Detection models: Experimental tests with electrocutaneous stimuli

GARY B. ROLLMAN[2]
UNIVERSITY OF PENNSYLVANIA

Contemporary psychophysical procedures were employed to determine the nature of the process for detecting rectangular electrical pulses delivered to the skin. Earlier experiments had shown that the somatosensory system is exceedingly responsive to small changes in the energy of an electrocutaneous stimulus, suggesting that sensory noise may be relatively unimportant in this modality. However, a rating scale study showed that both the high-threshold theory and Luce's low-threshold theory cannot describe the detection process. The form of the ROC curve is consistent with a signal-detection model that assumes a greater variance for the signal plus noise distribution than for the noise distribution. Additional support for signal detection theory came from an experiment that examined the likelihood of a correct second choice when the first was incorrect in a four-alternative forced-choice paradigm. Therefore, the sensory excitation caused by internal noise can sometimes be greater than the excitation produced by an added stimulus. Electrocutaneous detection differs from auditory or visual in that d' increases at a greater rate as stimulus intensity changes.

Classical studies of the detection process generally assume a high threshold that may be exceeded on some proportion of trials when a stimulus occurs but is never exceeded on trials containing no stimulus. Recent application of the theory of statistical decision to sensory detection has produced a different view that emphasizes the existence of sensory "noise." The two approaches receiving most attention are the signal-detection theory of Swets, Tanner, Birdsall, Green, and others (Swets, 1964), and the low-threshold theory of Luce (1963a, b). Both theories assume that on some trials when no stimulus occurs, sensory noise is high enough to be confused with the effects of a signal. Results from auditory and visual experiments with a number of methods indicate that classical high-threshold theory is incorrect (Swets, 1964). Evidently noise in these sensory systems is an important determinant of detection performance.

Recent experiments (Rollman, 1968, 1969) have indicated that in some respects the response of the somatosensory system to electrocutaneous stimulation differs greatly from the responses of the auditory and visual systems under analogous conditions. Psychometric functions obtained with electrical stimulation of the skin are exceedingly steep when compared to those obtained in the other modalities, and the limits on reciprocity of intensity and duration are considerably shorter. Thus, the somatosensory system is very responsive to small changes in the energy of electrocutaneous pulses. Such sensitivity suggests that noise may be minimal for electro-cutaneous stimulation. High-threshold theory might then describe the results of signal-detection procedures.

Several experimental paradigms are available for assessing detection behavior. Two were selected for these experiments to examine the somatosensory system. The first, employed in Experiment 1, is the rating-scale procedure (Egan, Schulman, & Greenberg, 1959; Nachmias & Steinman, 1963), and the other, in Experiment 2, is the four-alternative forced-choice procedure with second guesses (Swets, Tanner, & Birdsall, 1961). Both procedures permit the comparison of alternative models.

The data from rating scales, which show an observer's (O) certainty that a signal occurred, yield a receiver operating characteristic (ROC) curve upon summational treatment. Signal-detection theory predicts a monotonically increasing ROC function. High-threshold theory predicts that the data will follow two straight lines, the first of which is vertical at $P(Y \mid n) = 0.0$. Low-threshold theory denies the legitimacy of summating rating scale data but does predict how different response categories should be used (Nachmias & Steinman, 1963). Generally, ROC curves from several modalities favor signal-detection theory rather than high-threshold theory. In the first experiment to be reported, ROC curves were determined for electrical stimulation of the skin, and comparisons were made of the predictions of alternative theories concerning confidence ratings.

The data from second guesses in a four-alternative forced-choice paradigm also discriminate between some of the different detection models. Signal-detection theory implies that on trials when no stimulus occurred the magnitude of the internal "noise" in O could sometimes be greater than the representation of signal plus noise following a stimulus presentation. If asked which of four intervals contained the signal, O should sometimes select a blank interval. Since it is most likely, then, that the second largest level of sensory excitation followed the actual stimulus, O's second guess is likely to be correct with a probability greater than chance. High-threshold theory expects that O cannot distinguish among remaining intervals if he is incorrect with his first guess. Accordingly, in a four-alternative situation, 0.33 is the probability of being correct on the second guess given an incorrect first choice. Furthermore, the high-threshold model predicts chance probability independent of stimulus intensity. Signal-detection theory argues that the probability that the interval with the second largest internal event is the correct one should increase with the effectiveness of the stimulus. Swets, Tanner, and Birdsall (1961), Tanner, Swets, and Green (1956), and Eijkman and Vendrik (1963) obtained in various modalities second-guess data that favor signal-detection theory. Eijkman and Vendrik used electrical stimulation but determined only one point for each of two Os. The second experiment extended this research, determining proportions of correct second guesses at up to eight stimulus values for each O.

In the two experiments that follow, the five-point rating scale and the two-alternative forced-choice procedure with second guesses were employed to assess the role of sensory noise in the somatosensory system and to determine the characteristics of the model that best describes the detection process for electrocutaneous stimulation.

EXPERIMENT 1

Observers used a five-point rating scale on each trial to indicate their certainty that a stimulus was or was not presented.

Apparatus

Two Grass silver electrodes of 8-mm diam were filled with Grass electrode cream and taped over the left volar forearm, in the region of the ulnar nerve. About 2.5 cm separated the

centers of the electrodes, with the more proximal one connected to the cathode of an isolated, constant current stimulation system. Tektronix Series 160 waveform and pulse generators controlled the duration of the stimulus and the intertrial interval. For all sessions, the stimulus was a rectangular dc pulse of 0.20-msec duration. A more complete description of the stimulating equipment and the calibration procedures employed is found in Rollman (1969).

Observers

The Os were two paid undergraduate students (one male and one female) who were given extensive practice in psychophysical tasks.

Procedure

The O rested his arm on a foam-rubber pad. A red neon warning light was turned on for 1 sec and the stimulus presentation coincided with the offset of the light; O could also employ the clicks at the make and break of the relay for the warning light as auditory cues. A two-position silent switch permitted omission of a stimulus on "blank" trials. A large partition separated O from E and his control panel.

In order to combine data from several sessions, it was necessary to equate stimulus detectability. Unavoidable changes in electrode sites and arm positions caused sensitivity to vary from day to day. Consequently, E started each session with a two-alternative forced-choice staircase procedure (Rollman, 1969) to determine the stimulus current for a fixed level of detection, around 0.70 in a two-alternative situation. During a 15-min preliminary run about 25 to 30 reversals in the staircase procedure were obtained.

The usual experimental session had 348 detection trials, although a few had somewhat less. One O had five sessions and the other six. Equally probable signal and blank trials occurred randomly. Following cessation of the warning light, the O reported by a five-point rating scale his confidence that a stimulus had occurred on that trial. A "5" indicated that he was quite sure there had been a stimulus; a "4" that he thought there was one; a "3" that he was unsure; a "2," he believed there was no stimulus; and a "1," he felt quite certain there was no stimulus. Feedback following each response told O whether or not there actually had been a stimulus. Several practice sessions served to familiarize Os with the experimental procedures.

Results

The five individual ROC curves for C.B. were very similar, whereas the six from S.S. were somewhat more variable. Average ROC curves were calculated from all sessions for C.B. and for all but the two most extreme ones for S.S. These curves were plotted on both linear paper (Fig. 1) and on normal deviate axes. The latter are not shown, but are straight lines with slopes of 0.82 for C.B. and 0.62 for S.S. The asymmetry of the curves in linear coordinates and the slope of less than one on probability axes can be interpreted as indicating that the distribution of likelihood ratio for signal plus noise has a greater variance than the distribution for noise alone. A number of measures of sensitivity are available for ROC functions of unequal variances (Green & Swets, 1966), but the two most prominent are Δm and d_e'. The former is the z value on the abscissa corresponding to the intercept of the line of best visual fit on the normal deviate plot at $P(Y \mid s) = 0.50$. d_e' is twice the absolute value of z on either axis for the point where the ROC function crosses the negative diagonal. Table 1 contains the values of these parameters. Assuming normal distributions of unequal variance, the reciprocal of the slope in the probability axes plot yields an estimate of the ratio of the standard deviation of the signal plus noise distribution and noise distribution, $1/\text{slope} = \sigma_s/\sigma_n$. Table 1 also gives these values for the two averaged functions.

Signal-detection theory and Luce's low-threshold theory generate different predictions of the a posteriori probabilities of a signal given different ratings (Nachmias & Steinman, 1963). If the ratings 5 and 4 are used by the O only when he is in the "detect state" and 2 and 1 are used only when in the "nondetect state," low-threshold theory predicts $P(s \mid 5) = P(s \mid 4)$ and $P(s \mid 2) = P(s \mid 1)$. Signal-detection theory holds that $P(s \mid 5) > P(s \mid 4)$ and $P(s \mid 2) > P(s \mid 1)$. Table 2 shows the number of half-sessions for each S on which the conditional probabilities were in the direction indicated. The data from preliminary sessions with three Ss are similar. Clear support is shown for signal-detection theory.

Two-by-two contingency tables were established on the data of the rating experiment. Two tables were constructed for each O, with type of trial (signal or blank) heading the columns and response made (5 or 4 for one table, 2 or 1 for the other) heading the rows. Luce's low-threshold theory implies that these two ways of classification should be independent (Nachmias & Steinman, 1963). Trials were
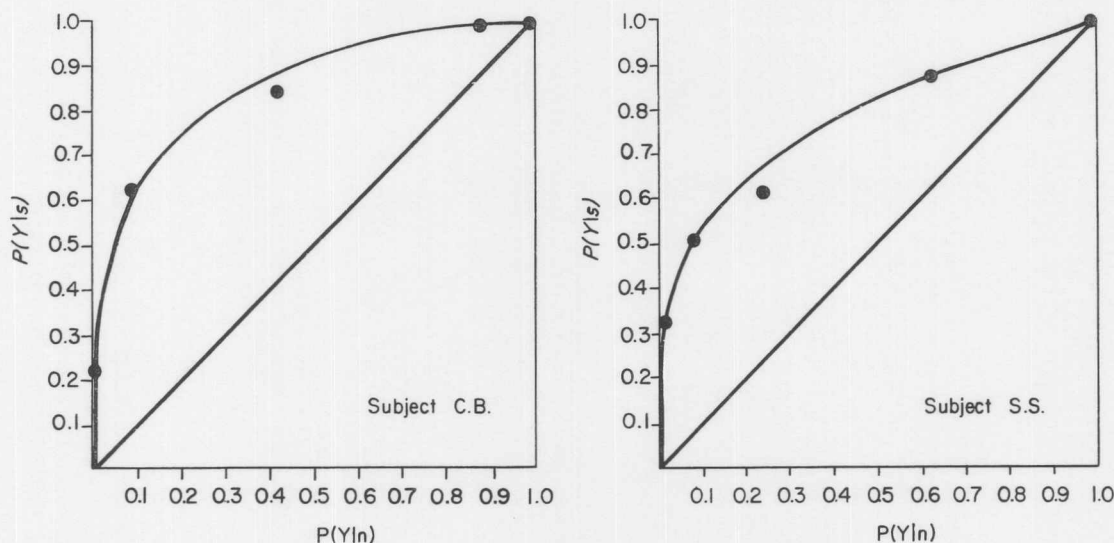

Fig. 1. Average ROC curves on linear coordinates.

Table 1
Measures of Sensitivity and Ratio of Variances of Underlying Distributions determined from ROC Functions on Normal Deviate Coordinates

Observer	Δm	d_e'	σ_s/σ_n
C.B.	1.60	1.48	1.22
S.S.	1.44	1.11	1.61

combined from all sessions for each O and χ^2 tests confirmed that the hypothesis of independence could be rejected for both high and low ratings at the 0.01 level or better. Consequently, an association exists between ratings and type of trial, as signal-detection theory expects.

EXPERIMENT 2

A further comparison of detection models was afforded by the four-alternative forced-choice procedure with second guesses.

Apparatus

The apparatus from Experiment 1 was used, with the addition of timing equipment to permit the signaling of four discrete intervals.

Observers

The Os were C.B. from the first experiment plus two male Os who had served in other psychophysical studies.

Procedure

Each O had from five to nine sessions of about 174 trials each. The 0.20-msec stimulus came after one of four successive 1-sec warning lights. A 1-sec interval separated the offset of each light from the onset of the following one. The order of presentation intervals for the cutaneous stimulus was random. The S indicated which of the four intervals he believed contained the stimulus and which interval was next most likely to be the correct one.

In a preliminary period in each experimental session, the intensity to be presented was determined. It was impossible to select in advance the current that would give a specific probability of choosing the correct interval on the first guess [$P_1(c)$]. Therefore, the threshold was approximated by the method of limits. On some sessions the intensity was raised or lowered slightly so that values of $P_1(c)$ ranging from about 0.30 to 0.80 could be obtained over the experiment. If a block of 29 trials showed that the current setting was inadequate (the O would get almost all correct or be correct on seven or fewer trials), E adjusted the intensity accordingly.

Results

The results of all sessions with each O were analyzed by combining data from blocks of 29 trials with similar values of $P_1(c)$. The number of correct first guesses was grouped so that $P_1(c)$ would have steps of 0.10. For each grouping, the number of correct second guesses and the maximum possible

Table 2
Half-Session Analysis of *a posteriori* Probabilities of a Stimulus given Different Ratings. The Figures Indicate the Number of Half-Sessions for which the Direction of Difference was as Noted

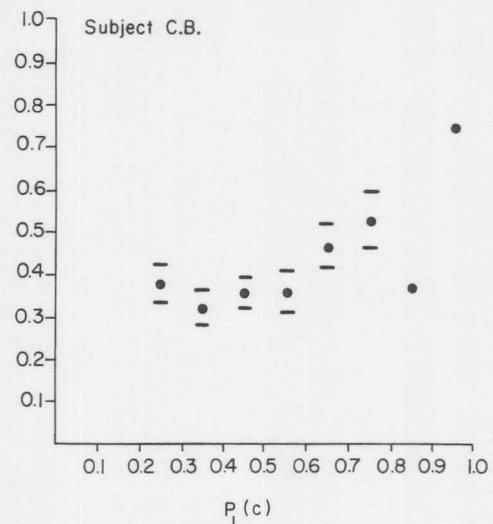| Observer	$P(s\|5)>$ $P(s\|4)$	$P(s\|4)>$ $P(s\|5)$	$P(s\|2)>$ $P(s\|1)$	$P(s\|1)>$ $P(s\|2)$
C.B.	10	0	10	0
S.S.	12	0	5	7
	22	0	15	7

Fig. 2. Typical results of four-alternative forced-choice experiment. Ordinate [$P_2(c)$] shows probability of correct second guess when first was incorrect. Abscissa [$P_1(c)$] shows probability of correct first guess. Horizontal marks indicate plus and minus one standard error of the mean for large sample size.

number of these were summed and $P_2(c)$, the probability of a correct second guess when the first was incorrect, was determined. Figure 2 is a typical plot of $P_2(c)$ as a function of $P_1(c)$. The horizontal marks indicate plus or minus one standard error of the mean when the sample size was large enough to permit support of the assumptions underlying the test. The differences between the observed $P_2(c)$ and the chance probability of 0.33 were examined with t-tests, the results of which are shown in Table 3. Of the 19 t-tests performed, 7 showed $P_2(c)$ significantly greater than chance level. Since high-threshold theory predicts $P_2(c)$ will be 0.33, it would expect only one or two significant differences to occur by chance.

DISCUSSION

The steepness of the psychometric function and the small range of intensity-duration interaction with electrocutaneous stimulation (Rollman, 1969) show that the somatosensory system is very responsive to small changes in the properties of electrical stimuli. These observations suggested that "noise" might be very low for such stimuli and that application of signal-detection experimental paradigms could yield unique results. Specifically, absence of noise should have provided data predicted by a high-threshold model. The ROC curves in Fig. 1 clearly do not confirm the predictions of the high-threshold theory. Instead, the points follow the monotonically increasing function that signal-detection theory predicts (Swets, Tanner, & Birdsall, 1961). They might also fit the two limbs specified by low-threshold theory (Luce, 1963a, b).

If signal-detection theory is appropriate for describing the data, several conclusions follow concerning the detection process and the nature of the underlying distributions. The slope of the straight line on normal deviate coordinates indicates a greater variance for the distribution of signal plus noise than for the noise distribution (Swets, Tanner, & Birdsall, 1961). In Table 1 these ratios are 1.22 for C.B. and 1.61 for S.S. The values resemble those reported for visual (Swets, Tanner, & Birdsall, 1961; Nachmias & Steinman, 1963), auditory (Tanner, Swets, & Green, 1956), and taste stimuli (Linker, Moore, & Galanter, 1964). Differences in the

Table 3

Probability of Correct Second Guess when the First Guess is Incorrect $[P_2 (c)]$ as a Function of Probability of Correct First Guess $[P_1 (c)]$ and the Results of One-Tailed t-Tests for Determining Frequencies Significantly Greater than 0.33 (0.05 level). N is the Number of Trials at Each Value of P_1 (c).

	Observer					
	C.B.		P.G-N.		M.B.	
P_1 (c)	P_2 (c)	(N)	P_2 (c)	(N)	P_2 (c)	(N)
0.15					0.37	(24)
0.25	0.38	(130)	0.43*	(131)	0.38	(42)
0.35	0.32	(131)	0.35	(152)	0.38	(56)
0.45	0.36	(175)	0.55*	(64)	0.25	(63)
0.55	0.36	(94)	0.58*	(26)	0.37	(51)
0.65	0.47*	(97)	0.41	(34)	0.38	(52)
0.75	0.53*	(57)	0.77†	(13)	0.55*	(31)
0.85	0.37†	(8)	0.75†	(4)	0.59*	(29)
0.95	0.75†	(4)				

*P_2 (c) significantly greater than 0.33.
† Too few trials for t-test.

shapes of the ROC curves and the relative variances between Os are also consistent with results in other modalities.

Luce's low-threshold theory assumes that O is in either a "detect state" or a "nondetect state" following a trial. Nachmias and Steinman (1963) have tested the concept of only two states for visual detection by making some assumptions about how the low-threshold model would be affected by rating scales. If one assumes that responses "4" and "5" are selected only when the threshold is exceeded and "1" and "2" only when the threshold is not exceeded, one can make two predictions about the probability that a signal was presented when a given response is made. If the presentation probability is 0.5, the two predictions are:

$$(1) \; P(s \mid 5) = P(s \mid 4) = c'$$
$$(2) \; P(s \mid 1) = P(s \mid 2) = c''$$

Essentially, the extension of the theory assumes that no ordering of events is possible within either state. If the data are pooled from Table 2 in the first experiment and from three Os in preliminary sessions we find that $P(s \mid 5)$ was greater than $P(s \mid 4)$ in 61 out of 62 half-sessions. $P(s \mid 2)$ exceeded $P(s \mid 1)$ in 43 out of 62 half-session blocks. Thus, both predictions from this form of low-threshold theory (Wickelgren, 1968) fail to receive support.

Further support for signal-detection theory comes from the results of the χ^2 test for independence. This was performed because the varying sample sizes may bias the estimates of a posteriori probabilities. The predictions of low-threshold theory imply that classifications of trials by rating response and by stimulus condition are independent. All χ^2 tests were significant at the 0.01 level, rejecting the hypothesis of independence for each of the tests. Therefore, Luce's low-threshold theory is unable to describe the results of this experiment. The evidence against the second prediction above argues against high-threshold theory as well.

The final approach to the assessment of sensory noise was the forced-choice second-guess experiment. The high-threshold theory prediction is that P_2 (c) = 0.33 for all values of P_1 (c). This prediction fails to receive confirmation, especially at the higher levels of detectability. The increase of P_2 (c) as P_1 (c) increases is predicted by signal-detection theory.

Two elements of signal-detection theory receive confirmation from these experiments: (1) the sensory excitation caused by internal noise alone can sometimes be greater than the excitation caused by stimulus plus noise, and (2) O can distinguish among his internal states, ranking them in order of signal likelihood. This provides evidence that the theory applies to a method of stimulation wholly unlike those previously studied.

As noted earlier, it was found (Rollman, 1969) that psychometric functions for electrocutaneous pulses are considerably steeper than those for visual or auditory stimuli and that the period of temporal summation is appreciably shorter. Most likely this is a consequence of bypassing the receptors with electrical stimulation and directly initiating the action potential at the peripheral nerve. The finding of a similar detection process for conditions that activate or bypass the receptors may indicate that transduction at the peripheral level is not an important source of variance for the distribution of signal plus noise. The steep ogive is compatible with the determination of the importance of sensory noise. In signal-detection-theory terms, the process of detection may be similar in all modalities; the somatosensory system differs from the others in that as the intensity of the electrocutaneous stimulus increases, the rate of increase in d' or Δm is greater than when corresponding increases in visual or auditory signals occur. Thus, although there may be important differences in the physiological mechanisms responsible for detection, the general framework of signal-detection theory seems to be supported by results of studies in audition, vision, taste, and the cutaneous senses.

REFERENCES

EGAN, J. P., SCHULMAN, A. I., & GREENBERG, G. Z. Operating characteristics determined by binary decisions and by ratings. Journal of the Acoustical Society of America, 1959, 31, 768-773.
EIJKMAN, E., & VENDRIK, A. J. H. Absolute sensitivity of sensory systems. Biophysical Journal, 1963, 3, 65-78.
GREEN, D. M., & SWETS, J. A. Signal detection theory and psychophysics. New York: Wiley, 1966.
LINKER, E., MOORE, M. E., & GALANTER, E. Taste thresholds, detection models, and disparate results. Journal of Experimental Psychology, 1964, 67, 59-66.
LUCE, R. D. A threshold theory for simple detection experiments. Psychological Review, 1963a, 70, 61-79.
LUCE, R. D. Detection and recognition. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), Handbook of mathematical psychology. New York: Wiley, 1963b. Pp. 103-189.
NACHMIAS, J., & STEINMAN, R. M. Study of absolute visual detection by the rating-scale method. Journal of the Optical Society of America, 1963, 53, 1206-1213.
ROLLMAN, G. B. Psychophysical detection of electrocutaneous pulses. (Ph.D. dissertation, University of Pennsylvania, 1967) Ann Arbor, Mich.: University Microfilms, 1968. No. 68-9231.
ROLLMAN, G. B. Electrocutaneous stimulation: Psychometric functions and temporal integration. Perception & Psychophysics. 1969, 5, 289-293.
SWETS, J. A. (Ed.), Signal detection and recognition by human observers. New York: Wiley, 1964.
SWETS, J. A., TANNER, W. P., JR., & BIRDSALL, T. G. Decision processes in perception. Psychological Review, 1961, 68, 301-340.
TANNER, W. P., JR., SWETS, J. A., & GREEN, D. M. Some general properties of the hearing mechanism. University of Michigan: Electronic Defense Group, Technical Report No. 30, 1956.
WICKELGREN, W. A. Testing two-state theories with operating characteristics and a posteriori probabilities. Psychological Bulletin, 1968, 69, 126-131.