The method of constant stimuli is efficient

WILLIAM A. SIMPSON

York University, North York, Ontario, Canada

Monte Carlo simulations show that data collected using the method of constant stimuli yield threshold estimates with the same variability as and less bias than estimates based on data collected using an adaptive trial placement rule. In both cases, the simulation data were analyzed using maximum likelihood; for 100 trials or less, the method of constant stimuli is the better trial placement rule.

An experimenter wishing to measure thresholds has a number of methods from which to choose. A distinction can be made between those methods in which the levels of the stimulus to be presented to the subject are selected before testing begins—nonadaptive methods—and those in which the stimulus levels are chosen while testing is in progress—adaptive methods.

The venerable method of constant stimuli is a nonadaptive method. In it, several (typically 5) stimulus levels in the vicinity of the estimated threshold are chosen, usually on the basis of pilot testing. Each stimulus level is presented many times (usually at least 20) in random order, and the proportion of responses to number of presentations (in two-alternative forced-choice [2AFC], the case considered in this paper, the proportion of correct responses) is obtained for each level. The threshold is obtained by fitting an ogive of some sort (cumulative normal, logistic) to the proportions.

Adaptive techniques include the classical method of limits, UDTR staircases (Levitt, 1970), PEST (Taylor & Creelman, 1967), and, lately, methods using maximum likelihood (Lieberman & Pentland, 1982) and Bayesian treatment of the data (Emerson, 1986). The basic idea behind these methods is that each trial should be placed at the current threshold estimate. It seems obvious that adaptive methods make better use of trials than does the method of constant stimuli. As Emerson (1984) puts it, "presentation of many trials far above or far below threshold is a waste of time, because the responses to such stimuli have little bearing on the question of the location of the threshold" (p. 199). Emerson goes on to state the widely held belief that adaptive methods "can be more efficient by a large factor than . . . constant stimuli." We shall see that in fact the method of constant stimuli is as efficient as (as measured by the standard deviation of the threshold estimate for a given number of trials) and less biased than the theoretically much more efficient adaptive maximum likelihood method.

In order to compare the constant stimuli and the adaptive maximum likelihood methods, it is necessary to analyze the data obtained with the two methods of stimulus placement. In the following Monte Carlo simulations, the data from both methods were analyzed using the same maximum likelihood technique (Lieberman & Pentland, 1982). Thus any differences are due to the stimulus placement rule used.

SIMULATIONS

The conditions of the simulations were as follows. For both constant stimuli and adaptive trial placement, the maximum likelihood estimate of the threshold's location was obtained by using a slight modification of Lieberman & Pentland's (1982) technique. The difference between the two methods was that in the method of constant stimuli the stimulus levels presented were preset but in the adaptive method each trial was presented at the current threshold estimate. The maximum likelihood data analysis technique calculated the likelihood that the threshold was at each of 21 stimulus levels; this array of likelihoods was updated after each trial. The stimulus level with the highest likelihood was the threshold estimate. An explicit equation for the psychometric function is required in order to calculate likelihoods; here the function was the 2AFC scaled logistic:

$$y = .5 + .5/\{1 + \exp[a*(b-x)]\},$$

where a , the slope, was 1, and b , the threshold, was 0 logits ($p = .75$). The subject's responses were modelled as conforming to this function and having Bernoulli variance.

Before each block of trials, an experimenter would have to supply an estimated upper and lower bound for the threshold region (Stimulus Levels 20 and 0). This was simulated by randomly choosing an upper bound between 0 and R logits and a lower bound between 0 and $-R$ logits (R was 2.5, 5, or 10; thus the range in which the threshold was constrained to lie was 5, 10, or 20 logits). The stimulus levels thus had an average range of R , with a minimum range approaching 0 and a maximum range of $2R$. On average, the trials were placed symmetrically about

Correspondence may be addressed to William A. Simpson, Department of Psychology, York University, 4700 Keele Street, North York, ON, Canada M3J 1P3.

the threshold (0 logits); at their most asymmetrical, the trials could be at threshold and above, or at threshold and below. Thus the experimenter's uncertainty about the width of the threshold region and the location of the threshold within this region were simulated.

In addition to a range of stimulus levels presumed to bracket the threshold, the Lieberman & Pentland method requires a slope estimate. The data analytical method is a grid search for the threshold (instead of a more complicated and time-consuming optimization method that would be required to fit both slope and threshold). For the adaptive method, the true slope and under- and overestimates were used (values of 1, .2, and 5, respectively). In the constant stimuli method the true slope was used, since in this method both parameters can be fit to the data after the block of trials is complete—no slope estimate is required for trial placement.

The only difference between the method of constant stimuli and the adaptive method involves trial placement. For the adaptive method, each trial was placed at the current threshold estimate (for the first trial this was the midpoint of the stimulus range). For the method of constant stimuli, trials were placed at 5 stimulus levels: at the upper and lower bounds, and evenly in between.

The random number generator (Amiga Basic) was reseeded prior to each simulation run of 500 blocks of trials. Each simulation data point was obtained in a separate run.

RESULTS

The simulation results for the bias and standard deviation of the threshold estimate are plotted in Figures 1 and 2 respectively. For improved resolution, the scale on the ordinate in both graphs increases with the stimulus range. Inspection of the bias results shows that bias increases with the range of stimuli used and decreases with the number of trials. The most important result for our purposes is that the method of constant stimuli is less biased than the adaptive method for all cases, except when the slope is underestimated by a factor of 5 (where it is about as biased). Note that although underestimating the slope reduces bias, it also reduces the efficiency of the method. As has been shown before (Emerson, 1986), the bias of maximum likelihood threshold estimation is negative. However, the use of the method of constant stimuli's trial placement rule results in far less bias than the adaptive method with true slope.

The results for the standard deviation of the threshold estimate (Figure 2) show a decrease in standard deviation with the number of trials and an increase in standard deviation with the range of the stimuli used. The standard deviation of the threshold estimate is about the same for the method of constant stimuli and the adaptive method with true slope. Lower standard deviations for the adaptive method are possible (if the slope is overestimated), but at the expense of unacceptable increases in bias (cf. Figure 1). The differences between the methods tend to decline with the number of trials. With large stimulus

Figure 1. Bias of the threshold estimate as a function of number of trials for data collected using the method of constant stimuli and using an adaptive trial placement rule. The adaptive method requires an estimate of the slope of the psychophysical function; the true value in these simulations was 1. Each data point is the mean of 500 simulation runs.

ranges, good slope estimates, and large numbers of trials, the adaptive method shows signs of becoming more efficient than the method of constant stimuli. In general, however, the method of constant stimuli is about as efficient as the adaptive method.

In summary, then, the method of constant stimuli is as efficient as an adaptive method using the same statistical treatment of the data, and it is less biased.

DISCUSSION

For threshold measurements consisting of a small number of trials (100 or less), the method of constant stimuli is as efficient as and less biased than the adaptive method of stimulus placement. It is likely that adaptive methods will prove to be more efficient for larger numbers of trials, but adaptive methods are normally prescribed for small n situations. Clearly, this prescription must change.

Figure 2. Standard deviation of the threshold estimate as a function of number of trials, for data collected using the method of constant stimuli and using an adaptive trial placement rule. The adaptive method requires an estimate of the slope of the psychophysical function; the true value in these simulations was 1. Each data point is the mean of 500 simulation runs.

In retrospect, it is not too difficult to explain the results of the simulations. It is already known that the statistical method of maximum likelihood is negatively biased for small n . In the adaptive method, the consistent placement of trials below threshold is not optimally efficient—in fact it turns out to be about as efficient as placing them uniformly above and below the estimated threshold. The method of constant stimuli's trial placement ensures uniform placement of the stimuli, resulting in less bias of the threshold estimate and comparable efficiency. As McKee, Klein, and Teller (1985) have stated, adaptive methods do not have any "magical power." They argue that the variability of estimates from adaptive methods "can never be less than those from the method of constant stimuli selected for the optimal deployment of trials" (p. 296). However, the present simulations of the method of constant stimuli did *not* optimally deploy trials (the

stimuli could have wide ranges and be placed asymmetrically about the threshold), and still the method of constant stimuli was found to give threshold estimates of variability comparable to those from the adaptive method.

The conclusions that these simulations suggest regarding the relative bias of threshold estimates from adaptive and nonadaptive methods are at odds with those of Taylor, Forbes, and Creelman (1983). These authors did some experiments comparing the methods and presented arguments that PEST was less biased than the method of constant stimuli. The fact remains, however, that no experiment can tell whether a psychophysical method is biased. Bias can only be determined in a simulation where the true threshold value is known.

If the method of constant stimuli is efficient, why does laboratory lore hold it to be inefficient? One explanation is simply that (1) the standard way of using the method of constant stimuli requires pretesting so that the levels used do not produce response probabilities of 0 or 1, and (2) normally, over 20 trials per stimulus level are used (with the usual 5 stimulus levels, this makes over 100 trials). Neither of these procedures is necessary. As in adaptive methods, merely the range that is thought to contain the threshold need be specified. The number of trials per stimulus level is traditionally large because the primitive methods of data analysis that previously existed fitted curves to the response probabilities, not to the 0,1 data points themselves. Any number of levels (greater than 1) or trials per level (even 1) can be used, and the data can be fitted to a logistic or cumulative normal using a least-squares or a maximum-likelihood criterion. The traditional version of the method of constant stimuli may have been seen as inefficient because of the large number of trials (instead of the true criterion of efficiency, standard deviation/number of trials). Adaptive methods may give a quick threshold estimate, which gives an impression of efficiency, but this estimate will have high variability. Another contributing factor is that many adaptive packages give what may be highly misleading variability estimates. For the UDTR staircase, one procedure uses the standard deviation of the last several reversals. This estimate is far lower than the true value, since the reversals are highly dependent upon each other. Variance estimates from standard nonlinear regression packages may also be grossly misleading (Gallant, 1975).

It may be that in experimental situations, the method of constant stimuli is indeed less efficient than adaptive methods are. Human subjects may give more reliable data when the stimuli are placed near the threshold. This seems unlikely, however, since subjects tend to complain about task difficulty if all trials are near threshold (Regan, 1982, p. 438; Watson & Pelli, 1983). Moreover, experiments by Hesse (1986), in which several psychophysical methods were compared, found that the method of constant stimuli gave threshold estimates with similar variability and with better consistency over time than did PEST. Hesse's results, besides corroborating the present simulations, also

show that the conclusions of the simulations can be generalized to apply to adaptive methods other than the maximum likelihood method.

The method of constant stimuli has other virtues besides the low bias and variability of thresholds resulting from its use. The extreme simplicity in implementing the trial placement rule is the most obvious advantage (which is offset, to some extent, by the necessity of using nonlinear regression on the resulting data). Another is the independence of the stimulus level presented on a given trial from that presented on previous trials. Finally, the method of constant stimuli requires no prior (fixed) assumptions about the shape of the psychometric function before proceeding. After the data have been collected, any of a number of functions can be fitted (logistic, cumulative normal, Weibull, arcsine, etc.). The maximum likelihood adaptive method requires both the type of function and its slope to be fixed in order to proceed. It is not clear what the consequences of fitting a poor model function will be. Assuming that the model function is appropriate, we have seen that errors in slope estimation lead to high biases and (depending on the nature of the error) high variability of the threshold estimate. The biasing is not so much of a problem in itself (since there is no way to know the "true" value of the threshold), but if different experimental conditions have corresponding psychometric functions with different true slopes, the threshold characteristic will be distorted (Laming, 1986, p. 38).

It is commonly asserted (e.g. Levitt, 1970) that the method of constant stimuli assumes a psychometric function that does not change during the experiment (note that the adaptive maximum likelihood method makes such an assumption). If one desires to track a changing psychometric function, this can, in fact, be easily done with data collected by using constant stimuli. One method is to fit the function successively, using the data points up to each trial (Data Points 1-2, 1-3, 1-4, ...). A better method would be to use a moving window of some trial width and to fit the psychometric function successively to the points within the window. In these ways a plot of the threshold versus trials can be obtained.

In conclusion, it is not true that the time-honored method of constant stimuli has outlived its usefulness. As

a trial placement rule, this method results in threshold estimates with low bias and variability. Data obtained using an adaptive placement rule result in higher bias and no smaller variability. For measurements of 100 trials or less, the method of constant stimuli is clearly superior to the adaptive maximum likelihood method. Since the maximum likelihood method is theoretically maximally efficient and empirically about as efficient as PEST and the UDTR staircase (Shelton, Picardi, & Green, 1982), we can conclude that these other adaptive procedures will give no better efficiency than constant stimuli.

REFERENCES

- EMERSON, P. L. (1986). Observations on maximum-likelihood and Bayesian methods of forced-choice sequential threshold estimation. *Perception & Psychophysics*, **39**, 151-153.
- EMERSON, P. L. (1984). Observations on a maximum likelihood method of sequential threshold estimation and a simplified approximation. *Perception & Psychophysics*, **36**, 199-203.
- GALLANT, A. R. (1975). Nonlinear regression. *American Statistician*, **29**, 73-81.
- HESSE, A. (1986). Comparison of several psychophysical procedures with respect to threshold estimates, reproducibility, and efficiency. *Acustica*, **59**, 263-273.
- LAMING, D. (1986). *Sensory analysis*. Toronto: Academic Press.
- LEVITT, H. L. (1970). Transformed up-down methods in psychophysics. *Journal of the Acoustical Society of America*, **49**, 467-477.
- LIEBERMAN, H. R., & PENTLAND, A. P. (1982). Microcomputer-based estimation of psychophysical thresholds: The best PEST. *Behavior Research Methods and Instrumentation*, **14**, 21-25.
- McKEE, S. P., KLEIN, S. A., & TELLER, D. Y. (1985). Statistical properties of forced-choice psychometric functions: Implications of probit analysis. *Perception & Psychophysics*, **37**, 286-298.
- REGAN, D. (1982). Visual information channeling in normal and disordered vision. *Psychological Review*, **89**, 407-444.
- SHELTON, B. R., PICARDI, M. C., & GREEN, D. M. (1982). Comparison of three adaptive psychophysical procedures. *Journal of the Acoustical Society of America*, **71**, 1527-1533.
- TAYLOR, M. M., & CREELMAN, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America*, **41**, 782-787.
- TAYLOR, M. M., FORBES, S. M., & CREELMAN, C. D. (1983). PEST reduces bias in forced choice psychophysics. *Journal of the Acoustical Society of America*, **74**, 1367-1374.
- WATSON, A. B., & PELLI, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, **33**, 113-120.

(Manuscript received September 21, 1987;
revision accepted for publication May 18, 1988.)