

METHODS & DESIGNS

A note on psychophysical measures¹

MICHAEL H. SIEGEL,² ALBION COLLEGE, Albion, Michigan 49224

The nature of liminal measurement is discussed, and the standard deviation is proposed as a suitable alternative measure to the limen.

In color-discrimination research, using the constant-stimulus method, I have encountered some difficulties posed by the use of the difference limen. This paper provides a review of these difficulties and a suggestion for an alternative solution. For purposes of discussion, it will be assumed that the distribution of response frequencies may be fitted by the normal ogive, and that the method of constant stimulus differences is used.

A requirement for psychophysical research is that the E have knowledge and control of the nature and the magnitude of the physical stimulus presented to the O. To calculate the absolute threshold, the E must know the physical stimulus difference between the point where no stimulus is presented and the place where the O reports detection of the stimulus 50% of the time. Similarly, to calculate the difference threshold, it is necessary to know the point where the two stimuli to be compared are physically equal and the point where they are judged to be different 50% of the time. There is, of course, little difficulty in establishing when no stimulus is being presented; however, it is most difficult to present two physically identical stimuli to the O. The difference limen will be spurious to the extent of the actual difference between true physical identity of the two stimuli and the assumed physical identity. Because the best available techniques are used to estimate physical identity of the stimuli, the extent of departure from identity is indeterminate. Because human discrimination ability is frequently extremely acute, even a small error becomes critical.

There are at least two alternatives available for the researcher interested in measuring differential sensitivity free from the indeterminate measuring error just discussed. One is the use of a mean rather than two separate thresholds from the point of assumed stimulus equality. The only additional assumption imposed by averaging is that the point of physical identity of the two stimuli under comparison must fall somewhere between the two 50% points. If this assumption is made, then the difference between the two 50% points must equal the sum of the thresholds in both directions from the standard. The mean threshold would then equal half the physical distance between the two 50% points.

The averaging procedure is not free from fault since all directional differences become obscured. If the E wishes to discover the difference limen as a function of its position to the standard, such a procedure would be inappropriate.

A second alternative is the use of the standard deviation of the distribution of response frequencies. This measure is inversely related to the precision and slope of the psychophysical curve (Guilford, 1954). The standard deviation may be obtained to the same accuracy and with the same ease as the limen. In practice, the standard deviation may be defined operationally as the physical distance between the points whose response frequencies fall at 50% and at about 84% when the response frequencies are placed on a normalized scale. Since the standard deviation reflects

the rate of change in the O's judgments, it is not influenced by an indeterminate point of physical equality of the two stimuli. The intuitive meaning of the standard deviation, the variability of the O's judgments, while different from that of the limen, seems equally appealing as an index of sensitivity.

If an "equals" category is used, the resulting limen will be in part a function of the number of times the O employs this response. But an O's willingness to use an "equals" category is at least partially dependent upon factors other than the experimental variables being studied, e.g., motivation. Thus, a highly confident O, who seldom used this category, would obtain a smaller limen than a second O, with comparable sensory ability, who used it frequently. In other cases, where only two categories are used and judgments are required on both sides of the standard stimulus, it is not possible to obtain the usual limen since both curves must cross at 50%, which here indicates a point of "subjective equality." Finally, if only two categories are used and the O is asked to judge whether or not a difference is detected in one direction, the curves will not necessarily cross at the 50% point, but, as in the three-category situation, the separations of the two curves at the 50% point can also be influenced by factors extraneous to the experiment.

The standard deviation, on the other hand, appears quite stable and remains more unaffected by these confounding factors (Kellogg, 1930; Woodworth, 1938). The measure is clearly less influenced by motivational factors and may be used with any of the three different category situations described above. While motivational influences may shift the entire position of the response curves with respect to the abscissa and hence alter the limen, they do not substantially affect the rate of change of these judgments. This finding has led Woodworth to write that "... the difference limen obtained by the method of Constant Stimuli is not a valid measure of differential sensitivity, though it may be of value as a measure of attitude. The SD is a true measure of differential sensitivity [Woodworth, 1938, p. 425]."

Recent work in color discrimination has also indicated that the standard deviation is a more satisfactory measure than the difference limen. Using highly trained Os, I found that when the experimental task was simple, the absolute values of the difference limen and the standard deviation were highly correlated, but that the latter was more reliable (Siegel & Dimmick, 1962). As the observational task increased in complexity and difficulty, the two measures became progressively dissimilar. Using split-half reliabilities, we have discovered far more stability and consistency among the standard deviations than among the limens (Siegel, 1964).

The comparative variability of the limens and the standard deviations can be illustrated by the following example. In our laboratory, typically, a total of 200 observations are required to calculate a limen and a standard deviation. It is possible to divide these observations in half, and thus obtain independently two limens and two standard deviations. Let L_1 be the larger of the two limens and L_2 , the smaller, σ_1 the larger of the two standard deviations and σ_2 , the smaller. The quantity, D, expresses the relative variability and is defined by the following equation:

$$D = L_1/L_2 - \sigma_1/\sigma_2$$

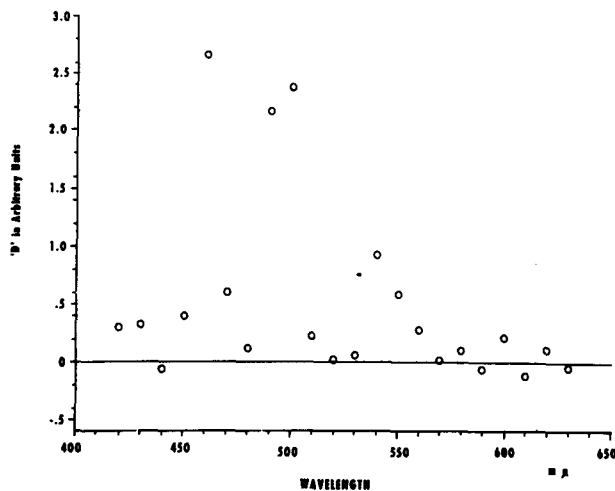


Fig. 1. "D" as a function of wavelength. The values in this figure are based upon several hundred thousand observations in a series of different experiments. See text for description.

A positive D would reflect more relative disparity between the two limens than between the standard deviations, and a negative D would indicate the converse. To test this, data from several of our early studies have been reexamined, and a D has been calculated at 22 separate wavelength locations. In 18 of the 22 points plotted, a positive D was obtained. The probability of obtaining an 18-to-4 split is .002 when the Sign Test is used. This effect is more pronounced in the blue and violet spectral region, where observing is more difficult than elsewhere.

There are two major difficulties inherent in the use of the standard deviation as the measure of sensitivity. The first, noted above, is the problem of comparing results reported as standard deviations with those reported as difference limens. These two measures are not directly comparable, so no real solution can be

found for this problem. The second difficulty concerns the notion of the just-noticeable difference (jnd) and its relation to the response measure. The limen has definite intuitive appeal since it may represent 1 jnd for the O. The standard deviation cannot be described in these terms. It should be pointed out, however, that only in the method of adjustments, where the O begins with supposed equality and adjusts the comparison field until it is just noticeably different from the standard, is the jnd obtained directly. In the constant-stimulus method, the limens are calculated after the data are obtained, and the O has no direct information about the size of the limen he is generating.

To conclude, the standard deviation is recommended as a measure of discrimination in place of the difference limen when the constant-stimulus method is used because of the following: (1) the difficulty of establishing a point of physical identity for the standard and variable stimuli, (2) the possibility of a spurious measurement due to incursion of an attitudinal effect, and (3) the empirical evidence demonstrating greater reliability for the standard deviation than for the limen. Where such replacement is impossible, both the standard deviation and the difference limen should be reported.

REFERENCES

- GUILFORD, J. P. *Psychometric methods*. (2nd ed.) New York: McGraw-Hill, 1954.
- KELLOGG, W. N. An experimental evaluation of equality judgments in psychophysics. *Archives of Psychology*, 1930, 112, 1-79.
- SIEGEL, M. H. Discrimination of color: IV. Sensitivity as a function of spectral wavelength, 410 through 500 mmu. *Journal of the Optical Society of America*, 1964, 54, 821-823.
- SIEGEL, M. H., & DIMMICK, F. L. Discrimination of color: II. Sensitivity as a function of spectral wavelength, 510 to 630 mmu. *Journal of the Optical Society of America*, 1962, 52, 1071-1074.
- WOODWORTH, R. S. *Experimental psychology*. New York: Holt, 1938.

NOTES

1. This paper was supported in part by Public Health Service Grant MH-14281.
2. Address: Department of Psychology, Albion College, Albion, Michigan 49224.