Perception & Psychophysics
1989, 45 (6), 572-576

The step method: A new adaptive psychophysical procedure

WILLIAM A. SIMPSON
York University, North York, Ontario, Canada

A new adaptive psychophysical method, the step method, is introduced. Simulations show the method to be less biased and more efficient than constant stimuli or Pentland's adaptive method for fewer than 40 trials. An experiment using discrimination of dot number, however, failed to find any differences among the three methods in either bias or efficiency.

In the last 20 years there has been considerable interest in improving the efficiency of psychophysical measurement (Emerson, 1986; Levitt, 1970; Lieberman & Pentland, 1982; Taylor & Creelman, 1967; Watson & Pelli, 1983). The main idea behind the new methods is that trials should be placed at the experimenter's current estimate of the threshold. Such methods are termed *adaptive* methods because the range of stimuli presented adapts to the subject's responses. In contrast, the classical method of constant stimuli presents the same set of stimuli no matter how the subject responds.

This paper has two aims. First, I will present a new adaptive method which has advantages over some others currently in use. Second, I will give both simulation and experimental data on the performance of this method as compared with constant stimuli (a nonadaptive method) and Pentland's maximum-likelihood adaptive method (Lieberman & Pentland, 1982; Pentland, 1980).

THE STEP METHOD

The *step method* is similar to Pentland's maximum-likelihood estimation (MLE) technique. However, there are two main differences. First, the assumed psychometric function is a step function instead of a logistic (Equation 1). Second, the step method uses least squares instead of maximum likelihood as the criterion for fitting a function to the response data. The rationale for using the step function is that Pentland's method fits only one of the two parameters of the logistic function; his method assumes a fixed value for the slope. If we are fitting only one parameter, we might as well use a psychometric function with just one parameter—a step has a value of 0 below threshold and 1 above, with the threshold being the only parameter. The advantage here is that there can be no errors in the assumed value of the slope. Such errors have been shown (Emerson, 1984; Simpson, 1988) to increase both the bias and the variability of the resulting threshold estimates. Because the step is a discontinuous function, maximum likelihood cannot be used; thus least squares, which is asymptotically equivalent to maximum likelihood, is used instead.

After each trial, which will result in a 0 or a 1 y value (in forced-choice methods, incorrect or correct responses, respectively), a step is fit by least squares to the data points gathered thus far. The location of the step is the threshold; this x value is presented on the next trial.

A BASIC program that implements the step method for two-alternative forced-choice (2AFC, the case considered in this paper) is given in Figure 1. The number of trials and the upper and lower limits of the estimated threshold region are input. In the program, 21 stimulus levels are used. First the x values for the 21 levels are calculated and stored in the X array; simultaneously the sum of the squared error (SSE) array is set to 0. The current threshold estimate is $X(M\%)$. Initially, $M\%$ is set to the midpoint of the range. The program presents $X(M\%)$, collects the response, then calculates $M\%$ for the next trial.

The main work is in the calculation of $M\%$. The program uses an update method similar to that of Lieberman and Pentland (1982). Each stimulus level is successively taken as a possible threshold location. If the just-tested x level is greater than this threshold location and the response is incorrect, then the SSE array at that level is incremented by $W1$. Conversely, if the just-tested x level is less than this threshold location and the response is correct, then the SSE array at that level is incremented by $W2$. Because the responses are 0 or 1, the squared deviation will be 1 in either case. The squared deviations are multiplied by weights. The weights will be $W1$ = target percentage and $W2 = 1 -$ target percentage. As long as the weights are in the proper proportion, any values can be used (although integers will be faster). Here the target percentage is 75%; therefore $W1 = .75$ and $W2 = .25$ (or 3 and 1, respectively). The level with the lowest SSE is indicated by $P1$. If the SSE function has a flat minimum, the other endpoint of the valley is indicated by $P2$. The value of $M\%$ is calculated as the point 90% of the way between $P1$ and $P2$. Simulations have shown that

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

Copyright 1989 Psychonomic Society, Inc.

```
' step.method
' psychophysical procedure based on least squares
' fit of step function to 1/0 data
' currently set up for 2AFC--change weights w1 and w2 for
' other target %s: 2 weights add up to 1.0 (or use integer ratio)
' e.g. .5,.5 = 1,1  .75,.25 = 3,1  .8,.2 = 4,1

DIM x(20), sse(20)
set.windows:
  WINDOW 1, "step.method", (0,0)-(617,185),31

nl% = 20 : w1=3 : w2=1
CLS
INPUT "# of trials";n%
INPUT "lower limit";ll
INPUT "upper limit";ul
j=(ul-ll)/nl%
    FOR i = 0 TO nl%
      x(i) = ll : ll = ll+j
      sse(i) = 0
    NEXT i

m% = nl%/2

      FOR tr = 1 TO n%
         PRINT  x(m%)
         INPUT "response (1/0)";ys%
         GOSUB calc
      NEXT tr

END

calc:
mn = 10000
FOR i = 0 TO nl%
  IF x(m%) >= x(i) AND ys% = 0 THEN sse(i) = sse(i) + w1
  IF x(m%) <= x(i) AND ys% = 1 THEN sse(i) = sse(i) + w2
  IF sse(i) < mn THEN mn = sse(i) : p1=i
  IF sse(i) = mn THEN p2 = i
NEXT i
 m = p1 + (p2-p1)*.9
 m% = INT(m+.5)
RETURN
```

Figure 1. An Amiga BASIC program that implements the step-adaptive method for two-alternative forced-choice (2AFC).

for 2AFC this value results in minimal bias. For a target percentage of 50%, use the value .5 (the value will always be between 0 and 1).

The algorithm is very fast (BASIC is more than adequate) and is not prey to rounding or out-of-range errors. Pentland's method can blow up due to either the exponential or the log functions. Emerson's (1986) Bayesian method is too slow to be implemented in BASIC. It is hard to evaluate Emerson's method, however, since he has not published the algorithm.

SIMULATIONS

The conditions of the simulations were as follows. The simulation data for constant stimuli and Pentland's MLE method were determined as previously described (Simpson, 1988). It was assumed that the psychometric function was the 2AFC scaled logistic:

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

where a (the slope) was 1 logit and b (the threshold) was 0 logits ($p = .75$). The subject's responses were modeled as conforming to this function and having Bernoulli variance. For constant stimuli and Pentland's method, a slope estimate is required to calculate likelihoods; the data here were obtained using the true slope (a value of 1).

In using any of the methods, an experimenter has to specify an upper and a lower bound for the threshold region. 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). The stimulus levels had an average range of R, with a minimum approaching 0 and a maximum range of $2R$ (the values of the range given in the figures). On average, the trials were placed symmetrically about the threshold; at their most asymmetrical they could all be on one side of threshold.

Results

The simulation results for the bias of the threshold estimates are given in Figure 2. For constant stimuli and Pentland's method, bias declines with the number of trials and increases with the range of stimuli. The step method is unbiased, showing only small, random fluctuations

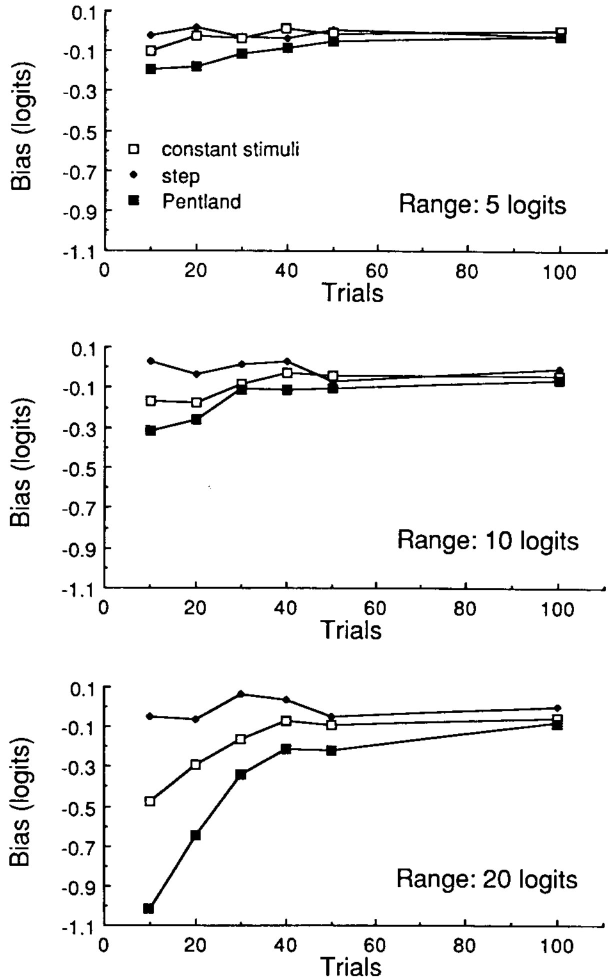

Figure 2. Bias of the threshold estimate (in logits) as a function of the number of trials and the stimulus range for constant stimuli, the step method, and Pentland's maximum-likelihood estimation (MLE) method. Each point is based on 500 simulation runs.

around the true threshold value; constant stimuli and Pentland's method show progressively more bias as the range increases. Pentland's method seems especially prone to negative bias with large stimulus ranges.

As can be seen in Figure 3, the standard deviation of the threshold estimate declines as the range decreases and as the number of trials increases. The variability of all three methods is quite similar. The step method has an advantage for runs of less than 40 trials. For runs of more than 40 trials, Pentland's method is most efficient, although it is not much better than constant stimuli.

Discussion

Overall, the methods are quite similar in performance. In a previous study (Simpson, 1988), I found that the constant stimuli method was about as efficient as Pentland's method. This finding is confirmed here using a different adaptive method. The step method has an edge in efficiency for small numbers of trials, and it is unbiased. In the latter respect it is similar to Emerson's (1986) Bayesian method. However, both Pentland's and Emerson's methods require an estimate of the psychometric function's slope. In the case of Pentland's method, it is known that errors in slope estimates will adversely affect the bias and variability of the threshold estimates (Emerson, 1984; Madigan & Williams, 1987; Simpson, 1988). The step method avoids these problems.

The step method's superior efficiency for small numbers of trials is probably due to the fact that the step function will fit the data perfectly for small numbers of trials, but the logistic will not. As the number of trials increases, however, the fit of the step becomes increasingly worse than the logistic. It is not clear whether this would be a problem in experimental situations, since a typical use of adaptive methods is to take the mean and standard deviation of several blocks of 40 or so trials. It is not a good idea to use larger blocks; if the stimulus range is inappropriate, the threshold will simply hit a ceiling or a floor.

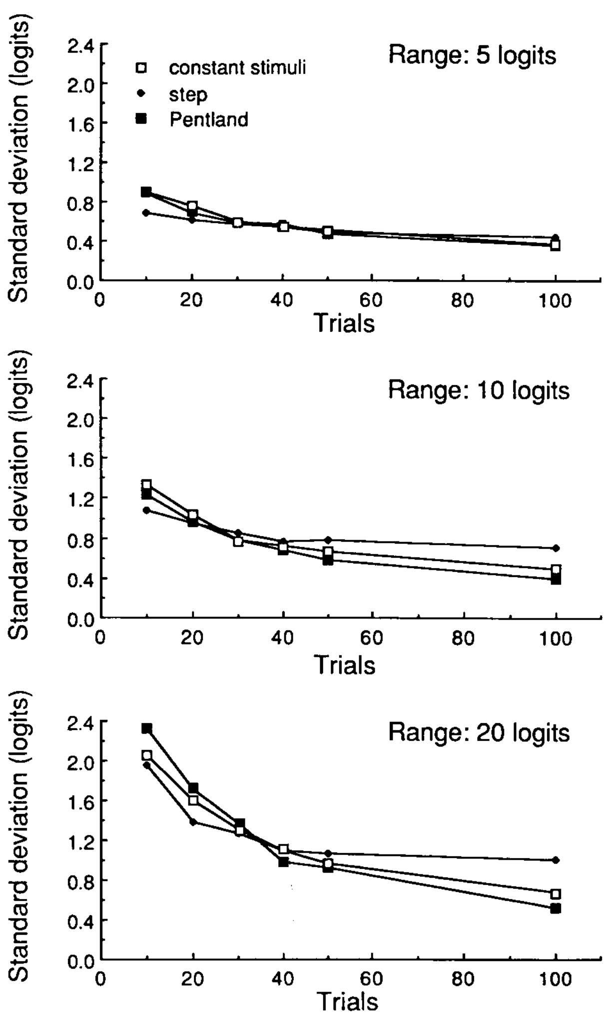

Figure 3. Standard deviation of the threshold estimate (in logits) as a function of the number of trials and the stimulus range for constant stimuli, the step method, and Pentland's MLE method. Each point is based on 500 simulation runs.

This same observation holds for constant stimuli; several blocks of about 40 trials should be used.

In order to test the performance of the different psychophysical methods in a real situation, the following experiment was performed.

EXPERIMENT

The step method, Pentland's MLE adaptive method, and constant stimuli were compared in an experiment on discrimination of the number of randomly placed dots. In other studies that compared different psychophysical methods (Hesse, 1986; Madigan & Williams, 1987; Shelton, Picardi, & Green, 1982), the threshold for only one experimental condition was obtained. Typically, however, we are interested in determining a threshold characteristic—the threshold as a function of some parameter. It is important to test psychophysical methods in their determination of a threshold characteristic because bias in a method can distort the shape of the characteristic (Laming, 1986). In this experiment, therefore, I determined the difference threshold for dot number at five points for each of the three methods.

In the present case of dot-number discrimination, it has been established by Burgess and Barlow (cited in Laming, 1986) that the threshold characteristic (here, the discrimination function) is a power function:

$$\text{difference threshold} = k * \text{dot number}^{e}. \quad (2)$$

The effects of bias will be reflected in the obtained values of k or e, or in departures of the data from a straight line on log-log coordinates. The efficiencies of the methods will be reflected in the standard deviations of their threshold estimates for a given number of trials.

Method

Subjects. The author and a naive observer, both with normal or corrected-to-normal vision, served as subjects.

Stimuli. On any given trial, the display consisted of two fields (left and right) of randomly placed dots on a CRT screen. Each field's area was 131 × 180 pixels. The subject viewed the display from a chinrest 57 cm from the screen; at this distance, the dots were .05° square and the fields were 7° (horizontal) × 9° (vertical). The subject fixated a .3°-square mark in the center of the screen. The inner edge of each dot field was 1° away from the fixation mark.

Five different numbers of dots (10, 20, 40, 80, and 160) were used as the standard. The comparison field in each trial contained more dots. The locations of the comparison and the standard were randomly varied between trials. The subject's task was to press a button corresponding to the field that contained the greater number of dots (the comparison).

Procedure. Each block consisted of 40 trials. On each trial, two dot fields were presented for 540 msec. The subject pressed a button corresponding to the field that contained the greater number of dots. There was a 1-sec pause between trials.

The three methods were given in random order. All values of the standard were given using one method before using the next. The level of the standard was determined randomly. In all, four measures were made for each level of the standard using each method. Each datum point is thus based on 160 trials.

Three psychophysical methods were used to determine the difference thresholds: constant stimuli, the step method, and Pentland's MLE method. Since Pentland's method requires an estimate of the slope of the psychometric function, two different slope estimates were used (1 and .2). Each method required an upper and a lower limit for the threshold estimate. The adaptive methods were as described above. The constant-stimuli method placed each trial randomly within the specified limits. The threshold estimate for the two adaptive methods was the value after the last trial. The threshold estimate for constant-stimuli method was calculated using a least-squares fit of a 2AFC logistic function to the data. The standard deviations of the thresholds were based on the variability of the values given by the four 40-trial blocks.

Results

The dot-number discrimination functions are plotted in Figure 4. For both subjects, the data are well fit by straight lines (least-squares fits) on the log-log plots. Hence, the data are described by the power function as given in Equation 2. The values of the parameters as determined by least squares were .2576 (A.H.) and .8444 (W.S.) for k, and .8878 (A.H.) and .5815 (W.S.) for e.

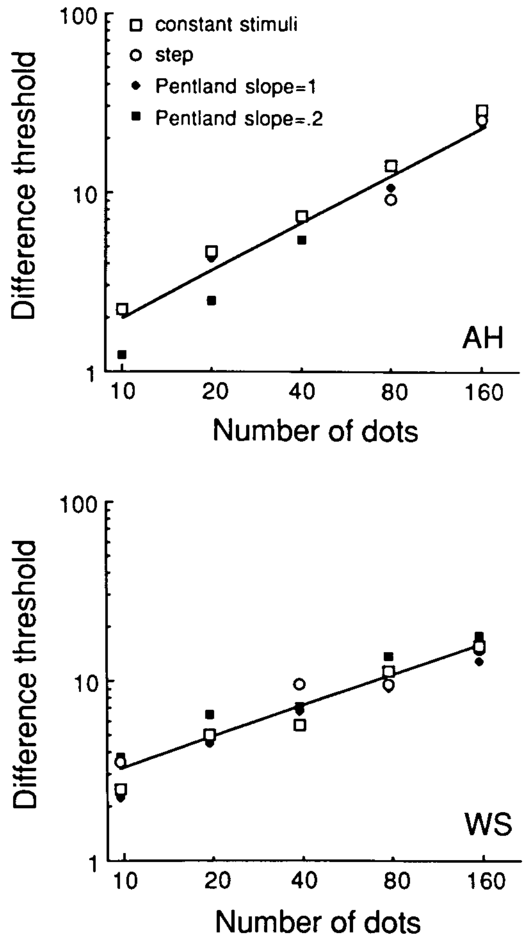

Figure 4. Difference thresholds for dot number obtained using constant stimuli, the step method, and Pentland's method. Each point is the mean of four 40-trial blocks. The straight lines are the least-squares fits of a power function to the data.

Table 1
Standard Deviation of Threshold Estimate as a Function of Psychophysical Method

	Psychophysical Method				
	Constant		Pentland		
Subject	Stimuli	Step	1.0	0.2	Mean
A.H.	3.66	2.70	3.15	3.48	3.25
W.S.	1.76	3.30	3.25	3.28	2.90

Note that the parameters are not identical for the 2 subjects (unlike Burgess & Barlow's [cited in Laming, 1986] result). As Burgess and Barlow found, the exponent of the discrimination function is just under 1 (an exponent of 1 would be predicted by Weber's law).

Only a single curve is plotted for each subject's data in Figure 4 because an ANOVA indicated no difference between the discrimination functions as measured by the different psychophysical methods. That is, method was neither a significant main effect [A.H., $F(3,36) = 1.482$, n.s.; W.S., $F(3,36) = 2.466$, n.s.] nor did it interact with number of dots [A.H., $F(12,36) = .344$, n.s.; W.S., $F(12,36) = .557$, n.s.]. Thus, there was no significant bias of the psychophysical methods relative to one another. An ANOVA comparing the standard deviations of the threshold estimates for the methods also found no significant differences [A.H., $F(3,12) = 1.585$, n.s.; W.S., $F(3,12) = 2.115$, n.s.]. As can be seen in Table 1, the standard deviations for the different methods are very similar.

Discussion

The three psychophysical methods all seem to give threshold estimates that are about the same in both bias and variability. No evidence was seen for any method being more biased or less efficient than any other. If this experiment can be regarded as typical, the differences in bias and variability between the methods as found in the simulations are hard to detect when using real subjects. Other experiments that have compared psychophysical methods also found the methods to be more-or-less equivalent (Hesse, 1986; Madigan & Williams, 1987; Shelton et al., 1982). In this experiment, the range of presented stimuli was probably small (as it likely is in all experiments). The simulation results for bias (Figure 2) and for standard deviation (Figure 3) with a range of 5 logits show only small differences between the methods at 40 trials.

The message from the experiment, then, as it was from the simulations, is that there is little, if any, difference in efficiency between constant stimuli and adaptive methods. Moreover, the differences in bias are probably too small to detect under normal (small-range) circumstances.

GENERAL DISCUSSION

The conclusion to be drawn both from the simulations and from the experiment is that there is little difference in efficiency between the classical method of constant stimuli and the newer, adaptive methods. According to the simulations, the step method is slightly more efficient than constant stimuli for small numbers of trials, and Pentland's method is more efficient than constant stimuli for large numbers of trials (the crossover point is about 40 trials). The experimental data, however, show no difference in efficiency for the methods. Although the step method was shown in the simulations to be less biased than constant stimuli or Pentland's method, the experiment revealed no differences in relative biasing. Both the simulations and the experiment show the step method to work quite well, despite its use of an unrealistic psychometric function.

It is heartening that one need not be overly concerned about which psychophysical method one chooses. The threshold estimate will take on the same value and have the same variability no matter which method is used. However, we should bear in mind that in an experiment we do not know the values of important variables, whereas these variables can be manipulated in simulations. In the simulations, the step method is shown to be less biased and to be slightly more efficient than constant stimuli or Pentland's method. Even in the simulations, though, these differences are small.

REFERENCES

Emerson, P. L. (1984). Observations on a maximum likelihood method of sequential threshold estimation and a simplified approximation. *Perception & Psychophysics*, **36**, 199-203.

Emerson, P. L. (1986). Observations on maximum-likelihood and Bayesian methods of forced-choice sequential threshold estimation. *Perception & Psychophysics*, **39**, 151-153.

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 & Instrumentation*, **14**, 21-25.

Madigan, R., & Williams, D. (1987). Maximum-likelihood procedures in two-alternative forced-choice: Evaluation and recommendations. *Perception & Psychophysics*, **42**, 240-249.

Pentland, A. (1980). Maximum likelihood estimation: The best PEST. *Perception & Psychophysics*, **28**, 377-379.

Shelton, B. R., Picardi, M. C., & Green, D. M. (1982). Comparison of three adaptive psychophysical methods. *Journal of the Acoustical Society of America*, **71**, 1527-1533.

Simpson, W. A. (1988). The method of constant stimuli is efficient. *Perception & Psychophysics*, **44**, 433-436.

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America*, **41**, 782-787.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, **33**, 113-120.

(Manuscript received July 11, 1988;
revision accepted for publication November 28, 1988.)