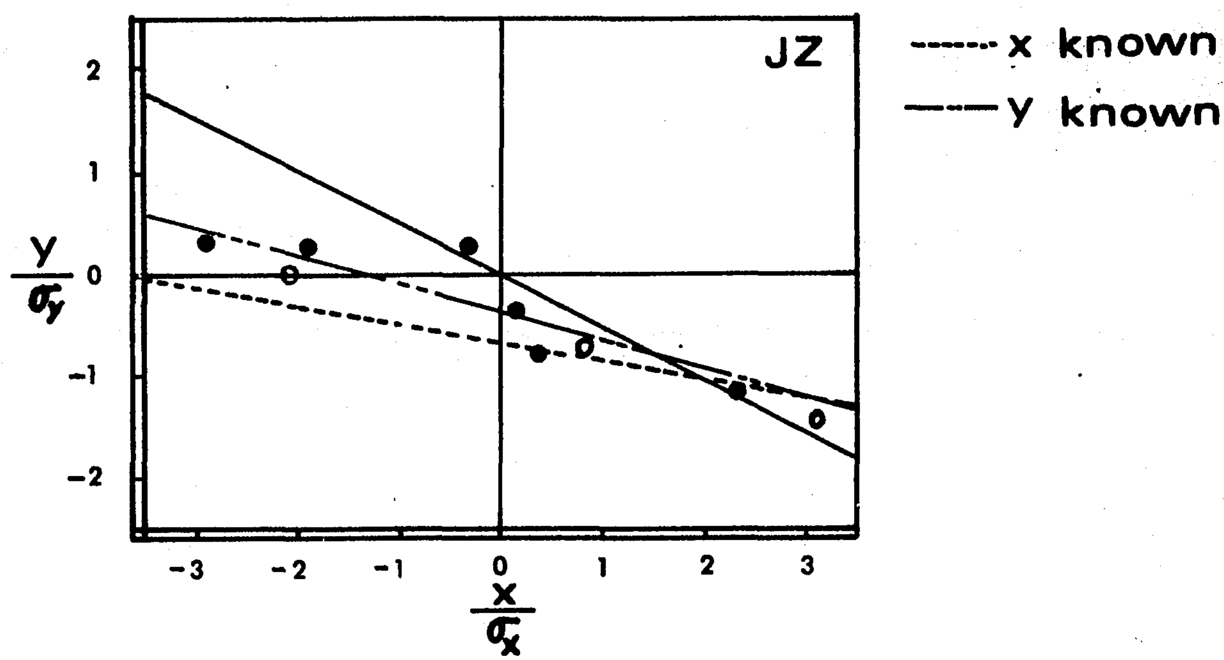
Table 4. Maximum likelihood estimation of σ_X and σ_Y on logarithmic scale.

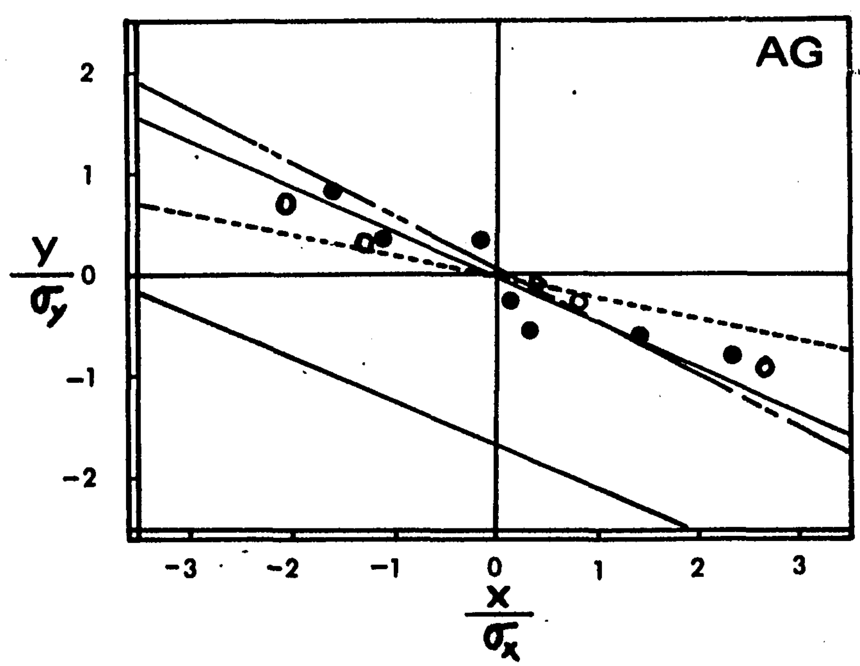
Ss	JZ	AG	RW	AK	OP
Estimated σ_X	0.0949	0.1577	0.1861	0.0809	0.0448
Estimated σ_Y	0.6986	1.1276	0.6181	0.6402	0.7045

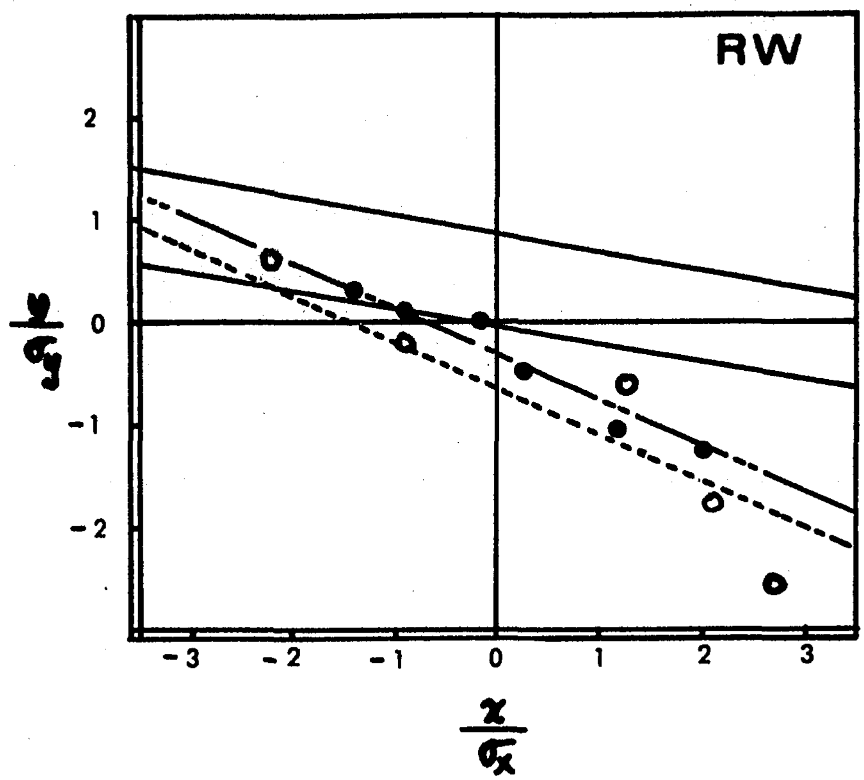
Table 5. Maximum likelihood estimation of σ_X and σ_Y on power function scale.

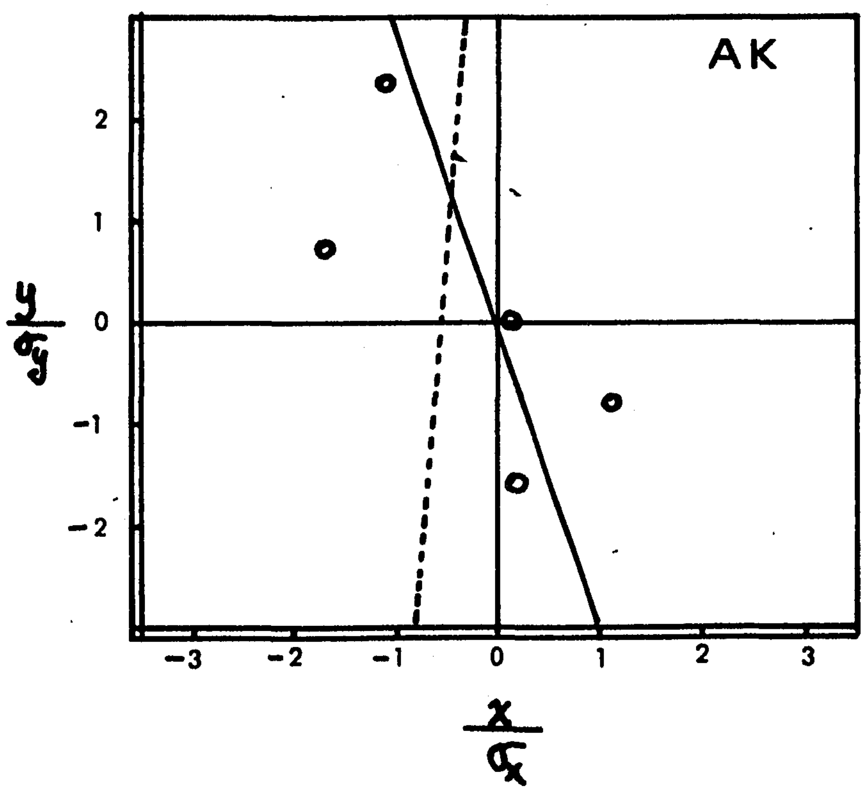
Ss	JZ	AG	RW	AK	OP
Estimated σ_X	0.2052	0.9342	0.4035	0.3861	0.0436
Estimated σ_Y	0.0909	0.2014	0.1116	0.3780	0.2159

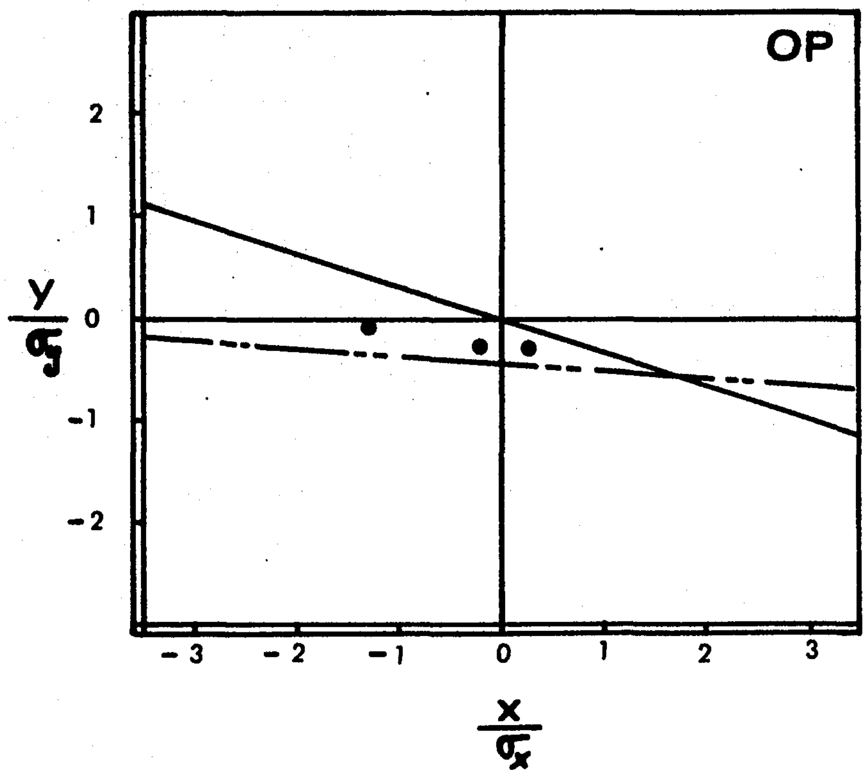
Figure 8. Theoretical decision line (solid line) and the fitted empirical 50% response points based on a logarithmic scale. The broken line is the one fitted to the 50% response points (filled points) which were obtained along the sound dimension at various fixed light intensity levels. The dashed line represents the 50% response points (open circles) along the light dimension at the various sound intensity levels. The light dimension is represented by the X-axis while the sound dimension is represented by the Y-axis. Becasue of the limitations of space, several data points were excluded from the graph.


JZ
x known
y known
2
1
0
-1
-2
-3
-2
-1
0
1
2
3
y/σy
x/σx


AG
2
1
0
-1
-2
-3
-2
-1
0
1
2
3
y/σy
x/σx


RW
2
1
0
−1
−2
$\frac{y}{\sigma_y}$
−3
−2
−1
0
1
2
3
$\frac{x}{\sigma_x}$


AK
2
1
0
−1
−2
$\frac{y}{\sigma_y}$
−3
−2
−1
0
1
2
3
$\frac{x}{\sigma_x}$


OP
2
1
0
-1
-2
y
σy
-3
-2
-1
0
1
2
3
x
σx

the theoretical lines. A comparison of the obtained 50% response points with the theoretical lines for unequal presentation ratio, shows clear discrepancies. However, the theoretical lines for equal presentation ratios are quite close to the obtained 50% points, suggesting that during generalization testing Ss AG and RW shifted their decision lines to the contour having K = 1.

A series of comparisons between the theoretical (K = 1) and the S-line and between the theoretical (K = 1) and the L-line was performed by asking the question whether each pair of lines was actually one line or indeed two separate lines. The hypothesis that the slope of the straight line fitted to the difference is zero was tested with a *t* test procedure suggested by Acton (1959). The results are shown in Table 6. The insignificant *t* value suggests that the pair of lines is actually one line.

The results reveal that, except for S RW, the theoretical line is statistically indistinguishable from one of the obtained sets of 50% response points. Thus the decision lines based on information from single dimensions provide a reasonably good approximation to actual data; particularly in view of the fact that there are no free parameters involved in generating the theoretical lines. However, the procedure used to estimate the parameters σ_X and σ_Y directly from the unidimensional gradients involves some of the difficulties which will be discussed later.

A. 2. Generating theoretical gradients based on logarithmic scales

Once the theoretical line is computed, the conditional response probability, $p(R \mid X_i, Y_j)$, given any compound stimulus with means equal to (μ_{X_i} , μ_{Y_j}) can be obtained directly from the table of normal distribution by a procedure described previously (page 11).

Table 6. Testing for "one line or two line" based on logarithmic scale.

S: JZ	t	df	S: AG	t	df	S: RW	t	df
T_o vs L	4.56	5	T_o vs L	5.08	5	T_o vs L	8.44	4
T_o vs S	2.41**	5	T_o vs S	0.67**	5	T_o vs S	3.19	5

S: AK	t	df	S: OP	t	df
T_o vs L	1.84**	5	T_o vs L	2.12**	5
T_o vs S	-	-	T_o vs S	19.59	5

T_o represents theoretical line when K = 1.
** not significant at 5 % level.

Table 7. Testing for "one line or two lines" based on power function scale.

S: JZ	t	df	S: AG	t	df	S: RW	t	df
T_o vs L	0.32**	5	T_o vs L	0.04** **	5	T_o vs L	9.75	3
T_o vs S	8.84	5	T_o vs S	1.92	5	T_o vs S	3.76	5

S:	t	df		t	df
T_o vs L	1.63**	5	T_o vs L	2.11**	5
T_o vs S	-	-	T_o vs S	8.35	5

The conditional response probability, $p(R \; X_i, Y_j)$, is equal to the area under standard normal distribution from - to r, where r is equal to

$$(\mu_{Y_j} - m \mu_{X_i}) / \sqrt{1 + m^2}$$

Figure 9 shows the S-shaped theoretical gradients (solid line) and the obtained response probability values (circle points) for each S. To avoid awkwardness, $p(L \mid X_i, Y_j)$ instead of $p(R \mid X_i, Y_j)$ was plotted against light intensity level in standard normal units. Each graph in the figure represents a gradient along the light dimension at a given level of sound intensity. Note that, for Ss AG and RW, K was assigned the value of 1 in computing the theoretical gradients. The numerical values of the theoretical and obtained conditional response probability L are tabulated in Appendix 5.

Inspection of these graphs indicates that the results of S AG are fitted best by the gradients. As a rough measure of agreement between the theoretical and the obtained gradients, the fomula

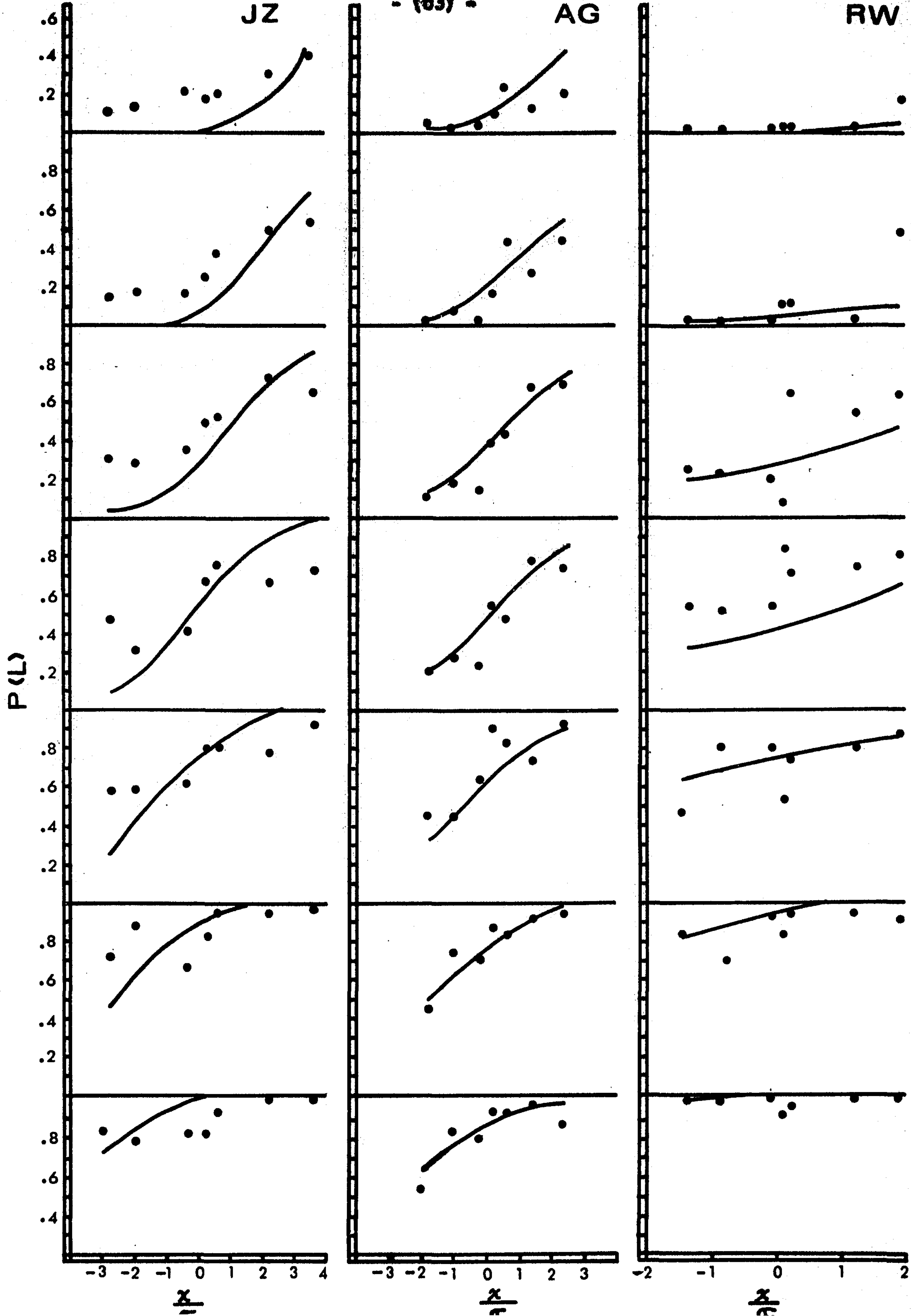
$$1 - \Sigma (y - x)^2 / \Sigma y^2$$

Was used as an index of the fraction of the total variance accounted for, where x and y denote the deviation of the theoretical and obtained values from their respective means. Table 8 shows the results of these computations. It can be seen that the theoretical gradients fit the obtained data reasonably well for all Ss except Ss JZ and OP.

A. 3. Deriving theoretical decision lines on the basis of a power function scale

The results of the magnitude estimation experiment were used to obtain the exponent n in $\psi = k \phi^n$. The logarithms of the mean estimates were plotted against the logarithms of the stimulus

Figure 9. The theoretical (solid line) and the obtained gradients (open circle) on logarithmic scales. Each column represents the gradients plotted along the light dimension in standard deviation units at various fixed intensity level of sound for each S.


JZ
AG
RW
P(L)
x/σ
x/σ
x/σ

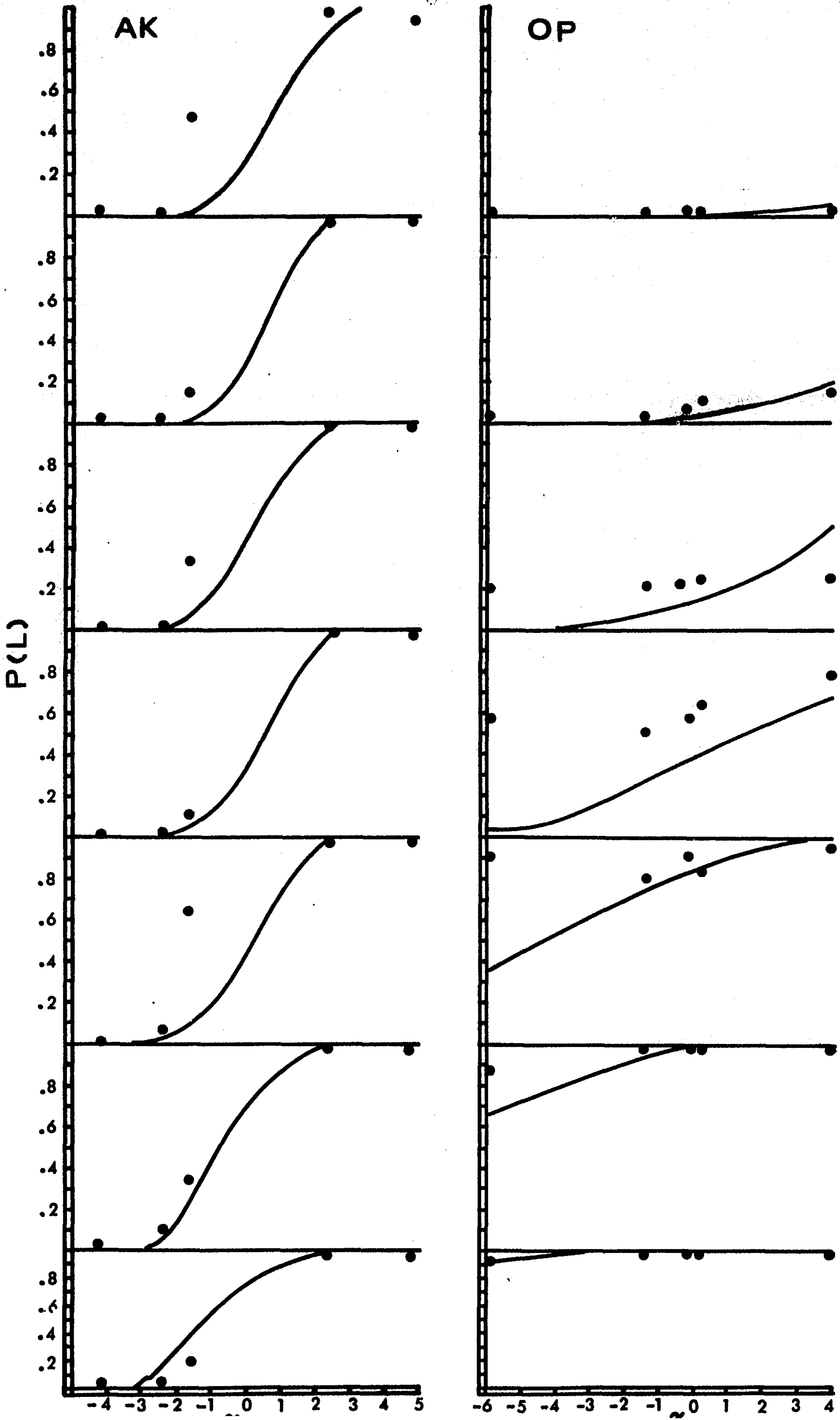

AK
OP
P(L)
.8
.6
.4
.2
-4 -3 -2 -1 0 1 2 3 4 5
-6 -5 -4 -3 -2 -1 0 1 2 3 4

Table 8. Percentage of total variance of the obtained response probabilities removed by the theoretical probabilities on logarithmic scale.

Ss	JZ	AG	RW	AK	OP
% Variance Removed	67	90	87	94	64

Table 9. Percentage of total variance of the obtained response probabilities removed by the theoretical probabilities on power function scale.

Ss	JZ	AG	RW	AK	OP
% Variance Removed	37	82	86	81	54

values and a straight line was fitted to the points. The slope of this line was then taken as the estimate of <u>n</u> (see Appendix 6). All stimuli were rescaled by taking ϕ^n.

After transformation of both stimulus scales, the entire procedure described in <u>A.1.</u> was repeated to obtain estimates of σ_X and σ_Y and, subsequently, the theoretical decision lines. Table 5 presents the estimated values of σ_X and σ_Y, and Figure 10 displays the theoretical lines along with fitted empirical 50% response lines. The difference between corresponding theoretical and empirical points was taken and the hypothesis that the slope of the line fitted to the differences is equal to zero was tested. The results of this test are similar to those found in <u>A.1.</u> (see Table 7). The results also indicate that the theoretical line assuming K = 1 fits much better to the obtained 50% response points lines based on the assumption that K = 1. The theoretical lines were found to depart insignificantly from one of two sets of obtained 50% points, except for <u>S</u> RW.

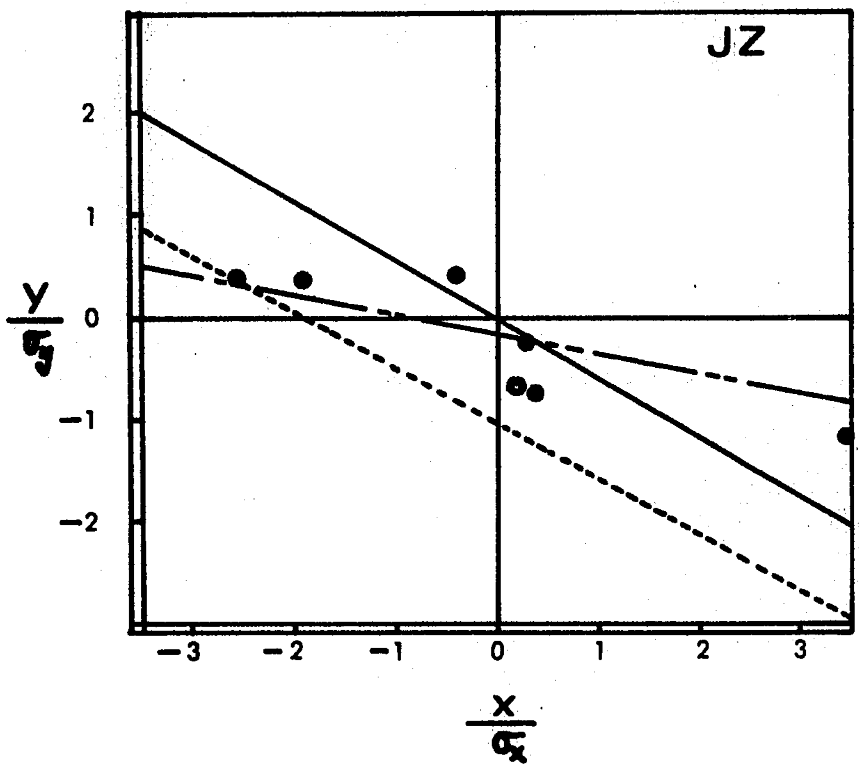
<u>A.4. Generating theoretical gradients based on power function scales</u>

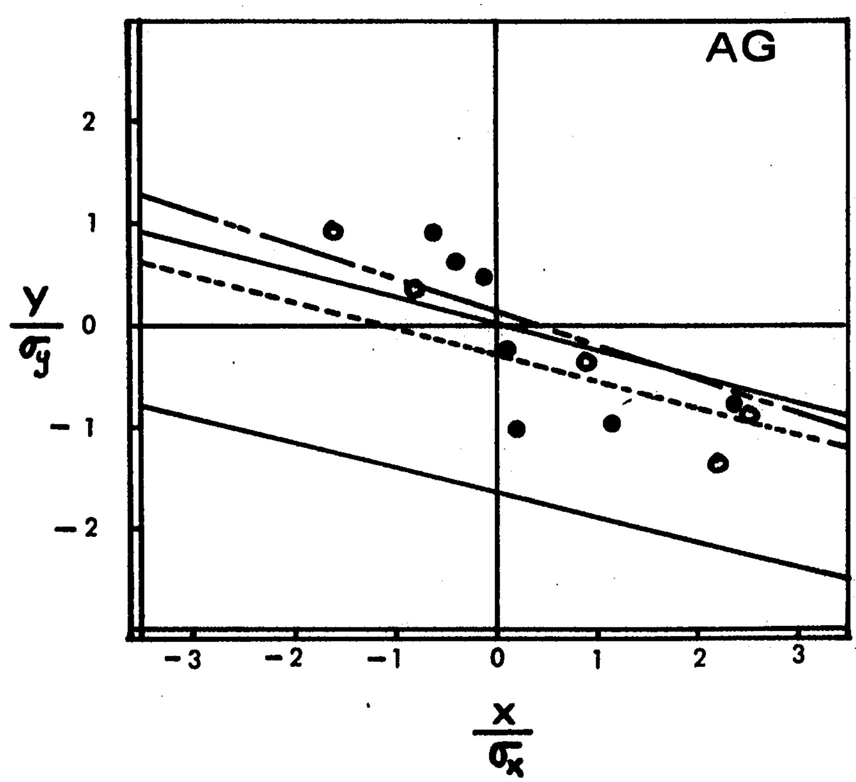
Using the same method described in <u>A.2.</u>, the theoretical gradients for all <u>S</u>s were computed (Appendix 7). They are shown in Figure 11. To assess the 'goodness of fit', the percentage of variance removed by the theoretical gradients was computed and is /shown in Table 9. The results indicate that the theoretical gradients generated on the power function scales deviate more from the obtained data than those generated on logarithmic scale.

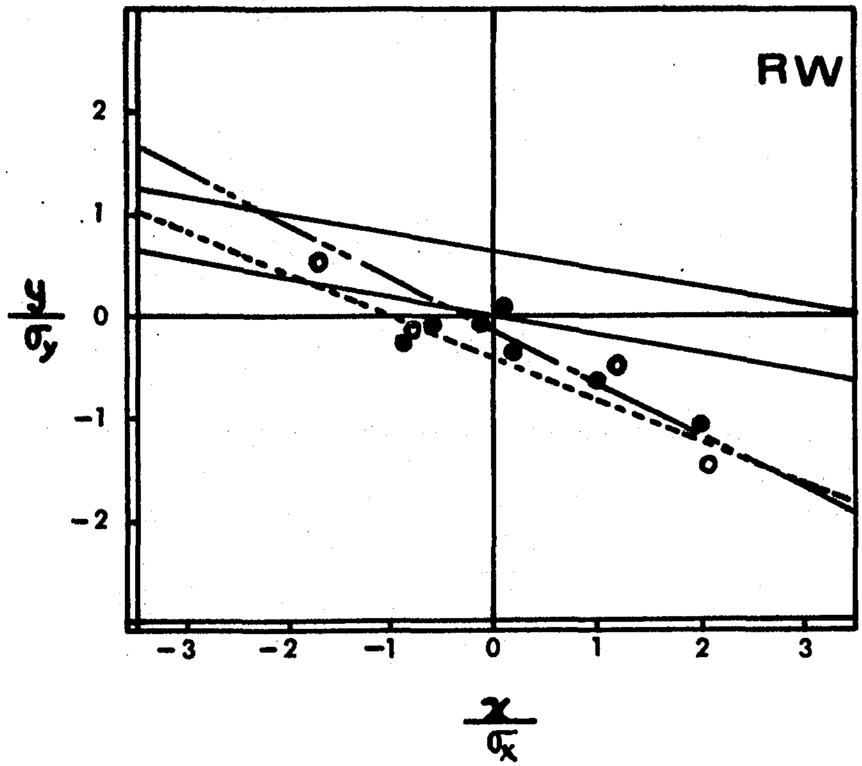
<u>B. Signal recognition theory as a descriptive model</u>

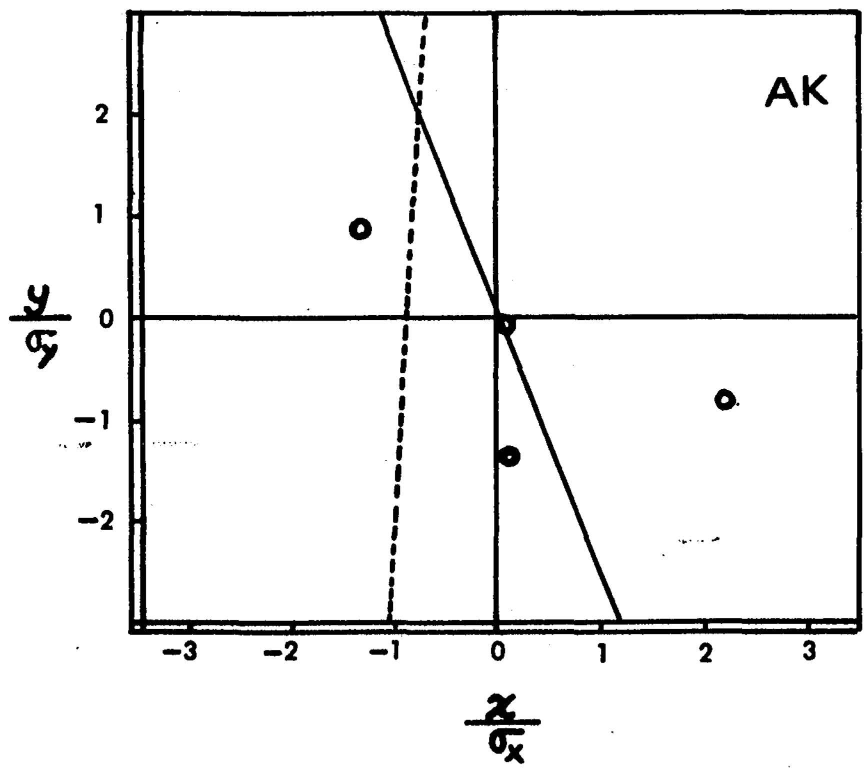
In the following analysis the parameters were not estimated from each of the unidimensional gradients. Instead, they were regarded

Figure 10. Theoretical decision line (solid line) and the fitted empirical 50% response points based on the power function scale. The broken line is the one fitted to the 50% response points (filled points) which were obtained along the sound dimension at various fixed light intensity levels. The dashed line represents the 50% response points (open circles) along the light dimension at the various sound intensity levels. The light dimension is represented by X-axis while the sound dimension is represented by Y-axis. Both axes are in standard deviation units. Because of the limitations of space, several points were excluded from the graph.


JZ
2
1
0
−1
−2
y/σy
−3
−2
−1
0
1
2
3
x/σx


AG
2
1
0
− 1
− 2
y/σy
− 3
− 2
− 1
0
1
2
3
x/σx


RW
2
1
0
−1
−2
−3
−2
−1
0
1
2
3
$\frac{y}{\sigma_y}$
$\frac{x}{\sigma_x}$


AK
2
1
0
−1
−2
−3
−2
−1
0
1
2
3
$\frac{y}{\sigma_y}$
$\frac{x}{\sigma_x}$

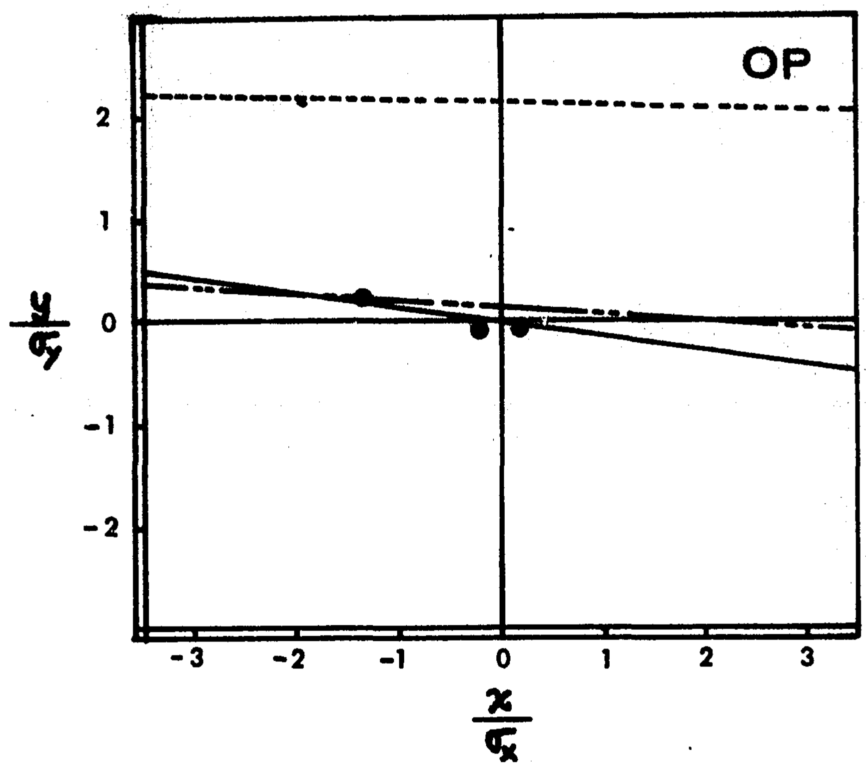
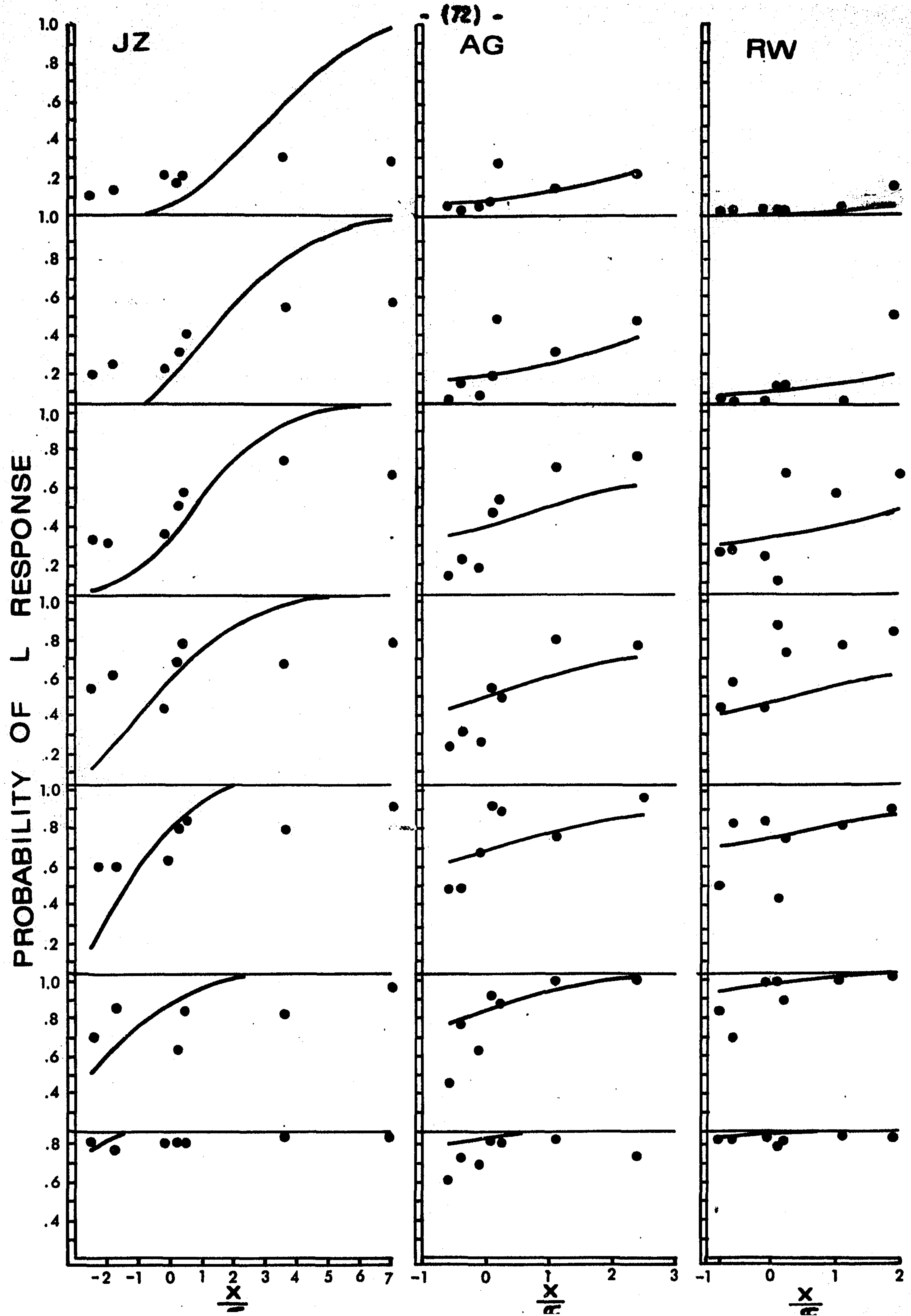
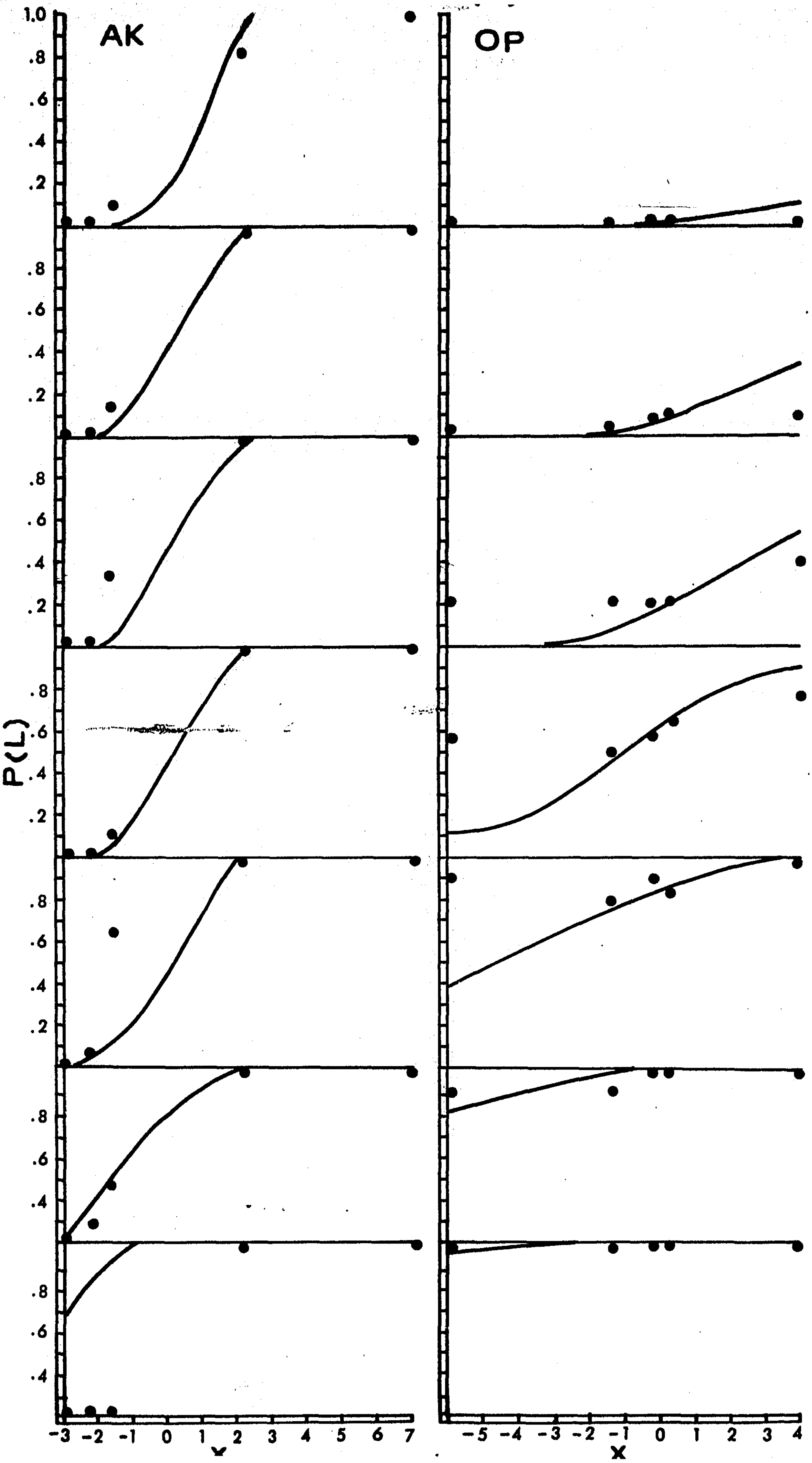

OP
2
1
0
-1
-2
$\frac{y}{\sigma_y}$
-3
-2
-1
0
1
2
3
$\frac{x}{\sigma_x}$

Figure 11. The theoretical (solid line) and the obtained gradients (open circle) on power function scale. Each column represents the gradients plotted along light dimension in standard deviation units at various fixed intensity level of sound for each S.

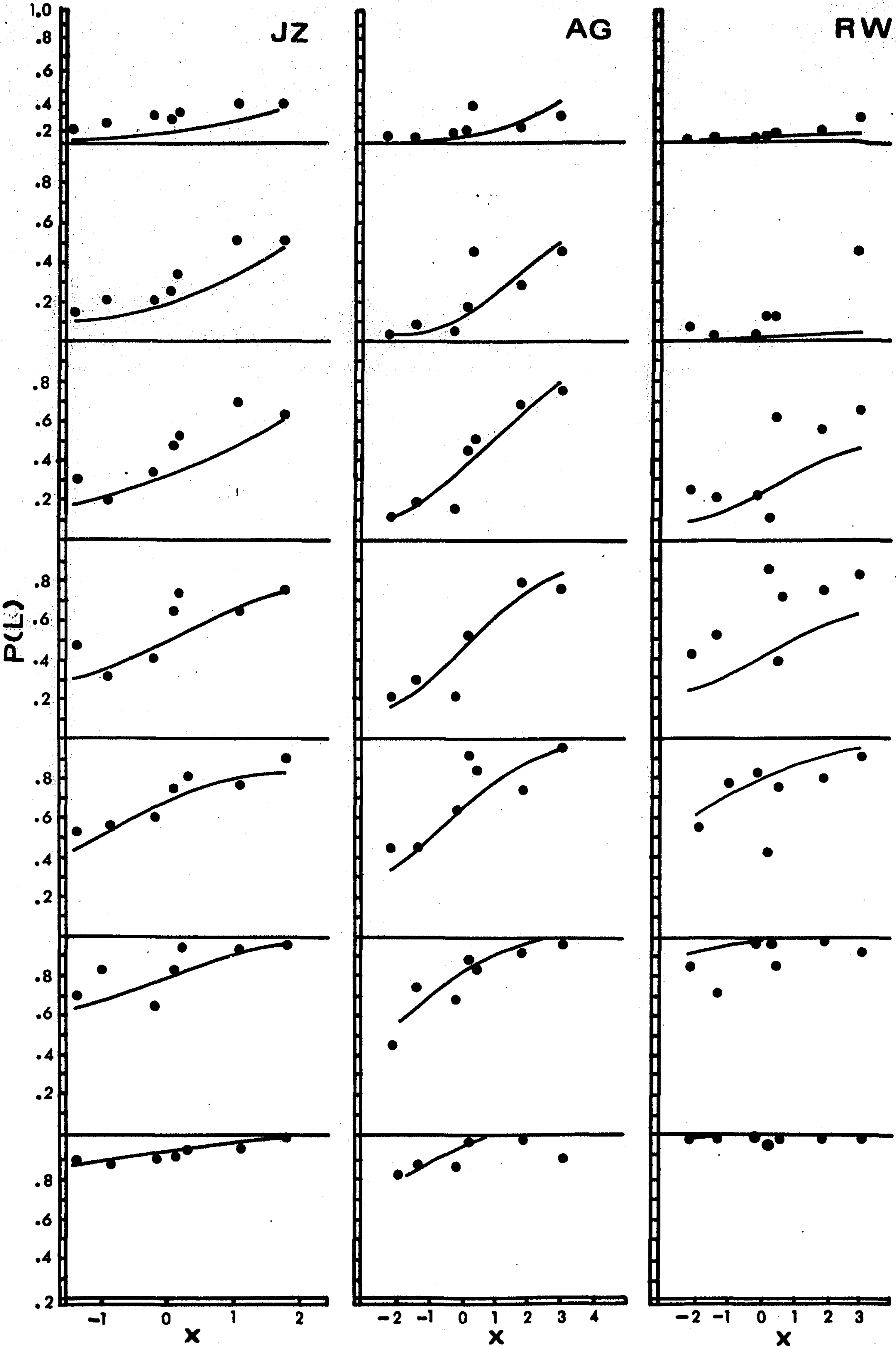

JZ
AG
RW
PROBABILITY OF L RESPONSE
1.0
.8
.6
.4
.2
-2 -1 0 1 2 3 4 5 6 7
-1 0 1 2 3
-1 0 1 2
X/σ



AK
OP
P(L)
1.0
.8
.6
.4
.2
-3 -2 -1 0 1 2 3 4 5 6 7
-5 -4 -3 -2 -1 0 1 2 3 4
X
X

as free parameters and recovered directly from the two-dimensional gradients by the method of the least sum of weighted squares (Chase and Heinemann, 1972). The method for fitting employed an iterative procedure implemented by a computer program which searched for σ_X and σ_Y such that they yielded theoretical gradients having the least sum of weighted squared deviations from the obtained gradients. Only the traditional logarithmic transformation was used to scale the stimuli. Figure 12 shows the fitted gradients (solid lines) which are plotted along the light dimension at various fixed levels of sound intensity. The percentage of variance removed by fitting ranged from 80 to 94 (Table 10). The fitting does not seem to improve greatly by applying attention correction procedure (Chase and Heinemann, 1972). For comparison purposes, the recovered values of σ_X and σ_Y are shown in Table 11 together with the values of σ_X and σ_Y estimated from the unidimensional gradients. It can be seen that the values of σ_X and σ_Y recovered from fitting were slightly greater on the average than those estimated from the unidimensional gradients.

Figure 12. Fitted theoretical gradients. The solid lines indicate the fitted gradients while filled points indicate the actual data points.


JZ
AG
RW
P(L)
1.0
.8
.6
.4
.2
x
-2 -1 0 1 2 3 4

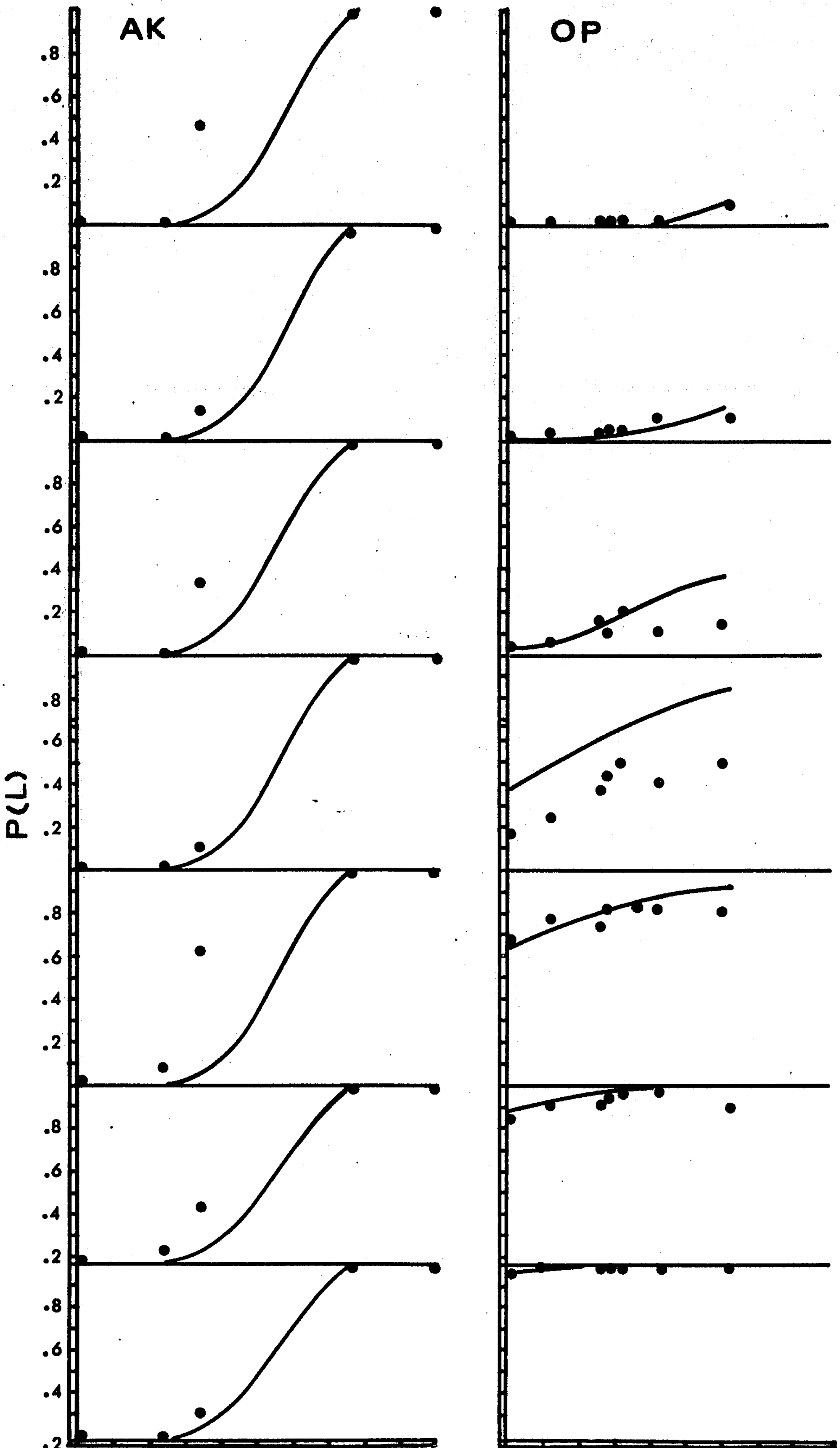

AK
OP
P(L)
.8
.6
.4
.2

Table 10. Percentage of total variance of the obtained response probabilities removed by fitting theoretical gradients.

Ss	JZ	AG	RW	AK	OP
% of Variance Removed	93	92	80	94	94

Table 11. The values of σ_X and σ_Y recovered from curve fitting, and σ_X and σ_Y estimated from the unidimensional gradients based on logarithmic scale.

	Ss	JZ	AG	RW	AK	OP
σ_X	Recovered Values	0.1394	0.0400	0.1861	0.1000	0.0714
	Estimated Values	0.0949	0.1577	0.1861	0.0809	0.0448
σ_Y	Recovered Values	1.0000	1.4398	1.0000	1.0000	0.6309
	Estimated Values	0.6986	1.1276	0.6181	0.6402	0.7045

Discussion

Although they are minor interest, the results of discriminative training clearly indicate the success of discrimination training between two compound stimuli. They also indicate that d' changes during the testing phase, but the direction of the change was not consistent among Ss. While Ss AG, AK and RW showed a continuing increase in d' throughout the testing phase, Ss JZ and OP showed initial decrement in d' when testing stimuli were introduced. The effect of introducing testing stimuli on discriminative performance has been reported by others. For example, Heinemann *et al.*(1969) using pigeons as Ss showed that the discriminative performance deteriorated after testing stimuli were introduced. LaBerge (1961), on the other hand, reported in an experiment done with college students that testing stimuli had little effect on discriminative performance.

It is not clear at this point what variables underlie these discrepancies. However, in the framework of the detection model, the changes in d' that occurred during the testing phase may be interpreted in terms of an increase or decrease in the variability of the criterion placement. It can be shown that an increase in criterion variability results in reduction of d' and *vice versa* (Tanner, 1964).

The generalization gradients obtained have an S-shaped form that differs from the traditional peaked one. These S-shaped gradients are qualitatively in agreement with what the detection model predicts -- the normal ogive. The S-shaped form of these gradients is of considerable theoretical significance. A number of proposals have been made regarding the variables that may be responsible for generating these two types of gradients, peaked and S-shaped ones. Heinemann et al. (1969) pointed out that the experimental method may be important. Usually, the peaked gradients have been obtained in experiments using the one-key free operant method, while S-shaped gradients have been obtained by the choice method. However, a further study conducted by Heinemann et al. (1970) showed that S-shaped gradients may also be obtained with the one-key free operant method, at least when the stimulus continuum is the intensity of white noise.

Cross and Lane (1962) offered another explanation which emphasized the characteristics of the stimulus. They asserted that the stimuli which generate peaked gradients are those which are "metathetic" rather than "prothetic". Prothetic sensory continua, on which discrimination is characteristically mediated by an additive process at the physiological level, may be those which produce S-shaped gradients. Metathetic continua, on which discrimination is mediated by a substitutive process, will not produce S-shaped gradients. This assertion, however, was not supported by LaBerge's findings. LaBerge's study, in which an obvious metathetic stimulus continuum -- position of a small light -- was employed yielded an S-shaped gradient.

Another explanation which also emphasizes the stimulus characteristics has been proposed by Chase and Heinemann (1972). They pointed out that stimulus continua which are circular in nature tend to generate peaked gradients. Angularity of a line is a typical example of such a continuum.

The present author tends to think of these two types of gradients as representations of two stages of a hierachical sensory process: The S-shaped gradient represents the results of a classification process among sensory inputs and the peaked gradient represents an absolute identification of the stimulus. If this is the case, it is argued that, after continuing and intensive training, the S-shaped grdients will ultimately result in peaked ones as the process shifts from classification to identification, if the experimental procedure so demands.

An attempt was made to generate theoretical two-dimensional gradients from the empirical unidimensional ones. The success in predicting the two-dimensional gradients from the unidimensional gradients depends upon several factors. One of these factors is the accuracy of the estimates of σ_X and σ_Y . This accuracy may be affected by the criterion variability in the unidimensional and two-dimensional situations. It has been pointed out that criterion variability is equivalent to an increase of distribution variability. Thus, the variance estimated from the unidimensional gradients may actually consist of two components, one component representing the 'true' distribution variability and the other criterion variability. A method of extracting the component of distribution variability has not been developed, and the rule according to which the criterion variability in the unidimensional situation combined to produce criterion variability in the two-dimensional situation, unfortunately, is not known. The obscurity of the relationships among all these components undoubtedly affects the accuracy of the estimation.

Another factor which also determines the success of predictions from unidimensional gradients is the degree of attention that the

S reveals in the unidimensional and two-dimensional situations. By "degree of attention" to a dimension is meant the relative number of trials on which the S uses the information from that dimension in making his choice.

Besides the afromentioned problem of estimation, stimulus spacing or scaling adds another major difficulty in using the otherwise mathematically simple model of signal recognition. To find a proper transformation law to scale stimuli is still an unsolved problem in psychophysics. However, if the detection model is used as a strictly descriptive one and parameters are estimated freely from the data, it is reasonable to assert that the model is a quite successful one.

According to the previous analysis, the decision line is the major factor in determining the form of generalization surface. The relative degree of control exercised by each dimension also depends upon the placement of the decision line. One of the factors which has been found to affect the location of the decision line is the discriminability between two values of training stimuli along each dimension. Infinite discriminability between two values of training stimuli from one dimension will result in placing the decision line perpendicular to that dimension; so that the other dimension has virtually no control over behavior. This is clearly shown in the results of Ss OP and AK. While the effect of the relative discriminability of stimuli from the two dimensions on the slope of the decision line is as predicted by the model, at least qualitatively, the presentation ratio of two compound training stimuli, which should determine the location of the decision line, failed to produce the expected effect during testing phase. The failure of the presentation ratio to affect behavior during the testing phase (but not during the training phase; see Figure 2) may be due to the arrangement of the testing situation, in which the stimuli were symmetrically spaced in a 7 x 7 array. The symmetrical spacing of testing stimuli may change the criterion to a location where K = 1.

<u>Latency-probability function</u>

One of the major findings in this experiment was that the probability of a given response was negatively correlated with its median latency. As response probability increases the median response latency decreases. These results are contrary to the latency data of pigeon reported by Heinemann <u>et</u> <u>al</u>. (1969). Their results showed that mean latency was independent of response probability. However, results reported by other investigators who employed human subjects seem in accord with the present results. An early study done by LaBerge (1961) showed that median latency was negatively correlated with response probability for a given response type. The response and latency gradients were found to cross each other.

In a similar situation, using pure tones as discriminative stimuli, Cross and Lane (1962) obtained similar results. That is, latencies of the stochastically dominant response were consistently shorter than the latencies of the non-dominant response.

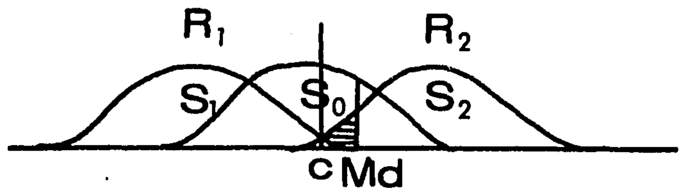
Audley and Mercer (1968) also reported data supporting these findings. Employing a two-choice, blue-green discrimination situation the response probability of blue was manipulated by varying the difficulty of the discrimination. Response-latency function were shown to be negatively accelerated with the shortest latencies associated with the highest probabilities.

Similar findings were also reported by Pickett (1967) Wollen (1963), Pierrel and Murray (1966), and Sekuler (1965, 1966). The present findings show that this inverse relationship also holds in a two-dimensional situation situation.

The relationship between latency and probabilities in the uni-dimensional and two-dimensional situation do not support the latency model of Stone (1960) according to which latencies of correct and incorrect decision are the same. However, Stone's model has not been refuted conclusively because previous experiments, including the present one, do not meet the optimal stopping assumption of Stone's model in which S̲ is presumed to adopt a strategy of minimizing the sample size. The model implies that the S will avoid unnessary costly extra samples and make responses as fast as possible. None of the studies described forced the S̲ to adopt this strategy, or provided any penalty for responding slowly. The present experiment, which limits the duration of the stimulus presentation, may lead the S̲ to respond quickly, but the reinforcement was contingent only on the accuracy of the response, not on the speed.

Stone's model is somewhat incomplete. It does not consider the a priori probability of the signal nor does it consider the payoffs. Also, the model does not specify the hypothetical distribution of likelihood ratio or their relationship to the stimulus features. A Bsysian version of Stone's model has been presented by Edwards (1965). Also, Laming's random walk model (1968) is actually an extension of Stone's basic idea.

Recently, several latency models based on signal detection theory have been proposed. They add one or more assumptions regarding the 'quality' or precision of the signal. Audley and Mercer (1968) proposed a model which postulates that the quality of the signal depends upon the signal strength and varies from trial to trial. An excellent signal leads to a fast response, and often to a correct decision, whereas a signal with poor quality needs a lengthy examination and often leads to wrong decisions. The model also implicitly postulates that the 'quality' of the signal is a function of the distance between the location of observation and criterion. To illustrate the basic idea of the model, consider the discrimination situation shown in the following figure. Suppose S was trained to respond R_1 when S_1 is presented and R_2 when S_2 is presented, and let S_0 be a testing stimulus. For simplicity, assume that all distributions are identical except for their means. Let the variance be unity and the decision criterion be equal to the mean of the distribution of S_0. Consider the distribution of S_0 and let Md be a point along the x-axis such that $p(x > Md) = 1/2 \cdot [\, p(x \geqslant c)]$. The latency t associated with Md is the median latency of R_2 given stimulus S_0. Now consider the distribution of S_2. The horizontally hatched area in the figure indicates the probability of R_2 with latencies longer than t, since these R_2 responses are based on poorer information than those requiring time Md to process. The area to the right of Md indicates the proportion of response having latencies shorter than t. As the test stimulus moves away to the right of c, the conditional response probability of R_2 will increase and the proportion of R_2 responses whose latencies are longer than t will decrease. One of the prediction of this model is that the response probability and latencies should be inversely related.



Thomas (1970) basically follows the same idea but with more sophisticated mathematical elaborations. Thomas' model also derives the sufficient conditions for obtaining a latency-probability function with negative slope. Similar models also have been proposed by Smith (1968), LaBerge (1962), Audley and Pike (1968). No attempt has been made to assess these models or choose among them. However, the latency data of the present experiment do seem to support the class of the models that predict an inverse relationship between response latency and probability.

Appendix 1. Method for computing the response probability of a type for any stimulus combination X_iY_j.

Given a decision line $Y = mX + b$, where

$$m = -\frac{(\mu_{X_5} - \mu_{X_3})\,\sigma_Y^2}{(\mu_{Y_5} - \mu_{Y_3})\,\sigma_X^2}$$

$$b = \frac{\sigma_Y^2(\mu_{X_5} - \mu_{X_3})}{2\sigma_X^2(\mu_{Y_5} - \mu_{Y_3})} + \frac{(\mu_{Y_5} - \mu_{Y_3})}{2}$$

for any stimulus combination (X_i, Y_j) with mean (μ_X^*, μ_Y^*), the response probability to the left side of the decision line is

$$p = \int_{-\infty}^{\infty}\int_{-\infty}^{Y=mX+b} \frac{1}{2\pi\sigma_X\sigma_Y} \exp -\frac{1}{2}\left[\left(\frac{X-\mu_X^*}{\sigma_X}\right)^2 + \left(\frac{Y-\mu_Y^*}{\sigma_y}\right)^2\right] dX\,dY$$

Transform to the unit normal distribution:

$$\frac{Y-\mu_Y^*}{\sigma_Y} = \frac{mX + b - \mu_Y^* + m\mu_X^* - m\mu_X^*}{\sigma_Y}$$

Let $Y - \mu_Y^* = y$, $X - \mu_X^* = x$ then,

$$\frac{y}{\sigma_X} = \frac{mx}{\sigma_Y} + \frac{b - \mu_Y^* + m\mu_X^*}{\sigma_X\,\sigma_y} \quad , \quad y = \frac{\sigma_X}{\sigma_y} mx + \left[\frac{b - \mu_Y^* + m\mu_X^*}{\sigma_Y}\right] = m'x + l$$

where $m' = \frac{\sigma_X}{\sigma_Y} m$, $b' = \frac{b - \mu_Y^* + m\mu_X^*}{\sigma_Y}$

Consequently, $$p = \int_{-\infty}^{\infty}\int_{-\infty}^{y=m'x+b'} \frac{1}{2\pi} \exp -\frac{1}{2}(x^2+y^2)\,dx\,dy$$

Figure A shows the geometrical relationship among m', b' and the distance , r, between the decision line $y = m'x + b'$ and $y = m'x$.

$$\sin\theta = \frac{b'm'}{m'b'\sqrt{1+m'^2}} = \frac{r}{b'}$$

$$r = \frac{b'}{\sqrt{1+m'}}$$

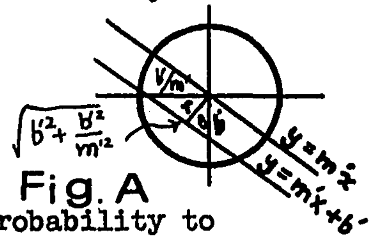


Fig. A

By rotating the decision line $y = m'x + b'$, the probability to the left side of the decision line $p(y < -r)$ can be read from standard normal distribution table.

Appendix 2. Median response latencies in msec. of all stimulus combinations for S JZ.

		X_1	X_2	X_3	X_4	X_5	X_6	X_7
	R	1200	1241	1183	1350	1300	1333	1346
Y_1	L	-	1500	1750	-	1675	1650	1450
	C	1242	1244	1350	1350	1375	1413	1375
	R	1225	1312	1325	1290	1325	1356	1476
Y_2	L	1675	1550	1475	1700	1550	1208	1350
	C	1288	1365	1300	1419	1330	1250	1375
	R	1338	1350	1297	1300	1483	1600	1400
Y_3	L	1625	1450	1561	1475	1512	1391	1200
	C	1300	1492	1400	1350	1350	1590	1413
	R	1169	1500	1319	1525	1675	1812	1550
Y_4	L	1385	1482	1550	1300	1281	1475	1300
	C	1300	1492	1400	1350	1350	1590	1413
	R	1338	1325	1562	1550	1568	1750	-
Y_5	L	1500	1442	1350	1265	1290	1325	1241
	C	1419	1417	1400	1292	1286	1385	1267
	R	1412	1612	1319	1825	-	-	-
Y_6	C	1268	1350	1300	1333	1310	1219	1285
	L	1257	1350	1300	1275	1291	1187	1270
	R	1200	1200	1500	1737	-	-	-
Y_7	L	1250	1275	1253	1250	1308	1254	1150
	C	1250	1275	1278	1292	1313	1254	1190

* $5 < N < 10$
\- $N < 5$

Appendix 2. Median response latencies in msec. of all stimulus combinations for S AG.

		X_1	X_2	X_3	X_4	X_5	X_6	X_7
	R	1441	1507	1541	1640	1570	1590	1775
Y_1	L	-	-	-	-	1900	2000	1675
	C	1446	1507	1547	1653	1600	1620	1700
	R	1625	1612	1570	1580	1821	1700	1700
Y_2	L	-	2350*	-	2075	1716	2100	1931
	C	1557	1650	1580	1622	1788	1750	1883
	R	1480	1550	1755	1975	1800	1900	2106
Y_3	L	2300*	2233	1858	1800	1850	1700	1670
	C	1517	1625	1775	1907	1875	1770	1800
	R	1637	1594	1630	1875	1950	2075	1975
Y_4	L	2100	2025	2250	1900	1725	1625	1662
	C	1783	1663	1738	1875	1741	1719	1775
	R	1682	1575	1869	-	2018	2175	-
Y_5	L	2187	1950	1850	1528	1587	1619	1500
	C	1688	1663	1869	1745	1651	1833	1569
	R	1775	1700	1875	-	2050	1575*	-
Y_6	L	2425	1825	1910	1740	1745	1569	1562
	C	2033	1842	1900	1729	1750	1570	1575
	R	1550	1675	1975	-	2475*	2075	-
Y_7	L	1931	1740	1808	1587	1679	1658	1531
	C	1942	1730	1850	1597	1700	1700	1571

Appendix 2. Median response latencies in msec. of all stimulus combinations for S RW.

		X_1	X_2	X_3	X_4	X_5	X_6	X_7
	R	1115	1058	1105	1187	1164	1131	1256
Y_1	L	-	-	-	-	-	-	1450
	C	1115	1058	1105	1187	1164	1131	1270
	R	1116	1115	1089	1242	1233	1118	1275
Y_2	L	-	-	-	1250*	1292	-	1356
	C	1125	1115	1089	1170	1263	1118	1318
	R	1179	1137	1151	1145	1300	1300	1325
Y_3	L	1325	1358	1262	1350	1332	1325	1329
	C	1204	1179	1152	1150	1318	1306	1327
	R	1261	1231	1182	1237	1300	1275	1316
Y_4	L	1300	1450	1405	1290	1343	1375	1567
	C	1298	1300	1236	1284	1343	1333	1378
	R	1225	1283	1338	1168	1227	1250	1500
Y_5	L	1281	1286	1354	1450	1347	1362	1317
	C	1250	1280	1358	1288	1319	1225	1339
	R	1200	1325	-	-	1250	-	-
Y_6	L	1331	1414	1283	1277	1283	1267	1250
	C	1330	1410	1291	1279	1282	1249	1267
	R	1208	-	-	-	-	-	-
Y_7	L	-	1269	1317	1237	1290	1228	1229
	C	1208	1272	1317	1239	1285	1228	1229

Appendix 2. Median response latencies in msec. of all stimulus combinations for S AK

		X_1	X_2	X_3	X_4	X_5	X_6	X_7
	R	1344	1325	1325	1450	-	-	-
Y_1	L	-	-	-	1800	1392	1358	1319
	C	1344	1325	1325	1775	1383	1358	1319
	R	1300	1275	1279	1550	-	-	-
Y_2	L	-	-	-	1725	1400	1338	1340
	C	1300	1275	1279	1671	1406	1338	1340
	R	1270	1280	1336	1406	-	-	-
Y_3	L	-	-	-	1700	1378	1358	1296
	C	1270	1280	1336	1524	1378	1313	-
	R	1235	1290	1288	1370	-	-	-
Y_4	L	-	-	-	1925*	1367	1286	1283
	C	1235	1290	1288	1425	1368	1286	1283
	R	1268	1367	1365	1625	-	-	-
Y_5	L	-	-	-	1616	1394	1319	1321
	C	1268	1367	1375	1620	1394	1319	1321
	R	1313	1375	1500	1460	-	-	-
Y_6	L	-	-	-	1500	1331	1319	1329
	C	1311	1375	1592	1491	1331	1319	1329
	R	1238	1336	1325	1520	-	-	-
Y_7	L	-	-	-	1600	1385	1221	1250
	C	1238	1336	1325	1541	1390	1221	1251

Appendix 2. Median response latencies in msec. of all stimulus combinations for S OP

	R	1090	1060	1147	1100	1121	1100	1085
Y_1	L	-	-	-	-	-	-	-
	C	1090	1060	1147	1159	1118	1110	1075
	R	1050	1075	1100	1097	1100	1075	1141
Y_2	L	1125	-	-	-	1200	1275	1225
	C	1115	1079	1106	1125	1100	1100	1150
	R	1071	1118	1163	1179	1125	1175	1080
Y_3	L	1250	1100	1450	1287	1287	1200	1300
	C	1100	1118	1194	1200	1166	1183	1181
	R	1108	1187	1256	1308	1225	1350	1125
Y_4	L	1275	1325	1319	1265	1300	1233	1217
	C	1183	1225	1285	1291	1281	1275	1200
	R	1100	-	1150	-	1300	-	-
Y_5	L	1218	1217	1133	1275	1236	1237	1160
	C	1211	1166	1135	1300	1238	1244	1150
	R	1187*	1225	-	-	-	-	-
Y_6	L	1211	1196	1128	1275	1160	1170	1185
	C	1209	1196	1178	1250	1216	1144	1172
	R	-	1179	-	-	-	-	-
Y_7	L	1169	-	1175	1145	1143	1079	1161
	C	1169	1179	1119	1140	1143	1079	1159

Appendix 3

Let decision contour be

$$Y = \left[-\frac{(\mu_{X_5}-\mu_{X_3})\sigma_Y^2}{(\mu_{Y_5}-\mu_{Y_3})\sigma_X^2}\right]X + \frac{\sigma_Y^2(\mu_{X_5}^2-\mu_{X_3}^2)}{2\sigma_X^2(\mu_{Y_5}-\mu_{Y_3})} + \frac{\mu_{Y_5}-\mu_{Y_3}}{2} \quad \ldots (1)$$

and let the straight line connecting points (μ_{X_3}, μ_{Y_3}), (μ_{X_5}, μ_{Y_5}) be

$Y = aX + b$ then,

$$\mu_{Y_3} = a\mu_{X_3} + b \quad \ldots (2), \qquad \mu_{Y_5} = a\mu_{X_5} + b \quad \ldots (3)$$

Solve for a and b

$$a = \frac{\mu_{Y_5}-\mu_{Y_3}}{\mu_{X_5}-\mu_{X_3}} \qquad b = \mu_{Y_3} - \frac{(\mu_{Y_5}-\mu_{Y_3})\mu_{X_3}}{(\mu_{X_5}-\mu_{X_3})}$$

i.e.,

$$Y = \left[\frac{\mu_{Y_5}-\mu_{Y_3}}{\mu_{X_5}-\mu_{X_3}}\right]X + \left[\mu_{Y_3} - \frac{(\mu_{Y_5}-\mu_{Y_3})\mu_{X_3}}{(\mu_{X_5}-\mu_{X_3})}\right] \quad \ldots (4)$$

Solve (1) and (4)

$$X = \frac{\mu_{X_5}+\mu_{X_3}}{2}, \quad Y = \frac{\mu_{Y_5}-\mu_{Y_3}}{2}$$

Let the origin of the coordinate be $\left(\frac{\mu_{X_5}+\mu_{X_3}}{2}, \frac{\mu_{Y_5}+\mu_{Y_3}}{2}\right)$, then

$$Y = Y' + \left(\frac{\mu_{Y_5}+\mu_{Y_3}}{2}\right) = \left[-\frac{(\mu_{X_5}-\mu_{X_3})\sigma_Y^2}{(\mu_{Y_5}-\mu_{Y_3})\sigma_X^2}\right]\left(X' + \frac{\mu_{X_5}-\mu_{X_3}}{2}\right) + \frac{\sigma_Y^2(\mu_{X_5}^2-\mu_{X_3}^2}{2\sigma_X^2(\mu_{Y_5}-\mu_{Y_3})} + \frac{\mu_{Y_5}+}{2}$$

Simplify, $Y' = -X'\left[\frac{(\mu_{X_5}-\mu_{X_3})\sigma_Y^2}{(\mu_{Y_5}-\mu_{Y_3})\sigma_X^2}\right]X'$, $\frac{Y'-0}{\sigma_Y} = \left[\frac{\frac{(\mu_{X_5}-\mu_{X_3})}{\sigma_X}}{\frac{(\mu_{Y_5}-\mu_{Y_3})}{\sigma_Y}}\right]\frac{X'-0}{\sigma_X}$, $Y'' = -\frac{d'_x}{d'_Y}X$

When presentation ratio is not equal to 1, the decision contour on the new coordinate system changed by a constant $\log K/\sigma_Y$

i.e., $Y'' = -\frac{d'_X}{d'_Y}X'' + \frac{\log K}{d'_y}$:

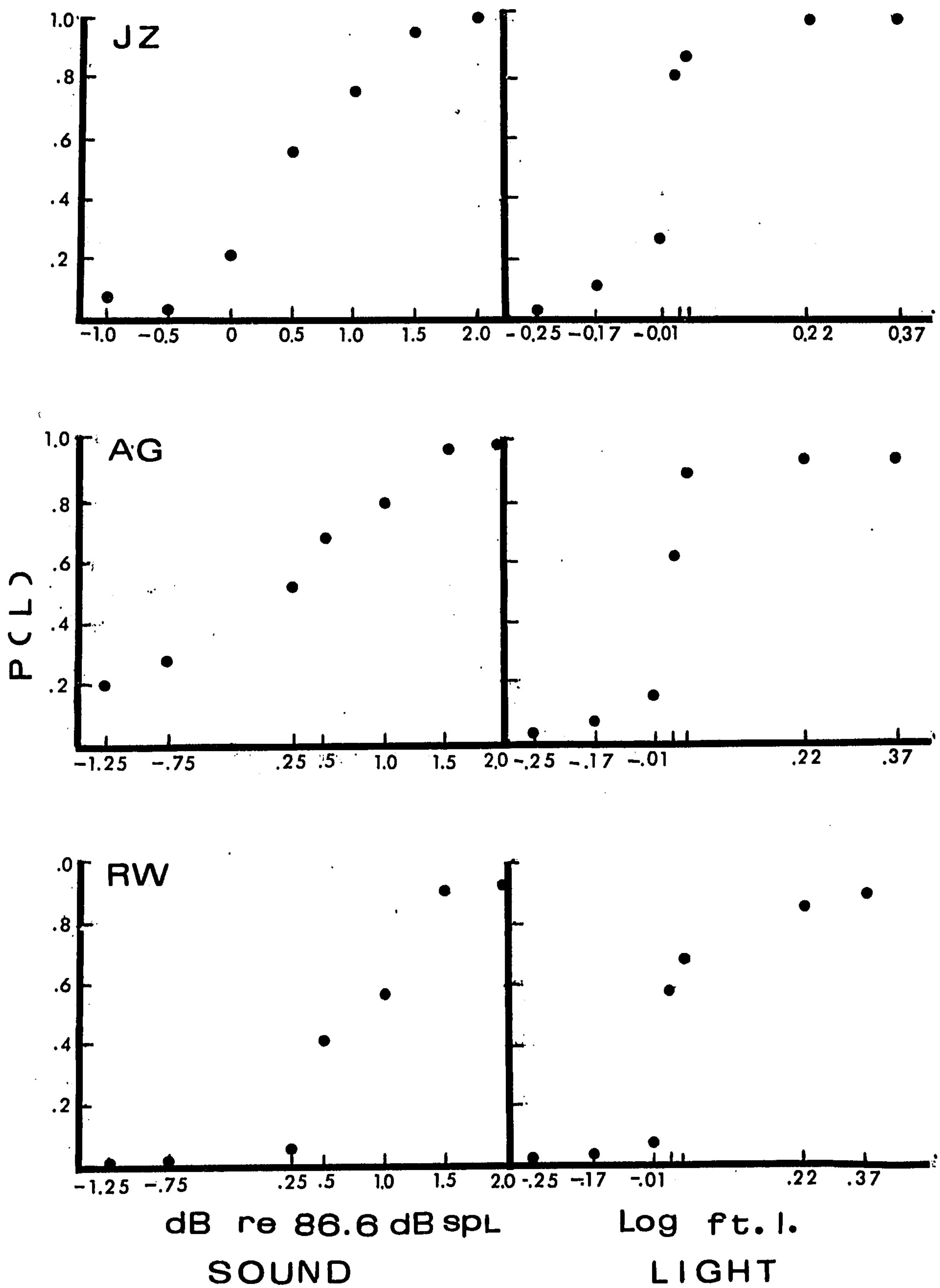
Let $Y = Y' + \frac{\mu_{Y_5}+\mu_{Y_3}}{2}$, $X = X' + \frac{\mu_{X_5}-\mu_{X_3}}{2}$

$$Y = Y' + \frac{(\mu_{Y_5}-\mu_{Y_3})}{2} = \left[-\frac{(\mu_{X_5}-\mu_{X_3})\sigma_Y^2}{(\mu_{Y_5}-\mu_{Y_3})\sigma_X^2}\right]\left(X' + \frac{\mu_{X_5}-\mu_{X_3}}{2}\right) + \frac{\sigma_Y^2(\mu_{X_5}^2-\mu_{X_3}^2)}{2\sigma_X^2(\mu_{Y_5}-\mu_{Y_3})} + \frac{\mu_{Y_5}-\mu_{Y_3}}{2} + \frac{\log K\,\sigma_Y^2}{(\mu_{Y_5}-\mu_{Y_3})}$$

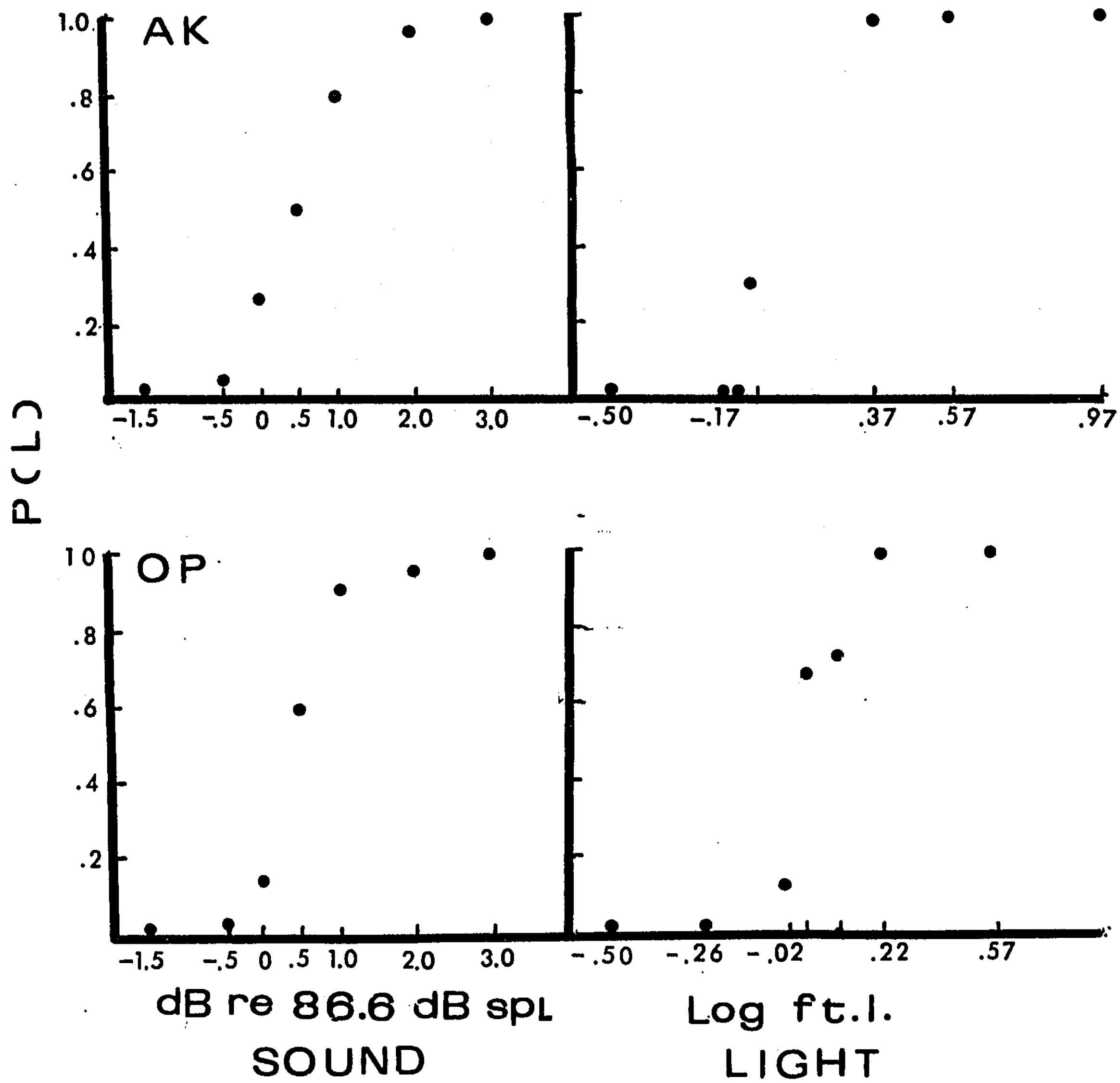
$$\therefore Y' = \left[-\frac{(\mu_{X_5}-\mu_{X_3})\sigma_Y^2}{(\mu_{Y_5}-\mu_{Y_3})\sigma_X^2}\right]X' + \frac{\log K\cdot\sigma_Y^2}{\mu_{Y_5}-\mu_{Y_3}}$$

$$\therefore Y'' = \frac{Y'-0}{\sigma_Y} = -\frac{d'_X}{d'_Y}X'' + \frac{\log K}{d'_Y}, \quad X'' = \frac{X'-0}{\sigma_X}.$$

Appendix 4. Unidimensional Gradients.



Appendix 4. Unidimensional Gradients.



Appendix 5. The numerical values of the theoretical and obtained conditional response probability L based on logarithmic scale.

S JZ		X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y_1	Theoretical p	0.99	0.99	0.98	0.97	0.96	0.81	0.58
	Obtained p	0.89	0.86	0.80	0.84	0.79	0.69	0.72
Y_2	Theoretical p	0.99	0.98	0.93	0.88	0.86	0.59	0.33
	Obtained p	0.85	0.80	0.81	0.73	0.64	0.51	0.48
Y_3	Theoretical p	0.98	0.94	0.79	0.71	0.68	0.34	0.14
	Obtained p	0.71	0.73	0.68	0.52	0.48	0.29	0.37
Y_4	Theoretical p	0.91	0.81	0.57	0.53	0.43	0.14	0.04
	Obtained p	0.51	0.69	0.59	0.34	0.25	0.34	0.28
Y_5	Theoretical p	0.76	0.60	0.30	0.24	0.21	0.05	0.01
	Obtained p	0.42	0.42	0.39	0.23	0.20	0.22	0.09
Y_6	Theoretical p	0.52	0.35	0.13	0.09	0.07	0.01	0.00
	Obtained p	0.29	0.14	0.36	0.18	0.06	0.07	0.03
Y_7	Theoretical p	0.28	0.15	0.04	0.02	0.01	0.00	0.00
	Obtained p	0.16	0.21	0.18	0.18	0.10	0.02	0.01

Appendix 5. The numerical values of the theoretical and obtained conditional response probability of L based on logarithmic scale.

S	AG	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y_1	Theoretical p	0.97	0.96	0.91	0.89	0.88	0.75	0.62
	Obtained p	0.95	1.00	0.94	0.91	0.72	0.87	0.81
Y_2	Theoretical p	0.94	0.91	0.83	0.79	0.78	0.61	0.46
	Obtained p	0.98	0.91	0.94	0.82	0.58	0.72	0.56
Y_3	Theoretical p	0.84	0.78	0.64	0.59	0.57	0.37	0.24
	Obtained p	0.90	0.82	0.86	0.58	0.51	0.32	0.28
Y_4	Theoretical p	0.79	0.71	0.56	0.51	0.49	0.30	0.18
	Obtained p	0.80	0.73	0.77	0.48	0.53	0.23	0.27
Y_5	Theoretical p	0.66	0.57	0.40	0.35	0.33	0.18	0.09
	Obtained p	0.58	0.57	0.36	0.10	0.15	0.25	0.04
Y_6	Theoretical p	0.50	0.41	0.26	0.22	0.20	0.09	0.04
	Obtained p	0.56	0.26	0.31	0.12	0.17	0.09	0.06
Y_7	Theoretical p	0.35	0.25	0.15	0.12	0.11	0.04	0.01
	Obtained p	0.30	0.17	0.19	0.08	0.07	0.07	0.15

Appendix 5. The numerical values of the theoretical and obtained conditional response probability L based on logarithmic scale.

S RW		X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y_1	Theoretical p	0.99	0.99	0.99	0.99	0.99	0.98	0.98
	Obtained p	1.00	1.00	1.00	0.98	1.00	0.98	0.85
Y_2	Theoretical p	0.98	0.97	0.96	0.95	0.95	0.93	0.91
	Obtained p	0.94	1.00	1.00	0.90	0.89	1.00	0.54
Y_3	Theoretical p	0.81	0.78	0.73	0.71	0.70	0.63	0.56
	Obtained p	0.75	0.78	0.80	0.93	0.38	0.45	0.37
Y_4	Theoretical p	0.69	0.65	0.58	0.56	0.55	0.47	0.41
	Obtained p	0.60	0.48	0.60	0.18	0.31	0.27	0.21
Y_5	Theoretical p	0.38	0.34	0.28	0.26	0.25	0.19	0.15
	Obtained p	0.52	0.21	0.19	0.59	0.27	0.21	0.13
Y_6	Theoretical p	0.14	0.11	0.08	0.07	0.07	0.05	0.03
	Obtained p	0.16	0.31	0.07	0.05	0.15	0.04	0.08
Y_7	Theoretical p	0.03	0.02	0.01	0.01	0.01	0.00	0.00
	Obtained p	0.02	0.02	0.00	0.07	0.05	0.00	0.00

Appendix 5. The numerical values of the theoretical and obtained conditional response probability of L based on logarithmic scale.

S	AK	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y_1	Theoretical p	1.00	1.00	0.99	0.99	0.09	0.00	0.00
	Obtained p	1.00	1.00	1.00	0.52	0.02	0.06	0.00
Y_2	Theoretical p	1.00	0.99	0.99	0.97	0.03	0.00	0.00
	Obtained p	1.00	1.00	1.00	0.85	0.02	0.00	0.00
Y_3	Theoretical p	1.00	0.99	0.99	0.95	0.02	0.00	0.00
	Obtained p	1.00	1.00	1.00	0.68	0.00	0.02	0.00
Y_4	Theoretical p	1.00	0.99	0.98	0.92	0.01	0.00	0.00
	Obtained p	1.00	1000	1.00	0.90	0.00	0.00	0.00
Y_5	Theoretical p	1.00	0.99	0.97	0.89	0.00	0.00	0.00
	Obtained p	0.98	1.00	0.93	0.37	0.00	0.00	0.02
Y_6	Theoretical p	1.00	0.99	0.94	0.77	0.00	0.00	0.00
	Obtained p	0.96	1.00	0.90	0.67	0.00	0.00	0.00
Y_7	Theoretical p	1.00	0.99	0.86	0.60	0.00	0.00	0.00
	Obtained p	0.98	0.98	1.00	0.84	0.00	0.00	0.00

Appendix 5. The numerical values of the theoretical and obtained conditional response probability of L based on logarithmic scale.

S	OP	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y_1	Theoretical p	0.99	0.99	0.99	0.99	0.99	0.98	0.75
	Obtained p	0.97	1.00	1.00	1.00	1.00	1.00	0.97
Y_2	Theoretical p	0.99	0.99	0.97	0.96	0.96	0.80	0.24
	Obtained p	0.85	0.97	.094	0.93	0.89	0.90	0.68
Y_3	Theoretical p	0.99	0.99	0.90	0.86	0.84	0.56	0.08
	Obtained p	0.79	0.81	0.81	0.81	0.77	0.78	0.67
Y_4	Theoretical p	0.99	0.96	0.76	0.68	0.65	0.33	0.02
	Obtained p	0.49	0.42	0.50	0.42	0.37	0.22	0.16
Y_5	Theoretical p	0.93	0.62	0.22	0.15	0.13	0.02	0.00
	Obtained p	0.15	0.10	0.19	0.10	0.16	0.06	0.03
Y_6	Theoretical p	0.78	0.35	0.07	0.04	0.03	0.00	0.00
	Obtained p	0.12	0.12	0.04	0.04	0.03	0.04	0.00
Y_7	Theoretical p	0.27	0.03	0.00	0.00	0.00	0.00	0.00
	Obtained p	0.08	0.00	0.03	0.03	0.00	0.00	0.03

Appendix 6. The obtained n in magnitude estimation experiment for sound and light dimensions.

Ss	Dimension	n
JZ	Sound	0.0546
	Light	1.0992
AG	Sound	0.0776
	Light	1.3979
RW	Sound	0.0656
	Light	0.6924
AK	Sound	0.2170
	Light	1.1715
OP	Sound	0.1039
	Light	0.4047

Appendix 7. The numerical values of the theoretical and obtained conditional response probability of L based on power function scale.

S	JZ	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y_1	Theoretical p	0.99	0.99	0.97	0.94	0.93	0.43	0.01
	Obtained p	0.89	0.86	0.80	0.84	0.79	0.69	0.72
Y_2	Theoretical p	0.97	0.94	0.83	0.67	0.65	0.10	0.00
	Obtained p	0.99	0.98	0.92	0.85	0.83	0.22	0.00
Y_3	Theoretical p	0.97	0.94	0.83	0.67	0.65	0.10	0.00
	Obtained p	0.71	0.73	0.68	0.52	0.48	0.29	0.37
Y_4	Theoretical p	0.91	0.84	0.59	0.44	0.41	0.11	0.00
	Obtained p	0.51	0.69	0.59	0.34	0.25	0.34	0.28
Y_5	Theoretical p	0.75	0.64	0.34	0.21	0.19	0.00	0.00
	Obtained p	0.42	0.42	0.39	0.23	0.20	0.22	0.09
Y_6	Theoretical p	0.49	0.36	0.13	0.06	0.05	0.00	0.00
	Obtained p	0.29	0.14	0.36	0.18	0.06	0.07	0.03
Y_7	Theoretical p	0.23	0.14	0.03	0.01	0.01	0.00	0.00
	Obtained p	0.16	0.21	0.18	0.18	0.10	0.02	0.01

Appendix 7. The numerical values of the theoretical and obtained conditional response probability of L based on power function scale.

S	AG	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y_1	Theoretical p	0.93	0.93	0.91	0.91	0.90	0.86	0.79
	Obtained p	0.95	1.00	0.94	0.91	0.73	0.87	0.81
Y_2	Theoretical p	0.87	0.86	0.84	0.83	0.83	0.76	0.66
	Obtained p	0.98	0.91	0.94	0.82	0.58	0.72	0.56
Y_3	Theoretical p	0.69	0.68	0.64	0.63	0.63	0.53	0.42
	Obtained p	0.90	0.82	0.86	0.58	0.51	0.32	0.28
Y_4	Theoretical p	0.60	0.59	0.55	0.54	0.53	0.44	0.33
	Obtained p	0.80	0.73	0.77	0.48	0.53	0.23	0.27
Y_5	Theoretical p	0.41	0.39	0.36	0.35	0.34	0.26	0.17
	Obtained p	0.58	0.57	0.36	0.10	0.15	0.25	0.04
Y_6	Theoretical p	0,22	0.21	0.18	0.17	0.17	0.12	0.07
	Obtained p	0.56	0.25	0.31	0.12	0.17	0.09	0.06
Y_7	Theoretical p	0.08	0.08	0.07	0.06	0.06	0.03	0.02
	Obtained p	0.30	0.17	0.19	0.08	0.07	0.07	0.15

Appendix 7. The numerical values of the theoretical and obtained conditional response probability of L based on power function scale.

S	RW	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y_1	Theoretical p	0.48	0.98	0.98	0.98	0.98	0.97	0.96
	Obtained p	1.00	1.00	1.00	0.98	1.00	0.98	0.85
Y_2	Theoretical p	0.95	0.94	0.93	0.93	0.93	0.90	0.88
	Obtained p	0.94	1.00	1.00	0.90	0.89	1.00	0.54
Y_3	Theoretical p	0.75	0.74	0.71	0.70	0.69	0.63	0.57
	Obtained p	0.75	0.78	0.80	0.93	0.38	0.45	0.37
Y_4	Theoretical p	0.63	0.61	0.58	0.56	0.56	0.49	0.43
	Obtained p	0.60	0.48	0.60	0.18	0.31	0.27	0.21
Y_5	Theoretical p	0.30	0.33	0.29	0.28	0.28	0.23	0.18
	Obtained p	0.52	0.21	0.19	0.59	0.27	0.21	0.13
Y_6	Theoretical p	0.11	0.10	0.09	0.08	0.08	0.06	0.04
	Obtained p	0.16	0.31	0.07	0.05	0.15	0.04	0.08
Y_7	Theoretical p	0.02	0.01	0.01	0.01	0.01	0.00	0.00
	Obtained p	0.02	0.02	0.00	0.07	0.05	0.00	0.00

Appendix 7. The numerical values of the theoretical and obtained conditional response probability of L based on power function scale.

S	AK	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y_1	Theoretical p	0.99	0.99	0.99	0.99	0.08	0.00	0.00
	Obtained p	1.00	1.00	1.00	0.52	0.02	0.06	0.00
Y_2	Theoretical p	0.99	0.99	0.99	0.98	0.05	0.00	0.00
	Obtained p	1.00	1.00	1.00	0.85	0.02	0.00	0.00
Y_3	Theoretical p	0.99	0.99	0.99	0.97	0.03	0.00	0.00
	Obtained p	1.00	1.00	1.00	0.68	0.00	0.02	0.00
Y_4	Theoretical p	0.99	0.99	0.98	0.95	0.01	0.00	0.00
	Obtained p	1.00	1.00	1.00	0.90	0.00	0.00	0.00
Y_5	Theoretical p	0.99	0.99	0.96	0.91	0.00	0.00	0.00
	Obtained p	0.98	1.00	0.93	0.37	0.00	0.00	0.02
Y_6	Theoretical p	0.99	0.94	0.80	0.64	0.00	0.00	0.00
	Obtained p	0.96	1,00	0.90	0.67	0.00	0.00	0.00
Y_7	Theoretical p	0.81	0.49	0.21	0.10	0.00	0.00	0.00
	Obtained p	0.98	0.98	1.00	0.84	0.00	0.00	0.00

Appendix 7. The numerical values of the theoretical and obtained conditional response probability of L based on power function scale.

S	OP	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y_1	Theoretical p	0.99	0.99	0.99	0.99	0.98	0.90	0.13
	Obtained p	0.97	1.00	1.00	1.00	1.00	1.00	0.97
Y_2	Theoretical p	0.99	0.99	0.96	0.93	0.92	0.67	0.02
	Obtained p	0.85	0.97	0.94	0.93	0.89	0.90	0.68
Y_3	Theoretical p	0.99	0.98	0.89	0.84	0.82	0.48	0.00
	Obtained p	0.79	0.81	0.81	0.81	0.77	0.78	0.67
Y_4	Theoretical p	0.97	0.85	0.52	0.42	0.38	0.10	0.00
	Obtained p	0.49	0.42	0.50	0.42	0.37	0.22	0.26
Y_5	Theoretical p	0.87	0.62	0.24	0.17	0.15	0.02	0.00
	Obtained p	0.15	0.10	0.19	0.10	0.16	0.06	0.03
Y_6	Theoretical p	0.63	0.31	0.06	0.04	0.03	0.00	0.00
	Obtained p	0.12	0.12	0.04	0.04	0.03	0.04	0.00
Y_7	Theoretical p	0.05	0.00	0.00	0.00	0.00	0.00	0.00
	Obtained p	0.08	0.00	0.03	0.03	0.00	0.00	0.03

References

Acton, F.S. Analysis of Straight-Line Data. New York: Wiley, 1959

Audley, R.J. and Mercer, A. The relation between decision time and relative response frequency in a blue-green discrimination The British Journal of Mathematical and Statistical Psychology, 1968, 21, 183-192.

Audley, R.J. and Pike, A.R. Some alternative stochastic models of choice. The British Journal of Mathematical and Statistical Psychology, 1965, 18, 207-225.

Bliss, C.L. The calculation of the dosage-mortality curve. Ann. Appl. Biol., 1935, 22, 134-167.

Boneau, C.A. and Cole, J.L. Decision theory, the pigeon, and the psychophysical function. Psychological Review, 1967, 74, 123-135.

Butter, C.M. Stimulus generalization along one and two dimensions in pigeon. Journal of Experimental Psychology, 1963, 65, 339-346.

Chase, S. Selectivity in multidimensional stimulus control. Journal of Comparative and Physiological Psychology, 1968,66, 787-792.

Chase, S. and Heinemann, E.G. Choise based on redundant information:An analysis of two-dimensional stimulus control. Journal of Experimental Psychology, 1972, 93, 161-175.

Cross, D.V. Metric properties of multidimensional stimulus generlization. In D.I. Mostofsky (Ed.), Stimulus Generalization. Stanford: Standford University Press, 1965.

Cross, D.V. and Lane, H.L. On the discriminative control of a concurrent response. Journal of the Experimental Analysis of Behavior, 1962, 5, 487-496.

Edwards, W. Optimal strategies for seeking information: Models for statistics, choice reaction time, and human information processing. Journal of Mathematical Psychology, 1965, 2, 312-329.

Finney, D.J. Probit Analysis. London: Cambridge University Press, 1964.

Green, D.M. and Swets, J. Signal Detection Theory and Psychophysics. New York: Wiley, 1966.

Guttman, N. and Kalish, H.I. Discrimination and stimulus generalization. Journal of Experimental Psychology, 1956, 51, 79-88.

Guttman, N. The pigeon and the spectrum and other perplexities. Psychological Report, 1956, 2, 449-460.

Heinemann, E.G. and Rudolph, R.L. The effect of discriminative training on the gradients of stimulus generalization. American Journal of Psychology, 1963, 76, 653-658.

Heinemann, E.G., Avin, E., Sullivan, M.A. and Chase, S. Analysis of stimulus generalization with a psychophysical method. Journal of Experimental Psychology, 1969, 80, 215-224.

Heinemann, E.G. and Chase, S. On the form of stimulus generalization curves for auditory intensity. Journal of Experimental Psychology, 1970, 84, 483-486.

Hull, C.L. Principle of Behavior. New York: Appleton-Century-Crofts, 1943.

Jenkins, H.M. and Harrison, R.H. Effect of discrimination training on auditory generalization. Journal of Experimental Psychology, 1960, 59, 246-253.

Johnson, D.F. Determiners of selective stimulus control in the pigeon. Journal of Comparative and Physiological Psychology, 1970,70, 298-307.

Jones, J. Stimulus generalization in two and three dimensions. Canadian Journal of Psychology, 1962, 16, 23-36.

Kamin, L.J. Selective association and conditioning. In N.J.Mackintosh and W.K. Honig (Eds.), Fundamental Issues in Associative Learning. Halifax: Dalhousie University Press, 1969.

LaBerge, D. Generalization gradients in a discrimination situation. Journal of Experimental Psychology, 1961, 62, 88-94.

LaBerge, D. A recruitment theory of simple behavior. Psychometrika, 1962, 27, 375-396.

Laming, D.R.J. Information Theory of Choice-Reaction Time. New York: Academic Press, 1968.

Lashley, K.S. The mechanism of vision: XV. Preliminary studies of the rat's capacity for detail vision. Journal of Genetic Psychology, 1938, 18, 123-193.

Lashley, K.S. An examination of the "continuity theory" as applied to discrimination learning. Journal of Genetic Psychology, 1942, 26, 241-265.

Lashley, K.S. and Wade, M. The Pavlovian theory of generalization. Psychological Review, 1946, 53, 72-87.

Luce, R.D. Discrimination. In R.D. Luce, R. R. Bush, and E. Galanter (Eds.), Handbook of Mathematical Psychology, Vol. I. New York: Wiley, 1963.

Mackintosh, N.J. Selective attention in animal discrimination learning. Psychological Bulletin, 1965, 64, 124-150.

Newman, F.L. and Baron, M.R. Stimulus generalization along the dimension of angularity. Journal of Comparative and Physiological Psychology, 1965, 60, 59-63.

Pavlov, I.P. Conditioned Reflexes. New York: Oxford University Press, 1927.

Pickett, B.M. Response latency in a pattern perception situation. Acta Psychologica, 1967, 27, 160-169.

Pierrel, R. and Sherman, J.G. Generalization of auditory intensity following dicrimination training. Journal of the Experimental Analysis of Behavior, 1960, 3, 313-327.

Pierrel, R. and Murray, C.S. Some relationships between comparative judgement, confidence, and decision-time in weight-lifting. American Journal of Psychology, 1963, 76, 28-38.

Purtle, R.B. and Newman, F.L. Multidimensional stimulus generalization of a tactile response along the dimension of angularity and texture. Journal of Experimental Psychology, 1969, 81, 566-571.

Reynolds, G.S. Attention in the pigeon. Journal of the Experimental Analysis of Behavior, 1961, 4, 203-208

Sekuler, R.W. Signal detection, choice response times, and visual backward masking. Canadian Journal of Psychology, 1965, 19, 118-132.

Sekuler, R.W. Choice times and detection with visual backward masking. Canadian Journal of Psychology, 1966, 20, 34-42.

Shepard, R.N. Attention and the metric structure of the stimulus. Journal of Mathematical Psychology, 1964, 1, 54-87.

Skinner, B.F. Science and Human Behavior. New York: Macmillian, 1953.

Spence, K.W. Behavior Theory and Conditioning. New Haven: Yale University Press, 1956.

Smith, E.E. Choice reaction times. Psychological Bulletin, 1968, 69, 77-110.

Smith, P.T. Cost, discriminability and response bias. The British Journal of Mathematical and Statistical Psychology, 1968, 21, 35-60.

Stone, M. Models for choice-reaction time. Psychometrika, 1960, 25, 251-260.

Stevens, S.S. On the psychophysical law. Psychological Review, 1957, 64, 153-181.

Stevens, S.S. The psychophysics of sensory function. In W.A. Rosenblith (Ed.), Sensory Communication. New York: MIT Press and Wiley, 1961, 1-33.

Tanner, W.P., Jr. Physiological implications of psychophysical data. In A. Swets (Ed.), Signal Detection and Recognition by Human Observer. New York: Wiley, 1964, 351-365.

Terrace, H.S. Stimulus control. In W.K. Honig (Ed.), Operant Behavior: Areas of Research and Application. New York: Appleton-Century-Crofts, 1966.

Thomas, E. Sufficient condition for monotone hazard rate- An application on latency-probability curve. Michigan Mathematical Psychology Program, 1970.

Wald, A. Sequential Analysis. New York: Wiley, 1947.

Weinstock, S., Robbins, D. and Chen, W. Successive brightness discrimination in T-maze. Journal of Experimental Psychology, 1973, in press.

Wollen, K.A. Relationships between choice time and frequency during discrimination training and generalization tests. Journal of Experimental Psychology, 1966, 66, 474-484.