THURSTONE'S DISCRIMINAL PROCESSES FIFTY YEARS LATER

R. DUNCAN LUCE

HARVARD UNIVERSITY

Four issues are discussed concerning Thurstone's discriminal processes: the distributions governing the representation, the nature of the response decision rules, the relation of the mean representation to physical characteristics of the stimulus, and factors affecting the variance of the representation. A neural schema underlying the representation is proposed which involves samples in time of pulse trains on individual neural fibers, estimators of parameters of the several pulse trains, samples of neural fibers, and an aggregation of the estimates over the sample. The resulting aggregated estimate is the Thurstonian representation. Two estimators of pulse rate, which is monotonic with signal intensity, are timing and counting ratios and two methods of aggregation are averaging and maximizing. These lead to very different predictions in a speed-accuracy experiment; data indicate that both estimators are available and the aggregation is by averaging. Magnitude estimation data are then used both to illustrate an unusual response rule and to study the psychophysical law. In addition, the pattern of variability and correlation of magnitude estimates on successive trials is interpreted in terms of the sample size over which the aggregation takes place. Neural sample size is equated with selective attention, and is an important factor affecting the variability of the representation. It accounts for the magical number seven phenomenon in absolute identification and predicts the impact of nonuniform distributions of intensities on the absolute identification of two frequencies.

Key words: Thurstonian scaling model, discriminal process, neural schema for auditory psychophysics, attention bands, attention as sampling, absolute identification, magnitude estimation, frequency discrimination.

This being the twenty-fifth anniversary of the L. L. Thurstone Psychometric Laboratory, I wished to relate my remarks to one or another aspect of Thurstone's scientific contributions—to either factor analysis, test theory, attitude measurement, or psychophysical analysis. With that list of topics and my interests, the choice was uniquely determined. But the choice is fortuitous since not only is this an anniversary of the Laboratory, but it also happens to be the fiftieth anniversary of Thurstone's three basic papers in psychophysics: "Psychophysical analysis", which appeared in the 1927 volume of *The Ameri-*

Table 1. A Comparison of Terms

Thurstone's description	Thurstone's term	Current term
Psychological values of a psychophysical stimulus	Discriminal process	Decision random variable; representation of signal
Fluctuations among discriminal processes of a signal	Discriminal dispersion	Standard deviation of random variable
Relation between the typical psychological value and the physical measure	Psychological scale	Psychophysical law

can Journal of Psychology, and "A law of comparative judgment" and "Three psychophysical laws", both of which were in the 1927 volume of the *Psychological Review.*

My goal, then, is to take stock of our progress in psychophysical analysis and to see how Thurstone's program has developed. I will not examine all of the other uses to which the idea of discriminal process has been put, both by Thurstone and others, nor will I describe and evaluate the various refined procedures for scaling so ably treated by Torgerson [1958] and by Bock and Jones [1968]. Rather, I shall accept the 1927 papers at face value as an approach to psychophysics, no more, no less, and ask how we have come to interpret and to modify these ideas in the ensuing fifty years.

Since some of the terms have changed over the years, let me remind you briefly of the ideas and terms (see Table 1). Thurstone's basic idea was the essence of simplicity: Each time a signal is presented, it is transduced in the nervous system into what he called a discriminal process. Few if any psychologists now use that term; rather, the random variable language is pervasive. Of course, we frequently speak of the value of the random variable associated with a particular signal presentation as a psychological, or internal, representation of the signal. The representation captures some relevant aspect of the signal, but does so in some language of the nervous system and only imperfectly because of variability inherent in the transduction, transmission, and processing. Decisions about a signal or about its relation to other signals are assumed to be based entirely upon these representations.

This conception of internal representations of signals is so simple and so intuitively compelling that no one ever really manages to escape from it. No matter how one thinks about psychophysical phenomena, one seems to come back to it. Our concern in recent years has not been so much to replace it as to make it specific—to transform it from a theoretical schema into a psychophysical theory.

Four Questions About the Representation

Four questions must ultimately be resolved about Thurstone's discriminal processes before one can be said to have a fully articulated psychophysical theory.

Question 1. What is the family of distributions governing the random variables that represent signals?

Thurstone assumed normal distributions. But recall his caution: "The normal probability curve has been so generally abused in psychological and educational measurement that one has reason to be fearful of criticism from the very start in even mentioning it. The only valid justification for bringing in the probability curve in this connection is that its presence can be experimentally tested." [Thurstone, 1927a, p. 373] He was equally explicit that he would not attempt to derive it from other considerations. "I shall not assume that the process by which an organism differentiates between two stimuli is either psychic or physiological. . . . I shall try not to disturb the main argument with systematic irrelevances or with my personal notions regarding the psychic or physiological nature of the psychophysical judgment." [Thurstone, 1927a, p. 368]

History has not followed him closely on either of these positions. Testing the assumption of normality has proved rather less easy and direct than he seemed to imply and many of us feel that it is quite profitable to take account of physiological results. We will return to both points later.

Question 2. What is the nature of the decision rules that lead to responses?

Although Thurstone invoked a decision rule—namely, when discriminating between two signals, select as the larger the one having the larger representation—he hardly mentioned this fact. Later when others generalized his model to absolute identification and category judgments involving two or more signals, they simply assumed that the range of the several random variables is partitioned into as many intervals as there are responses. Again, they made little of this choice. In particular, both he and many of his followers failed to emphasize that the boundaries of the response categories might be manipulated experimentally.

In contrast, the theory of signal detectability, which grew out of developments concerning sonar and radar detection during World War II, clarified the significance of the response criterion. In addition, that theory postulated the likelihood ratio to be the decision random variable, and showed that under some assumptions its logarithm is normally distributed under the presentation of either signal in noise or noise alone. Formally, such a model is Thurstonian, but with a variable response criterion. The empirically important discovery resulting from these ideas is the fact that subjects really can vary their response criterion over a very large range, thereby sweeping out a whole locus of error

pairs. This is the same error tradeoff that a statistician considers when making inferences involving two hypotheses. The statistician calls it the power of the test; the psychophysicist calls it the receiver operator characteristic (ROC).

The tradeoff of errors has proved to be an exceedingly powerful tool, as I shall illustrate below. To my mind, it is the major residue of signal detectability theory in psychology. The interpretation of the random variables in terms of likelihood ratio, although much stressed by some advocates of the theory, has not proved to be particularly important, and it is virtually impossible to generalize the likelihood formulation to responses to more than two signals.

Was Thurstone aware of the ROC tradeoff? He neither discussed it explicitly nor emphasized its importance. But he did write: "The writer has found experimentally that the normal probability curve was not applicable for certain stimuli. In most of the experiments the distributions are reasonably close to normal." [Thurstone, 1927a, p. 373] How did he draw these conclusions? He did not say. Is it possible that he generated ROC data and plotted it in z-scores, as we do today? I find it difficult to understand the quote otherwise; but I also find it remarkable that he failed to write about it.

Beyond variable boundaries, there are other aspects to the question about decision rules. For example, to encorporate the widely used methods of magnitude estimation, magnitude production, and sensory matching into psychophysical theory, we must also arrive at decision rules for these designs. I shall pursue this at length below.

Question 3. What is the psychophysical law, i.e., how does the central tendency of the representation depend upon the physical characteristic of the signal that is being varied?

Thurstone was much concerned with this question, as have been many others from Fechner through Stevens. Thurstone pointed out explicitly that the answer to this question may be quite independent of the nature of the variability of the random variable, a point much emphasized by Stevens. However, in accounting for some data a change in one can often compensate for a change in the other. For example, Thurstone explicitly contrasted Fechner's logarithmic law with a special case of Stevens' power law, namely, the square; and he showed how each is compatible with Weber's law provided that one makes the correct assumption about how standard deviation varies with intensity.

Stevens [1957, 1975] argued that an easy way to get at the psychophysical function is simply to ask subjects to report numbers in proportion to the subjective sensations arising from signals, and then to average these numbers. This seems to suggest the following simple decision rule: respond with a number proportional to the value of the internal representation of the signal. If it were that simple, then magnitude estimation data would provide us not only with the psychophysical function, but with the entire distribution of Thurstone's random variables. As we shall see, this response rule is overly simple, but it may not be terribly far from the truth.

Question 4. Upon what does the standard deviation of the random variable depend?

Thurstone explicitly acknowledged that the standard deviation might depend upon whatever physical attribute is being varied; however, the special Case V of constant standard deviations attracted a great deal of attention both because of its mathematical tractability and because of its close relation to Fechner's postulate. Aside from that, it is implicit in Thurstone's discussion that a unique random variable, or process, is associated with each signal, independent of the experimental context within which the signal appears. It does not matter whether the presentation is embedded in a discrimination or an absolute identification or a magnitude estimation experiment, it will always be represented by the same random variable. Apparently it was not until some time after World War II that it was fully recognized how erroneous this implicit assumption actually is.

Miller [1956] cited psychophysical data showing that responses in an absolute identification experiment could not possibly arise from the Thurstonian model in which a unique random variable is associated with each signal. As this result is extremely important, let me remind you of the details. If one carries out a two-signal absolute identification of loudness, it suffices to separate the two signals by 5 dB in order to achieve perfect identifiability. Thus, the two representations cannot overlap to any substantial degree. So, if both the Thurstonian model and the implicit assumption were true, we should be able to identify without error 20 signals equally spaced in dB over a range of 95 dB. The facts are otherwise. A total of 7 signals, adjacent ones being separated by a little more than 15 dB, is about the limit of perfect identifiability for a range of 95 dB, as can be seen in Garner's [1953] data reproduced in Fig. 1.

Another way to view the same phenomenon is to fix the number of signals and vary the range. Braida and Durlach [1972] proceeded in this manner, and found that performance improves linearly with range up to about 15 or 20 dB, at which point it begins to slacken off just when the Thurstonian model predicts improvement. Such data will be shown below (Fig. 15).

Clearly, Thurstone's implicit assumption of a unique representation must go. In some sense the representation must vary with the experimental context. Both Durlach and Braida [1969] and, independently, Gravetter and Lockhead [1973] have pointed out that one can fit the variable range data quite accurately by assuming that the Thurstonian variance grows linearly with the square of the range in dB. In my view, this assumption is *ad hoc* and implausible, and thus demands a deeper explanation.

Let us now turn to what is known about the answers to these four questions: the form of the distribution, the nature of the response rule, the form of the psychophysical law, and the factors that affect the variability of the representation. In attempting to formulate answers, I find it helpful first to outline a neural schema that D. M. Green and I have evolved as possibly underlying the Thurstonian representation.
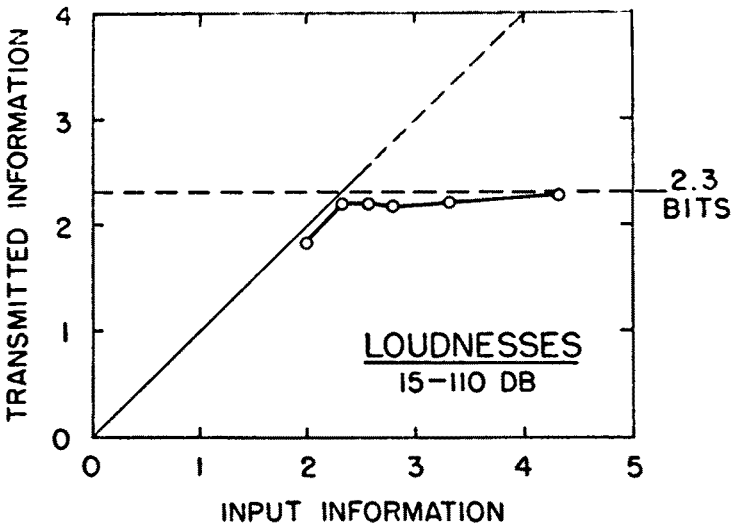
FIGURE 1

Information transmitted vs. information input (= log₂ number of signals) in absolute identification of auditory intensity. Signals were equally spaced in *dB* over a 95 *dB* range and they were presented equally often. The data are from Garner (1953); this is Fig. 2 of Miller (1956).

A Neural Schema

Any sensory transducer has the role of transforming sensory signals into the internal language of the nervous system, presumably into neural pulse trains. In at least the case of the eye and the ear, the neural structure immediately following the transduction is an elaborate system of parallel nerve fibers. In the case of the human ear, there are some 30,000 peripheral fibers. One can examine the activity of these fibers in at least two quite different ways.

If one looks at the whole bundle of fibers, each signal activates some characteristic subset of the fibers. Theories which assume that the central nervous system (CNS) identifies the signal by identifying the active subset of fibers are called place theories. One should note that it is not completely trivial to devise mechanisms to identify more or less arbitrary subsets of 30,000 fibers—that problem has blocked the development of satisfactory readers for computers.

If instead one examines an individual fiber, one can ask both which signals activate it and, when it is active, what is encoded by the neural pulse train. In audition, the plot of the firing rate of a single fiber as a function of the intensity and frequency of a pure tone is known as the tuning curve of that fiber [Kiang, 1965]. Typically, at any frequency there is a relatively narrow range of intensities, say 15 to 20 dB, for which the fiber is neither at its resting level nor its maximum firing level.

Turning to the code, for pure tones of less than 2000 Hz, Rose *et al.* [1967]

have established that the pulse train on a single fiber encodes both frequency and intensity information. This surprising result opens up the possibility of statistical theories of a character quite different from place theories. The only difficulty is that in order to achieve a quality of statistical inference needed to account for psychophysical results a very long sample of pulses is needed from a single fiber. Not only is this unsatisfactory because signals do not usually stay put for very long, but it would surely prove lethal to the organism. Perhaps, then, some of the role of the parallel structure of nerve fibers is to increase the sample size on which inferences are made without introducing excessive delays. According to such a view, the place information is simply an artifact of two factors: the replication of fibers in order to get adequate samples quickly and the inability of a single fiber to span the full dynamic range of intensities and frequencies.

Assuming this point of view, three key ideas must be developed and studied. First, consider some feature of the signal, say intensity or frequency, which is encoded in the pulse train. Some *estimator* of that parameter must be used to decode it from the sample of pulses observed on an individual fiber. For example, single unit records show that pulse rate increases with the intensity of a pure tone. So when considering questions about intensity, the CNS is forced to use estimators of rate. Two extreme ones, which we have been investigating, are called *counting* [McGill, 1967] and *timing* [Luce & Green, 1972; Uttal and Krissof, 1965] estimators. In each case, the name corresponds to the random variable observed by the CNS: for the counting estimator, the observable is the count of pulses obtained in a fixed period of time; for the timing one, it is the time required to complete a fixed count. In each case the rate is estimated by the ratio of counts to time. Both obviously require that in some sense the CNS be able to count pulses and to time intervals, thus suggesting that maybe both estimators are available under the proper conditions.

A second feature of the schema is that these individual estimates must be *aggregated* over some sample of fibers in order to upgrade their statistical quality. If the samples are at all large, as we have reason to believe, then the resulting distribution is largely determined by the aggregation scheme employed and not much influenced by the exact nature of the estimator used. Such decoupling is to our advantage.

A third feature of the schema is the *sample* over which the aggregation is carried out. One's initial inclination is to assume that all of the fibers activated by the signal are in the sample, but the data have forced us to conclude otherwise. Indeed, after the fact, it seems quite plausible to me that it is unlikely for a system to evolve in which all peripheral fibers are monitored simultaneously. That would entail a great deal more computational capacity—higher-order neural cells—than if the monitoring is more selective. In any event, we have been led to suppose that there is some mechanism that focuses attention, i.e., a mechanism that monitors fully just one group of substantially

identical fibers, so that signals activating this group are represented as accurately as possible, while keeping tabs on all other groups of similar fibers only sufficiently well to attract attention when one of them is activated. When that happens, the CNS may decide—we know not how—to shift attention to that group. By "shift attention" I simply mean to monitor it fully. In this view of sensory information processing, attention is treated as identical to increasing the sample size underlying the inference being made by the CNS.

In sum, then, our neural schema involves four ingredients:

1. a system of parallel peripheral fibers of which only a relatively small subset is activated by a particular signal;

2. an estimator of a parameter of the pulse train which operates on a time sample of that pulse train;

3. a sample of fibers whose size depends upon where attention is focussed; and

4. a scheme for aggregating over the sample of fibers the estimates obtained from individual fibers.

I shall treat the aggregated estimate as the random variable in Thurstone's theory. As will soon be evident, we use our neural schema as a means for arriving at experiments which help to answer the four questions originally posed about the Thurstonian representation.

The Form of the Distribution

The first decision one ought to make about the representation of a signal is whether it is a discrete or continuous random variable. Of course, strictly speaking it must be discrete, but the issue is whether or not the discreteness is sufficiently coarse—well above the molecular level—so that it shows up in some behavioral observations. The theory of such coarse representations is called neural quantum theory, and there has been much controversy about it. (Several of the major publications are: Corso, 1956; Krantz, 1969; Luce, 1963; Norman, 1964; Stevens, 1972; Stevens, et al., 1941; Swets, 1961.) Although the consensus seems to be against the discrete theory, I do not feel that the issue has really been settled, in part because no decisive experiment has yet been devised.

Despite my nagging doubts, I shall nonetheless proceed as if the random variables are continuous. Having accepted that, should we assume with Thurstone that they are normally distributed? Turning back to the neural schema just outlined, the answer centers primarily on the nature of the aggregation across fibers. If the CNS averages the estimates from the individual fibers and if these estimates are of comparable magnitude, the Central Limit Theorem tells us that the resulting estimate is approximately normally distributed.

Is there any serious alternative? A priori, there are many possibilities, of which one of the most inviting is to suppose that decisions are based on an extreme observation, in particular, the largest of the several estimates. Accord-

ing to classical work, summarized in detail by Gumbel [1958], if the estimates from single fibers are identically and independently distributed with the upper tails being asymptotically exponential, then the limit distribution of the maximum value is the double exponential,

$$\exp\left[-e^{-\alpha(x-\beta)}\right].$$

I find this model attractive, in part, because it is known [McFadden, 1974; Yellott, 1977] that a Thurstonian model with a shift family of distributions is equivalent to the choice axiom [Luce, 1959] if and only if the Thurstonian random variables are distributed as the double exponential.

Our reasoning so far leaves us with two ideas for the estimator, namely, counting and timing, and two ideas for the aggregation, namely, averaging and maximizing. The question then is how to put them together into an experiment capable of selecting among the alternatives. Initially, one is not optimistic because it has been shown by computations that the double exponential and normal distributions do not differ greatly for appropriate choices of parameters. Despite this, I will show that there is a simple way to decide.

Consider a Yes-No design in which, on each trial, one of two levels of intensity is presented, denoted s and n, with s more intense than n. The subject is to say either Yes if he thinks it is s or No if he thinks it is n, and let the signals be terminated when the subject responds. Green and Luce [1973] argued that if an experimenter-controlled time deadline is imposed on all trials of such an experiment, it is optimal for the subject to establish an observation time, just short enough so that most responses beat the deadline, and to base decisions on the number of pulses occurring in that interval. However, if the deadline is imposed only on s trials, then it is optimal to invoke the timing estimator. Later, Wandell [1977] pointed out an interesting asymmetry of prediction: when the deadline is imposed only on n trials counting, not timing, is optimal. The reasoning in all three cases rests entirely on seeing which estimator manages both to satisfy the deadline and achieves the larger sample size.

The two kinds of estimators and the two methods of aggregation lead to different predictions about both reaction times and the error tradeoff embodied in the ROC curve. Assuming Poisson pulse trains, the predictions [Green & Luce, 1973; Wandell & Luce, Note 1] are summarized in Table 2. Let me point out four important facts about these predictions. First, by choosing different coordinates for each estimator and each method of aggregation, each ROC can be made into a straight line and consequently one can speak of its slope. Second, for the counting estimator, the slope of the ROC is always less than unity, and for the timing estimator, it is always greater than unity. Third, for the timing estimator, if the deadline is varied, the plot of the mean reaction time calculated over n trials is a linear function of the mean reaction time calculated over s trials. And fourth, for the timing estimator, the slope of the mean reaction time line is identical to that of the ROC in the coordinates appropriate to the aggregation method being used.

Table 2. Slopes of ROC Curve and Mean Reaction Time Relations. All ROC curves are linear in the coordinates shown; for timing, mean reaction time to noise is linear with mean reaction time to signal. The Poisson pulse rates are $\mu(s)$ and $\mu(n)$, with $\mu(s) > \mu(n)$.

| | Coordinates | | Slope | |
Estimator	Average (Normal)	Maximum (Double Exponential)	ROC	Mean RT
Counting	$z(p)$	$\ln \ln 1/(1-p)$	$[\mu(n)/\mu(s)]^{1/2}$	—
Timing	$z(p)$	$-\ln \ln 1/p$	$\mu(s)/\mu(n)$	$\mu(s)/\mu(n)$

Turning to the auditory intensity data of Green and Luce [1973], consider first those obtained when the deadline was applied to all trials. This is plotted in normal coordinates in Fig. 2 and in double exponential coordinates in Fig. 3. Note that both slopes are less than unity, as they should be for the counting estimator, and that it is impossible by eye to decide which fit is the better. Next, the data obtained when the deadline is applied only on s trials is shown in normal coordinates in Fig. 4 and in double exponential coordinates in Fig. 5. Both slopes are appreciably greater than unity, as they should be for the timing estimator. The evidence, therefore, supports the belief that, no matter how the estimates are aggregated, applying the deadline to all trials induces the counting estimator and applying it only to s trials induces the timing estimator.
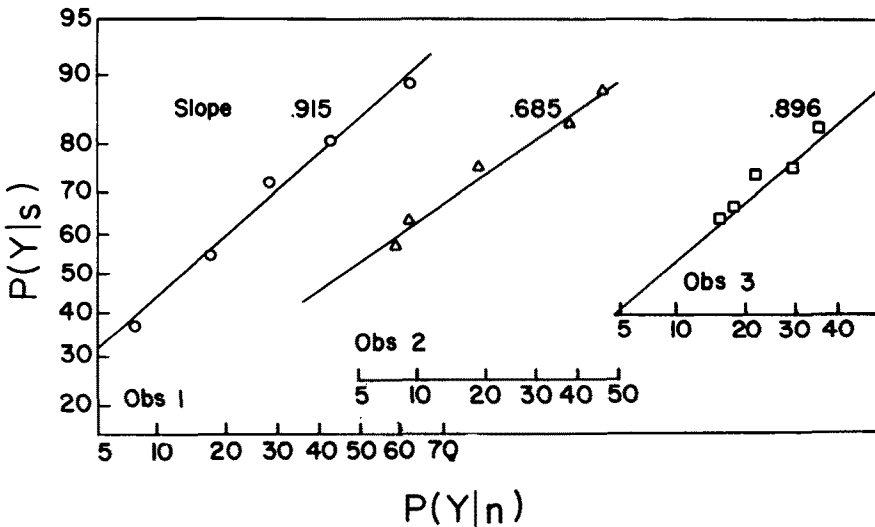

FIGURE 2

ROC curves, plotted in z-score coordinates, for a 600 msec deadline on all trials and a signal to noise ratio $10 \log P/N_0 = 20$. Five different payoff matrices were used. This is Fig. 4 of Green and Luce (1973).
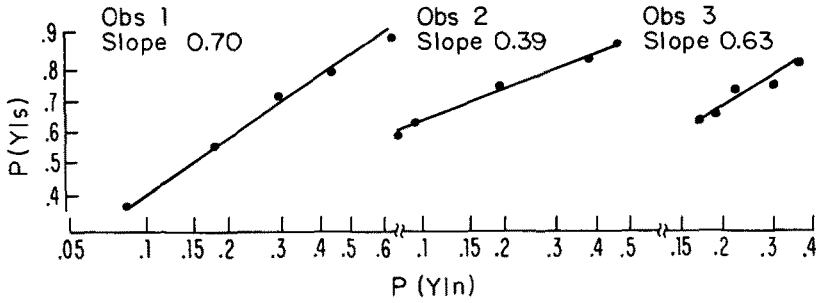
FIGURE 3

These are the same data as in Fig. 2 but plotted in the coordinates ln ln $1/(1 - p)$. This is Fig. 6 of Wandell and Luce (1977).

The corresponding auditory data for the deadline applied just to *n* trials were not collected. But Wandell [1977], using visual intensity, replicated the whole experiment including this third condition. He not only found results paralleling the auditory ones for the *sn* and *s*-deadline conditions, but for the *n*-deadline condition the slopes were well below unity, suggesting the counting estimator as predicted.

To select between the two aggregation schemes we turn to the reaction time data for the *s*-deadline condition, shown in Fig. 6. Observe that the slopes are nearly identical to those of Fig. 4 and about half the value of those of Fig.
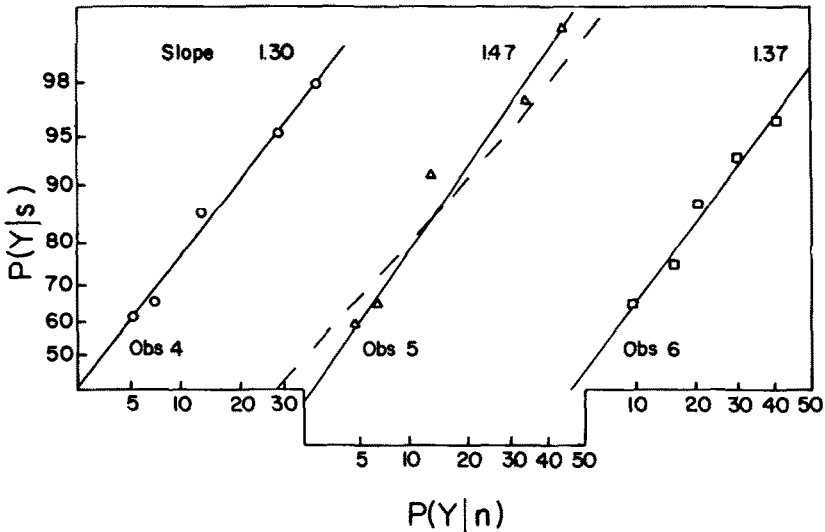

FIGURE 4

ROC curve, plotted in *z*-score coordinates, for a 600 msec deadline on *s* trials only. This is Fig. 9 of Green and Luce (1973).
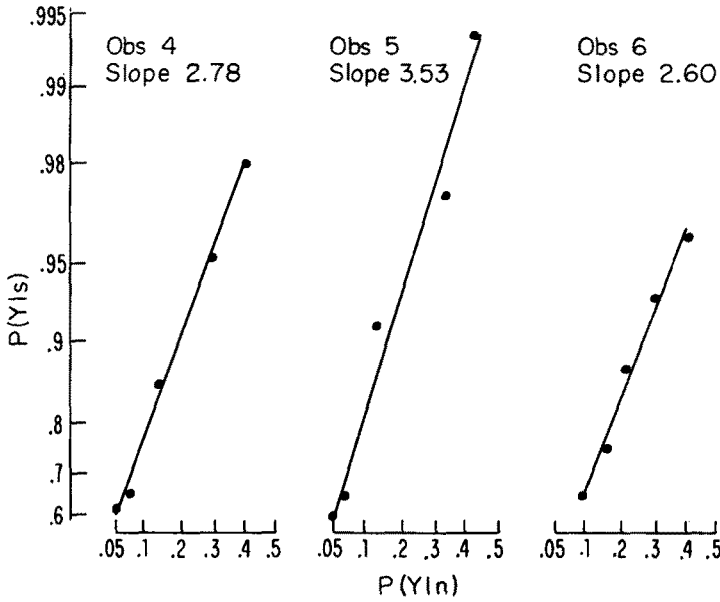
FIGURE 5

These are the same data as Fig. 4 but plotted in the coordinates $-\ln \ln 1/p$. This is Fig. 7 of Wandell and Luce (Note 1).
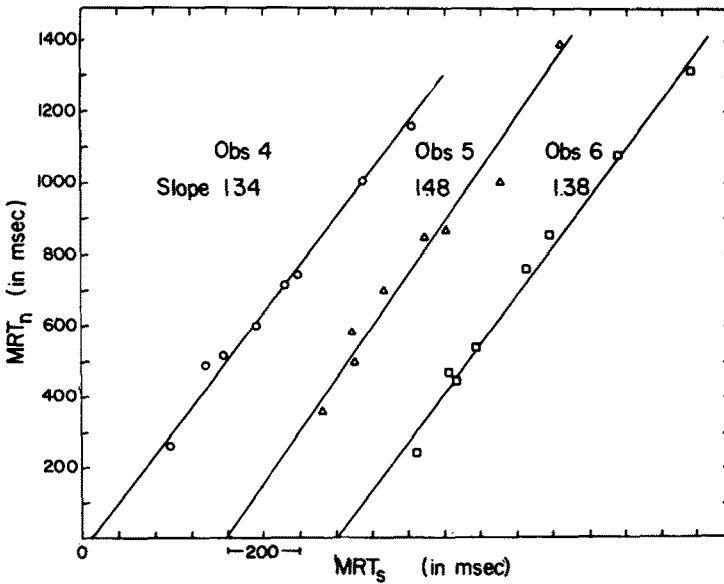

FIGURE 6

Mean reaction time to n trials versus mean reaction time to s trials where the deadline is on s trials only. The deadlines used were 250, 300, 400, 500, 600, 800, 1000, and 1500 msec. This is Fig. 7 of Green and Luce (1973).

5. Thus, of the two methods of aggregation, the averaging one is strongly favored.

This last argument is not without methodological interest. From the pairs of figures, 2 versus 3 and 4 versus 5, it is apparent that one cannot decide on the form of the distribution using the ROC data alone; indeed, these plots make clear that it really is idle to defend an hypothesis about the distribution on the basis of the fit to ROC data alone. In contrast, the comparison of ROC slopes to mean reaction time slopes appears to be extremely sensitive to the underlying distribution, so much so that no statistical test was required to select between averaging and maximizing.

At this point I conclude that Thurstone's choice of the normal distribution was correct, although I rather doubt that it was well-justified at the time. Moreover, we have demonstrated that we can experimentally induce either counting or timing estimators. Since most psychophysical procedures involve observation intervals of fixed, brief duration, which amounts to deadlines on all types of trials, we anticipate that experienced subjects will resort to the counting estimator. But probably they do not normally use counting because such constraints are rare in natural environments; moreover, the fact that response times increase significantly with decreased signal intensity strongly suggests timing, not counting, estimators. Thus, our experiments probably force subjects to switch from timing to counting, which is at least one factor in what we mean by a well-practiced observer. Moreover, if most of our psychophysical data reflect, in part, the counting estimator, then they do not generalize immediately to natural situations where timing estimators are most likely used. The latter point may sometimes be important.

For those who find the differences between timing and counting of interest, there is another striking way of exhibiting these differences. Suppose we vary the value of the deadline in the two procedures and observe how the accuracy of the performance is affected. We measure the latter in terms of the normalized difference d' between the Thurstone distributions. The result, shown in Fig. 7, is that d' grows much more rapidly when the deadline applies only to signal trials than when it applies to both kinds of trials.

The Response Decision Rule

Having discussed the distribution of the Thurstonian representation, let us turn now to the question of how it is used in arriving at responses. For all procedures that can be thought of as category experiments—those for which each signal presentation is uniquely associated with a response, though the converse need not be so—one decision rule is usually assumed: Partition the range of the possible representation into as many intervals as there are ordered responses, and respond according to the interval within which the representation falls. I implicitly assumed this rule in the above analysis of the response terminated, Yes-No design. Although experiments of this type are common, there are other types. For one, there are those with more than one signal
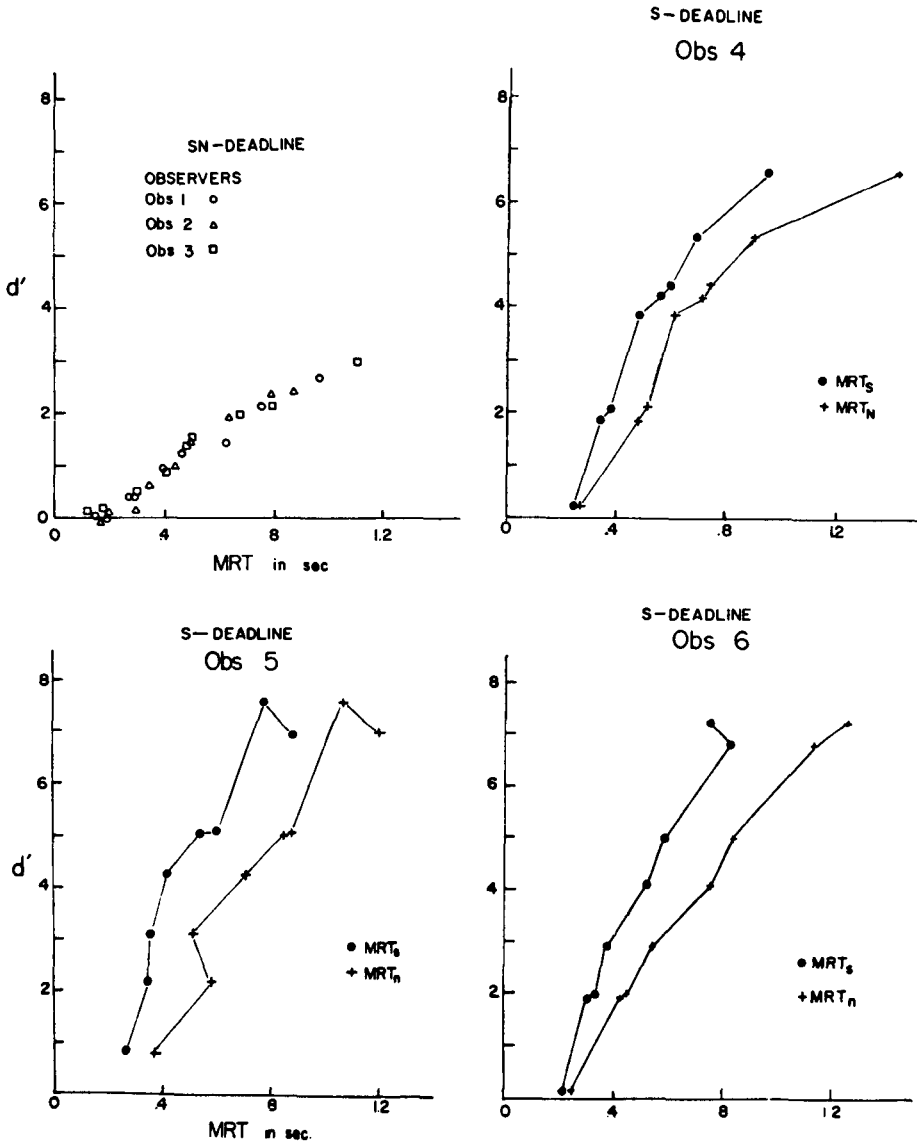
FIGURE 7

d' as a function of mean reaction time generated by varying the deadline. The upper left panel contains the data from three subjects where the deadline applied to all trials. The other three panels are for individual subjects where the deadline applied only to s trials. This is Fig. 11 of Green and Luce (1973).

interval, such as the two-interval forced-choice design. It is usual to assume the subject transforms the problem into one of the category type by looking at a new random variable, either the difference or ratio of the random variables

arising in the two intervals. I shall not go into this in any detail here because modeling of this type is widely accepted and, so far as I know, we do not have any experimental designs directly aimed at testing this assumption.

Other types of experiments which do not readily transform into the category mode, and in my opinion are really quite different, are magnitude estimation, magnitude production, and cross modality matching. In estimation no response is prescribed as correct; the subject simply reports whatever number he or she feels is appropriate to the signal. In magnitude production, the stimulus is a number and the subject selects the signal which seems appropriate to that number. And in matching, the stimulus is a signal in one modality and the response is that one in a different modality which is felt to match the first one. Although there are no correct responses, exceedingly regular data are obtained.

Presumably, whatever a subject does, it involves some use of the internal representation of the signal. One simple idea for magnitude estimation [Luce & Green, 1972] is that the number emitted is proportional to the representation. Symbolically, if S_n is the random variable describing the signal presentation on trial n, $X(s)$ is the internal representation (random variable) of signal s when it is presented, and R_n is the response random variable on trial n, then for some constant C,

$$(1) \qquad\qquad R_n = CX(S_n).$$

Without going into the details here, Green and Luce [1974] cast doubt upon this by showing that the observed distribution of responses to signals of relatively long duration is not what one would have expected if (1) were true and X is the representation due to a timing estimator operating on a Poisson pulse train.

A second suggestion [Luce & Green, 1974a], which arises in part from the fact that instructions in magnitude estimation urge the subject to reflect in his responses the subjective ratios of intensities, is that

$$(2) \qquad\qquad \frac{R_n}{R_{n-1}} = C\,\frac{X(S_n)}{X^*(S_{n-1})}\,.$$

This is called the *response-ratio hypothesis*. The starred random variable in the denominator reflects the assumption that independent samples are used on the successive trials, which is needed to account for sequential effects in the responses. Although this is important, I shall not go into it here.

To test the response-ratio hypothesis, Jesteadt, Luce and Green [1977] noted that taking the logarithm of (2) suggests fitting the following regression equation to the data:

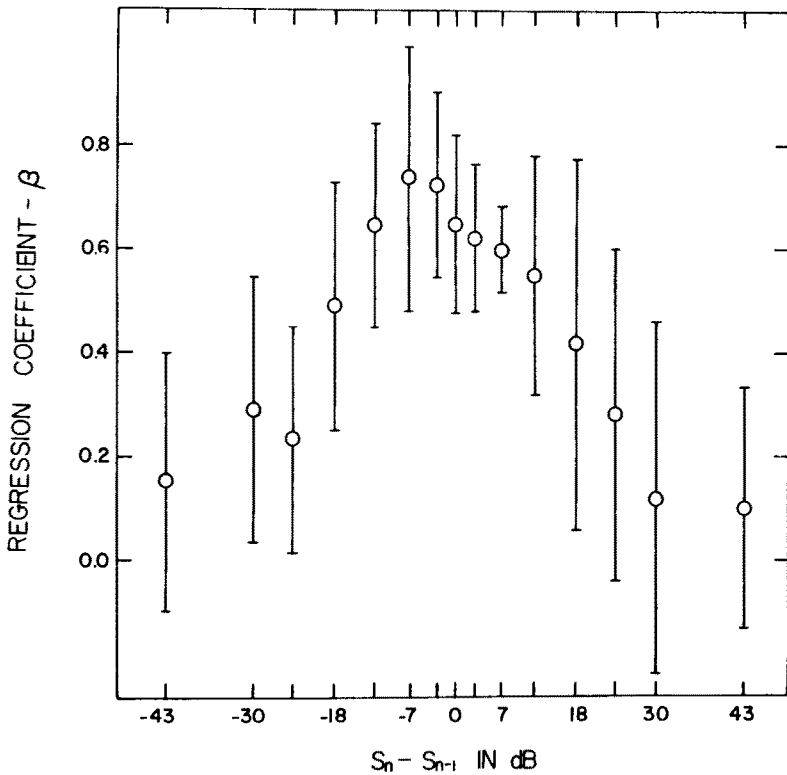$$\ln R_n = \beta \ln R_{n-1} + \gamma \ln I(S_n) + \alpha \underline{\ln I(S_{n-1})} + \delta.$$

FIGURE 8

Sequential effects in magnitude estimation. The regression coefficient of log R_n on log R_{n-1} as a function of $S_n - S_{n-1}$, in dB, averaged over 4 subjects. This is Fig. 4 of Jesteadt, Luce, and Green (1977).

Note that if the response-ratio hypothesis is true, then the parameter β should be 1. Our estimate of β was about .3, which left us just about nowhere until we calculated it for each stimulus pair separately and averaged over all pairs with constant differences in dB. The result is shown in Fig. 8. A similar result, shown in Fig. 9, holds for magnitude production [Green, Luce, & Duncan, Note 2]. These data give considerable support to the response ratio hypothesis, (2), when the successive stimuli are not too far apart. But when they are quite far apart, the response appears to be directly proportional to the representation of the signal, as in (1).

A possible neural interpretation is as follows: If the signals on successive trials activate the same, or nearly the same, set of fibers, then the response is determined by calculating the ratio of the successive representations and multiplying it by the preceding response. When, however, the signals fall in two different groups of fibers, the preceding trial is utterly ignored and the response is based only on the current representation. It is as though one kind of

calculation is possible within a fiber which is not available across distinct fibers. This is an interesting possibility and surely bears careful investigation.

The Psychophysical Law

Fechner initiated his psychophysical studies in order to discover how sensation grows with physical variables or, in the language of Stevens, to determine the psychophysical law. On the assumptions that Weber's law is correct and that subjective confusions are all equal, he arrived at a logarithmic relation, now called Fechner's law. One trouble with this approach is that Weber's law is not correct for important modalities; e.g., Fig. 10 gives the latest data [Jesteadt, Wier, & Green, 1977] on the Weber fraction for auditory intensity, and it is hardly constant. Another trouble is that there really is no justification for assuming the equality of subjective confusions.

Stevens [1957, 1975] attacked this approach, pointing out quite cogently, as Thurstone had done earlier, that there is no necessary interlock between the rate of growth of the mean of a random variable and its variance. One has to study each aspect independently. Stevens proposed studying the mean via magnitude estimation, magnitude production, and matching. Implicitly he assumed. (1), and so
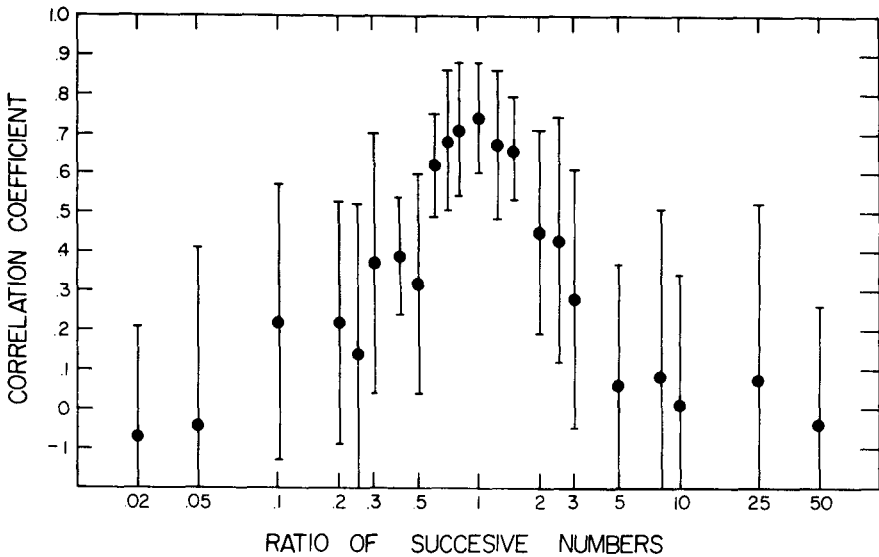
$$E(R) = CE[X(s)],$$

FIGURE 9

Sequential effects in magnitude production. The correlation coefficient of the signal setting in *dB* on trial *n* with the setting on trial *n* − 1 versus the logarithm of the ratio of the corresponding numerical stimuli, averaged over 5 subjects. This is Fig. 4 of Green, Luce, and Duncan (Note 2).
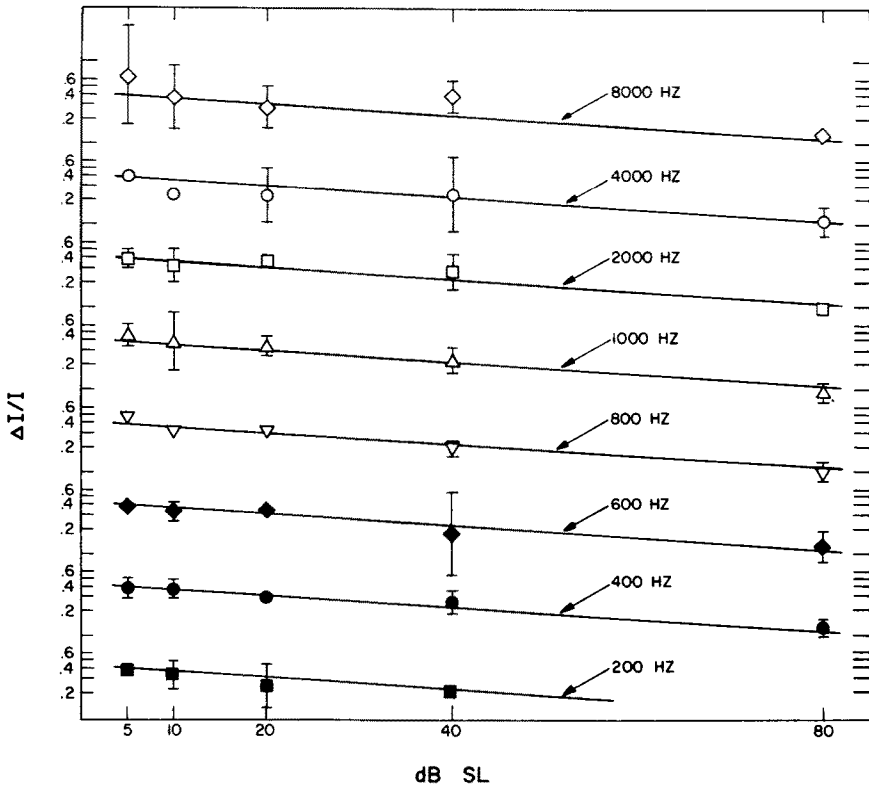
FIGURE 10

Weber fraction for intensity as a function of intensity with frequency as a parameter. The data are averaged over subjects and replications; the error bars are for ±3 standard errors of the mean based on variability among subjects. The function fit to the data is

$$\Delta I/I = 463(I/I_0)^{-.072}$$

This is Fig. 1 of Jesteadt, Wier, and Green (1977).

where the expectation is taken over the trials for which $S_n = s$. His data, based upon a few observations per stimulus per subject and averaged over a number of subjects, suggested that $E(R)$ grows approximately as a power function of $I(s)$, where I is the physical intensity of signal s. These data led him to urge that the psychophysical law is a power function, which has since been dubbed Stevens' law.

Data from individual subjects appear to deviate systematically, but idio-syncratically, from power functions by amounts up to about 5 dB. Examples are shown in Fig. 11. In addition, there are considerable differences in the "exponent" from subject to subject. For example, it is usually said that the typical exponent for loudness (with intensity measured in power) is .30, but my
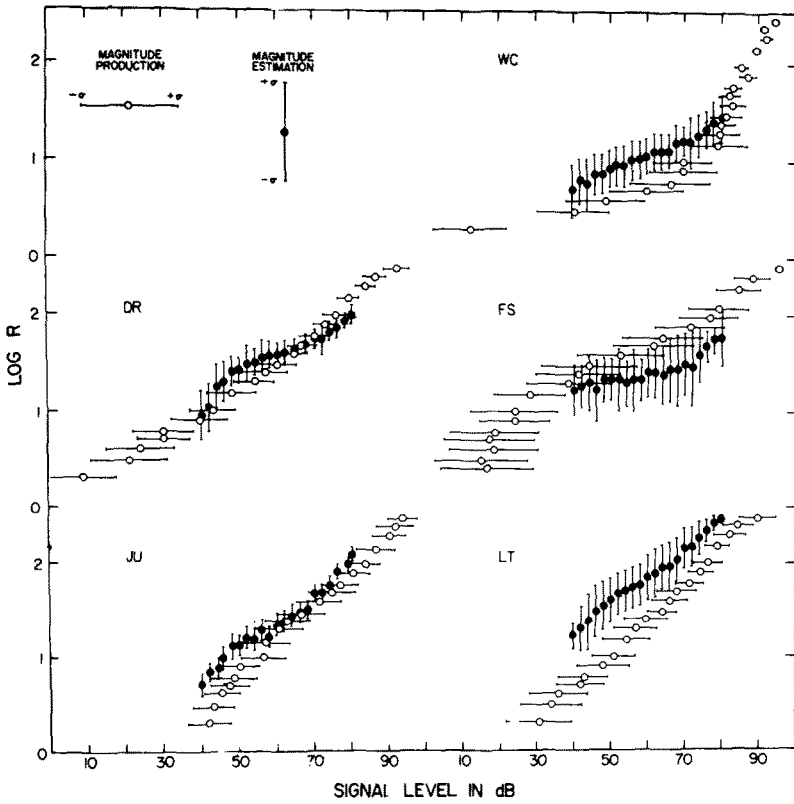
FIGURE 11

Magnitude estimates and productions in logarithmic coordinates. The ordinate is numbers; the abcissa is sound pressure level in *dB*. The bars show estimated one standard deviation bands. This is Fig. 1 of Green, Luce, and Duncan (Note 2).
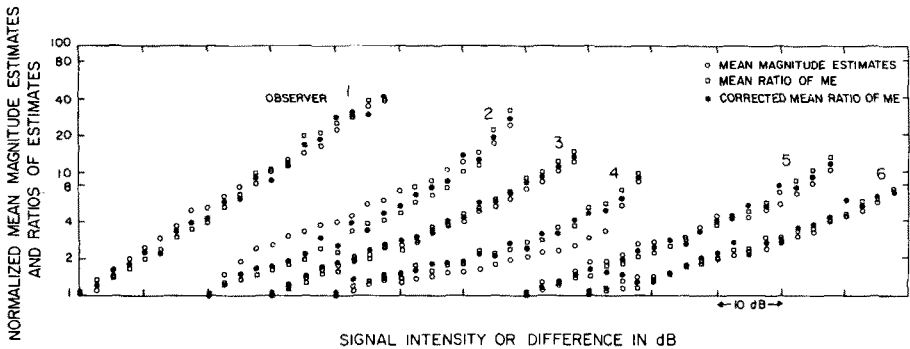

FIGURE 12

Mean magnitude estimates (○), mean of the ratio of successive estimates (⊡), and a corrected mean ratio not discussed here (●) versus signal intensity in *dB*. The former has been normalized to be 1 at the lowest signal and the latter two to be 1 at 0 *dB* difference. This is Fig. 3 of Luce and Green (1974a).

experience is that the observed values run from about .10 to .60 with .25 or .27 as the typical value.

The implicit assumption of (1), the proportionality of responses to the representation, does not appear to be crucial. It we suppose the response-ratio hypothesis, (2), then we should look at the mean response ratio as a function of the signal ratio, and according to Fig. 12 there do not appear to be any more or less difficulties in the form of the function.

If the psychophysical function describes some sort of physiological growth, I do not see any reason to expect a particular simple class of mathematical functions to fit the data exactly. Stevens' law appears to be a good first approximation to the growth, and certainly it is a plausible assumption to insert into various psychophysical models. There are, however, considerable difficulties in knowing exactly how to interpret it vis-a-vis what we know about neural activity. For example, in the speed-accuracy work, which led us to timing and counting estimators and to accept averaging as the aggregation schema, we interpreted the mean of the random variable as proportional to neural pulse rate. That, together with the magnitude data, suggests the pulse rate grows as a power function with an exponent of about .3 for loudness. If the signals span 10 log units, then the pulse rate spans 3 orders of magnitude. However, auditory pulse data on single nerve fibers show a change in responsiveness over only a range of about two logarithmic units, the exact location of the change being a function both of the fiber and of signal frequency. The amount of responsiveness is from one to two logarithmic units—a factor of between 10 and 100. What seems to be involved are two intensity codes—a gross one that is the total number of fibers firing at full rate and a more refined vernier that is the exact rate of those fibers that are neither at rest nor at full rate. Presumably, the magnitude data tell us about the overall pattern and discrimination data about the vernier. Our models have not maintained a careful separation between the two, and much work needs to be done.

For example, we should be able to draw upon what we know about the representations of signals to derive the Weber functions of Fig. 10—that is, in a sense, to reverse the program of Fechner. Luce and Green [1974b] and, independently and somewhat differently Sanderson [1975, 1977] attempted to do this for both auditory intensity and frequency. Although Luce and Green achieved tolerable fits to the data, their model is surely wrong in detail because it does not maintain adequately the distinction between the gross and vernier codes just discussed.

Factors Affecting the Variance

As we have proceeded, we have, almost incidentally, touched on two factors affecting the variance of the representation. One, which was known to Thurstone, is signal intensity. Although this factor is important, it is hardly novel and I shall not dwell on it here.

A second factor is how the CNS estimates pulse rate, by counting or by

timing. I did not point it out explicitly earlier, but the reason that the ROC curves corresponding to these estimators have very different slopes is that the estimators have very different variance structures.

A third factor, of still a different character, appears to be as important as the first two. As I pointed out earlier, there is good reason to study the ratios of successive responses in magnitude estimation, and so we were led to examine the relative variability of these ratios, i.e., the coefficient of variation

$$\frac{\sigma(R_n/R_{n-1})}{m(R_n/R_{n-1})}.$$

Green and Luce [1974] predicted from a Poisson model of neural pulses that this quantity should be constant independent of the stimulus pair used. But as a matter of fact, when plotted as a function of signal difference on successive trials, the coefficient of variability for most subjects exhibits the V-shaped pattern seen in the four and five subjects in Fig. 13. Note that this phenomenon holds for both estimation data (left panel) and for production data (right panel) and, except for one subject, a factor of about three is involved.

At first one is inclined to suspect that this is nothing but a manifestation of the first factor, signal intensity; however, that probably is not the explanation. For one thing, I have never been able to devise any argument to account for how variability that depends on intensity produces the V-shaped pattern. For another, and more persuasively, changing the frequency distribution of signal differences inverted the V-shaped pattern, so that small signal differences had the largest coefficient of variation and large differences the smallest value, for two out of six subjects in Green and Luce [1974].

Our conjecture about the source of this phenomenon is the limited capacity of the CNS to attend to peripheral neural activity. This was discussed above as a concept of attention which is identified with a larger sample size over which aggregation takes place. If so, then a signal does not have a unique Thurstonian representation. Rather, the representation is an estimate based either on a large sample when the signal happens to activate the bundle of fully monitored fibers or on a small sample when it activates any other bundle of fibers.

If we look at this phenomenon from the point of view of one stimulus variable, say intensity, with the other (frequency) fixed, then there will be a narrow band of intensities, about 15 or 20 dB wide according to the neurophysiological data, which will receive the large sample size. The exact location of this attention band will depend both on the frequency of the signal and the bundle to which attention is currently being paid. In the frequency dimension the band, or something very much like it, has been called the critical band [Fletcher, 1940; Scharf, 1970].

Now, if we assume that such an attention band exists, that the samples within and outside the band differ by about an order of magnitude, and that in
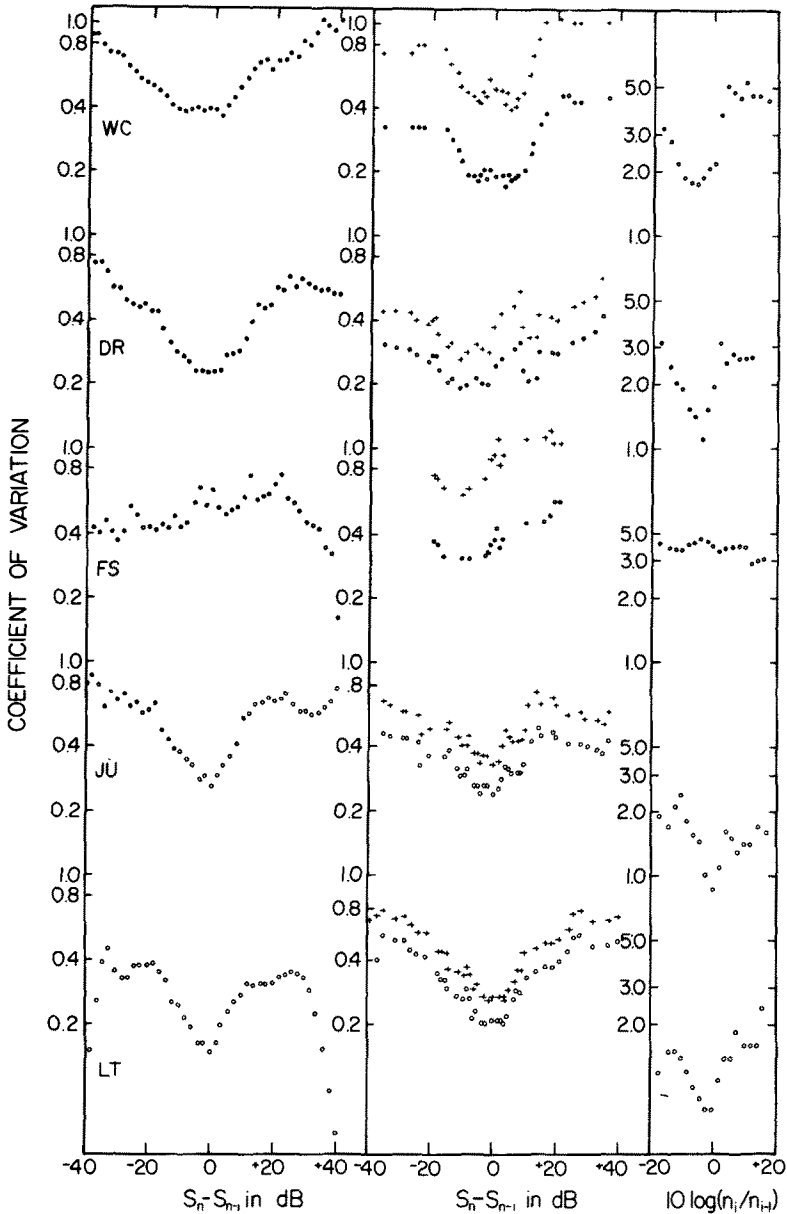
FIGURE 13

Coefficient of variation of R_n/R_{n-1} from magnitude estimation (left panel) and from transformed magnitude productions (center panel), and the coefficient of variation of I_n/I_{n-1} from magnitude production (right panel). The center panel involves inverting the production data through the mean magnitude estimate and mean magnitude production functions so as to make the variability comparable to that of magnitude estimation. This is Fig. 2 of Green, Luce, and Duncan (Note 2).

the ordinary magnitude estimation experiment the band tends to be located in the region of the last signal, then we predict the V-shaped pattern with a factor of three in level.

Do we predict anything else? It is evident that whenever the signal range exceeds the width of the band, some signals will fall outside the band [Luce, Green, & Weber, 1976]. That seems to be just the idea needed to explain why the variance of the representation appears to grow with range in the absolute identification experiment. Indeed, if we accept the model and suppose that the psychophysical function is a power function, the Garner data are well fit with reasonable parameters, as seen in Fig. 14. Similarly, data in which the range is varied, Fig. 15, are nicely accounted for by the solid curve; the dotted curve, which is equally good, arises from assuming the variance of the representation grows linearly with the square of the range in dB.

If we are correct about the differential monitoring of fibers, another prediction follows. Suppose the intensity of signals varies, so that sometimes a signal lies within the region of attention and sometimes not; in this case a subject's performance both on frequency and intensity should be affected. In particular, suppose we ask the subject to identify absolutely one of two frequencies, f_1 or f_2, but with randomized intensities to which he does not
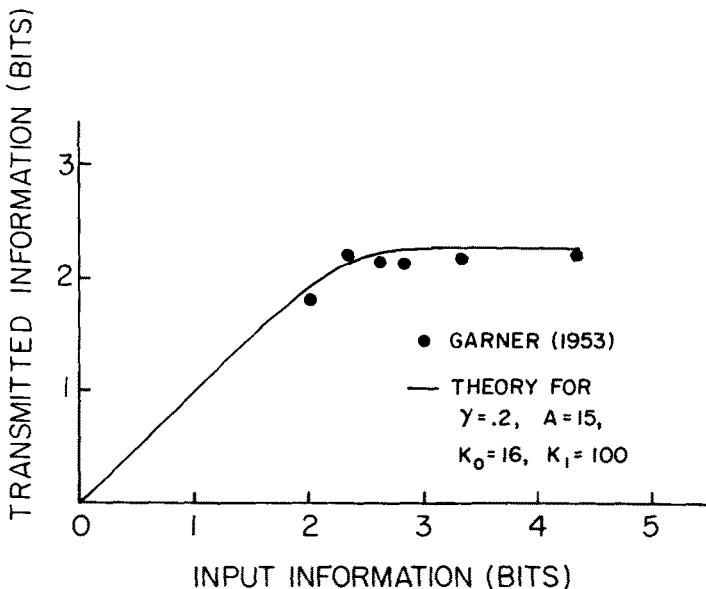

FIGURE 14

Garner's (1953) data fit by an attention band, Thurstonian model in which the mean random variable is proportional to Γ, the band is A dB wide, and the sample size outside the band is K_0 and inside the band is K_1. Values of the parameters are shown. This is Fig. 3 of Luce, Green, and Weber (1976).

FIGURE 15

Information transmitted as a function of range for four observers. The solid line is for the attention band model described in the caption to Fig. 14. The dotted line is for the theory of Braida and Durlach (1972). This is Fig. 5 of Luce, Green, and Weber (1976).

FIGURE 16

Design of frequency identification experiment with clustered intensities. Each of the three panels represents one experimental condition. Each point in a panel represents a possible signal presentation, and they are equally likely to occur. The subject's task is to identify the frequency presented.
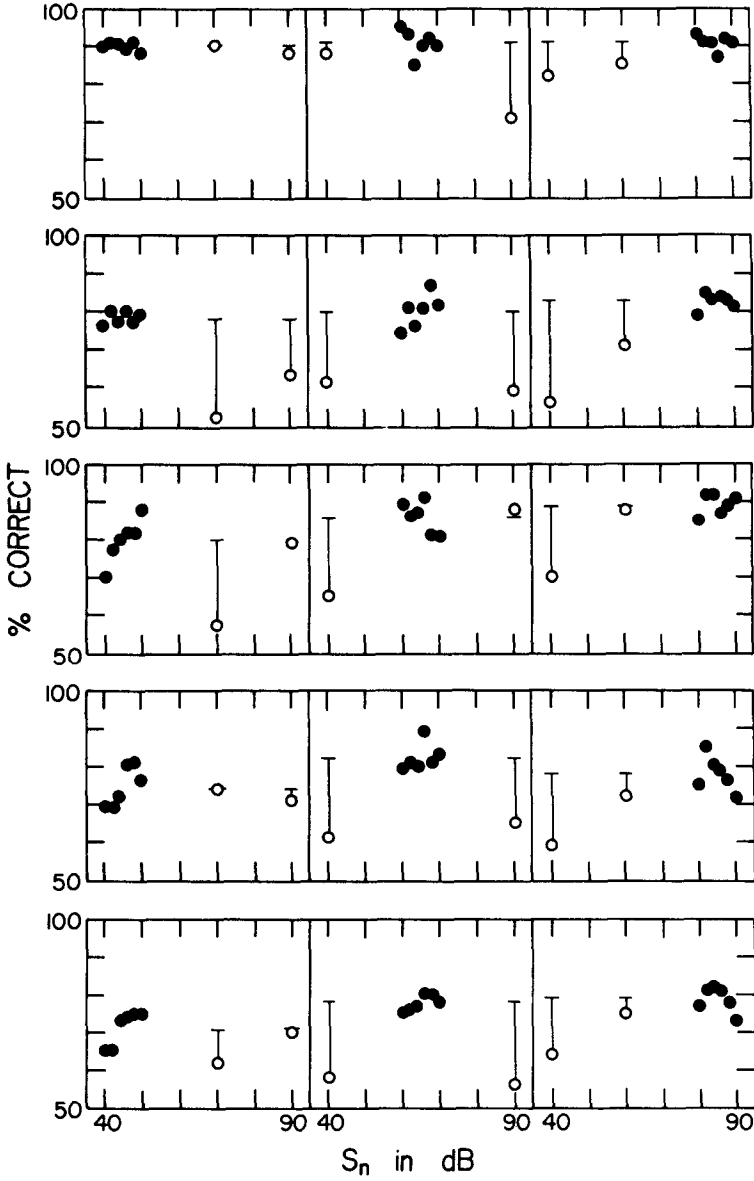
FIGURE 17

Percent correct frequency identifications as a function of signal intensity for each of the three conditions. The vertical line to each outlier intensity shows the drop from the mean percent correct for the cluster. Note that the intensities 40, 60, 70, and 90 appear both in a cluster and as an outlier.

respond. Further, suppose we group the intensities so that most, but not all, fall in a narrow cluster, as shown in Fig. 16 for three possible locations of the cluster. Then, it is to the subject's advantage in identifying frequency to locate his attention wherever the cluster is. Thus, we should find performance on signals within the cluster somewhat better than for the outliers. In fact, if the sample sizes on which the decisions are based do differ by an order of magnitude, it is possible to select frequency differences so that the one probability of correct identification is .8 while the other is .6. The data from this experiment [Luce & Green, Note 3] are shown in Fig. 17. The prediction is well confirmed when the cluster is in the midrange of intensities, which is about the normal range for voices; it is somewhat less clear for very soft and for very loud clusters. Despite the discrepancies from the prediction, I feel the evidence is, on the whole, quite supportive of the attention hypothesis.

Our major problem in working with this idea is that we do not really understand what controls the focus of attention. Everything that seems to have some effect—the location of the last signal, the clustering of signals, the most intense signals—is less than totally controlling. Thus, it is very difficult to isolate the phenomena for detailed study.

Conclusions

According to the view I am urging, the Thurstonian representation of a signal arises as follows. There is some sort of aggregation, calculated over a sample of nerve fibers, of estimates of a parameter of neural pulse trains. These estimates result from estimators being applied to limited time samples of individual neural pulse trains. This view has four key concepts—samples of individual pulse trains in time, estimators, samples of fibers, and an aggregation of the estimates. For each, several alternative possibilities seem plausible. I have tried to show how auditory and visual intensity data can be used to select among these alternatives on the well-justified assumption that intensity is encoded as pulse rate.

First, the nature of the time sample of pulses and the estimator are closely interlocked, at least for timing and counting estimators. These led to different predictions for both mean reaction times and ROC curves. Experimental procedures have been devised which appear to show that both kinds of estimators are available. Interestingly, timing seems to be the natural one, whereas typical psychophysical procedures with brief signals of fixed duration appear to cause people to shift over to counting.

Second, in aggregating these estimates over fibers, which is done in order to upgrade the statistical quality of the responses, we considered two alternatives: averaging and maximizing. Comparing ROC and mean reaction time data, we were led to favor the averaging model and hence, by the Central Limit Theorem, we are led to normally distributed representations.

Third, we considered carefully the sample of fibers over which the aggregation occurs because this may well not be an invariant of the nervous system.

Rather, because of a limited capacity of the central nervous system to process peripheral information, the size of the sample depends on whether or not the signal activates fibers to which the system is attending. If we vary intensity alone, the result is a band, about 15-20 *dB* wide in auditory intensity, of variable location such that when a signal falls within the band it receives a sample whose size is about an order of magnitude larger than when it falls outside the band. This assumption was shown to account for three things: in absolute identification, the limiting effect on performance of either increasing signal range with a fixed number of signals or of increasing the number of signals for a fixed large range; in magnitude estimation, the reduced relative variability of the ratio of responses to successive signals when they are close to one another; and in frequency discrimination, the effect of intensity distribution on the probability of correct identification.

Of the four questions we originally asked about Thurstone's model—the form of the distribution, the nature of the decision rule, the psychophysical law, and factors affecting the variance—we have obtained satisfying answers about both the form of the distribution and the factors affecting the variance. Concerning the decision rule, there is no evidence against simple partitions of the range of the representation when it is appropriate; however, for methods such as magnitude estimation we have encountered complexities suggesting quite different rules depending on whether successive signals activate the same or different groups of fibers. Similarly, the superficial simplicity of Stevens' psychophysical law seems to break down into different answers depending on whether the experiment lies within or across groups of similar fibers.

From my perspective, our progress has led us both to some understanding and to some unexplained, but interesting, phenomena which clearly complicate the Thurstonian model without, however, destroying its basic spirit.

REFERENCE NOTES

1. Wandell, B., & Luce, R. D. Pooling peripheral information: averages versus extreme values. Unpublished manuscript, 1977.
2. Green, D. M., Luce, R. D., & Duncan, J. E. Variability and sequential effects in magnitude production and estimation of auditory intensity. *Perception & Psychophysics.* Manuscript submitted for publication, 1977.
3. Luce, R. D., & Green, D. M. Manuscript in preparation, 1977.

REFERENCES

Bock, R. D., & Jones, L. V. *The measurement and prediction of judgment and choice.* San Francisco: Holden-Day, 1968.
Braida, L. D., & Durlach, N. I. Intensity perception. II. Resolution in one interval paradigms. *Journal of the Acoustical Society of America,* 1972, *51,* 483–502.
Corso, J. F. The neural quantum theory of sensory discrimination. *Psychological Bulletin,* 1956, *53,* 371–393.
Durlach, N. I., & Braida, L. D. Intensity perception. I. Preliminary theory of intensity perception. *Journal of the Acoustical Society of America,* 1969, *46,* 372–383.
Fletcher, H. Auditory patterns. *Review of Modern Physics,* 1940, *12,* 47–65.

Garner, W. R. An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology*, 1953, *46*, 373–380.

Gravetter, F., & Lockhead, G. R. Criterial range as a frame of reference for stimulus judgment. *Psychological Review*, 1973, *80*, 203–216.

Green, D. M., & Luce, R. D. Speed-accuracy tradeoff in auditory detection. In S. Kornblum (ed.) *Attention and performance, IV*. New York: Academic Press, 1973. Pp. 347–569.

Green, D. M., & Luce, R. D. Variability of magnitude estimates: A timing theory analysis. *Perception & Psychophysics*, 1974, *15*, 291–300.

Gumbel, E. J. *Statistics of extremes*. New York: Columbia University Press, 1958.

Jesteadt, W., Luce, R. D., & Green, D. M. Sequential effects in judgments of loudness. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, *3*, 92–104.

Jesteadt, W., Wier, C. C., & Green, D. M. Intensity discrimination as a function of frequency and sensation level. *Journal of the Acoustical Society of America*, 1977, *61*, 169–177.

Kiang, N. Y.-S. *Discharge patterns of single fibers in the cat's auditory nerve*. Cambridge, Mass.: M.I.T. Press, 1965.

Krantz, D. H. Threshold theories of signal detection. *Psychological Review*, 1969, *76*, 308–324.

Luce, R. D. *Individual choice behavior: A theoretical analysis*. New York: Wiley, 1959.

Luce, R. D. A threshold theory for simple detection experiments. *Psychological Review*, 1963, *70*, 61–79.

Luce, R. D., & Green, D. M. A neural timing theory for response times and the psychophysics of intensity. *Psychological Review*, 1972, *79*, 14–57.

Luce, R. D., & Green, D. M. The response ratio hypothesis for magnitude estimation. *Journal of Mathematical Psychology*, 1974, *11*, 1–14. (a)

Luce, R. D., & Green, D. M. Neural coding and psychophysical discrimination data. *Journal of the Acoustical Society of America*, 1974, *56*, 1554–1564. (b)

Luce, R. D., Green, D. M., & Weber, D. L. Attention bands in absolute identification. *Perception & Psychophysics*, 1976, *20*, 49–54.

McFadden, D. Conditional logit analysis of quantitative choice behavior. In P. Zarembka (ed.) *Frontiers in econometrics*. New York: Academic Press, 1974. Pp. 105–142.

McGill, W. J. Neural counting mechanisms and energy detection in audition. *Journal of Mathematical Psychology*, 1967, *4*, 351–376.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, *63*, 81–97.

Norman, D. A. Sensory thresholds, response biases, and the neural quantum theory. *Journal of Mathematical Psychology*, 1964, *1*, 88–120.

Rose, J. E., Brugge, J. F., Anderson, D. J., & Hind, J. E. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *Journal of Neurophysiology*, 1967, *30*, 769–793.

Sanderson, A. C. Discrimination of neural coding parameters in the auditory system. *IEEE Transactions, Systems, Man, and Cybernetics*, 1975, *SMC-5*, 533–542.

Sanderson, A. C. Comments on the paper 'Discrimination of neural coding parameters in the auditory system.' *IEEE Transactions, Systems, Man, and Cybernetics*, 1977, in press.

Scharf, B. Critical bands. In J. V. Tobias (ed.) *Foundations of modern auditory theory*. Vol. 1. New York: Academic Press, 1970. Pp. 159–202.

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

Stevens, S. S. A neural quantum in sensory discrimination. *Science*, 1972, *177*, 749–762.

Stevens, S. S. *Psychophysics*. New York: Wiley, 1975.

Stevens, S. S., Morgan, C. T., & Volkman, J. Theory of the neural quantum in the discrimination of loudness and pitch. *American Journal of Psychology*, 1941, *54*, 315–335.

Swets, J. A. Is there a sensory threshold? *Science*, 1961, *134*, 168–177.

Thurstone, L. L. Psychophysical analysis. *American Journal of Psychology*, 1927, *38*, 368–389. (a)

Thurstone, L. L. A law of comparative judgment. *Psychological Review*, 1927, *34*, 273–286. (b)

Thurstone, L. L. Three psychophysical laws. *Psychological Review,* 1927, *34,* 424–432. (c)

Torgerson, W. S. *Theory and methods of scaling.* New York: Wiley, 1968.

Uttal, W. R., & Krissoff, M. Response of the somethetic system to patterned trains of electrical stimuli: an approach to the problem of sensory coding. In D. R. Kenshalo (ed.) *The Skin Senses.* Springfield, Ill.: Thomas, 1965.

Wandell, B. A. Speed-accuracy tradeoff in visual detection: Applications of neural counting and timing. *Vision Research,* 1977, *17,* 217–225.

Yellott, J. I., Jr. The relationship between Luce's choice axiom, Thurstone's theory of comparative judgment, and the double exponential distribution. *Journal of Mathematical Psychology,* 1977, *15,* 109–144.