Advances in Pain Research and Therapy, Vol. 1,
edited by J. J. Bonica and D. Albe-Fessard.
Raven Press, New York © 1976.

Signal Detection Theory Assessment of Pain Modulation: A Critique

Gary B. Rollman

Department of Psychology, University of Western Ontario,
London, Ontario, Canada

When pain researchers attempt to modulate acute or chronic pain, they vary greatly in the methods they use to test the efficacy of their treatments. The persistent problem of pain measurement (1,11,20,25) centers on the difficulty in assessing the degree of discomfort felt by a patient and the need to disentangle sensory and emotional contributions to the pain response (18). Traditional pain and tolerance thresholds are not well suited for understanding the nature of pain modulation since they confound the sensory and motivational components.

A relatively new psychophysical methodology, signal detection theory or sensory decision theory (16), has held out the hope that some of these problems might be overcome. The theory presents a means by which one can independently measure the sensitivity and response bias of an observer, thus separating the sensory and cognitive factors underlying his performance.

Recognizing the desirability of independently assessing a subject's sensory experience and his proclivity for describing such an experience as painful, a number of pain researchers have adopted signal detection theory procedures in their investigations. This chapter summarizes and critically reviews this existing literature, suggesting that the use of the signal detection model is generally inappropriate and that the conclusions drawn by the authors may be unwarranted.

SIGNAL DETECTION THEORY

Clark (8) has provided an excellent account of the fundamental aspects of signal detection theory in the study of pain, and several other publications (e.g., 17,19,27) have demonstrated its wide applicability. The model assumes that the detection of a weak stimulus requires the observer to make a statistical decision. Since the stimulus, if it is there, is superimposed on a randomly varying background of internal "noise" arising from spontaneous neural discharges, each trial involves a decision as to whether the activity at that time is more likely to result from the signal added to the noise or from the internal noise alone. This decision is based on the magnitude of the sensory experience and the criterion adopted by the observer for making a positive decision.

In the simplest case, described in Fig. 1, the model assumes that both the noise and the signal plus noise distributions are normal or gaussian and of equal variance. When the signal is weak, the two distributions show considerable overlap. This is reduced as the signal strength increases, since the signal plus noise distribution then moves to the right. A parameter, d′, indicates the separation of the two hypothetical underlying curves and thus represents the observer's true sensitivity for detection of their difference.

An experiment involves a large number of noise or blank trials and a large number of trials presenting signal plus noise. Because of the overlap, the observer can never be certain whether or not a stimulus occurred. Therefore, he establishes a criterion for making an affirmative response, setting it to the right if he is conservative (thus minimizing "false alarms" at the cost of a reduced "hit" rate) and setting it to the left if he is more liberal (thus maximizing hits at the cost of increased false alarms). An independent parameter, described by various names, represents the observer's response bias or criterion. Changes in the subject's motivational state can shift the criterion and alter his hit rate (and thus his "threshold"), even though his sensitivity remains unaltered. However, d′ would remain constant.

The same arguments can be advanced to describe discrimination between two more intense stimuli, each of which is added to the internal noise. If stimuli are easily confused, the signal detection calculations (8) yield a low value of d′. Since pain studies require strong stimulus levels, investigators have typically examined discrimination after a variety of modulation procedures.

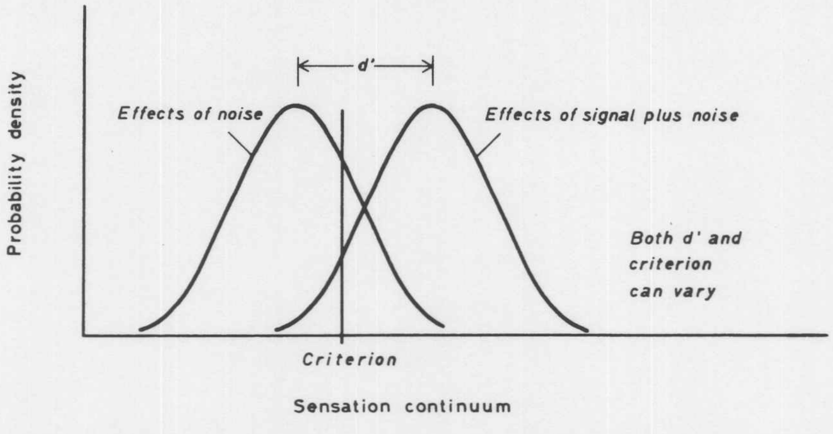

FIG. 1. Representation of the overlapping effects of noise and signal plus noise assumed by the signal detection theory model.

EXPERIMENTAL STUDIES

The typical pain experiment presented several intensities of a noxious stimulus, such as radiant heat or electric shock, which the observer was asked to rate on a scale ranging from "nothing" to "extremely painful." Rating scales have often been used in the signal detection framework since they are assumed to provide information based on a number of simultaneously held criteria rather than on a single one. Although Clark and Dillon (9) recommended that data from binary decisions (such as no-yes or low-high) rather than sensory intensity ratings be used to compute d′, all researchers have maintained the rating procedure.

Most of the studies have deviated from standard practice in a number of potentially important ways: subjects were asked to report the intensity of strong stimuli rather than confidence in their detection or discrimination ability; the scales often included as many as 13 categories; up to four different stimulus levels were intermixed with blank trials; and generally few trials per intensity were presented to each observer.

Clark (7) found that an orally administered placebo, described as a potent analgesic, increased the pain threshold but left the discriminability parameter unaltered. The effect of the placebo was to increase the criterion for issuing a report of pain. Clark and Mehl (12) concluded that older men and women show higher pain criteria than younger ones, and that older women also exhibit a reduction in sensitivity. Clark and Goodman (10) gave instructions to increase or decrease pain or tolerance thresholds, modifying behavior in the desired manner. Changes in the criterion, but not discriminability, caused them to conclude that suggestion (and, perhaps, hypnosis and counterirritation) forces subjects to alter the level of sensory activity required before they will verbally report pain, even though their sensations are unchanged.

Clark and Yang (13) applied signal detection theory to test the effects of a putative therapeutic agent, acupuncture, finding no significant differences in d′ between needled and control arms or between the periods before, during, or after the 15- to 20-min stimulation period. The criterion parameter was significantly higher in the experimental arm during acupuncture than in the control arm, leading to their conclusion that acupuncture altered the likelihood that subjects would admit to feeling pain, but did not change the pain experience.

Some Critical Remarks on Clark's Hypotheses

Clark proposed that sensory modulation alters the neural effects of a noxious stimulus, thus reducing the d′ in the same manner as if the physical intensity of that signal were weakened. Figure 2 demonstrates that this is not necessarily the case. If acupuncture truly functions as an analgesic by

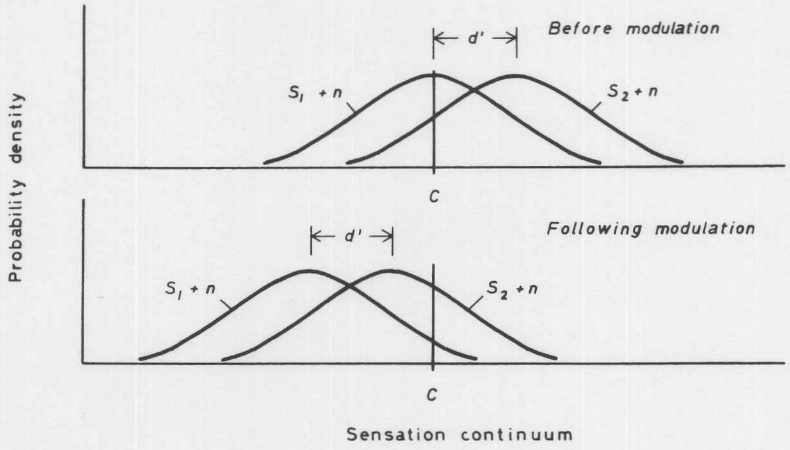

FIG. 2. Representation of the change produced by an analgesic modulation procedure. Sensory effects of noise plus both the weaker (S_1) and stronger (S_2) stimuli are reduced, but d′ remains constant.

modulating the neural response to potentially painful stimuli, the activity produced by both of the intense heat levels should be reduced. In that case, the distributions of both signal levels will be shifted to the left, but the discriminability index, d′, which represents the distance between them, could be unaffected. The sensory pain components of both stimuli could be greatly reduced, yet d′ would not change. The shift that would be observed is one in criterion level, but this would be illusory.

Clark's sensory magnitude criterion could be constant with respect to the level of activity required to elicit a report of pain, but one would mistakenly conclude from performing the signal detection calculations that it had shifted in a conservative direction. The basis for this problem is that no fixed distribution exists in pain modulation—both distributions are free to vary. As a result, acupuncture might act as a powerful analgesic, yet signal detection analysis would erroneously infer that only the subject's predilection for reporting pain had been affected. The same problem of interpretation exists for Clark's earlier studies as well. Suggestion (10) or placebo (7) or advancing age (12) could have affected a central biasing mechanism, inhibiting the response to all noxious stimuli, but the signal detection analysis would show that only the response bias had been affected.

Additional Studies

These difficulties plague other signal detection investigations of pain as well. Bloedel et al. (2) analyzed the modulating effects of peripheral nerve and spinal cord stimulation, reporting that dorsal column stimulation yielded both a lower d′ and a shift to a more conservative criterion, whereas transcutaneous stimulation primarily shifted the bias parameter. Again, it may be more appropriate to conclude that stimulation of the cord and nerve influenced the sensory impact of the thermal pulses and not, as they suggested, the criterion adopted for issuing pain reports.

Craig and Coren (14) exposed naive subjects to social models who simulated different levels of discomfort and pain susceptibility. Subjects used a 10-point scale to describe the effects of a random series of five levels of electric shock after being paired with a tolerant, intolerant, or control model. An index, related to d′, for the discriminability of adjacent pairs of stimuli was significantly increased for the subjects exposed to the intolerant model but was not altered for those in the tolerant modeling ground, even though both groups showed ratings that differed significantly from those of controls. They concluded that exposure to an intolerant model produced a change in the sensory experience of pain, whereas exposure to a tolerant one changed the response bias.

In addition to the difficulty of unambiguously interpreting a constant d′ in pain modulation experiments, Craig and Coren's study (14) presents another problem of data interpretation. It appears that they calculated d′ based on individual data points rather than on a receiver operating characteristic (ROC) curve fit to all the data. The sensitivity change occurred primarily at intermediate ratings. This point is too complex to receive much elaboration here, but it is important. For any pair of hit and false alarm rates, one can calculate d′ or look it up in published tables (e.g., 26). Such calculations and tables are based on a vital assumption: that the variance of the two underlying distributions is equal. Clark (7) showed that for thermal stimuli this may well be the case, but Rollman (22,23) found that it is not true for electrical or mechanical stimulation of the skin. When such signals are used the data must be plotted on ROC curves, and then special conventions (16) must be followed to determine d′ for distributions of unequal variance. Failure to do so can cause an experimenter to report a d′ shift when actually a criterion change occurred, or vice versa. As Theodor (28) has stated, "d′ is independent of response bias only when [the ratio of the variances of the underlying distributions] has been taken into account; . . . it is necessary to know [the ratio] in order to properly calculate d′."

Some of these same comments apply to the research of Chapman and his colleagues. Feather et al. (15), in a study similar to Clark's (7), found that d′ for discriminating two levels of radiant heat was uninfluenced by an oral placebo, although criterion was shifted in a more conservative direction. Again, a shift in sensory response to both the "warm" and the "painful" stimuli could have yielded this outcome. The same is true in a study by Chapman and Feather (4) on the analgesic effects of an oral tranquilizer, diazepam (10 mg). The drug significantly increased pain tolerance times to tourniquet-induced ischemia in comparison to both a matching placebo and

600 mg of aspirin. However, d′ for the discrimination of adjacent levels of radiant heat was not affected by the drug, nor was the criterion for pain response.

In a study cited earlier, Clark and Yang (13) concluded that, "the sole effect of acupuncture was to cause the subjects to raise their pain criterion in response to the expectation that acupuncture works." Chapman et al. (5) came to different conclusions in their study of the effects of bilateral needling of the hand on the discrimination of electrical stimuli at the dental pulp. The pain-attenuating effects of acupuncture were weak, but significant declines in d′ were observed. Since the criterion for describing their most intense stimulus as painful also increased, Chapman et al. (5) concluded that acupuncture's effects involve both sensory loss and alterations in response bias. They suggested a number of procedural differences that might result in their finding a sensitivity shift when Clark and Yang (13) failed to find one. However, they also determined d′ in a different manner from that used by Clark and Yang. Chapman et al. (5) used a technique presented by Richards and Thornton (21), which calculates a least squares fit to a straight-line ROC function on normal deviate coordinates. Even if Chapman et al. assumed unequal variances in their calculations, their values of d′ are subject to some error. A least squares procedure inaccurately assumes that the abscissa, the false alarm rate, is not free to vary. In addition, Richards and Thornton's (21) estimate of d′ for unequal variances calculates d′ at the point where the ROC curve intersects the 50% false alarm rate rather than where it intersects the negative diagonal (16).

OBJECTIONS TO THE RATIONALE OF SIGNAL DETECTION STUDIES OF PAIN MODULATION

Objections have been raised above to hypotheses about what should happen to the discriminability index if a modulation procedure were truly analgesic and to the methods sometimes used in making d′ calculations. These criticisms apply if it is valid to use the signal detection model to study pain. Is such validity apparent?

All of the cited studies do not, in fact, determine a bias-free estimate of how painful a stimulus seems to be and if the modulation procedure reduces that painfulness. Although this is what they purport to investigate, they ask merely if the strongest stimulus changes in discriminability from the next strongest. Is it critical that an analgesic, which drastically reduces the painfulness of a series of strong impulses, also reduces the ability to discriminate adjacent levels? If not, the premise of all of the studies is untenable.

One might suggest that d′ should be calculated between the potentially noxious levels and the zero intensity of the blank or noise stimulus. In fact, Chapman et al. (6) showed that 33% nitrous oxide reduced d′ for three levels of radiant heat when calculated against the zero intensity responses, but the gas "did *not* change sensitivity to the differences between adjacent nonzero stimulus pairs." Thus they concluded, "33 percent nitrous oxide did not interfere with the ability of the subjects to discriminate among nonzero stimuli, even though the gas did reduce absolute sensory sensitivity to each of the nonzero stimuli." This point has been overlooked in each of the later pain papers which refers to this experiment. There are good reasons for not examing d′ between a strong stimulus and a blank. But no good rationale has been advanced for studying d′ between adjacent nonzero intensities, as all other investigators have done.

It has been argued here that an analgesic need not reduce the discriminability of adjacent stimuli. It can also be argued that discriminability can be reduced in the absence of analgesic effects. Some intervention could render it more difficult to discriminate between two shocks or heat levels, thereby reducing the d′ between them, yet have little effect on the painfulness of the stimuli.

A procedure which claims that motivational changes have occurred when, in fact, they were sensory, which claims that sensory changes have occurred when, in fact, they were motivational, and which examines the discriminability of adjacent stimuli without measuring their noxiousness has no utility in the study of pain. Chapman's (3) statement that "the threshold has had its day in court, and it has proven to be largely unreliable" seems to be accurate. Signal detection theory, on the other hand, has permitted great conceptual and substantive advances in the study of human decision processes. But it has not yet proven that it can offer a bias-free estimate of the sensory and motivational changes produced by pain modulation. Until it does, pain researchers might be well advised to heed the caveat of Rosner and Goff (24), "Beware of psychophysicists bearing gifts."

ACKNOWLEDGMENTS

These investigations were supported by Grant AO-392 from the National Research Council of Canada and by a Leave Fellowship from the Canada Council. This chapter was prepared while the author was on leave at the Psychological Laboratories, University of Stockholm.

REFERENCES

1. Beecher, H. K. (1959): *Measurement of Subjective Responses: Quantitative Effects of drugs.* Oxford University Press, New York.
2. Bloedel, J. R., Erickson, D., and McCreery, D. B. (1974): *Abstracts of the Annual Meeting of the Neuroelectric Society,* 7:54.
3. Chapman, C. R. (1974): In: *Advances in Neurology, Vol. 4: International Symposium on Pain,* edited by J. J. Bonica, pp. 115–122. Raven Press, New York.
4. Chapman, C. R., and Feather, B. W. (1973): *Psychosom. Med.,* 35:330–340.
5. Chapman, C. R., Gehrig, J. D., and Wilson, M. E. (1975): *Anesthesiology,* 42:532–537.

6. Chapman, C. R., Murphy, J. M., and Butler, S. H. (1973): *Science,* 179:1246–1248.
7. Clark, W. C. (1969): *J. Abnorm. Psychol.,* 74:363–371.
8. Clark, W. C. (1974): *Anesthesiology,* 40:272–287.
9. Clark, W. C., and Dillon, D. J. (1973): *Percept. Psychophys.,* 13:491–493.
10. Clark, W. C., and Goodman, J. S. (1974): *J. Abnorm. Psychol.,* 83:364–372.
11. Clark, W. C., and Hunt, H. F. (1971): In: *Physiological Basis of Rehabilitation Medicine,* edited by J. A. Downey and R. C. Darling. W. B. Saunders Co., Philadelphia.
12. Clark, W. C., and Mehl, L. (1971): *J. Abnorm. Psychol.,* 78:202–212.
13. Clark, W. C., and Yang, J. C. (1974): *Science,* 184:1096–1098.
14. Craig, K. D., and Coren, S. (1975): *J. Psychosom. Res.,* 19:105–112.
15. Feather, B. W., Chapman, C. R., and Fisher, S. B. (1972): *Psychosom. Med.,* 34:290–294.
16. Green, D. M., and Swets, J. A. (1966): *Signal Detection Theory and Psychophysics.* John Wiley & Sons, New York.
17. Lusted, L. B. (1971): *Science,* 171:1217–1219.
18. Melzack, R., and Casey, K. L. (1968): In: *The Skin Senses,* edited by D. Kenshalo. Charles C Thomas, Publisher, Springfield, Illinois.
19. Pastore, R. E., and Scheirer, C. J. (1974): *Psychol. Bull.,* 81:945–958.
20. Procacci, P., Della Corte, M., Zoppi, M., Romano, S., Maresca, M., and Voegelin, M. R. (1974): In: *Recent Advances in Pain,* edited by J. J. Bonica, P. Procacci, and C. Pagni. Charles C Thomas, Publisher, Springfield, Illinois.
21. Richards, B. L., and Thornton, C. L. (1970): *Educ. Psychol. Meas.,* 30:855–859.
22. Rollman, G. B. (1969): *Percept. Psychophys.,* 5:377–380.
23. Rollman, G. B. (1974): *Percept. Psychophys.,* 16:291–294.
24. Rosner, B. S., and Goff, W. R. (1967): In: *Contributions to Sensory Physiology, Vol. 2.* Academic Press, New York.
25. Sternbach, R. A. (1974): *Pain Patients—Traits and Treatment.* Academic Press, New York.
26. Swets, J. A. (Ed.) (1964): *Signal Detection and Recognition by Human Observers.* John Wiley & Sons, New York.
27. Swets, J. A. (1973): *Science,* 182:993–1000.
28. Theodor, L. H. (1972): *Psychol. Bull.,* 78:260–261.