

A comparison of six methods to estimate thresholds from psychometric functions

TREFFORD LEE SIMPSON

University of Waterloo, Waterloo, Ontario, Canada

There are many ways in which to estimate thresholds from psychometric functions. However, almost nothing is known about the relationships between these estimates. In the present experiment, Monte Carlo techniques were used to compare psychometric thresholds obtained using six methods. Three psychometric functions were simulated using Naka-Rushton and Weibull functions and a probit/logit function combination. Thresholds were estimated using probit, logit, and normit analyses and least-squares regressions of untransformed or z -score and logit-transformed probabilities versus stimulus strength. Histograms were derived from 100 thresholds using each of the six methods for various sampling strategies of each psychometric function. Thresholds from probit, logit, and normit analyses were remarkably similar. Thresholds from z -score- and logit-transformed regressions were more variable, and linear regression produced biased threshold estimates under some circumstances. Considering the similarity of thresholds, the speed of computation, and the ease of implementation, logit and normit analyses provide effective alternatives to the current “gold standard”—probit analysis—for the estimation of psychometric thresholds.

Researchers and clinicians have a variety of ways in which to estimate thresholds from data comprising psychometric functions. A particular problem arises because the relationship between the dependent probability variable and the independent stimulus-strength variable is nonlinear and usually sigmoidal. Solutions range from estimating a threshold from a curve fitted by eye to nonlinear curve-fitting procedures (e.g., Bliss, 1935). Although the formal underpinnings of some of the latter procedures have been examined, little is known about the relationships between the estimates obtained using these methods.

The “gold standard” appears to be the maximum-likelihood procedure, probit analysis (Finney, 1971). Probability levels are converted to probits [z score [$p(\text{correct}) + 5$], weighted by the variance expected (assuming a binomial distribution), and a threshold is obtained from the straight line fitted to these weighted scores and the independent variable. This procedure is repeated until the fit has converged to its maximum likelihood. Two other rigorous (and quite similar) procedures are based on minimizing the chi-square: Normit analysis¹ (Berkson, 1955) assumes that probability is linearized with a z -score transformation, and logit analysis (Berkson, 1944, 1953) assumes that it is linearized with a logit² transformation. These two procedures have the major computational advantage of being noniterative.

There are, in addition, computationally simpler least-squares regression procedures that may be performed to analyze the psychometric-function data. Probability can be linearized with either a z -score transformation or a logit transformation (e.g., Gescheider, 1985), and the threshold calculated using the linear least-squares regression equation obtained from the transformed dependent and independent variables. This corresponds approximately to plotting probabilities in z -score space (on probability paper) or in logit space and fitting a straight line to this function. It is also possible to estimate a threshold using a linear least-squares regression equation from the untransformed probabilities and the independent variable.³

Each of the procedures described carries with it the baggage of theoretical assumptions about the nature of the data. For probit and normit analyses, a list of these assumptions would include (1) the distribution of the underlying instantaneous thresholds being Gaussian, (2) the responses being quantal and binary (e.g., seen/invisible), and (3) the variability of responses at a specific stimulus strength being binomial (Berkson, 1955; Finney, 1971). A similar list for logit analysis would include (2) and (3) above, as well as the distribution of the underlying responses being the derivative of a logistic function (Berkson, 1944, 1953). Performing a linear regression of the z score of $p(\text{correct})$ versus stimulus strength would require an assumption of a Gaussian underlying distribution, so that taking the z score of $p(\text{correct})$ would linearize the psychometric function. If a linear regression of $p(\text{correct})$ versus stimulus strength were being conducted because of some theoretical belief that the responses were in accord with neural quantum theory (Stevens, 1972), this would require an assumption of quantal underlying responses. There are addi-

This work was conducted while the author was a postdoctoral researcher at Toronto Hospital (Western Division), supported by MRC and NEI grants to D. Regan. I would like to thank David Regan, Alex Vincent, Stan Hamstra, and Merton Flom for their help with early versions of this manuscript. Correspondence should be addressed to T. L. Simpson, University of Waterloo, School of Optometry, Centre for Contact Lens Research, Waterloo, ON, Canada N2L 3G1 (e-mail: tsimpson@sciborg.uwaterloo.ca).

tional assumptions about the statistical procedures conducted. For example, using the least-squares method to derive a linear regression equation to estimate a threshold would necessitate assumptions of linear regression such as homoscedastic, Gaussian distributions of the dependent variable [$p(\text{correct})$ or z score of $p(\text{correct})$] for each fixed stimulus strength (e.g., Kleinbaum, Kupper, & Miller, 1988).

Any comparison between the thresholds obtained using the various procedures, therefore, is in a sense a question about what happens when these assumptions are, at best, ignored and, at worst, violated. The results of such a comparison are of value, however, because if computationally simpler procedures produce unbiased thresholds, these might be used instead of the more cumbersome ones. The simpler methods could, for example, be incorporated into adaptive psychophysics and easily used in widely available computer tools such as spreadsheets.

A Monte Carlo study was therefore carried out to examine the relationships between six threshold-estimate methods, including probit analysis. The effects of the shape of the psychometric function, the number of data points sampled, and the position of those samples were evaluated.

METHOD

Three different psychometric functions were simulated. The first two were used because they are expressions of plausible transducer functions at some level in the visual system. The first was a Weibull function (Quick, 1974; Weibull, 1951), probability of a correct response [$p(\text{correct})$] = $1 - 2^{-x^2}$. The second was a Naka-Rushton function (Naka & Rushton, 1966), $p(\text{correct}) = x^2/(1+x^2)$. For each of these, illustrated in Figure 1, when $x = 1.0$, $p(\text{correct}) = .5$ (the threshold as defined in this study).

The third psychometric function was a combination of probit and logit functions. Probit analysis is based on the underlying distribution of responses being Gaussian. Similarly, in logit analysis, it is assumed that the underlying responses form a derivative of a logistic distribution. Generating a psychometric function from either a cumulative Gaussian or logistic distribution would therefore bias the estimates in favor of the threshold-estimation procedure with the simulated psychometric function most like the expected underlying distribution. To counteract this problem, a simulated psychometric function was generated which was a combination of a cumulative Gaussian distribution and a logistic distribution. In transformed space, the equation was either $z \text{ score}_p$ or $\text{logit}_p = x$, with $z \text{ score}_p$ or logit_p determined randomly. The consequence of this was that a probability for each dependent variable level could usually take on two values. If, for example, $x = +1.0$, using the z -score equation, $p(\text{correct}) = .84$, whereas using the logit equation, $p(\text{correct}) = .73$. Nevertheless, when $x = 0.0$, $p(\text{correct}) = .5$ using either equation. Figure 2 depicts this simulated psychometric function.

For the Weibull and Naka-Rushton psychometric functions, variability in the responses was mimicked by jittering the x values randomly, whereas for the logit/probit function, the transformed p values were jittered. Thresholds were obtained from the former two psychometric functions using $\log(x)$ versus $p(\text{correct})$, and from the latter using x versus $p(\text{correct})$, and were defined as $p(\text{correct}) = .5$.

In addition to the effect of the shape of the psychometric function on calculated thresholds, two questions were examined. First, the effect of the number of levels of the dependent variable was

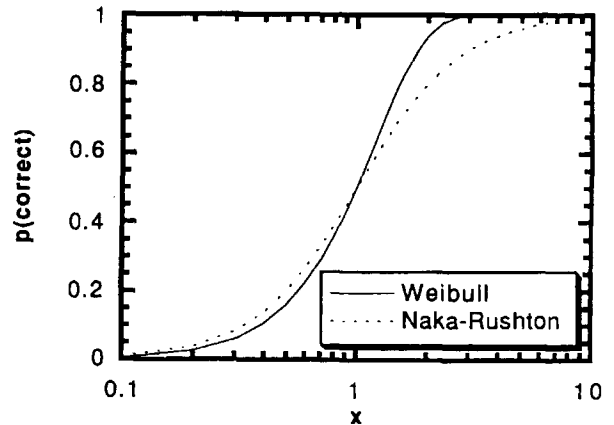


Figure 1. The relationship between intensity (x , the abscissa) and performance [$p(\text{correct})$, the ordinate] for two psychometric functions. The solid line is a Weibull function; the dotted line is a Naka-Rushton function.

simulated by calculating $p(\text{correct})$ at five, seven, and nine intensity levels on the psychometric function, symmetrically distributed around $p(\text{correct}) = .5$. Second, the effect of asymmetrically sampling the psychometric function was examined by truncating the sampling at the lower or upper end. Seven intensity levels were compared, with the points symmetrical about $p(\text{correct}) = .5$, or with four above and two below or four below and two above $p(\text{correct}) = .5$.

The probit analysis algorithm used was from Finney (1971). It consisted essentially of using a least-squares regression of z scores versus stimulus strength to obtain a provisional estimation of the relationship between probits and stimulus strength. The expected probits for each stimulus strength were then computed, and these were used to derive weights used to perform a weighted linear regression. This weighted regression was in turn used to re-estimate expected probits, and the process was repeated until successive differences in estimates differed by very little.

The normit and logit analysis algorithms used are very similar, both originating from Berkson (1955 and 1953, respectively). Each is a weighted linear regression, the weights being derived from $p(\text{correct})$, $1 - p(\text{correct})$, and either the z score [$p(\text{correct})$] or logit [$p(\text{correct})$] for normit and logit analysis, respectively. Details of the methods, as well as computed examples, were provided in Berkson's papers.

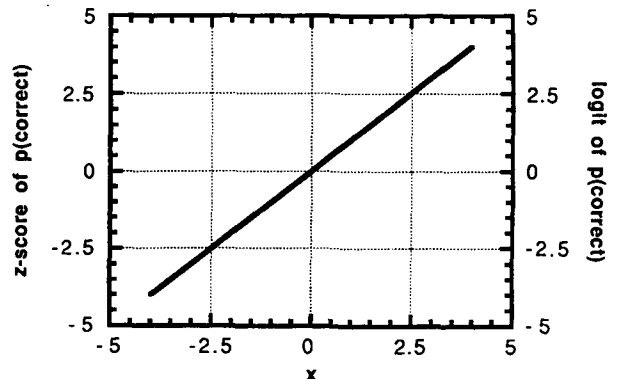


Figure 2. The relationship between intensity (x , the abscissa) and performance [z score or logit [$p(\text{correct})$], the ordinate] for the third psychometric function (the solid line).

The software was written in Modula-2 (Version 4.1, MetroWerks, Hudson Heights, Canada) on a Macintosh SE and Macintosh Powerbook. The probit, logit, and normit procedures were tested with examples from Finney (1971), Berkson (1944, 1953), and Lieberman (1983), as well as with ideal psychometric functions that were linear in z -score or logit space. The z -score transforms and inverse transforms were approximations obtained from Abromowitz and Stegun (1964).

Because the probabilities were transformed to z scores or logits, each of which extends from $-\infty$ [$p(\text{correct}) = .0$] to $+\infty$ [$p(\text{correct}) = 1.0$], explicit methods were used to deal with this problem. The probit analysis algorithm disregarded the data points at $\pm\infty$ for the first iteration (as outlined in Finney, 1971). Logit and normit analyses and the z -score and logit regression algorithms incorporated the method suggested by Berkson (1955). The probabilities at $p(\text{correct}) = 1.0$ became $(2n-1)/2n$, and at $p(\text{correct}) = .0$ became $1/2n$. It was assumed that n (the number of trials) = 100.

RESULTS

The data were summarized in two ways. First, the thresholds estimated by each procedure were binned and histograms were derived. These histograms were fitted with the Gaussian function

$$n = n_{\max} e^{-\frac{(x-\mu)^2}{2V}}, \quad (1)$$

where n_{\max} = number at peak of function, μ = mean of function, and V = variance of function.

Second, because the yardstick used for comparison was probit analysis, correlations between thresholds obtained using this and each of the other five procedures were derived, and these were illustrated in scatterplots.

Figures 3, 4, and 5 are histograms showing the distributions of log thresholds obtained from Weibull and Naka-Rushton psychometric functions and thresholds estimated from the logit/probit psychometric function, respectively. The smooth curves are the best-fitting Gaussian functions (Equation 1). Figure 6 contains scatterplots showing the relationships between thresholds obtained using probit analysis and the other methods from Weibull, Naka-Rushton, and logit/probit psychometric functions, respectively.

The histograms illustrate that the threshold estimates were generally similar, regardless of the methods used. The differences between the peaks of the Gaussians of the z -score or logit regression data and the probit data in Figure 3 do, however, suggest that there were biases in the threshold estimates from asymmetrical psychometric functions (0.07 and 0.09 log units for z -score and logit regression, respectively). This bias is also evident in the separation of the z -score and logit regression data from the equality line in the top graph in Figure 6. The data in Figures 3–6 were obtained using five intensity levels on the psychometric function, symmetrical around $p(\text{correct}) = .5$. Figure 7 contains scatterplots similar to Figure 6 obtained after sampling the psychometric functions using nine points. The vertical distance between the dashed and continuous lines on the upper two scat-

terplots illustrates the biased thresholds estimated using linear regression.

In Figure 8, the scatterplots show the relationships between thresholds estimated using probit analysis and the other procedures using data from five-, seven-, and nine-point Weibull psychometric functions. Sampling at the extreme ranges of the independent variable did little to increase the precision of the threshold computations, and in the case of linear regressions, introduced large biases. The histograms in Figure 9 illustrate that using nine-point psychometric functions *decreased* the precision of the estimates in comparison with using five and seven points. This reduction in precision as more extreme points were sampled was obtained regardless of the procedure. The upper half of Figure 9 illustrates the difference using estimates from z -score regressions, and the lower half illustrates the difference using probit analysis. Of note also are the similarities of the positions of the peaks of the z -score regression and probit analysis histograms. Figure 10 shows similar histograms obtained from logit/probit psychometric functions. The data in Figure 10 suggest that the reduction in precision revealed in Figure 9 was a characteristic of the (asymmetrical) Weibull psychometric functions that was not reflected in the threshold estimates obtained from the (symmetrical) logit/probit psychometric functions. The histograms of Figure 10 show that thresholds were similar regardless of the number of points defining the logit/probit psychometric function and the method used to estimate them.

The effect of the position of the samples was also dependent on the nature of the underlying psychometric function. When thresholds were estimated using logit/probit psychometric functions, there were no differences between methods, regardless of whether the samples were taken from around $p(\text{correct}) = .5$ symmetrically. A similar conclusion for probit, logit, and normit analysis applied when a Weibull psychometric function was used (Figure 11, top, illustrating the probit analysis results). Similar, although somewhat more dispersed, distributions of threshold estimates were obtained with logit or z -score regressions (Figure 11, middle, illustrating the logit-regression results). If, however, linear regression were used to estimate thresholds, the histogram (Figure 11, bottom) illustrates that biased thresholds were obtained, with the bias depending on whether the sample was primarily from above $p(\text{correct}) = .5$ or below it. The average bias (illustrated by the separation of the means of the histograms) was approximately 0.05 log units.

Because in many of the simulations thresholds were similar, other yardsticks such as size and speed might provide information with which to decide on the advantage of one procedure over another. There should be caution in comparing software that has not been optimized because differences in performance might reflect the programmer's style more than they reflect the benefits of a particular algorithm. This danger was minimized in this study because all of the procedures compared were linear regressions. The maximum-likelihood and mini-

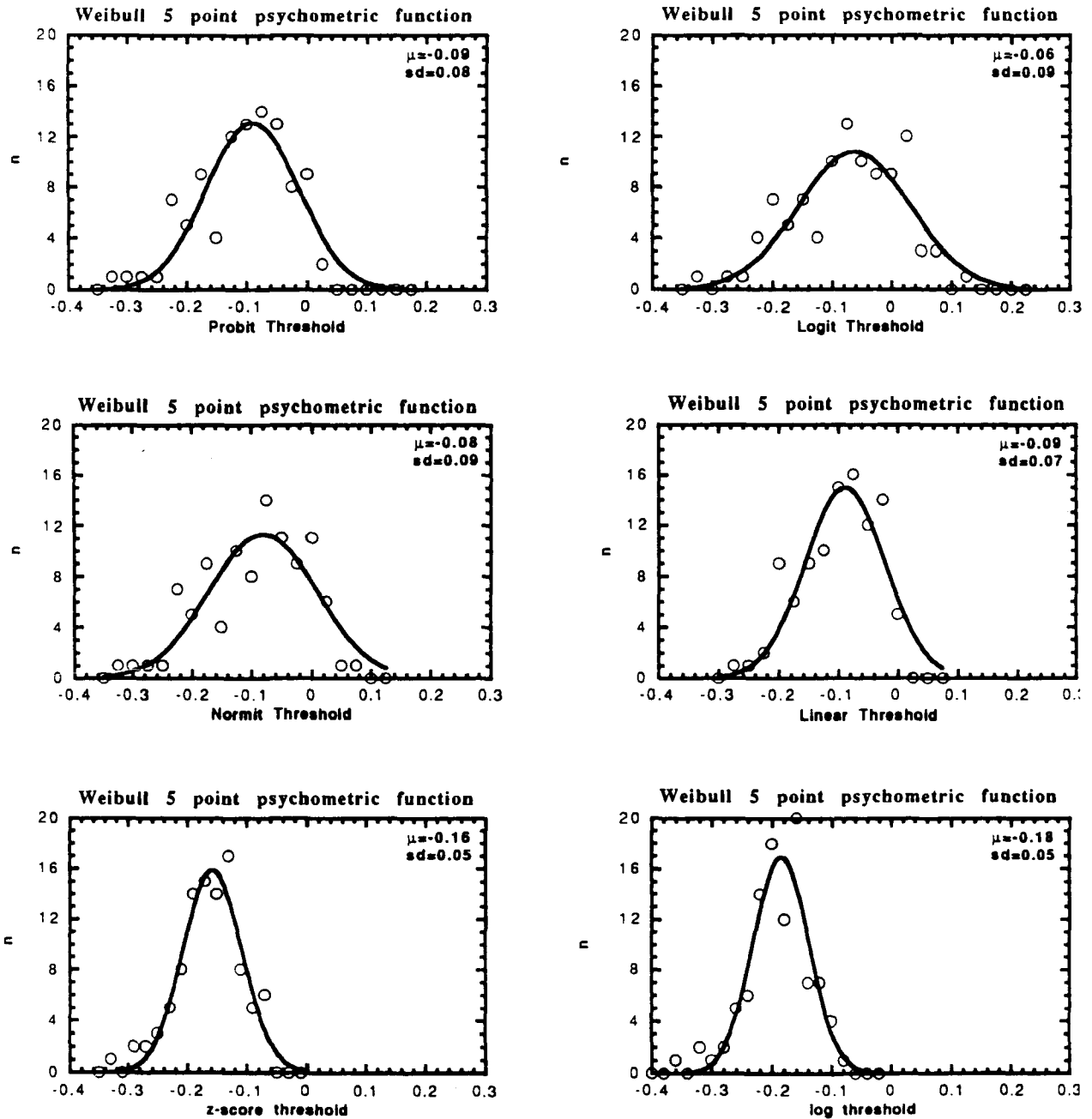


Figure 3. Histograms of thresholds estimated from 5-point Weibull psychometric functions obtained using six procedures. The ordinates are thresholds and the abscissas are counts. The solid lines are best-fitting Gaussians, and the means (μ) and standard deviations (sd) are in the upper right corner of each graph.

imum chi-square methods were weighted linear regressions and the others were least-squares linear regressions. The basic algorithms were therefore very similar. The source-code file sizes (including the curve-fitting and input-output modules) were 21K, 14K, 14K, 5K, 6K, and 6K for probit, logit, normit, linear-regression, z-score-regression, and logit-regression modules, respectively. Because the file sizes are relatively small, procedure speed might be a more important criterion. To quantify this, the time to calculate 100 thresholds was

measured for each of the threshold computation modules, accurate to ± 17 msec (using a Macintosh "Tick-Count"). The upper graph in Figure 12 shows that the fastest was the linear-regression module, with each of the other regressions and logit analysis slightly slower. The lower graph in Figure 12 shows the relative speed of the modules, normalized to linear regression. As is apparent, probit analysis is approximately 7-14 times slower than linear regression, logit analysis, and logit regression. Probit analysis is about 4 times slower than normit

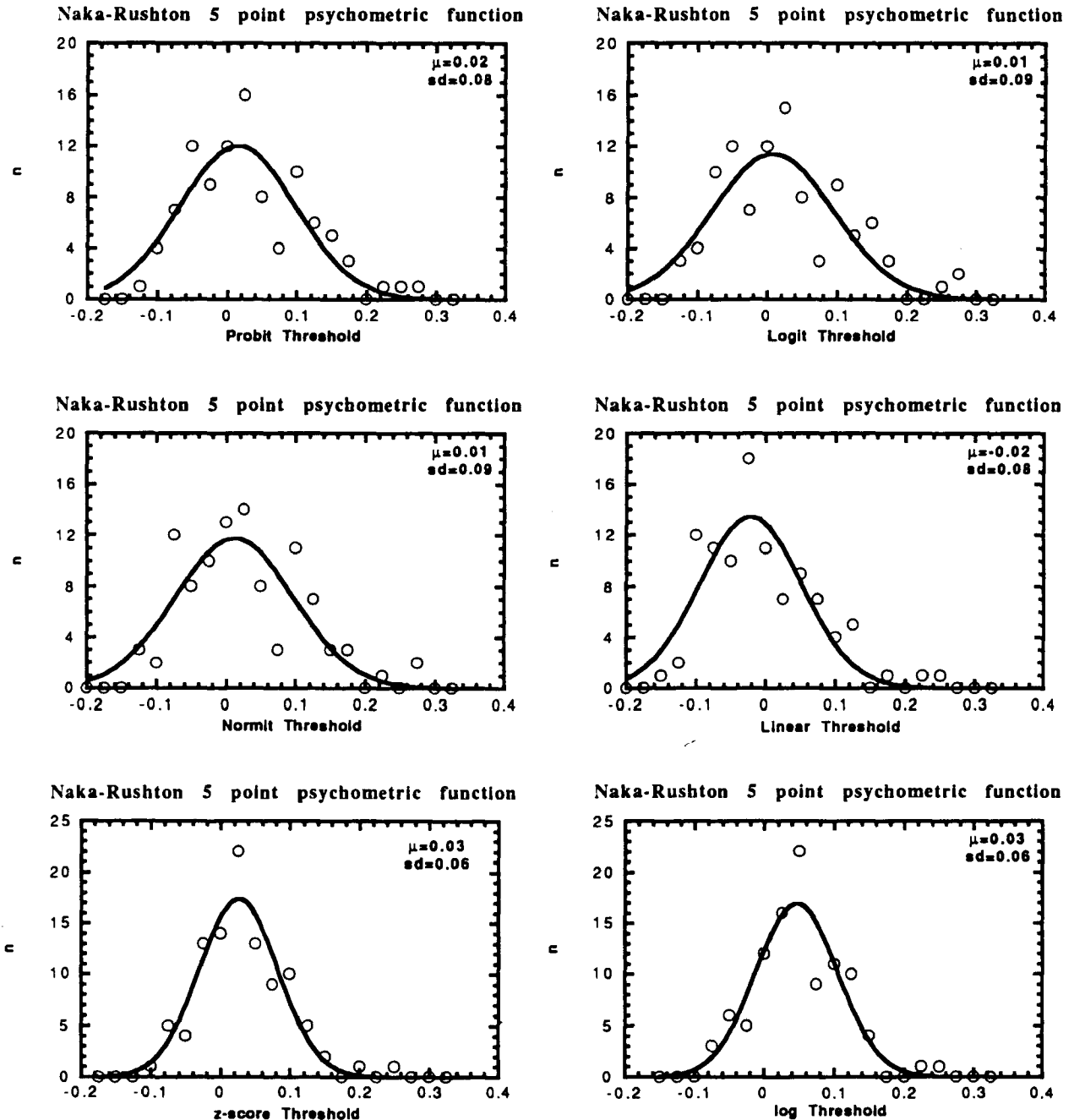


Figure 4. Histograms of the thresholds estimated from Naka-Rushton psychometric functions obtained using six procedures. The ordinates are thresholds and the abscissas are counts. The solid lines are best-fitting Gaussians, and the means (μ) and standard deviations (sd) are in the upper right corner of each graph.

analysis, and between 4 and 7 times slower than z-score regression.

DISCUSSION

The simulations showed that there were remarkable similarities between thresholds estimated using probit, logit, and normit analyses. The correlation coefficients between these three procedures were always greater than

.95 and, as is illustrated by these data clustering along the equality lines in the scatterplots, the logit and normit thresholds were unbiased and very similar to the probit thresholds. Perhaps the similarity of estimates of thresholds should not be surprising, since each procedure was derived using asymptotic theory, although probit analysis is a maximum-likelihood procedure, whereas in logit and normit analysis, the chi-square is minimized. Because the minimum chi-square procedures provide sim-

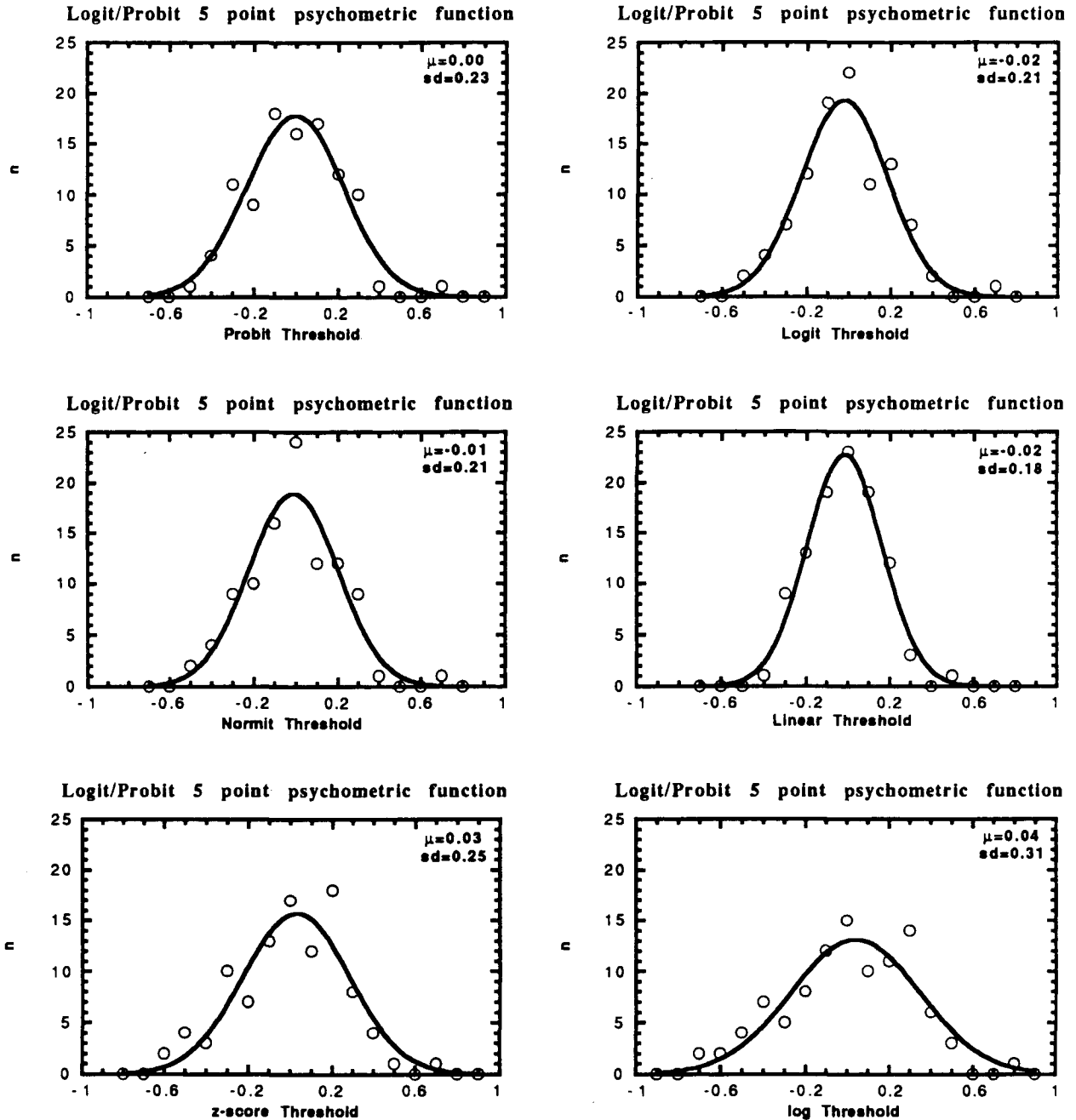


Figure 5. Histograms of thresholds estimated from logit/probit psychometric functions using six procedures. The ordinates are thresholds and the abscissas are counts. The solid lines are best-fitting Gaussians, and the means (μ) and standard deviations (sd) are in the upper right corner of each graph.

ilar thresholds to probit analysis, will always converge on an estimate of threshold (something not true for probit analysis), and are faster, much can be gained by using them in time-critical situations (such as while running an experiment). In addition, because these two procedures are simple weighted linear regressions, they may also be more easily implemented on spreadsheets or other tools that can compute linear regression.

A decision about the relative advantage of logit analysis over normit analysis appears to be based mainly on time. As implemented in these simulations, logit analysis was about twice as fast. This may have arisen because of a nonoptimal $p(\text{correct})$ -to-z-score algorithm. An indication that this was the case is in the comparison between the time taken to compute thresholds using logit and z-score regression. All aspects of these modules were iden-

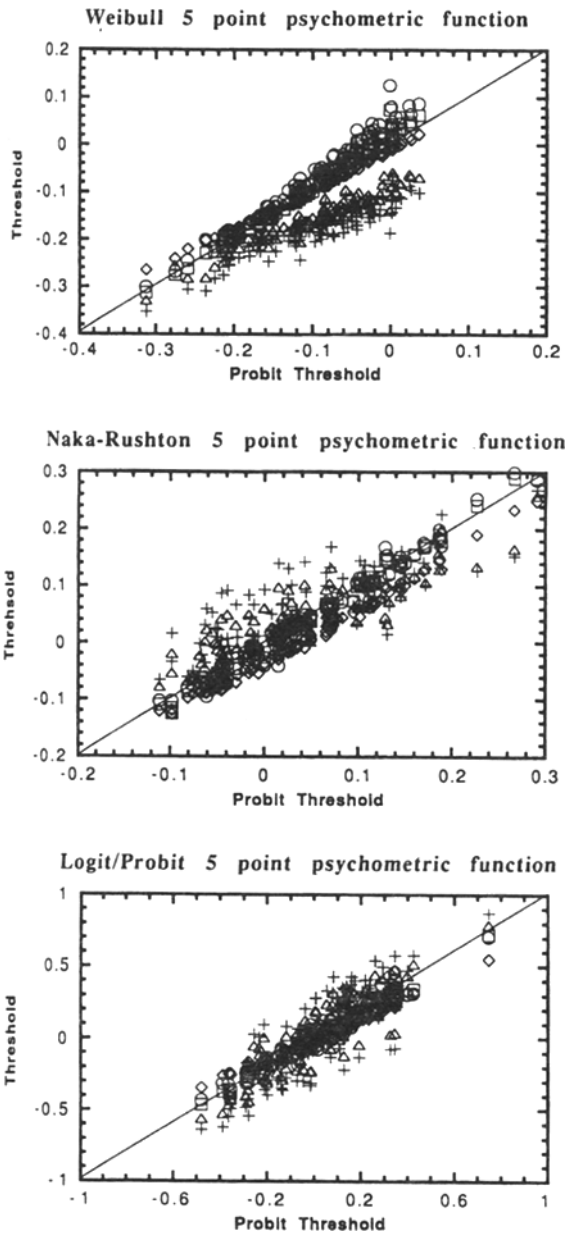


Figure 6. Scatterplots showing the relationships between thresholds estimated using probit analysis (abscissas) and logit analysis (circles), normit analysis (squares), linear regression (diamonds), z-score regression (triangles), and logit regression (crosses). The upper, middle, and lower graphs show data derived using 5 points from, respectively, Weibull, Naka-Rushton, and logit/probit psychometric functions. The oblique solid lines on each graph are lines of equality.

tical except for the z-score- or logit-transform procedures, and yet the logit-regression thresholds were still computed more quickly. The relative slowness of the $p(\text{correct})$ -to-z-score algorithm might be overcome, thereby reducing the difference in speed between logit and normit analysis, by, for example, using look-up tables in order to perform the transformations. Similar look-up tables might also be used to further speed up logit analysis.

Regressions using z-score- and logit-transformed probabilities were also used to derive thresholds. Generally these estimates were more variable than probit, logit, or normit analysis, but were highly correlated with each other (Figure 13). Considering that the threshold estimations were generally unbiased, these procedures pro-

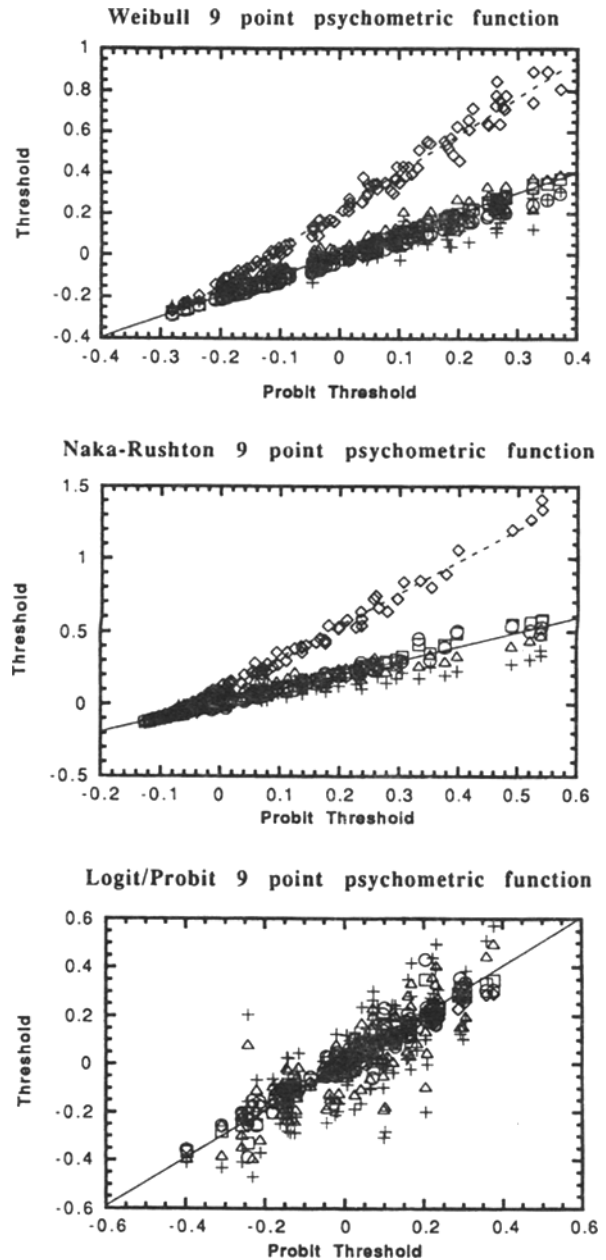


Figure 7. Scatterplots showing the relationships between thresholds estimated using probit analysis (abscissas) and logit analysis (circles), normit analysis (squares), linear regression (diamonds), z-score regression (triangles), and logit regression (crosses). The upper, middle, and lower graphs show data derived using 9 points from, respectively, Weibull, Naka-Rushton, and logit/probit psychometric functions. The oblique solid lines on each graph are lines of equality and the dashed lines are the least-squares fits to the linear-regression thresholds.

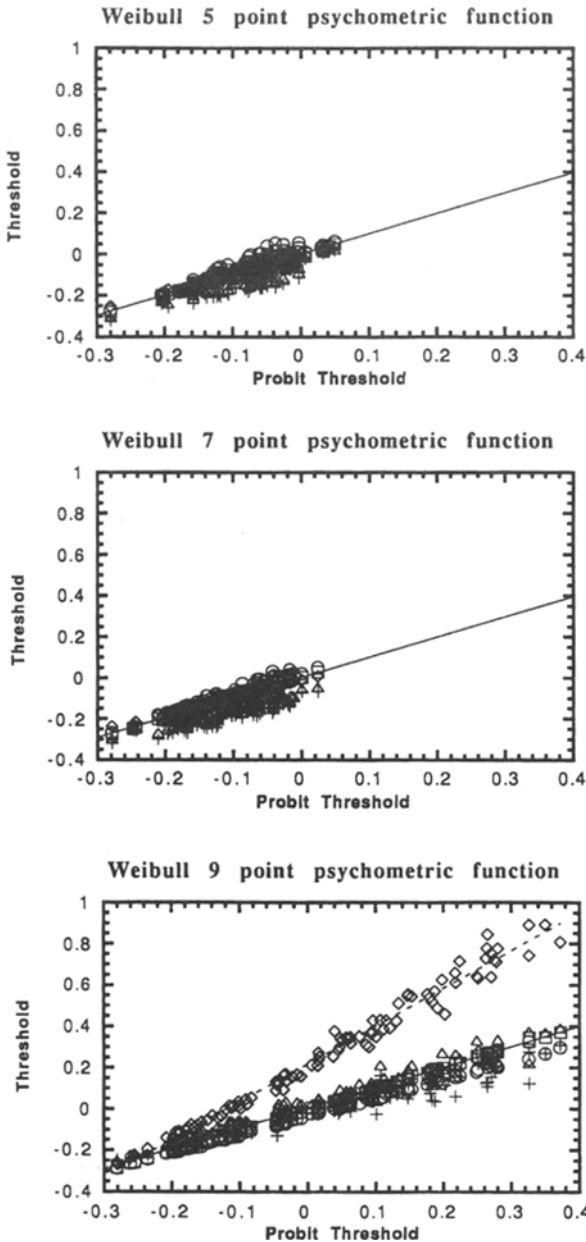


Figure 8. Scatterplots showing the relationships between thresholds estimated using probit analysts (abscissas) and logit analysis (circles), normit analysis (squares), linear regression (diamonds), z-score regression (triangles), and logit regression (crosses). The upper, middle, and lower graphs show data derived using, respectively, 5, 7, and 9 points from a Weibull psychometric function. The oblique solid lines on each graph are lines of equality and the dashed lines are the least-squares fits to the linear-regression data.

vide practical alternatives particularly because they can be easily implemented—for example, using the hardware and software available in many pocket calculators.

Simple linear regression was examined because it was the equivalent of fitting a straight psychometric function by eye and using this to obtain a threshold. This method potentially presents many problems. For example, a straight-line fit to probabilities can produce impossible

thresholds (e.g., negative luminance) and/or stimulus levels associated with impossible probabilities (i.e., $.0 > p > 1.0$; Collett, 1991). In addition, the simulations showed that extremely biased thresholds were estimated from some psychometric functions and that somewhat less biased estimates were obtained when the points on the psychometric functions were not symmetrical around $p(\text{correct}) = .5$. The former point is particularly germane because it appears that some psychometric functions obtained in vision experiments are asymmetrical and “Weibull-like” (e.g., Harvey, 1986).

Nevertheless, the linear-regression algorithm performed well in many simulations, especially in comparison with z-score or logit regression. There are a number of reasons why this might be so. As is illustrated in Figure 14, the central portion of the z-score- and logit-transformed $p(\text{correct})$ function is linear (although the range of this linearity differs for logit and z-score transformations). Therefore, comparisons between thresholds from untransformed and transformed modules

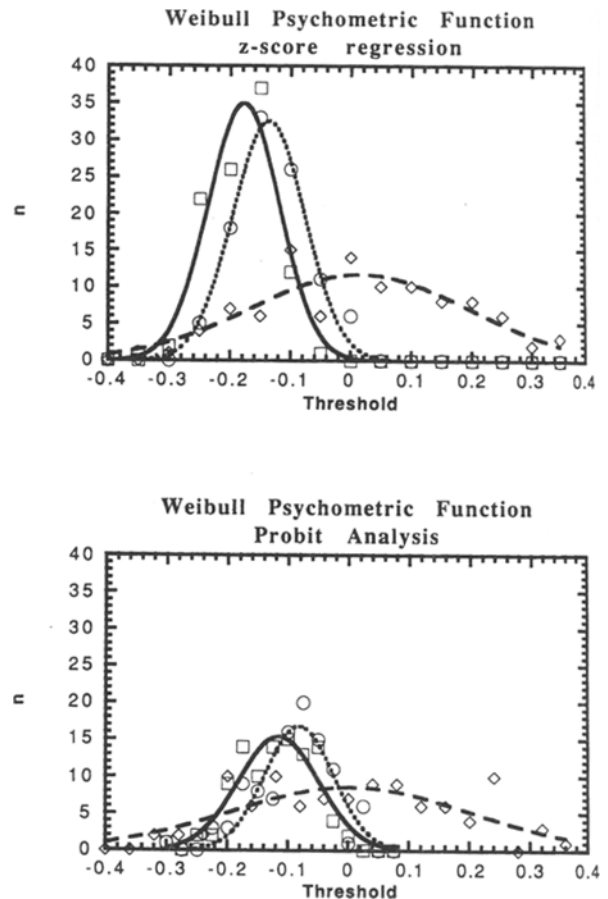


Figure 9. Histograms showing the distributions of thresholds estimated using 5 points (circles, dotted line), 7 points (squares, solid line), and 9 points (diamonds, dashed line) from a Weibull psychometric function. The abscissas are thresholds and the ordinates are counts. The upper and lower panels depict thresholds estimated using, respectively, z-score regression and probit analysis. The smooth curves are best-fitting Gaussians.

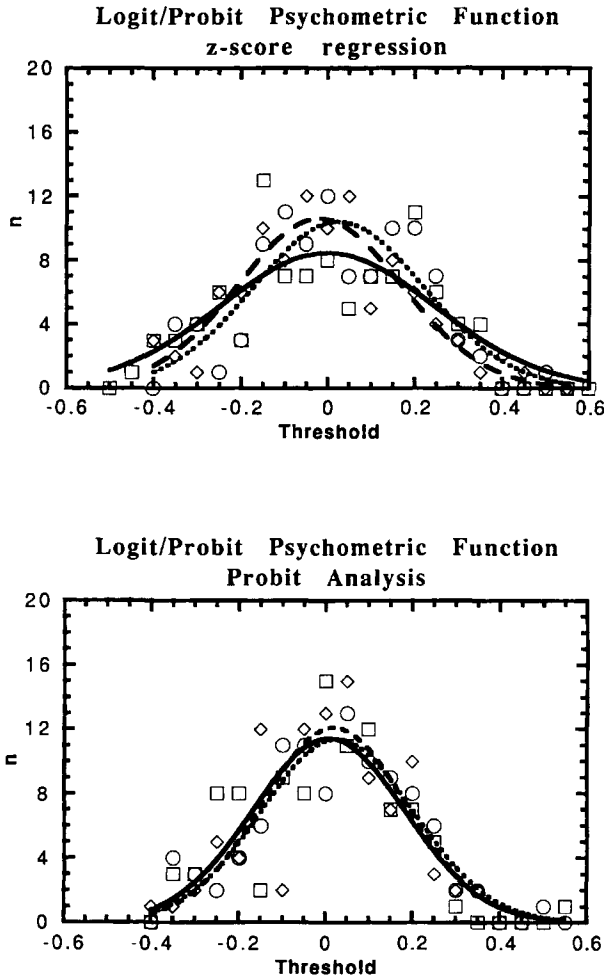


Figure 10. Histograms showing the distributions of thresholds estimated using 5 points (circles, dotted line), 7 points (squares, solid line), and 9 points (diamonds, dashed line) from logit/probit psychometric functions. The abscissas are thresholds and the ordinates are counts. The upper and lower panels depict thresholds estimated using z-score regression and probit analysis, respectively. The smooth curves are best-fitting Gaussians.

would be quite similar, provided that the psychometric functions were sampled from points closer to the middle. The biased linear-regression estimates occurred when points were sampled closer to the asymptotes, where the psychometric functions were nonlinear.

As was previously stated, linear-regression thresholds were especially impressive in comparison with z-score and/or logit regressions. This probably occurred because the maximum-likelihood and minimum chi-squared procedures weighted probabilities closer to the asymptotes progressively less often. On the other hand, the z-score and logit regressions effectively weighted these points more because the transformed probabilities increased/decreased greatly as the asymptotes were approached. Because linear regression weighted all of the points equally, the thresholds obtained in many instances were closer to those obtained using probit, logit, or normit analysis than were the other regression procedures.

It should be noted that the relatively good performance of linear regression was partially a consequence of having defined threshold as $p(\text{correct}) = .5$. This value would serve as a point around which the inappropriate linear function could pivot, thereby providing reasonable threshold estimates. Under circumstances in which this did not apply, the linear-regression thresholds

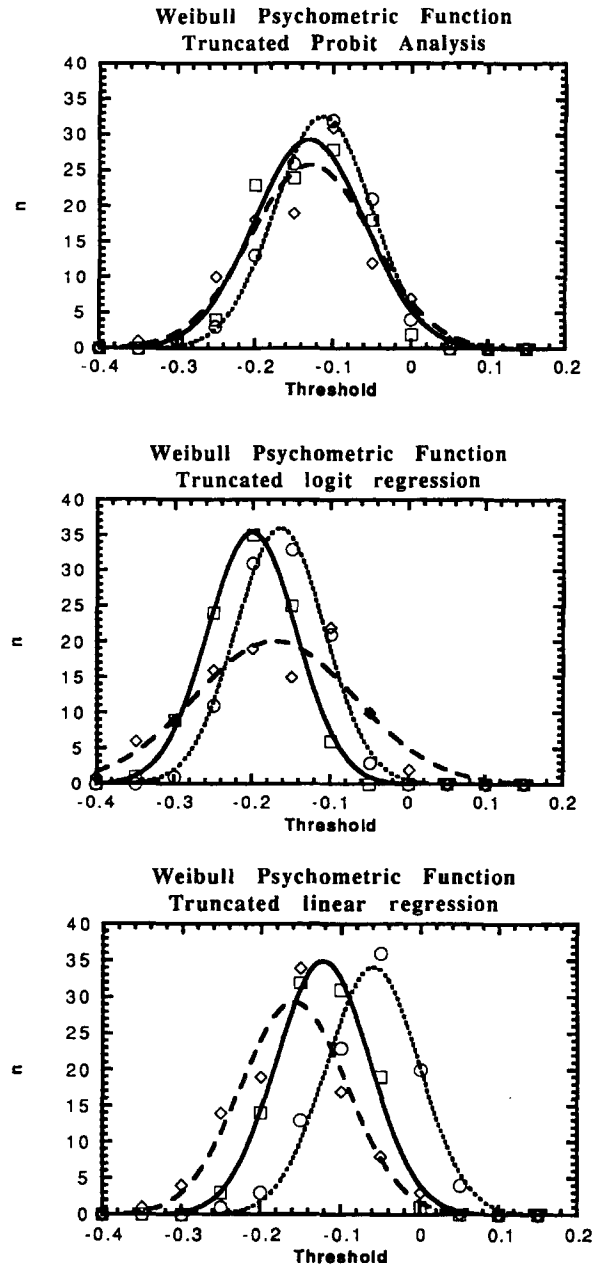


Figure 11. Histograms showing the distributions of thresholds estimated using 7 points from the middle (squares, solid line), upper (circles, dotted line), and lower (diamonds, dashed line) positions of a Weibull psychometric function. The abscissas are thresholds and the ordinates are counts. The upper, middle, and lower graphs depict thresholds estimated using probit analysis, z-score regression, and linear regression, respectively. The smooth curves are best-fitting Gaussians.

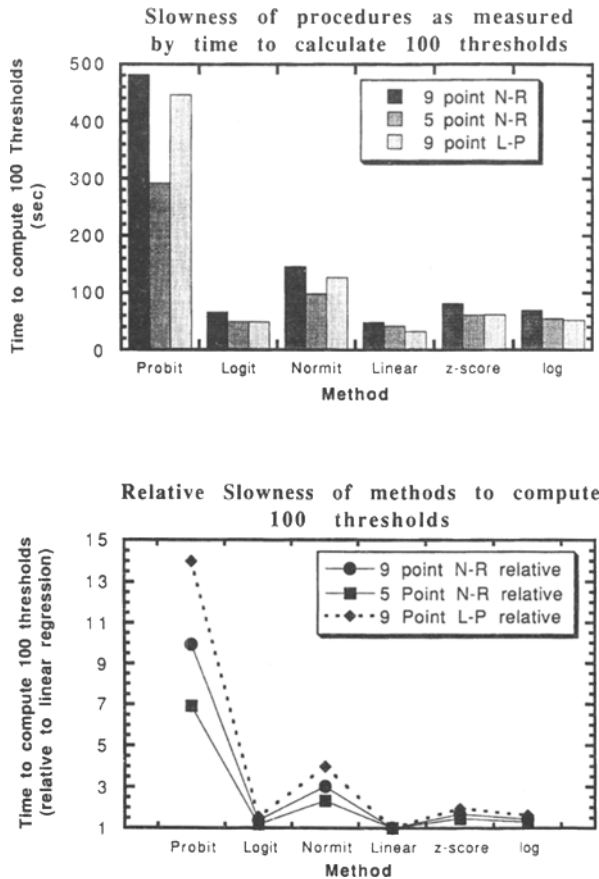


Figure 12. Graphs illustrating the amount of time to compute 100 thresholds (ordinate) using the various methods (abscissas). In the upper graph, the ordinate is in seconds and on the lower graph, it is time relative to linear regression. The dark, intermediate, and light bars in the upper graph and the circles, squares, and triangles in the lower graph represent time taken using data from, respectively, 9-point Naka-Rushton, 5-point Naka-Rushton, and 9-point logit/probit psychometric functions.

were biased (e.g., with Naka-Rushton and Weibull psychometric functions and when points were sampled from the upper and lower psychometric functions). The generally good performance of the linear regression reported here would therefore most likely not apply to thresholds defined at different probability levels, such as $p(\text{correct}) = .8$ (e.g., Flom, 1966).

Although not directly addressed, the simulations reported here provide support for the empirical suggestion by Berkson (1955) of dealing with $\pm\infty$. As was outlined in the Method section, Berkson suggested that an arbitrary probability level based on the number of trials be assigned to $p(\text{correct}) = 1.0$ or $.0$. This was used in normit and logit analyses, but not in probit analysis. Despite this, the thresholds were consistently similar, even though in some of the simulations, as many as 50% of the simulated psychometric functions included points with $p(\text{correct}) = 1.0$ or $.0$. This result illustrates another potential advantage of logit and normit analysis over probit analysis. If a psychometric function consisted of just

three points, two of which were at $p(\text{correct}) = 1.0$ and/or $p(\text{correct}) = .0$, the method outlined by Finney (1971) to obtain a threshold using probit analysis could not be performed because to obtain a provisional line, $p(\text{correct}) = 1.0$ and $p(\text{correct}) = .0$ would be disregarded. As undesirable as it might be, a threshold with these three data points may be derived using logit and normit analysis.

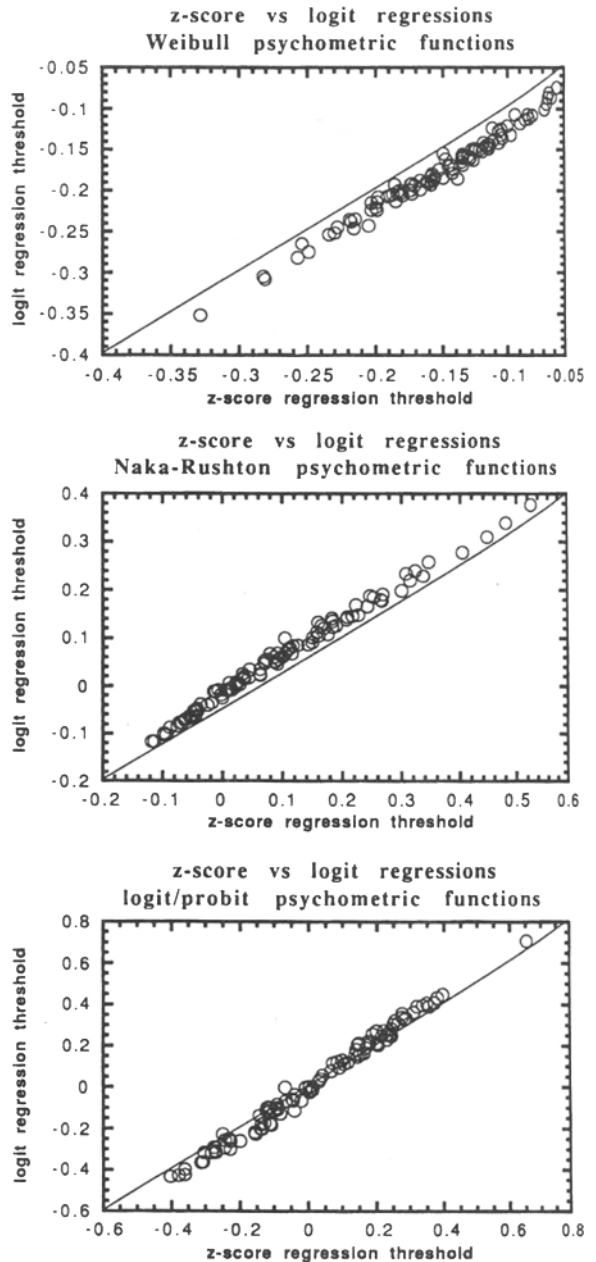


Figure 13. Scatterplots showing the relationships between thresholds estimated using z-score regressions (abscissas) and logit regressions (ordinates). The upper graph depicts thresholds estimated from five-point Weibull psychometric functions, the middle graph depicts thresholds estimated from 7-point Naka-Rushton psychometric functions, and the lower graph depicts thresholds estimated from 9-point logit/probit psychometric functions. The oblique solid lines on each graph are lines of equality.

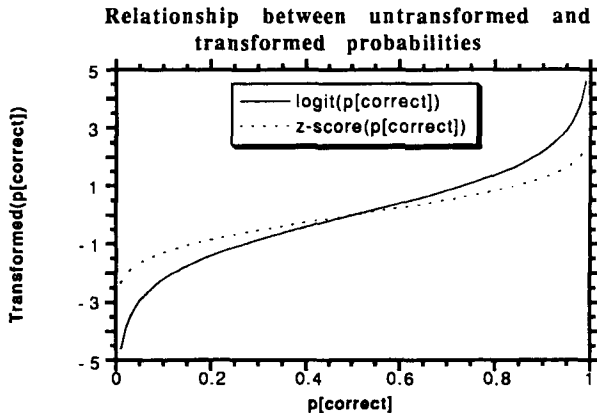


Figure 14. The relationships between the untransformed and transformed probabilities (ordinate and abscissa, respectively) for logit (solid line) and probit (dotted line) functions.

The question as to whether to include points at $p(\text{correct}) = 1.0$ and $p(\text{correct}) = .0$ (and if so, how many) is an empirical one not addressed here.

The results of this series of simulations showed that, with the exception of linear regression, the methods were generally immune to peculiarities of the psychometric functions (e.g., truncated sampling). This is not a trivial observation, since these types of peculiar psychometric functions may be obtained especially when dealing with clinical patients and when the luxury of repeated measurements is unavailable. It remains to be determined whether the different threshold methods used in the simulations provide similar threshold estimates using real data, particularly data from clinical subjects.

In conclusion, the different methods provided similar estimates of thresholds, except for linear regression, which produced biased thresholds under certain circumstances. The thresholds obtained using probit, logit, and normit analyses were nearly identical, and because probit analysis was substantially slower, the logit/normit algorithms seem particularly useful, especially in time-critical situations.

REFERENCES

ABROMAWITZ, M., & STEGUN, I. A. (1964). *Handbook of mathematical functions*. Washington, DC: U.S. Government Printing Office.

- BERKSON, J. (1944). Application of logistic function to bio-assay. *Journal of the American Statistical Association*, **39**, 357-365.
- BERKSON, J. (1953). A statistically precise and relatively simple method of estimating the bio-assay with quantal response, based on the logistic function. *Journal of the American Statistical Association*, **48**, 565-599.
- BERKSON, J. (1955). Estimate of the integrated normal curve by minimum normit chi-square with particular reference to bio-assay. *Journal of the American Statistical Association*, **50**, 529-549.
- BLISS, C. I. (1935). The calculation of the dosage mortality curve. *Annals of Applied Biology*, **22**, 134-167.
- COLLETT, D. (1991). *The modelling of binary data*. London: Chapman & Hall.
- FINNEY, D. J. (1971). *Probit analysis* (3rd ed.). Cambridge: Cambridge University Press.
- FLOM, M. C. (1966, July). New concepts on visual acuity. *Optometric Weekly*, **57**, 63-68.
- GESCHIEDER, G. A. (1985). *Psychophysics: Method, theory, and application*. Hillsdale, NJ: Erlbaum.
- HARVEY, L. O., JR. (1986). Efficient estimation of sensory thresholds. *Behavior Research Methods, Instruments, & Computers*, **18**, 623-632.
- KLEINBAUM, D. G., KUPPER, L. L., & MULLER, K. E. (1988). *Applied regression analysis and other multivariable methods*. Boston: PWS-Kent Publishing.
- LIEBERMAN, H. R. (1983). Computation of psychophysical thresholds using the probit technique. *Behavior Research Methods & Instrumentation*, **15**, 446-448.
- NAKA, K.-I., & RUSHTON, W. A. H. (1966). S-potentials from colour units in the retina of fish (Cyprinidae). *Journal of Physiology*, **185**, 587-599.
- QUICK, R. F. (1974). A vector magnitude model of contrast detection. *Kybernetik*, **16**, 65-67.
- STEVENS, S. S. (1972, September). A neural quantum in sensory discrimination. *Science*, **177**, 749-762.
- WEIBULL, W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics*, **18**, 293-297.

NOTES

1. Normit is an abbreviation of *normal deviate*, and is a synonym of z score. A normit is therefore a probit - 5.0.
2. Logits are transformed probabilities, where $\text{logit}_p = \log_e [p/(1-p)]$.
3. For brevity, z-score- and logit-transformed probability least-squares regressions and linear probability regression will be referred to simply as z-score regressions, logit regressions, and linear regressions, respectively.

(Manuscript received September 29, 1993;
revision accepted for publication September 13, 1994.)