Signal Detection Theory and the Psychophysics of Pain: An Introduction and Review

MARY ANNE LLOYD AND JAMES B. APPEL, PHD

A brief explanation of signal detection theory is presented, followed by a review of the literature in which differences in pain report were analyzed to determine if changes were in sensitivity (physiological processes) or in response bias (a subject's willingness to report that a signal occurred). Three kinds of studies are reviewed: modification, procedural, and comparative and normative. The advantages of using a signal detection approach in pain research are emphasized. It is concluded that experimenters applying signal detection analysis to pain research should give greater consideration to methodological procedures and be cognizant of all possibilities for shifts in sensitivity (d').

INTRODUCTION

Signal detection or sensory decision theory (SDT) has become invaluable as a method for analyzing situations calling for the discrimination of stimuli (perception). The value of SDT is that it separates response determinants into those that involve detection, and are presumably a function of physiological variables (sensitivity), from those that involve decision and reflect the subject's tendency to report that a given event has occurred (response criterion, likelihood, or bias). Thus, SDT has been applied to research in vigilance, perceptual selectivity (word-frequency experiments), memory, and other areas where discriminations are required from subjects. Recently, researchers have applied SDT to the analysis of pain report.

Researchers concerned with pain threshold became interested in SDT for two basic reasons. First, even in experiments in which all major variables were well-controlled, there was little agreement as to how to quantify pain threshold (1). Second, the traditional 50% threshold measure (which defines pain threshold as the stimulus intensity at which a subject reports that pain is present 50% of the time) was judged to be unreliable, because it cannot account for at least one crucial determinant of behavior, response likelihood. Using the 50% method, a low proportion of responses in which pain is reported indicates a high threshold *by definition*, and a high proportion of pain reports indicates a low threshold. Both high and low thresholds are said to reflect physiological factors, that is, the amount of pain that can be detected. Such a conclusion fails to deal with biasing factors, i.e., factors other than stimulus intensity that influence the subject's decision to report that pain is present (e.g., response consequences or cost). A good example of the importance of bias is the situation of a radar operator who has to both *detect* a blip on a radar screen and *decide* whether or

From the Behavioral Pharmacology Laboratory, Department of Psychology, University of South Carolina, Columbia, South Carolina 29208.

Address reprint requests to: Dr. James B. Appel

Received for publication April 28, 1975; final revision received October 1, 1975.

not an enemy aircraft is present. During World War II, a British operator would be biased toward reporting that an enemy aircraft was present because the consequences of "missing" an enemy would be very grave. However, the same operator in England today would be reluctant to call a blip a dangerous enemy aircraft and to take retaliatory actions. Researchers in pain suspected that some of the variability found in pain threshold or change in pain threshold following various modification procedures might be due to the uncontrolled influence of response bias factors and therefore turned to signal detection analysis (2). A natural step after applying SDT to pain threshold research was to apply SDT to the analysis of tolerance to pain to determine if changes in tolerance reflected altered sensitivity or bias.

The change from traditional threshold methods to signal detection analysis of pain research data has been profoundly important in the field of analgesia. With SDT, it is possible to determine (1) if "analgesic" effects following the administration of various drugs, or of other procedures such as suggestion or acupuncture, are due to procedurally-induced changes in sensitivity or to biasing factors of the experimental situation, which might affect a subject's tendency to respond, and (2) to determine if variables such as age, sex, or classification label affect sensitivity or bias regarding pain responses. In other words, SDT analysis allows one to more precisely separate sources of variance in pain report.

The present review has two purposes. The first is to present a simple mathematical explanation of SDT to acquaint readers with some of the concepts involved in this kind of analysis and to provide an introduction for some who hopefully will begin experimentation with procedures appropriate for SDT analysis. The second purpose is to review the literature wherein SDT has been applied to pain research to elucidate the strengths of SDT analysis and to point out weaknesses in designs so that future studies of pain research might yield more clear-cut results.

EXPLANATION OF SIGNAL DETECTION THEORY

Generally speaking, SDT provides two measures of the subject's performance in a discrimination situation. These two parameters, sensitivity and response bias, are statistically independent and reflect, respectively, the ability to detect a stimulus and the predisposition to respond that a stimulus has or has not occurred (3,4).

The theory assumes that a given stimulus leads to a sensory experience X (the evidence variable), which can be scaled, rated, or ranked on a continuum. Given an infinite number of presentations of a "noise" stimulus, there results a Gaussian (normal) distribution of sensory experience events on the continuum. Infinite presentations of a second stimulus consisting of noise + signal leads to a Gaussian distribution on the same continuum but further along, since the stimulus is of higher intensity. Thus, any discrimination experiment has at least two distributions of noise and noise + signal (hereafter referred to as signal), with the signal distribution further along the continuum of subjective experience (5) (see Fig. 1a).

If the signal distribution is well separated from the noise distribution, there will be very little overlap between the two curves and the discrimination will be easy. However, as the distributions become progressively closer and have a greater de-

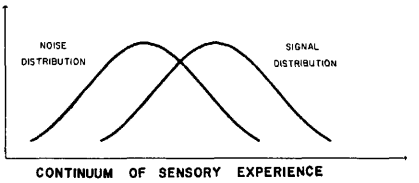

Fig. 1a. Hypothetical noise and signal distributions assumed in SDT.

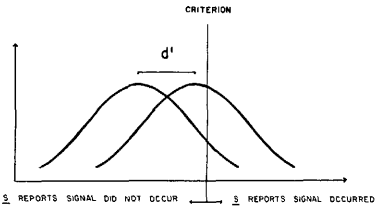

Fig. 1b. Representation of the criterion and d'.

gree of overlap, the discrimination becomes correspondingly more difficult. The sensitivity measure, usually called d', is intrinsically related to the signal and noise distributions; it is equal to the distance between the distribution means and is measured in Z scores of the noise distribution (below).

Response bias is represented by the subject's criterion for responding either that noise alone has occurred (signal is absent) or that a signal has occurred. The criterion can fall anywhere on the continuum and is expressed as β, the likelihood ratio (height of the signal distribution divided by the height of the noise distribution) at the criterion point. If X falls to the right of the criterion (i.e., exceeds the criterion), the subject will report that a signal has occurred; if X falls to the left (or below) of the criterion, the subject will report that a signal has not occurred, or that noise alone has occurred (see Fig. 1b).

If a person shifts his criterion to the right, he is more biased toward reporting that a signal did not occur, or he is using a stricter criterion. If a person shifts his criterion to the left, he is more biased toward reporting that a signal did occur and is using a more lenient criterion. If $0 < \beta < 1$, the subject is biased toward reporting a signal; if $\beta > 1$, the subject is biased toward reporting a signal did not occur. It can be seen from Fig. 1b that the subject could shift his criterion while d' remained the same. If the degree of overlap of the noise and signal distributions changes, it is possible for d' to shift and for β to remain unchanged (6).

In a discrimination experiment one can determine d' and β from the relative frequencies of a subject's responses. If a subject's task is to detect a signal, there are essentially four possible stimulus-response combinations. The subject can call a stimulus a signal when in fact a signal was presented (hit) or when a signal was in fact not presented (false alarm); he can call a stimulus noise when in fact no signal was presented (correct rejection) or he can call a stimulus noise when in fact a signal was presented (miss) (see Table 1).

TABLE 1. Four Combinations of Stimulus-Response Events in a Discrimination Task

	Response	
Stimulus	S	N
s	Hit	Miss
n	False alarm	Correct rejection

For purposes of determining d′ and β, only hit and false alarm rates are needed, from which the probabilities of a hit and false alarm are determined.

Probability of hit is determined by dividing the number of signal responses (S) made to signals by the number of signal presentations (s); false alarm probability is determined by dividing the number of signal responses (S) made to noise stimuli by the number of noise presentations (n), yielding conditional probabilities of S (S /s) given s (S /s), and S given n (S /n) (6) (see Fig. 1c).

From Fig. 1c, one can observe how d′ and β can be determined from a given set of hit and false alarm probabilities. By definition, $d' = [(\bar{X}_s - \bar{X}_n)/\delta n]$, where $\bar{X}_s$ = the mean of the signal distribution, $\bar{X}_n$ = the mean of the noise distribution, and δ_n is the standard deviation of the noise distribution. When the noise and signal distributions are Gaussian with equal variance, d′ can be determined from hit and false alarm probabilities by subtracting the Z score corresponding to the hit rate [Z (S /s)] from the Z score corresponding to the false alarm rate [Z (S /n)]: d′ = [Z (S /n) − Z (S /s)]. β is the height of the signal distribution divided by the height of the noise distribution. To determine β, the ordinate (y) of the respective signal (hit) and noise (false alarm) probabilities can be found in a normal curve table; the signal ordinate (ys) is then divided by the noise ordinate (yn): β = ys /yn.

The above is an oversimplified explanation of signal detection theory; there are many other measures of sensitivity and bias that might be better than d′ or β, depending on the nature of noise and signal distributions, the number of noise and signal presentations, and the type of experimental procedure used. For instance, in a discrimination experiment, more than one set of hit and false alarm rates is usually obtained for a given experimental manipulation. With many sets of hit and

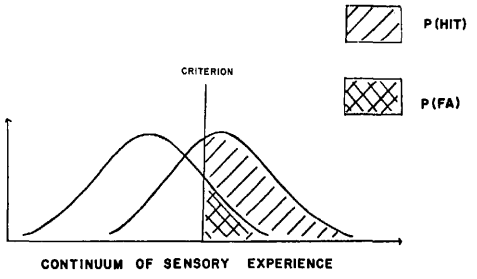

Fig. 1c. Hit and false alarm probabilities.

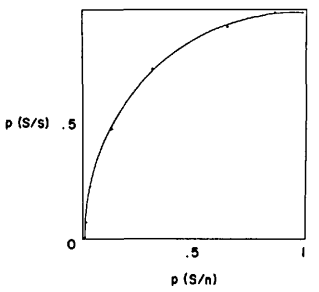

Fig. 2. Hypothetical ROC curve.

false alarm probabilities, a graph can be drawn in which p (S /s) is plotted as a function of p (S /n). Such a graph is called a receiver operating characteristic (ROC) curve (see Fig. 2). Points on the curve give an index of bias because they represent a range of criteria from very strict to very lenient, the criteria becoming more lenient as the points move upward and to the right. Another measure of sensitivity, called P (A), is represented by the area under the ROC curve (7) and is often a safer measure of sensitivity than d' if one is unable to present a large number of noise and signal stimuli (8).

In addition, if the variance of the signal distribution differs from the variance of the noise distribution, Δm, which is the value of Z (S /n) when Z (S /s) = 0 (8), or d'_e (9), which equals 2 Z (S /n) or 2 Z (S /s) at the point where a plot of Z (S /s) and Z (S /n) scores (double probability plot) crosses the negative diagonal are better indices of sensitivity than d' (8). If the signal variance differs from the noise variance, ys must be multiplied by the slope of the double probability plot before being divided by yn to find β (8) (Fig. 3).

There is also a measure of bias that can be used more safely than β when there are few noise and stimulus presentations. B, a nonparametric measure of bias, is expressed in category scales and corresponds to the rating scale category (below) where the p (S /s) = p (S /n) = 1 (8). B can only be used if the signal and noise distributions have equal variances.

There are many different procedures for obtaining data that can be analyzed within the framework of SDT. For example, sub-

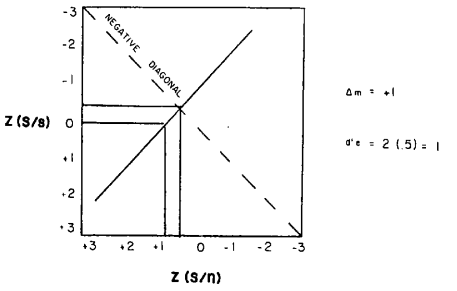

Fig. 3. Double probability plot showing Δm and d'_e. Hypothetical plot of Z scores from hit and false alarm rates.

jects might be asked to place pain intensities in rating categories such as "no pain," "warmth," "faint pain," and "extremely painful" (8). By calculating cumulative probabilities for each stimulus intensity across rating scales, ROC curves for each adjacent stimulus pair can be obtained. Probability of a hit is the cumulative probability of the higher intensity stimulus of a pair being placed in a given category; probability of a false alarm is the cumulative probability of the lower stimulus being placed in the same category.

Despite the profusion of SDT analyses and procedures, the basic advantage of signal detection remains the same; one is able to separate how well a subject detects a stimulus (sensitivity) from his tendency to report that he did or did not detect a stimulus (response bias).

APPLICATION OF SDT TO PAIN RESEARCH

The experiments to be reviewed are organized (below) into three categories (modification, procedural, and normative and comparative). The modification category includes studies concerning placebos, drugs, suggestion, modelling, and acupuncture; the procedural category includes studies that investigate different procedures employing the radiant heat pain induction method; and the normative and comparative category includes studies in which normative discriminability is determined, or in which different populations of subjects are compared.

Modification Studies

The first study in which SDT was used was that of W. C. Clark (1), who suspected that response bias rather than sensitivity is altered when pain report varies following administration of placebos.

Each of 22 subjects participated in control and placebo sessions counterbalanced to control for order. Five radiant heat stimuli were randomly presented to each subject; the stimuli were tailored to each subject's predetermined threshold level, with the average levels of stimuli being: $S_0 = 0$, $S_1 = 94.4$, $S_2 = 208.6$, $S_3 = 271.0$, and $S_4 = 308.8$ mcal/sec/cm^2. For the control session, Ss were told that the experiment was to determine ability to detect warmth, heat, and faint pain. The Ss were given a list of 13 categories ranging from no pain to extremely painful, by which to rate each stimulus they experienced. During the placebo session, the Ss swallowed a bitter tasting white powder that was described as an analgesic previously used only for internal pathological pain. Ss were given a drug reaction checklist, and all but one S reported having side effects following administration of the placebo (e.g., numbness of tongue, faint nausea, confusion, euphoria).

From the rating scale data, four different noise and signal distributions were generated, i.e., S_4 was a signal when S_3 was noise; S_3 was a signal when S_2 was noise, and so forth. Values of d' and β were calculated for each pair of noise and signal distributions for both control and placebo sessions. Values of d' were not altered by the placebo nor were there any order effects, although relative sensitivity increased at higher intensity levels. However, placebo administration significantly increased β for each pair of stimuli. In other words, following placebo administration, Ss raised their criteria for or were less predisposed to report pain, heat, or warmth.

Feather, Chapman, and Fisher (5) con-

ducted a study that was basically a replication of the work of Clark (1). These investigators used only two radiant heat stimuli (250 and 310 mcal/sec/cm^2). The Ss participated in two sessions during which they assigned the intensity of stimuli to one of four categories (just noticeable to painful). Before the second session, Ss were given a placebo described as a fast-acting analgesic. Ss reported side effects after placebo administration.

Values of d' and standard scores associated with hit rate (response bias) were calculated for both control and placebo sessions relative to warm, hot, and painful categories. Values of d' were not significantly different during control and placebo sessions. Bias increased significantly for the painful category but, unlike the results of Clark (1), increases in β were not significant in the hot or warm categories.

SDT has been used to analyze drug-induced changes in pain report in two studies (10,11). Chapman, Murphy, and Butler (10) administered 33% nitrous oxide to 14 male Ss randomly assigned to two groups differing with respect to order of control and test sessions. Four radiant heat intensities, tailored for each subject, were assigned to one of six categories ranging from no pain to strong pain. During each session, 50 stimuli of each intensity were presented to each S who wore an anesthesia mask through which either room air or 33% nitrous oxide was delivered, depending on the session.

The data were analyzed as follows: values of d' were determined for each nonzero stimulus as the distance between means of the nonzero distributions and the zero distribution for both control and test sessions. Values of d' changed significantly for the three most intense stimuli; thus, 33% nitrous oxide reduced sensitivity overall. However, d' values between adjacent nonzero stimuli were also determined; when this was done, it was found that the gas did not change sensitivity relative to differences between adjacent stimuli.

Response bias was quantified as percent bias (12) for rating categories for each S. Subjects were significantly less willing to report moderate and faint pain, although willingness to label a stimulus as hot did not change. Thus, administration of 33% nitrous oxide in contrast to placebo procedures, changed both sensitivity, and response bias.

A second study was conducted by Chapman and Feather (11) to evaluate the effects of diazepam, Valium (10 mg), on pain report. In two separate experiments, it was found that diazepam increased tolerance (willingness to endure relatively high levels of pain stimuli) relative to aspirin and placebo if an S had had some previous experience with a given pain condition. A third study was carried out to determine whether during this tolerance sensitivity or bias is modified by administration of diazepam.

Two groups of subjects, drug and placebo, were apparently exposed to two conditions, drug (or placebo) capsule and no capsule. Five radiant heat stimulus intensities were presented, tailored to each S. Each stimulus was presented 50 times, and Ss responded by rating their sensory experience in categories ranging from no pain to strong pain. Sensitivity and bias were calculated for four pairs of adjacent stimuli. Diazepam (10 mg) did not affect sensitivity or response bias relative to placebo.

Experimental designs similar to those already described have been used to assess the effects of such variables as suggestion, modelling, and acupuncture on tendency to respond to pain. For example, the effects

of suggestion on a withdrawal response to painful radiant heat stimuli have been analyzed by Clark (2). During each of two periods, six stimuli (0-435 mcal/sec/cm²) were presented 12 times to 10 paid college students who were asked to rate the stimuli in categories from no pain to intense pain. During an "initial period," Ss were told to remove the projector from the skin if the stimulus was "getting too hot (p. 275)." Before the "Post-suggestion Period," Ss were told that the previous stimulation had fatigued their pain receptors and that they should be able to endure more pain before withdrawing. They were also told that the second test was a test of "ability to endure maximal pain (p. 275)." Thus, Clark manipulated suggestion ("pain receptor change") and relative permissiveness of instructions (permissive to nonpermissive).

In this experiment, mean values of d' and response bias relative to only two pairs of adjacent stimuli (370 and 435 mcal/sec/cm² and 305 and 370 mcal/sec/cm²) were determined. The withdrawal responses (tolerance) occurred only following the higher intensity pair; reports of very paint pain (threshold) occurred following the lower pair of intensities. Experimental manipulations did not increase d' significantly for either the lower or higher pair; however, they did significantly increase the criterion for withdrawal, although not for very faint pain. Therefore, Clark concluded that sensory experience (d') was not altered by suggestion or permissiveness of instructions, i.e., by instructional set.

A second study by Clark and Goodman (13) was conducted to determine the effects of suggestion on both raising and lowering the report of pain. In addition to the question of whether suggestion alters sensitivity or bias, the authors were interested in whether suggestion had more of an effect on report of very faint pain or on report of intolerable pain. Forty paid college students were randomly assigned to one of four groups; two groups received instructions to either raise or lower pain threshold, and two received instructions to either raise or lower pain tolerance. "Pre-suggestion Period" instructions were identical to those in Clark's study (2); once again, both suggestion and permissiveness of instructions were varied.

In analyzing the data, sensitivity and bias were calculated for only two adjacent stimulus pairs associated with very faint pain and with withdrawal from pain. There were no significant differences in d' changes for the two intervals studied. Suggestion of decreased sensitivity significantly raised criteria for both very faint pain and withdrawal; however, suggestion of increased sensitivity did not lower criteria for either.

Craig and Cohen (14) used signal detection to determine if groups who had models tolerant or intolerant to pain differed in respect to sensitivity. Sixty shocks, 12 at each of five intensity levels, were randomly delivered to subjects who rated the intensities of a 10-category scale. Subjects were or had been paired with a confederate who modelled either tolerance or intolerance to the painful stimuli. Results indicated that sensitivity was significantly greater for the group with the intolerant model than for both the tolerant and control group. Response criteria changes (β) were not computed. The authors conclude that an intolerant modelling situation can affect physiological processes associated with perception of electrical stimulation and actually increase a person's vulnerability to the sensory expeience of pain. The results are very interesting, particularly in light of Clark and Goodman's (13) results,

which indicate that suggestion of increased vulnerability to pain did not affect sensitivity. The discrepancy of results could be due to the different experimental manipulations used (modelling vs. instructions) or to procedural differences, which will be discussed in detail at the end of the modification section.

There have been two studies designed to determine if acupuncture affects sensitivity or response bias. In an experiment by Clark and Yang (15) 24 radiant heat stimuli were presented on the forearms of 12 Ss who placed the stimuli into 12 categories ranging from no pain to withdrawal. Following initial testing, acupuncture needles were inserted at sites that are used in surgery of the arm. One-half of the Ss received needles in their left arm and the other half in their right. The needles were electrically stimulated by standard acupuncture apparatus for 15 to 20 min, during which time thermal sensitivities were again measured. After the needles were removed, a final sensitivity measure was taken.

Values of d′ and β were determined for only one adjacent stimulus pair, 370 and 425 mcal/sec/cm^2. There were no significant differences in d′ for control and acupunctured arms or between periods before and after acupuncture treatment. Subjects were less willing to report pain in the treated arm during acupuncture application, although there were no significant differences in bias between periods before and after treatment. Thus, acupuncture raised the criterion for reporting pain but did not actually reduce sensitivity (i.e., have an "analgesic" effect). Presumably, Ss experienced the same amount of physical stimulation in both arms, but were less predisposed to admit that pain occurred in the acupunctured arm during acupuncture treatment. It was concluded that speculations concerning the physiological mechanisms underlying the process of acupuncture are premature until the role of response bias can be made clear.

Chapman, Gehrig, and Wilson (16) obtained somewhat different results when they compared the effects of acupuncture and 33% nitrous oxide to appropriate controls. Four levels of electrical stimulation (tailored to each S) were applied to the tooth pulps of 42 male volunteers; zero was the lowest stimulus level. All Ss received 100 practice trials with all four stimuli and returned for two further sessions. During test sessions, each S received 300 stimulus presentations, 75 at each intensity level. During all sessions, Ss rated each stimulus on a seven-category scale ranging from nothing to moderate and strong pain.

During the second test session (treatment), Ss given 33% nitrous oxide inhaled the gas through a nasal mask. For acupuncture Ss, needles were inserted 2 cm bilaterally between the thumb and first finger of each hand. The needles were then electrically stimulated. Subjects in treatment groups were tested after a 20-min period; controls were tested under conditions of the first test session.

Values of d′ were calculated for all adjacent stimulus pairs. Both acupuncture and nitrous oxide significantly reduced sensitivity. Although the acupuncture group did not significantly differ from the nitrous oxide group with regard to the total sum of change in d′ across levels, acupuncture consistently lowered sensitivity at all three levels, whereas nitrous oxide primarily affected lower levels. Response bias for labeling the high intensity stimulus as "very faint pain" was calculated; nitrous oxide and acupuncture Ss were significantly less willing to report very faint pain than were control Ss. Thus,

acupuncture and nitrous oxide reportedly reduced both sensitivity to painful stimuli and the predisposition to report pain.

Chapman et al. (16) found that acupuncture changed sensitivity, whereas Clark and Yang (15) found that it did not. The discrepancy could be due to the different pain induction measures used, the relative noxiousness of the stimulation used (Clark and Yang used seemingly relatively more noxious stimulation than did Chapman et al.), or that the Ss in Clark and Yang's study received no practice trials. This discrepancy, as well as others, brings up an important point: the results of the modification experiments we have been discussing are difficult to compare because of the use of different standards of SDT analysis.

First, in some cases (1, 2, 13–15) each stimulus was presented very few times. In these studies researchers were generally concerned with whether sensitivity and bias differed in control and experimental sessions, so that paucity of stimulus presentations was not as important as it would have been had they been attempting to determine precise, parametic, normative data. Nevertheless, McNicol (6) warns that at least 50 presentations are needed to realistically estimate ROC curves in any design; moreover, he argues that in situations with few stimulus presentations, P (A) rather than d′ should be used to estimate sensitivity and B rather than β should be used as an estimate of bias.

Second, in some studies (1, 2, 5, 11, 13–15) the slopes of double probability plots were not delineated. This means that the variances of signal and noise distributions could have been unequal, in which case Δm or d'_e would have been more appropriate measures of sensitivity than d′. The methodology of Richards and Thornton (17), which provides for the averaging of data with the slope of the plot taken into account, was used in some cases and is appropriate. However, many researchers failed to account for slope in any manner.

Finally, some researchers (5, 14, 15) failed to use, or failed to report the use of, a zero stimulus (no physical pain stimulus). Without a zero stimulus, there can be no analysis of changes in d′ between non-zero distributions (pain stimulus present) and the zero distribution (no pain stimulus). It is possible for all nonzero distributions to shift significantly toward the zero distribution without shifting significantly in relation to each other; therefore, a zero stimulus is a necessary control in studies involving the detection of pain. Along the same lines, in analgesia research, it is necessary to analyze changes between all adjacent stimulus pairs, even though interest may center around pairs concerned with certain criteria such as very faint pain or intolerable pain. In the absence of such analysis, a shift in sensitivity can occur somewhere along the sensory continuum without necessarily being detected. The criticisms are summarized in Table 2.

To briefly summarize studies involving pain modification, placebos and suggestion appear to modify bias but not sensitivity relative to adjacent stimulus pairs; an intolerant model reportedly increases sensitivity. Nitrous oxide reportedly changes bias and sensitivity across all stimuli but not between all adjacent stimulus pairs. Diazepam reportedly modifies neither sensitivity nor bias. Results concerning the effects of acupuncture on pain sensitivity and bias conflict.

Procedural Studies

Two studies have been concerned specifically with analyzing the relative merits of different procedures that can be

TABLE 2. Summary of Criticisms: Modification Studies

Modification study	Modification procedure	Change in sensitivity	Change in response bias	Sufficient stimulus presentations	Variances of distributions accounted for	Zero stimulus reported	Analysis of all adjacent pairs reported
Clark, 1969	Placebo	No	Yes	No	No	Yes	Yes
Feather, Chapman, and Fisher, 1972	Placebo	No	Yes	Yes	No	No	—
Chapman, Murphy, and Butler, 1973	Nitrous oxide	Yes	Yes	Yes	Yes	Yes	Yes
Chapman and Feather, 1973	Diazepam (Valium®)	No	No	Yes	No	Yes	Yes
Clark, 1974	Suggestion	No	Yes	No	No	Yes	No
Clark and Goodman, 1974	Suggestion	No	Yes	No	No	Yes	No
Craig and Cohen, 1975	Modelling	Yes	—	No	No	No	—
Clark and Yang, 1974	Acupuncture	No	Yes	No	No	No	No[a]
Chapman, Gehrig, and Wilson, 1975	Acupuncture	Yes	Yes	Yes	Yes	Yes	Yes

[a] Clark and Yang have since reported that they used a zero stimulus and analyzed changes in all adjacent pairs (Science 189:66–68, 1975).

used in SDT analysis of pain modification data. Although rating procedures (used in all of the studies thus far reviewed) permit the calculation of many criteria and an ROC curve to be obtained in a single session, the rating procedure tends to yield a lower d' than corresponding binary (e.g., yes-no) decision procedures (6). Therefore, Clark and Dillon (18) hypothesized that the best approach might be to have each subject concurrently make both binary and rating scale decisions.

Three radiant heat stimuli (one null and two at higher intensities) were presented to 24 college students. Three sessions counterbalanced for order effects were used, i.e., during half of the session, the reference and "low" stimulus were presented, and during the other half, the reference and "high" stimulus were presented. Each stimulus was presented 25 times. The Ss had to decide whether the reference or a higher intensity had been presented by responding either "high" or "low" (yes-no task). During another session (sensory intensity rating), Ss rated the stimuli on a scale ranging from no pain to very painful and withdrawal. The final (concurrent) session was a combination of the two (single) procedures; Ss made a binary decision followed by a rating scale decision.

There were no differences between single and concurrent measures for sensitivity and response bias; thus, a concurrent procedure is feasible in pain research. For both single and concurrent procedures, the binary decision process yielded significantly higher d' values. It was concluded that a concurrent procedure should be used; the yes-no task is better for the determination of d' since higher values of this parameter are presumably more accurate indices of sensitivity; the rating scale should be used to estimate bias since many criteria can thereby be determined at one time.

A second procedural study was conducted by Clark and Mehl (19). These investigators were concerned with the assumption that the invariance of SDT parameters across different procedures is based on research with few subjects, told to detect audio or visual stimuli for relatively few sessions. The purpose of their study was to compare d' values and ROC curves obtained with various procedures over a wide range of thermal heat stimulation.

During a 10-week period, each of four Ss participated in 34 sessions controlled for order effects. During each session, Ss discriminated between two radiant heat stimuli that differed in intensity by 25 mcal /sec /cm.^{2} In all, there were 17 pairs of stimuli ranging from 0-25 to 400-425 mcal /sec /cm.^{2} After each stimulus presentation, the Ss made a number of sensory decisions: (1) one-interval binary or yes-no task (S responded either "high" or "low"); (2) one-interval confidence rating from 1 to 12 (S rated how confident he was that a higher stimulus had been presented); (3) one-interval magnitude rating (S rated his sensory experience by means of an 11-category scale from no pain to excruciatingly painful); and (4) two-interval binary or two alternate forced choice task (following an additional stimulus presentation, S decided which of the two was the higher intensity stimulus).

Values of d' were analyzed for the different methods used across all stimulus intensities. These values varied significantly over method and intensity, i.e., d' was significantly higher for the binary procedures than for the confidence rating procedure or for the magnitude rating procedure. Values of d' for the different binary procedures did not differ

significantly from each other. Because there was no method × intensity interaction, it was concluded that differences in method were consistent over the entire range of stimuli used. It was also concluded that the one-interval binary and two-interval procedures are superior to rating procedures because they yield higher d′ values and that, when two thermal intensities are used in pain research, the most valuable procedures for estimating d′ are the one- or two-interval binary. However, because pain research generally necessitates obtaining many criterion measures, they recommend, as did Clark and Dillon (18), using a binary procedure to estimate d′ in conjunction with a rating procedure to estimate β or pain criteria.

One criticism of both of the otherwise excellent studies (18,19) is that only rating scales with 10 categories were used. It would have been interesting had a scale with fewer categories been included to determine whether it yielded values closer to the binary d′ values. In addition, the researchers did not provide a sufficient number of stimulus presentations (above); Clark and Dillon (18) did not mention double probability plot, whereas Clark and Mehl (19) both provided a sufficient number of stimulus presentations and mentioned double probability plot.

Normative and Comparative Studies

The following normative and comparative studies may be mentioned briefly to acquaint the reader with the varied applications of SDT to pain research. Clark (20) conducted a normative study to determine the ability of trained subjects to discriminate over a range of thermal intensities from 0 to 425 mcal/sec/cm.^{2} The procedure and methodology of Clark and Mehl (19) were used. As in the Clark and Mehl study, values of d′ varied with method and stimulus intensity. The results also indicated that ability to discriminate between stimuli separated by 25 mcal/sec/cm^2 increased between 0 and 75 mcal/sec/cm^2, then decreased until it reached the initial level at 150 mcal/sec/cm^2. From 150 to 425 mcal/sec/cm^2, discriminability increased in a linear fashion; the slope of the line was significantly different from zero.

A study by Clark and Marmor (21) is important because it is the only one in which SDT was used in conjunction with a physiological measure. In this experiment, d′ values for verbal report were compared with values of the palmer skin potential (PSP). A PSP above a certain level was calculated as a hit when a signal was present and as a false alarm when noise was present.

Electric shocks tailored individually to each of eight Ss were randomly presented during a 1½-hr session. Eight nonzero intensities were presented eight times; zero intensity was presented 16 times. Ten seconds after the stimulus presentation, each S rated the magnitude of the sensation from 0 to 10. If he rated a stimulus 0, he also rated certainty of his answer from 0.1 to 0.9. Values of d′ for PSPs were calculated from the absolute sums of the largest negative and positive deflections, which occurred within 10 sec after a stimulus presentation. No comparison statistics were used; however, a descriptive table of the results indicated that verbal report was a more sensitive measure (yielded higher d′ values) than PSP responses.

Clark and Rubin (22) used SDT to compare normal controls and psychiatric patients with regard to possible differences in sensory capacity. Control Ss were 16 male and 16 female college students divided into a younger (19.0 years) and older (22.9 years) group. Psychiatric pa-

tients were 8 males and 24 females from various diagnostic categories; they were also divided into a younger (19.8 years) and older (39.1 years) group. Procedures for presenting noxious stimuli were the same as those used by Clark and Dillon (19), except that each stimulus was presented 22 times.

Values of d′ and β for each stimulus intensity relative to zero were determined. Neither d′ nor β varied with knowledge of results; d′ did not vary between groups nor with order of stimulus presentation. However, for the high intensity stimulus, older patients set a significantly higher criterion than normals and younger patients, and Ss experiencing the higher stimulus first set a lower criterion. Thus, the SDT analysis revealed that the higher threshold of the older patients and the lower threshold of Ss who experienced the higher intensity stimulus first were due to changes in bias and not in sensitivity. Considering that the older psychiatric patients differed significantly from the other three groups, it would have been more appropriate had the Es used "old" controls closer in age to the older psychiatric patients.

Clark and Mehl (23) carried out a study to investigate the effects of age and sex on d′ and three measures of response bias: Cvfp (distance of the criteria from the noise mean in Z scores of the noise distribution); Lvfp (β of the criterion for very faint pain); and CVFP (mcal/sec/cm^2 per d′). The subjects were 32 females and 32 males whose mean age for groups of young and old were female, 22.2 and 49.4 years and male, 23.9 and 43.8 years. Six radiant heat stimuli ranging from 0 to 300 mcal/sec/cm^2 were presented 16 times to each S who rated intensities on a 10-category scale.

Values of d′ and β were calculated for each adjacent stimulus pair. Generally speaking, young Ss and older males reported moderate pain equally, but young Ss were more willing to label the experience as painful. Older women also set a high criterion for reporting pain, but the criterion was accompanied by a decrease in d′, indicating that lower pain responses in older women were due to decreased pain sensitivity as well as a higher criterion.

Comparison of different measures of response bias indicated that values of Cvfp basically agreed with values of Lvfp; this is to be expected since both are statistically independent of d′. However, the values differed somewhat, and the Es suggest that until more is known about the two measures, both should be reported. Values of CVFP were obtained by combining d′ and the response criteria; they were not in agreement with either Lvfp or Cvfp.

SUMMARY AND CONCLUSIONS

The foregoing has been a review of studies that involved the use of SDT in pain research. The studies concerned modification of pain threshold and tolerance by placebos, drugs, suggestion and acupuncture, questions of procedure, and comparative and normative research. Conclusions were that placebos and suggestion modify bias but not sensitivity; an intolerant model reportedly increases sensitivity. The effects of the few drugs thus far studied vary—diazepam modifies neither sensitivity nor bias; nitrous oxide reportedly modifies both sensitivity and bias. The results concerning acupuncture were inconclusive; the meaning of d′ must be qualified relative to procedural and methodological details. Procedural studies suggest that the best SDT experi-

mental design relative to pain research is to obtain both binary and rating scale data concurrently.

Signal detection has been used to analyze normative data concerning discriminability across noxious intensities. It was determined that verbal report yields higher d′ values when compared with an autonomic measure (palmer skin potential). Results of other comparative studies indicate that age and sex affect bias with the exception that older women may be less sensitive to noxious stimuli than younger subjects and older men. Psychiatric patients differ from normal controls only in willingness to report pain.

There are methodological standards that were not adhered to in all cases. Future researchers using SDT should be cognizant of problems involved in averaging data, presenting sufficiently large numbers of stimuli, and estimating variances from double probability slopes. Most importantly, researchers should be aware of the necessity for a zero stimulus and analyses for changes in d′ between all adjacent stimulus pairs. The failure to analyze all possible d′ changes could lead to many Type II errors if, in fact, an experimental manipulation changes sensitivity differently relative to all noxious stimuli presented.

It is clear that SDT has proven to be an invaluable method of analysis in pain research, because it separates changes in sensitivity from changes in response bias. It is probably the most powerful statistical tool yet developed for analyzing pain research data. Thus far it has enabled researchers to demonstrate the effectiveness of various modification techniques on physiological processes independently of their effects on predisposition to respond to pain stimuli. In addition, discrepancies between groups regarding pain report have been separated into those resulting from sensitivity differences and those resulting from bias differences. Hopefully, future researchers in the area of pain will incorporate signal detection methodology into their work in order to more efficiently account for variance in pain report.

REFERENCES

1. Clark WC: Sensory-decision theory analysis of the placebo effect on the criterion for pain and thermal sensitivity (d′). J Abnorm Psychol 74:363–372, 1969
2. Clark WC: Pain sensitivity and the report of pain: an introduction to Sensory Decision Theory. Anesthesiology 40:272–287, 1974
3. Dykstra LA, Appel JB: Effects of LSD on auditory perception: a signal detection analysis. Psychopharmacologia (Berl), 34:289–307, 1974
4. Haber RN, Hershenson M: The Psychology of Visual Perception. New York, Holt, Rinehart, and Winston, 1973
5. Feather BW, Chapman CR, Fisher SB: The effects of a placebo on the perception of painful radiant heat stimuli. Psychosom Med 34:290–299, 1972
6. McNicol D: A Primer of Signal Detection Theory. London, George Allen & Unwin, 1973
7. Pollack I, Norman DA: A non-parametric analysis of recognition experiments. Psychonom Sci 1:125–126, 1964
8. Green DM, Swets JA: Signal Detection Theory and Psychophysics. New York, Wiley, 1966
9. Egan JP, Clarke FC: Psychophysics and signal detection, in Experimental Methods and Instrumentation in Psychology (edited by JB Sidowski). New York, McGraw-Hill, 1966

10. Chapman CR, Murphy TM, Butler SH: Analgesic strength of 33% nitrous oxide: a signal detection theory evaluation. Science 179:1246–1248, 1973
11. Chapman CR, Feather BW: Effects of diazepam on human pain tolerance and pain sensitivity. Psychosom Med 35:330–340, 1973
12. Hodos W: Non-parametric index of response bias for use in detection and recognition experiments. Psychol Bull 74:351–356, 1970
13. Clark WC, Goodman JS: The effect of suggestion on pain sensitivity and pain report: a sensory decision theory analysis. J Abnorm Psychol 83:364–372, 1974
14. Craig KD, Cohen S: Signal detection analysis of social modelling influences on pain expressions. J Psychosom Res 19:105–112, 1975
15. Clark WC, Yang JC: Acupunctural analgesia? Evaluation by signal detection theory. Science 184:1096–1097, 1974
16. Chapman CR, Gehrig JD, Wilson ME: Acupuncture and 33% nitrous oxide modify detectability of painful dental stimuli. Anesthesiology 42:532–537, 1975
17. Richards BL, Thornton CL: Quantitative methods of calculating the d′ of signal detection theory. Educ Psychol Measurements 30:855–859, 1970
18. Clark WC, Dillon DJ: SDT analysis of binary decisions and sensory intensity ratings to noxious thermal stimuli. Percept Psychol 13:491–493, 1973
19. Clark WC, Mehl L: Signal detection procedures are not equivalent when thermal stimuli are judged. J Exp Psychol 97:148–153, 1973
20. Clark WC: d′ and the Weber ratio for warmth, heat, and pain. Paper presented to the Psychonomic Society, 1971
21. Clark WC, Marmor E: Comparison of sensory sensitivities (d′) determined from verbal and autonomic responses to electrical stimulation. Paper presented at the Eastern Psychological Association, 1969
22. Clark WC, Rubin LH: Signal detection theory analysis of threshold differences in psychiatric patients and normal controls. Paper presented at the 77th Annual APA Convention, 1969
23. Clark WC, Mehl L: Thermal pain: a sensory decision theory analysis of the effect of age and sex on d′, various response criteria, and 50% pain threshold. J Abnorm Psychol 78:202–212, 1971