

# The Application of an Unfolding Model of the PIRT Type to the Measurement of Attitude

David Andrich  
Murdoch University, Australia

Unfolding data for unidimensional variables constructed from direct responses (e.g., agreement or disagreement) are characterized by single peaked functions involving the locations of each person and each stimulus. A continuous discrimininal process, of the form postulated by Thurstone when he proposed his Law of Comparative Judgment, is suggested. This process is transformed to a qualitative dichotomous response in which the probability of endorsement is governed by the square of the distance between the locations of the person and the stimulus. Maximum likelihood estimates of the parameters are derived, and it is shown that the information associated with any re-

sponse is a bimodal function of the difference between the person and stimulus locations. The feasibility of parameter estimation is demonstrated with a limited simulation study. The model is applied to a set of statements designed to measure attitudes toward capital punishment and scaled by the methods of Thurstone. The responses conformed to the unfolding mechanism, and the scale values of the statements are statistically equivalent to those obtained by Thurstone's methods. *Index terms: Attitude measurement, Developmental data, Discriminal process, Item response theory, Person response theory, Thurstone scaling, Unfolding data, Unidimensional scaling.*

Thurstone's formulation of the Law of Comparative Judgment (Thurstone, 1927a) and his related work on the measurement of traits with no physical counterparts introduced two clearly distinguishable response mechanisms or principles for psychological scaling, now generally known as the *cumulative* and *unfolding* mechanisms. In the cumulative mechanism, the probability of a positive response takes a cumulative sigmoid shape as a function of the location of the person on the scale, while in the unfolding mechanism, it takes a single peaked shape with probabilities tending to 0 as attitude values tend to extremes in either direction. In both, persons and items are presumed located on a unidimensional continuum. Although extensions to multidimensional continua for both types of data have been made, only unidimensional scaling is considered in this paper.

For each of these two scaling principles there are two basic data collection designs. The first design, in which persons react directly to stimuli, will be termed generally the *direct response design*. The second, in which persons make comparisons among a set of stimuli and select one or more of these, will be termed generally the *comparison response design*. The simplest and most familiar direct response designs are (1) those in achievement or intelligence testing, in which persons respond correctly or incorrectly to

test items; and (2) those concerned with attitude measurement, in which persons respond by either agreeing or disagreeing with statements. The basic and most familiar comparison response design is that of pair comparisons, in which persons compare each pair of stimuli from a set on some property and select one of them. Thurstone's Law of Comparative Judgment (Thurstone, 1927b) is based on this design.

Thurstone's own work on attitude measurement (Thurstone, 1928) involved both the direct and comparison designs and both the cumulative and unfolding data. For the purpose of constructing a scale, Thurstone articulated the case for placing statements on a linear continuum in such a way as to reflect not only the order of their intensity, called their *affective values*, but also the distance between their positions.

For this stage of scale construction, various comparison designs (including those where statements were compared and sorted into classes or ranked; Thurstone, 1931) were used, and the associated models proposed were based on the cumulative mechanism. A feature of this mechanism is that if statement A is further along the direction of intensity of the affective continuum than statement B, then it is more likely to be endorsed as such by every member of a population *irrespective of that person's own attitude*. A clarification of the cumulative model for the direct response design is provided below.

For the stage of person measurement, the persons responded directly by agreeing or disagreeing with the scaled statements, and their measures were taken simply as the median or mean of the scale values of the statements with which they agreed. The assumption behind this approach was that a person would tend to agree with those statements which reflected and were close to his or her attitude, and that a person would tend to disagree with those statements which were at either extreme of that attitude. Such an assumption thus involved the mechanism of unfolding.

There is another important distinction between the direct and comparison response designs in Thurstone's work. For the stage of scale construction associated with the latter, Thurstone specified explicitly formal probabilistic models based on the Law of Comparative Judgment, while in the person measurement stage, the procedure described above was essentially ad hoc.

The most advanced developments in explicit probabilistic models are in the area of cumulative models for both direct and comparison designs. Research in modern person-item response theory (PIRT) is devoted essentially to direct response designs with cumulative data (Baker, 1977; Bock & Wood, 1971; Lumsden, 1976; Traub & Wolfe, 1981), and these are applied to both achievement and attitude data. Models for the pair comparison design have also been studied extensively (Bock & Jones, 1968; David, 1963); Davidson and Farquhar (1976) presented a bibliography of about 350 papers on the topic.

Explicit probabilistic models for unfolding data with respect to pair comparison designs have only infrequently been considered (Sixtl, 1973; Zinnes & Griggs, 1974), while those with a direct response design seem to have had even less consideration.

The purpose of this paper is to consider a simple probabilistic model for unfolding data collected by a direct response design in which the responses are scored dichotomously. The model provides a formal probabilistic expression of the kind found in PIRT, yet it is based on the same principles proposed in the correlational-factor analytic approach.

In any development of a new mathematical model for data, two particular issues are invariably of interest. The first is whether, in a nontrivial way, certain results conform to those obtained by traditional methods. The second is whether any new insights may be obtained from the specification of the model. An example of each of these issues is considered.

Because the proposed model returns the topic to first principles, principles which have gained currency in areas other than attitude measurement (Coombs & Avrunin, 1977), and because the unfolding mechanism is not often applied in psychological research, the next section briefly elaborates on the unfolding principle. This is followed by an introduction of the probabilistic model for which maximum likelihood estimates

and their standard errors are derived; the remaining sections discuss the feasibility of parameter estimation in a small simulation study, and present the analysis of a real dataset.

### Unfolding Data

As indicated above, Thurstone's measurement of attitude on a single continuum involved the concept of unfolding in that a person was expected to endorse statements close to the person's attitude and to reject statements not close, irrespective of direction. Taking the median or mean of the scale values of statements endorsed, which follows from the above principle, was reasonable and manageable when the statements had been scaled independently using a comparison design and the cumulative principle.

The concept of unfolding has been found to be appropriate in areas governed by the general concept of development along a hypothesized continuum in which qualitatively distinct manifestations of the positions on the continuum can be identified (Coombs & Smith, 1973). These stages are generally seen to proceed in an invariant sequence. An early example in this field was provided by Leik and Matthews (1968); the example concerned the degree of independence of wives from husbands.

Conceiving of independence as being acquired in developmental stages, and following the usual scalogram approach, we assumed that as a wife becomes independent in new spheres, such as economic, she retains whatever earlier independent behavior she had exhibited, such as independence in family decision-making. The data, unfortunately, did not share our enthusiasm for the hypothesized pattern. Careful examination of individual response patterns showed that one aspect of the Guttman approach was inappropriate: Many of the women evidently ceased exhibiting earlier types of independence once they had attained more "advanced" independence. (p. 64)

As an alternative to the cumulative response principle, it is evident that these authors had the unfolding principle in mind. The same principles have been formalized by Davison, Robbins, and Swanson (1978) with respect to stages of moral development in the sense of Kohlberg (1969).

To clarify the distinction between the principles of the cumulative scale and the unfolding scale (Guttman, 1950, 1954) for a direct response design, a perfect cumulative Guttman scale and an unfolding scale are shown in Table 1. Similar tables have been presented in Leik and Matthews (1968) and Coombs (1964).

In addition to the examples cited above, it is possible that the unfolding principle might sometimes be appropriate in what would seem initially to be an unlikely domain: that of attainment or achievement, where the cumulative principle is applied routinely. For example, it is conceivable that as a person progresses through a field of study, as further learning and greater understanding of deeper relationships takes place, certain surface skills used and taught at an earlier stage of development are not retained in the same way. Thus college students studying algebra may find difficulty in solving certain problems which they could solve readily when they were in high school.

### A Model for Unfolding Data

The early work on unfolding data involved primarily deterministic models (Coombs, 1950, 1964), with various algorithms being proposed for uncovering any unfolding structure. Many of these algorithms were based on the absolute distance,  $d_{ni} = |\beta_n - \delta_i|$ , between the locations  $\beta_n$  of person  $n$  and  $\delta_i$  of stimulus  $i$ . Clearly, these approaches had to contend, as did the attempted applications of the Guttman scale, with data that were stochastic; as with the Guttman scale, the structures of correlations have been used to establish whether or not the unfolding principle underpins the observed data. The analyses of such correlations are based on the traditional true score models (Gulliksen, 1950).

Table 1  
Response Patterns for Ideal Cumulative and  
Unfolding Scales

Type of Scale	Stimulus				
	1	2	3	4	5
Cumulative: All possible responses					
	0	0	0	0	0
	1	0	0	0	0
	1	1	0	0	0
	1	1	1	0	0
	1	1	1	1	0
	1	1	1	1	1
Unfolding: Example of possible responses					
	0	0	0	0	0
	1	0	0	0	0
	1	1	0	0	0
	0	1	1	0	0
	0	0	1	1	0
	0	0	0	1	1
	0	0	0	0	1
	0	0	0	0	0

For a continuous variable, Davison (1977) postulated an explicit model for the generation of unfolding data according to the form

$$z_{ni} = a_i + b_i(\beta_n - \delta_i)^2 \quad (1)$$

where  $a_i$  and  $b_i$  are constants pertaining to stimulus  $i$ .

The reaction  $z_{ni}$  when person  $n$  encounters stimulus  $i$  is seen to be governed by the square of the difference between the person and the stimulus locations, which is the usual alternative to the absolute value of this difference when distance, rather than direction, is the relevant feature. Based on the above formulation, Davison (1977) established and illustrated the factor structure among stimuli which results from such data.

The form of the expression  $z_{ni}$  is reminiscent of Thurstone's specification of the discriminial process as

$$z_{ni}^* = \delta_i^* + e_{ni}^* \quad (2)$$

where  $\delta_i^*$  is the location value of stimulus  $i$  and  $e_{ni}^*$  is the error associated with person  $n$ . The asterisk in the  $z_{ni}^*$  term denotes that no person parameter is postulated;  $z_{ni}$  (without the asterisk) is used when a person parameter is postulated. Without loss of generality, and with some further insights into the discriminial process thus conceived,  $z_{ni}$  and  $z_{ni}^*$  may be related by resolving  $e_{ni}^*$  into a location parameter  $\beta_n$  for person  $n$  and an error component within that person (Andrich, 1978a; Lumsden, 1977) according to

$$e_{ni}^* = \beta_n + e_{ni} \quad (3)$$

Then, again without loss of generality, setting  $\delta_i^* = -\delta_i$ , Thurstone's discriminial process may be written as

$$z_{ni} = \beta_n - \delta_i + e_{ni} \quad (4)$$

with  $E(z_{ni}) = \beta_n - \delta_i$  and  $\text{Var}(z_{ni}) = \sigma_{ni}^2$ . If  $z_{ni} > 0$ , a positive response is observed, otherwise it is negative.

The case where  $\sigma_{ni}^2 = \sigma^2$  for all  $n$  ( $n = 1, 2, \dots, N$ ) and all  $i$  ( $i = 1, 2, \dots, I$ ) effectively gives Thurstone's Case V of the Law of Comparative Judgment. In Thurstone's work, the error was taken to be normally distributed, such that if  $X$  is a random variable which takes the value  $x = 1$  for a positive response and  $x = 0$  otherwise, the probability

$$p(X = x; \beta, \delta) = F[x(\beta - \delta)] \quad (5)$$

is obtained, where  $F$  is the cumulative normal distribution with mean 0 and variance unity. In Case V of the Law of Comparative Judgment, the homogeneous variance  $\sigma^2$  is absorbed into the location parameters  $\beta$  and  $\delta$ . The alternative distribution to the normal is the logistic, which takes the form

$$p(X = x; \beta, \delta) = G[x(\beta - \delta)] \quad (6)$$

where  $G$  is the function  $G(y) = \exp(y)/[1 + \exp(y)]$ . The latter form provides the well-discussed simple logistic model (SLM) of Rasch (Rasch, 1960/1980) and is, apart from any other justifications, more attractive than the normal because it is more tractable.

Paralleling the above development, a discriminial process  $z_{ni}$  will be postulated for the unfolding principle according to

$$z_{ni} = (\beta_n - \delta_i)^2 + e_{ni} \quad (7)$$

with  $E(z_{ni}) = (\beta_n - \delta_i)^2$  and  $\text{Var}(z_{ni}) = \sigma_{ni}^2$ . In this case, because a positive response is more likely when  $\beta_n - \delta_i$  is small, a positive response will be assumed to occur when  $z_{ni} < 0$ , and a negative response otherwise. Now, if  $\sigma_{ni}^2 = \sigma^2$  ( $n = 1, 2, \dots, N; i = 1, 2, \dots, I$ ), and if the logistic distribution is presumed to characterize the cumulative error distribution, then the probability of the response  $X = x$ , where again  $x = 1$  corresponds to a positive response and  $x = 0$  otherwise, may be written as

$$p(X = x; \beta, \delta) = G[-x(\beta - \delta)^2] \quad (8)$$

More explicitly,

$$p_{xni} = p(x_{ni}; \beta_n, \delta_i) = \frac{1}{\gamma_{ni}} \exp[-x_{ni}(\beta_n - \delta_i)^2] \quad (9)$$

where

$$\gamma_{ni} = 1 + \exp[-(\beta_n - \delta_i)^2] \quad (10)$$

is a normalizing factor, is the model proposed for the unfolding response mechanism. Clearly, as the distance between  $\beta_n$  and  $\delta_i$  increases, the probability of an endorsed response decreases.

The model proposed here, termed the squared simple logistic model (SSLM), is intended to begin to bring the simplest principles of unfolding within the general framework of PIRT, where a considerable body of literature has developed for explicit probabilistic models. One possible limitation of the model, that the maximum probability of a response when  $\beta = \delta$  is only .5, may be noted immediately. There may, however, be data of the unfolding type in which the likelihood of a response is greater than .5, when the person and the item (or stage and location) are matched perfectly. Another possible limitation is the symmetry of the probability distribution. The rate of transition to a maximum probability of exhibiting the stage may not be the same as the rate of transition proceeding beyond that stage. In practice, modifications to the model may be required in order to allow for these variations.

Because the SSLM involves bringing an unfolding model within the framework of PIRT, the term "item" will be used in subsequent sections, rather than "stimulus".

### Estimation Equations

The extensive literature on estimation in PIRT immediately indicates a number of procedures that may be entertained. Here the simplest will be used, that of joint maximum likelihood in which the person

and item parameters are estimated jointly or simultaneously (Andrich, 1978b; Jensema, 1974; Lord, 1968; Wright & Panchapakesan, 1969). Other newer procedures such as marginal maximum likelihood (Bock & Aitkin, 1981; Thissen, 1982) are not considered, nor are the conditional estimation methods used with the SLM (Andersen, 1973; Fischer, 1973; Rasch, 1961; Wright & Douglas, 1977). These latter procedures are not appropriate because no simple sufficient statistics exist for the parameters, as is generally the case with Rasch models for measurement (Douglas, 1982; Rasch, 1968). However, some connections with the SLM and the principles of separability of parameters associated with the Rasch models will be discussed below.

Based on other results with respect to the Rasch models, it is known (and therefore anticipated) that for a fixed number of items and with an increasing number of persons, the item parameter estimates will not be consistent (Andersen, 1973). Considering the vast literature on the estimation issues even for the simple logistic model, it is expected that a similar array of procedures may be canvassed with this model if it proves to be of substantive interest.

The matrix of responses  $\mathbf{X}$  has elements  $x_{ni}$  with dimensions  $N \times I$ , where  $N$  is the number of persons and  $I$  is the number of items. The joint likelihood of  $\mathbf{X}$  is given by

$$L = \prod_i \prod_n \prod_{\gamma_{ni}} \frac{1}{\gamma_{ni}} \exp[-x_{ni}(\beta_n - \delta_i)^2]$$

$$= \frac{\exp\left[-\sum_n \sum_i x_{ni}(\beta_n - \delta_i)^2\right]}{\prod_i \prod_n \gamma_{ni}} \quad (11)$$

Therefore

$$\begin{aligned} \log L &= -\sum_n \sum_i x_{ni}(\beta_n - \delta_i)^2 - \sum_n \sum_i \log \gamma_{ni} \\ &= -\sum_n \sum_i x_{ni}(\beta_n^2 + \delta_i^2 - 2\beta_n \delta_i) - \sum_n \sum_i \log \gamma_{ni} \\ &= -\sum_n \sum_i x_{ni} \beta_n^2 - \sum_n \sum_i x_{ni} \delta_i^2 + 2 \sum_n \sum_i x_{ni} \beta_n \delta_i - \sum_n \sum_i \log \gamma_{ni} \\ &= -\sum_n r_n \beta_n^2 - \sum_i s_i \delta_i^2 + 2 \sum_n \sum_i x_{ni} \beta_n \delta_i - \sum_n \sum_i \log \gamma_{ni} \quad (12) \end{aligned}$$

where  $r_n = \sum_i x_{ni}$  and  $s_i = \sum_n x_{ni}$ .

Let  $p_{ni}$  be the probability  $p_{xni}$  with  $x = 1$ . Then, noting that

$$\frac{\partial \log \gamma_{ni}}{\partial \beta_n} = -2(\beta_n - \delta_i)p_{ni} \quad (13)$$

differentiating  $\log L$  with respect to  $\beta_n$  and equating the result to 0 gives

$$\begin{aligned} \phi(\beta_n) &= \frac{\partial \log L}{\partial \beta_n} = -2r_n \beta_n + 2 \sum_i x_{ni} \delta_i + 2 \sum_i (\beta_n - \delta_i) p_{ni} = 0 \\ &= \left(-r_n + \sum_i p_{ni}\right) \beta_n + \sum_i \delta_i (x_{ni} - p_{ni}) = 0 \quad (14) \end{aligned}$$

Analogously,

$$\phi(\delta_i) = \frac{\partial \log L}{\partial \delta_i} = \left(-s_i + \sum_n p_{ni}\right) \delta_i + \sum_n \beta_n (x_{ni} - p_{ni}) = 0 \quad (15)$$

To obtain a solution to Equations 14 and 15, a constraint, the need for which can be demonstrated in various ways, must be imposed. Intuitively, it is evident that if a constant is added to each  $\beta_n$ , then the same constant can be added to each  $\delta_i$  without changing the probability of an endorsed response. The usual constraint that

$$\sum_i \hat{\delta}_i = 0 \quad , \quad (16)$$

which effectively specifies the origin, may be imposed.

The simultaneous solution of Equations 14, 15, and 16 must be obtained iteratively. This may be effected readily by the Newton-Raphson algorithm, which requires considering the second derivative of the likelihood function with respect to  $\beta_n$  and  $\delta_i$ . Focusing first on the solution to Equation 14, it can be shown readily that

$$\frac{\partial^2 \log L}{\partial \beta^2} = -2r_n + 2 \sum_i [2(\beta_n - \delta_i)^2 p_{ni} - 2(\beta_n - \delta_i)^2 p_{ni} + p_{ni}] \quad . \quad (17)$$

An important characteristic of Equation 17 is that it contains data, in particular the term  $-2r_n$ . Accordingly, for the purposes of parameter estimation and for subsequent estimation of the standard errors of estimates, the expectation of Equation 17, which leads to the Fisher information function, is employed explicitly, rather than implicitly as in the SLM where no data appear in the second derivatives.

Because  $E(r_n) = \sum_i p_{ni}$ , this expectation is given by

$$\phi(\beta) = E\left(\frac{\partial^2 \log L}{\partial \beta_n^2}\right) = 4 \sum_i (\beta_n - \delta_i)^2 p_{ni} (p_{ni} - 1) \quad . \quad (18)$$

The analogous equation with respect to item  $i$  is given by

$$\phi(\delta) = E\left(\frac{\partial^2 \log L}{\partial \delta_i^2}\right) = 4 \sum_n (\beta_n - \delta_i)^2 p_{ni} (p_{ni} - 1) \quad . \quad (19)$$

The form of these equations contains the usual factor related to the variance of a Bernoulli trial,  $p_{ni}(p_{ni} - 1)$ , together with the coefficient  $4(\beta_n - \delta_i)^2$ . Equations 18 and 19, which are straightforward, will be discussed in more detail below.

### A Solution Algorithm

In multiparameter estimation of parameters, the direct Newton-Raphson algorithm would take the form

$$I - 1 \begin{matrix} \hat{\delta} \\ \vdots \\ \hat{\beta} \end{matrix} \Big|_{(q+1)} = \begin{matrix} \hat{\delta} \\ \vdots \\ \hat{\beta} \end{matrix} \Big|_{(q)} - \begin{matrix} \Psi_{\delta\delta} & \vdots & \Psi_{\delta\beta} \\ \vdots & \ddots & \vdots \\ \Psi_{\beta\delta} & \vdots & \Psi_{\beta\beta} \end{matrix} \Big|_{(q)}^{-1} \begin{matrix} \Phi_{(\delta)} \\ \vdots \\ \Phi_{(\beta)} \end{matrix} \Big|_{(q)} \quad (20)$$

(1)                      (1)                      (k+N-1)                      (1)

where

$$\Psi_{\delta\delta} = E\left(\frac{\partial^2 \log L}{\partial \delta_i \partial \delta_i}\right) \quad , \quad (21)$$

$$\Psi_{\beta\beta} = E\left(\frac{\partial^2 \log L}{\partial \beta_n \partial \beta_n}\right) \quad , \quad (22)$$

and

$$\Psi_{\delta\beta} = E\left(\frac{\partial^2 \log L}{\partial \delta_i \partial \beta_n}\right) \quad (23)$$

provide the submatrices of the expected values of the matrix of second derivatives, and  $[\delta; \beta]_{(q)}$  is the vector of provisional estimates at the  $q$ th iteration.

However, in the present case this algorithm is not feasible for two reasons. First, as  $N$  increases, the matrices of second derivatives become too large for inversion and implementation. Second, because the addition of each person introduces a new person parameter, the matrix of expected values of second derivatives is not based on replications with respect to the same parameters and therefore does not provide the limiting variance-covariance matrix of the parameter estimates. These problems, practical and theoretical, are analogous to those with respect to the binomial logistic model (Andrich, 1978b).

In the alternative algorithm, the provisional item parameters are regarded as fixed, and the Newton-Raphson algorithm is employed to improve the estimate of the parameter of the persons according to

$$[\hat{\beta}]_{(q+1)} = [\hat{\beta}]_{(q)} - [\Psi_{\beta\beta}]_{(q)}^{-1} [\phi(\hat{\beta})]_{(q)} \quad (24)$$

where the matrix  $[\Psi_{\beta\beta}] = \phi(\beta)$  is readily inverted because it is simply diagonal. Following the estimation of  $\beta$  to a chosen degree of accuracy, these are considered fixed and improved item parameter estimates are obtained, using the result  $[\Psi_{\delta\delta}] = \psi(\delta)$ , according to

$$[\hat{\delta}]_{(q+1)} = [\hat{\delta}]_{(q)} - [\Psi_{\delta\delta}]_{(q)}^{-1} [\phi(\hat{\delta})]_{(q)} \quad (25)$$

Now, following the improved estimates of  $\delta$  to the chosen degree of accuracy, the restriction  $\sum_i \hat{\delta}_i = 0$  is imposed and the entire process is repeated until the item parameters as a set converge to a chosen level.

Unlike solution algorithms for other PIRT models, no collapsing of data with respect to persons is implemented. Instead, and because the pattern of responses for each person must be considered, each person's parameter is estimated separately. In principle it would be possible, especially with a small number of items, to consider all possible response patterns and to group persons according to those patterns (Bock, 1972). Such grouping was not implemented in the program used for the analyses in this paper.

Because of the unfolding nature of the data, initial estimates for the items are found in two stages. First, the magnitudes of the estimates for the item parameters are obtained from the assumption that  $\beta_n = 0$  ( $n = 1, 2, \dots, N$ ). This gives

$$\hat{\delta}^{(0)} = \{\log [(N - s_i)/s_i]\}^{1/2} \quad (26)$$

The sign for  $\hat{\delta}^{(0)}$  is obtained from the signs of the item loadings on the first factor in a factor analysis. As shown by Davison (1977), if the data exhibit an unfolding pattern, then the factor loadings will exhibit a corresponding identifiable structure in which the signs reflect the direction of the item. Initial estimates for the person parameters are retained as 0.

It should be stressed that the existence and uniqueness of the solution for Equations 14, 15, and 16 has not been considered here, although interpretable solutions have always been found in a series of simulation and real data examples.

#### Standard Errors of the Estimates and Information

According to maximum likelihood theory, variances of the estimates of  $\beta_n$  and  $\delta_i$  are approximated by



$$\hat{\sigma}_{\beta_n}^2 = - \left[ E \left( \frac{\partial^2 \log L}{\partial \beta_n^2} \right) \right]^{-1} = - \left[ 4 \sum_i (\beta_n - \delta_i)^2 p_{ni} (p_{ni} - 1) \right]^{-1} \quad (27)$$

and

$$\hat{\sigma}_{\delta_i}^2 = - \left[ E \left( \frac{\partial^2 \log L}{\partial \delta_i^2} \right) \right]^{-1} = - \left[ 4 \sum_n (\beta_n - \delta_i)^2 p_{ni} (p_{ni} - 1) \right]^{-1} \quad (28)$$

respectively. In order to explore the implications of these estimates, the expressions for information which are their inverses will be considered in more detail. In addition, to gain more insight into the response mechanism, only one person-item combination representing the above two expressions will be considered, giving

$$I = I_{\beta\delta} = 4(\beta - \delta)^2 p(1 - p) \quad (29)$$

For convenience, this expression will be written without the person and item subscripts. Because information with respect to a person or item parameter is simply additive, the effect of having either many persons or many items can be determined immediately.

Of specific interest is the value  $I = 0$  when  $\beta = \delta$ . This contrasts with symmetric cumulative response models where maximum information is obtained when  $\beta = \delta$ . Although this result may be surprising in the first instance, on further reflection it can be construed to make intuitive sense; this will be elucidated below.

Given that  $I = 0$  when  $\beta = \delta$ , it is important to understand how this information changes as  $(\beta - \delta)^2$  deviates from 0. Setting  $a = (\beta - \delta)^2$ ,  $a > 0$ , and recognizing then that

$$p = \exp(a) / [1 + \exp(a)] \quad (30)$$

gives on simplification

$$I = 4a / [2 + \exp(a) + \exp(-a)] \quad (31)$$

Of significance now is the maximum value of  $I$ , if indeed there is a maximum. Obtaining  $\partial I / \partial a$ , setting it equal to 0, and simplifying, gives the implicit solution equation

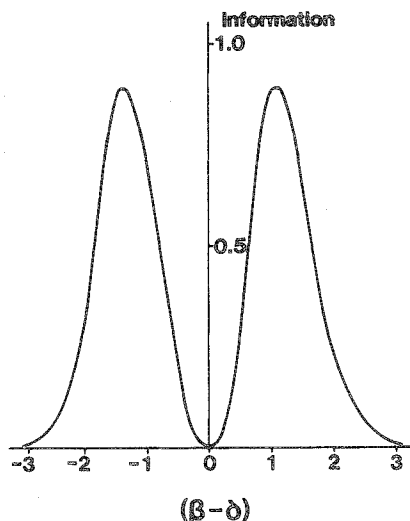
$$(a - 1)\exp(a) - \exp(-a)(a + 1) - 2 = 0 \quad (32)$$

for  $a$ . The solution to this equation is  $a = (\beta - \delta)^2 = 1.54340$  correct to five decimal places. A graphical display of  $I$  as a function of  $(\beta - \delta)$  is shown in Figure 1. From the figure it is evident that the information  $I$  with respect to  $(\beta - \delta)$  has a maximum at  $\beta - \delta = \pm 1.24234$ , correct to five decimal places.

As indicated earlier, the fact that the information for either  $\beta$  or  $\delta$  is zero when  $\beta - \delta = 0$  might seem surprising. That it has a maximum at a finite distance greater than 0 is, however, reassuring. Some possible insights regarding responses in unfolding mechanisms, which might have different mathematical formulations, can be gleaned from this result.

It is important to recognize that the response of agreement (endorsement) and the alternative response of disagreement (non-endorsement) are asymmetrical in information ("information" being used here in an informal sense). In particular, a lack of endorsement of an item is very uninformative. All that is learned is that the person is not close to the item. There is no information regarding the direction. A positive response indicates that the person is close to the item, but again, there is no indication of direction. Such asymmetry contrasts with responses in the cumulative models, where a positive response implies that the person is located at a position *greater* than the item and a negative response implies exactly the opposite. Seemingly, in certain cumulative models it is the symmetry in the evidence from a positive and negative response which provides maximum information when  $\beta - \delta = 0$ . In unfolding data, the asymmetry has the effect that information is maximum at a value  $|\beta - \delta| \neq 0$ .

Figure 1  
 Information in a Single Response as a Function of  $(\beta - \delta)$



This asymmetry is also reflected in the estimation. In the SLM, no finite parameter estimate can be obtained for a person with either a score of 0 or a maximum score of  $I$ , where  $I$  is the number of items. In the SSLM, no finite estimate can be obtained for a score of 0; however, a finite estimate can be obtained for a maximum score of  $I$  (Alan Lyne, personal communication, 1983).

**Relationship to the SLM**

The model of Equation 9 is closely related to the SLM (Rasch, 1960/1980; Wright & Stone, 1979). A feature of the SLM is that the item and person parameters are separable, where this separability is achieved through sufficient statistics (Andersen, 1977). Another important manifestation of separability is the additivity of the parameters. Although the parameters in the model of Equation 9 do not have simple sufficient statistics, there is a sense in which the parameters are separable. To illustrate this feature, write

$$p_{ni} = \frac{1}{\gamma_{ni}} \exp[-(\beta_n - \delta_i)^2] \quad (33)$$

The ratio giving the odds of non-endorsement is then

$$\frac{1-p_{ni}}{p_{ni}} = \exp[(\beta_n - \delta_i)^2] \quad (34)$$

from which

$$\log \left( \frac{1-p_{ni}}{p_{ni}} \right) = (\beta_n - \delta_i)^2 \quad (35)$$

Upon taking the square root,

$$\pm \left[ \log \left( \frac{1-p_{ni}}{p_{ni}} \right) \right]^{1/2} = \beta_n - \delta_i \quad (36)$$

While the choice of sign is a problem, it is nevertheless evident that the relationship derived above is related to the SLM in which

$$\log \left( \frac{p_{ni}}{1-p_{ni}} \right) = \beta_n - \delta_i \quad (37)$$

That is, in both cases the probability statement can be transformed so that it is simply the difference between the two relevant parameters, with the transformation for the unfolding model being more complex than the one for the cumulative model.

### A Simulation Study

It is important in any proposed model that the estimation of parameters be feasible. One purpose of this simulation study was to demonstrate that it is feasible for the SSLM for realistic datasets. A second purpose was to draw attention to any special features of the estimates. This study was not intended to be exhaustive regarding the properties of the estimators.

The simulation involved 20 items; this is a relatively large number of items for unfolding data, but is by no means atypical in attitude measurement. The items were uniformly spread over the range  $(-3.325, 3.325)$ , while the 200 person parameters were drawn from a normal distribution with mean  $\mu = 0.0$  and standard deviation  $\sigma = 2.5$ . The simulations were repeated five times with a convergence criterion of .01.

#### Recovery of Item Parameters

Table 2 shows the generating parameters of the items, and their estimates in each replication. Three features of this table are particularly relevant. First, the parameter estimates can be obtained readily by the chosen procedure. It is important to have initial estimates which are reasonable; otherwise, estimates corresponding to local maxima can be obtained. For example, if the initial estimates of both the item and person parameters are taken to be 0, then generally a local maximum is found.

Second, the variance of the estimates of the item parameters is consistently greater than the variance of the generating parameters. This reflects the inconsistency in the estimates anticipated earlier, and parallels the inconsistency in estimates for the SLM. In the SLM, the estimates can be corrected for inconsistency by multiplying them by  $(I - 1)/I$  (Wright & Douglas, 1977). No simple correction factor has been identified on the basis of this limited number of simulation studies.

Third and most important, the correlation between the generating parameters and the estimates is 1.00 correct to two decimal places, indicating that the inconsistency in estimates is essentially a matter of scale with respect to any set of data. Thus the inconsistency may be no handicap with respect to the analysis of any one set of real data.

#### Recovery of Person Parameters

Table 3 shows the summary statistics in the comparison between the generating and estimated parameters of the persons. Statistics shown in Table 3 are

1.  $r(\hat{\beta}, \beta)$ , the correlation between the estimated and the generating parameters,
2.  $r(\hat{\beta}, \hat{\beta}_T)$ , the correlation between the estimated parameters and the Thurstone estimates,
3.  $\bar{\beta}$  and  $\bar{\hat{\beta}}$ , the means of the generated and estimated parameters, respectively,
4.  $\sigma_{\hat{\beta}}^2$  and  $\sigma_{\beta}^2$ , the variances of the estimated and generating parameters, respectively,
5.  $\bar{\sigma}_{\hat{\beta}_n}^2$ , the average error variance within each person parameter estimate,

Table 2  
Item Parameter Estimates ( $\hat{\delta}$ ) in Five Simulation Studies

Item Number	Generating Parameter	Study				
		1	2	3	4	5
1	-3.325	-3.543	-3.351	-4.114	-3.695	-3.688
2	-2.975	-3.438	-3.982	-3.400	-3.073	-3.323
3	-2.625	-2.888	-2.780	-2.461	-2.989	-2.893
4	-2.275	-2.226	-2.630	-2.495	-2.338	-2.438
5	-1.925	-2.085	-1.993	-1.989	-2.172	-2.160
6	-1.575	-1.679	-1.791	-1.829	-1.882	-1.747
7	-1.225	-1.352	-1.095	-1.187	-1.428	-1.225
8	-.875	-.998	-1.036	-.720	-.825	-.737
9	-.525	-.664	-.701	-.646	-.450	-.480
10	-.175	-.163	-.081	-.542	-.174	-.484
11	.175	.239	.276	.194	.227	.330
12	.525	.544	.498	.758	.566	.288
13	.875	.874	.678	.846	.871	1.032
14	1.225	1.363	1.405	1.424	1.369	1.566
15	1.575	1.812	1.612	1.719	1.531	1.578
16	1.935	1.893	1.976	2.333	2.021	2.176
17	2.275	2.746	2.335	2.676	2.331	2.337
18	2.625	3.001	2.790	2.971	2.827	3.046
19	2.975	3.069	3.476	2.907	3.522	3.493
20	3.325	3.494	3.492	3.555	3.761	3.291
Mean	0.0	0.0	0.0	0.0	0.0	0.0
Variance	4.073	4.807	4.605	4.977	4.906	4.904
Correlation of $\delta$ and $\hat{\delta}$		1.0	1.0	1.0	1.0	1.0
Regression of $\delta$ and $\hat{\delta}$		.92	.94	.90	.91	.91

6.  $\hat{\sigma}_\beta^2 = \hat{\sigma}_\beta^2 - \bar{\sigma}_{\beta_n}^2$ , an estimate of the true variance among person parameters, and
7.  $\hat{\sigma}_\beta^2 - \sigma_\beta^2$ , the difference between the variance of the generating parameters and the estimate of the true variance of the generating parameters.

As Table 3 shows, the means of the estimated and generating parameters are very similar. Also, a pattern similar to that observed in Table 2 may be discerned in the variances. Thus the variance of the estimated parameters,  $\hat{\sigma}_\beta^2$ , is greater than the variance of the generating parameters,  $\sigma_\beta^2$ . This excess is more than would arise from the error associated with the estimated parameters. The excess may be calculated by the difference between the variance of the estimates and the average error variance  $\bar{\sigma}_\beta^2$  of these estimates, shown as

$$\hat{\sigma}_\beta^2 = \hat{\sigma}_\beta^2 - \bar{\sigma}_\beta^2 \quad (38)$$

As Table 3 indicates, this difference is greater than the variance of the generating parameters, and the difference is shown as  $\hat{\sigma}_\beta^2 - \sigma_\beta^2$ . Thus, the inconsistency of estimates is manifested again.

Second, the correlation  $r(\hat{\beta}, \beta)$  between the generating parameters and the estimated parameters is no less than .96 to two decimal places, indicating again that the effect is primarily one of scale. Especially important are the correlations between the estimated parameters and those obtained as the mean of the items endorsed, the method used by Thurstone. In each case the correlation  $r(\hat{\beta}, \beta_r)$  is no less than .99 correct to two decimal places. The last of these correlations indicates that the model is in agreement with that entertained by Thurstone for the measurement of attitude when the scale values of the items were presumed known, having been obtained by an independent scale procedure.

Table 3  
Generated and Estimated Person Parameters in  
Five Simulation Studies

Statistic	Study				
	1	2	3	4	5
$r(\hat{\beta}, \beta)$	.96	.97	.96	.97	.97
$r(\hat{\beta}, \beta_T)$	1.00	1.00	.99	1.00	1.00
$\bar{\beta}$	-.002	.112	-.046	-.048	-.037
$\bar{\hat{\beta}}$	.038	.096	.010	-.064	.004
$\sigma_{\beta}^2$	3.596	4.384	4.108	4.043	3.868
$\hat{\sigma}_{\beta}^2$	4.366	5.224	4.903	4.548	4.882
$\hat{\sigma}_{\beta n}^2$	.437	.459	.454	.446	.449
$\hat{\sigma}_{\beta}^2 - \hat{\sigma}_{\beta}^2 - \hat{\sigma}_{\beta n}^2$	3.929	4.765	4.449	4.102	4.433
$\hat{\sigma}_{\beta}^2 - \sigma_{\beta}^2$	.333	.381	.341	.059	.565

### Frequency of Endorsed Responses

One of the consequences of the maximum probability of an endorsed response to any stimulus being .5 is that there are relatively few endorsed responses. To illustrate this feature, Table 4 displays three entire response patterns associated with each total score. The patterns were selected from one simulation with three different locations for each total score: one with a clear negative location, one with a clear positive location, and one in between. The maximum total score from 20 stimuli is only 6. Whether real data conform to the model with respect to this feature is, of course, an empirical question; however, it does provide a specific perspective from which to examine real sets of responses anticipated to be of the unfolding form.

### An Example With Real Data

The real example concerns a scale designed to measure attitudes toward capital punishment, using the methods of Thurstone (Wohlwill, 1963). Thus a judgment group was employed to decide the relative intensities or affective values of the statements, and a model based on the cumulative principle was employed to transform these judgments into scale values. According to Thurstone's methodology, persons whose attitude is to be measured would respond directly by agreeing or disagreeing with these statements, and their attitude value would be taken as the mean or median of the scale values of the statements given a positive response.

Wohlwill (1963) scaled the statements immediately from the direct Agree or Disagree responses using a method proposed by Mostellar (1949). In this method, there was a restriction to these responses: The respondents were asked to agree with exactly three statements. Thus two sets of scale estimates for the statements were already available.

In the scaling described by Wohlwill, nine statements were employed. However, only eight statements were used for the present scaling because one of the statements was considered not to conform to the scale as well as the others. The eight statements are shown in Table 5. For convenience, the statements

Table 4  
Three Response Patterns Associated With Each Observed Total Score Greater Than 0

Case	Total Score	Generating Parameter	Thurstone Estimate	MLE	Standard Error	Response Pattern
1	1	-2.88	-3.44	-4.23	.91	0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
2	1	1.57	1.36	1.43	.60	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0
3	1	4.58	3.49	4.32	.79	0 0
4	2	-2.58	-2.56	-2.77	.69	0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
5	2	.11	.62	.61	.61	0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0
6	2	2.92	2.28	2.26	.63	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0
7	3	-2.74	-2.37	-2.49	.64	1 0 0 1 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
8	3	.42	.52	.72	.60	0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 1 0 0 0 0 0 0
9	3	2.869	3.10	3.36	.90	0 1 0 1
10	4	-1.85	-2.05	-2.12	.63	0 0 1 1 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
11	4	.26	.50	.49	.60	0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 1 0 1 0 0 0 0 0 0
12	4	2.77	2.73	2.82	.77	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 1
13	5	-.88	-1.03	-1.01	.60	0 0 1 0 0 0 0 1 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0
14	5	-.02	-.04	.04	.59	0 0 0 0 0 0 0 1 1 0 0 1 1 0 1 0 1 0 0 0 0 0 0 0
15	5	.83	1.17	1.19	.62	0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 1 1 1 0 0 0 0 0
16	6	-1.27	-.96	-.93	.59	0 1 0 0 0 0 0 1 1 1 1 0 0 1 0 0 1 0 0 0 0 0 0 0
17	6	.46	.97	1.00	.62	0 0 0 0 0 0 0 0 0 0 1 0 0 1 1 1 1 1 1 0 0 0 0
18	6	2.35	2.40	2.39	.67	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 1

are presented in the order of their scale values, although this was not the order in which the statements were presented to the persons responding.

The persons who responded to the statements for the analysis according to the SSLM were 54 graduate students taking an introductory course in educational measurement and statistics. They responded simply by agreeing or disagreeing with each statement, with no restrictions placed on how many statements should receive an Agree response.

Table 5 shows the scale values of the statements obtained using the methods of Thurstone and quoted by Wohlwill, the method used by Wohlwill, and the method described above using the SSLM. To facilitate interpretation of the relationships between the scale values, the last set of estimates has been transformed linearly to have the same mean and standard deviation as the first set. According to the methods and models of Thurstone, the relative scale values of any pair of statements should not be affected by the presence of other statements in any data collection and scaling exercise. Therefore no artificial effect should be present on the scale values because of the use of eight statements rather than nine. Table 5 also shows the estimates of the standard errors obtained from the SSLM analysis and the 68% confidence limits based on one standard error either side of the estimate. It is clear from Table 5 that the results from the three analyses are essentially equivalent, with statements against capital punishment at the top of the table and statements supporting it at the bottom.

Figure 2 shows the plot of the scale values obtained from Thurstone's method and the SSLM. This plot confirms the equivalence of the two sets of estimates. The most impressive feature of the above results is that the equivalence exists even though the data were obtained in two different countries (Australia and the United States) a half century apart in time, according to different response formats, and analyzed by correspondingly different methods. This shows, therefore, not only that the SSLM is an appropriate model for these unfolding direct-response data, but also that there is a substantial degree of invariance for the scale across relevant differences in context.

Table 5  
Scale Values of Statements About Capital Punishment  
(Statements are Ordered According to the Scale Values Obtained  
by Thurstone's Method)

Statement	Thurstone's Method	Wohlwill	SSLM SE [68% limits]
1. Capital punishment is one of the most hideous practices of our time.	.6	.6	1.7 1.3 [.4, 3.0]
2. The state cannot teach the sacredness of human life by destroying it.	2.0	1.6	1.9 1.2 [.7, 3.1]
3. Capital punishment is not an effective deterrent to crime.	2.8	1.9	1.9 1.2 [.7, 3.1]
4. I don't believe in capital punishment but I am not sure it isn't necessary.	5.4	3.6	4.1 1.1 [3.0, 5.2]
5. I think capital punishment is necessary but I wish it were not.	6.2	6.6	7.4 1.2 [6.2, 8.6]
6. Until we find a more civilized way to prevent crime we must have capital punishment.	7.2	7.4	7.9 1.3 [6.6, 9.2]
7. Capital punishment is justified because it does act as a deterrent to crime.	8.4	8.5	8.7 1.3 [6.4, 10.0]
8. Capital punishment gives the criminal what he deserves.	9.4	9.4	8.4 1.3 [7.1, 9.7]

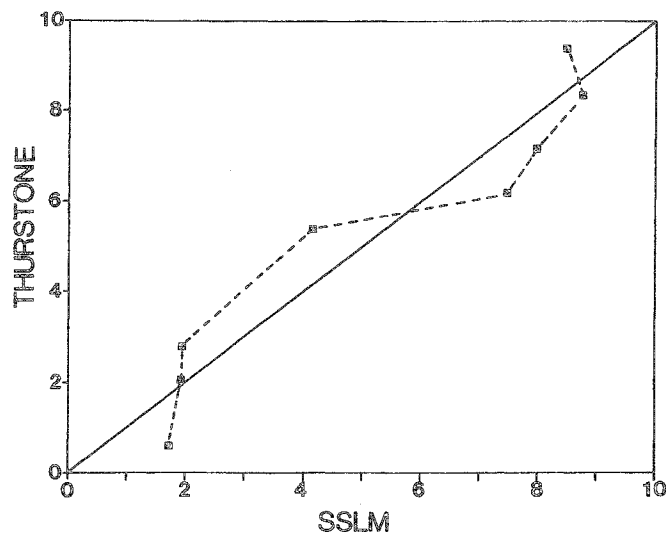
It is also interesting to see the distribution of attitudes in the sample of 54 persons who responded to the statements. This distribution is shown in Table 6, in which the statement order is the same as in Table 5 (i.e., in the order of their affective values).

First, it is clear that a whole range of attitudes is shown in the sample, from very much against to very much in favor of capital punishment. Second, the patterns are essentially as expected when persons and statements are ordered according to their scale values. Thus, Disagree responses are evident in the top right and the bottom left parts of the response matrix.

Some small anomalies are also evident. For example, it might be expected that the first response pattern in Table 6 (01100000) would be associated with a greater attitude value than the second response pattern (11100000). That it is not is attributed at this stage to the level of precision chosen for terminating the estimation process. This level was set at .01; that is, when the improvement to the *average* discrepancy between the successive scale estimates of the statements was less than .01, then the process was terminated. As a result, some individual estimates may not have reached the same level of convergence. Another kind of anomaly is the observed response pattern (10001111) which was produced by two persons. This anomaly is diagnosed as a response pattern which does not fit the model, though no formal fit statistic has been calculated. The pattern may involve a coding error by the persons to the first statement; perhaps negatives in the statement confused them into coding Agree rather than Disagree when responding.

Overall, however, the results applying the SSLM seem quite satisfactory. As application of the SSLM to direct response unfolding data is much more straightforward than any of the other methods proposed, the results suggest that the model should be tried routinely as a first step in analyzing such data.

Figure 2  
 Scale Values From Thurstone's Method and the SSLM



Discussion

One possible limitation of the model as presented is that the maximum probability of endorsement, when the person and stimulus have identical scale values, is .5. It seems reasonable to suggest that the maximum probability may be greater in some circumstances, even though the answer to this question is an empirical one in any given situation. However, most formulations for unfolding consider that the response depends on the distance, or square of the distance, between the person and stimulus locations, and the development above may add insight into assumptions behind the unfolding response mechanism.

The dominant approach to the analysis of attitude data, which includes the extremely common approach based on Likert scaling (Likert, 1932), includes cumulative models. However, the common data collection design for attitude measurement, in which persons are asked simply to agree or disagree with statements (with or without degrees of intensity), does not imply direction as required by these cumulative models. Thus, if a person disagrees with a statement, there is no evidence as to which direction the person perceives himself or herself to be, relative to the location of the statement. Therefore, the usual data collection design is entirely consistent with an unfolding response mechanism, and not the cumulative one.

Because the solution algorithm to the model is applicable to large numbers of stimuli, as are found in achievement testing and attitude measurement, the model can be applied to real datasets and the results compared to an analysis based on models subscribing to the cumulative principle. In particular, data which are analyzed according to the simple logistic model of Rasch (1960/1980) may be readily analyzed according to the model presented here. Because the two models are of identical form, with the probability of endorsement of one depending on the directed distance between the person and the stimulus locations and of the other on the square of this distance, a determination of which principle seems to be operating may then be performed by comparing the two analyses and checking which model the data conformed to more closely. Of course, a researcher should have a prior rationale for selecting the particular model, but in some cases, such as that of Leik and Matthews mentioned earlier, the other mechanism may conform when the first chosen does not.



Table 6  
Distribution of Attitude Scale Values

Attitude Estimate	Standard Error	Total Score	Response Pattern	Frequency
-1.37	.89	2	01100000	4
-1.17	.89	3	11100000	10
-.92	.82	3	01110000	3
-.90	.82	4	11110000	8
-.48	.71	4	01111000	1
-.40	.71	4	01110010	1
-.36	.70	6	11111100	2
-.21	.70	5	10111100	1
-.15	.70	4	01101010	1
.00	.71	6	01111110	1
.22	.73	5	10011011	1
.23	.72	5	00111101	2
.31	.74	6	01011111	1
.35	.74	4	01001110	2
.37	.75	4	00101101	2
.38	.74	4	01001101	1
.49	.76	5	10001111	2
.64	.79	5	00011111	3
.64	.80	3	00010011	1
.95	.88	4	00001111	5
1.09	.91	2	00001100	1
1.11	.91	3	00000111	1

Another point may be made by the observation of the information associated with the unfolding and cumulative response mechanisms and the way it clarifies features which distinguish the two mechanisms. This clarification may explain the element of surprise that newcomers to person-item response theory seem to show when they are confronted with the result that in most response models, the maximum information occurs when the person and item parameters are identical. Analogies with physical measurement, such as the measurement of height, are usually invoked in order to explain this result.

A common analogy is to suppose that each of a class of persons is asked to indicate whether he or she is taller than each of some set of objects. However, for a cumulative model, which provides maximum information when the person and object are of identical height, it is important that the response required in the analogy is indeed in answer to the question "Are you taller than the object?". A question such as "Are you close in height to the object?", which is consistent with the unfolding mechanism, does not provide maximum information when the person and object are of equal height. As it happens, the questions used in attitude measurement in particular are generally of the latter type. Thus the element of surprise may arise in part because the response mechanism supposed intuitively is the unfolding and not the cumulative one.

Finally, the model was applied to responses to a set of statements designed to measure an attitude toward capital punishment and originally scaled by one of the methods of Thurstone. These responses were made by 54 students in 1984 and the scale values of the statements agreed very closely with those obtained by the earlier method. The viability of the model, therefore, is confirmed by this real example, and perhaps many other attitudes could be studied using the SSLM model.

## References

- Andersen, E. B. (1973). Conditional inference for multiple choice questionnaires. *British Journal of Mathematical and Statistical Psychology*, 26, 31–44.
- Andersen, E. B. (1977). Sufficient statistics and latent trait models. *Psychometrika*, 42, 69–81.
- Andrich, D. (1978a). Relationships between the Thurstone and Rasch approaches to item scaling. *Applied Psychological Measurement*, 2, 451–462.
- Andrich, D. (1978b). A binomial latent trait model for the study of Likert-style attitude questionnaires. *British Journal of Mathematical and Statistical Psychology*, 31, 84–98.
- Baker, F. B. (1977). Advances in item analysis. *Review of Educational Research*, 47, 151–178.
- Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, 37, 29–51.
- Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, 46, 443–459.
- Bock, R. D., & Jones, L. V. (1968). *The measurement and prediction of judgment and choice*. San Francisco: Holden Day.
- Bock, R. D., & Wood, R. (1971). Test theory. *Annual Review of Psychology*, 22, 193–224.
- Coombs, C. H. (1950). Psychological scaling without a unit of measurement. *Psychological Review*, 57, 145–158.
- Coombs, C. H. (1964). *A theory of data*. New York: Wiley.
- Coombs, C. H., & Avrunin, G. S. (1977). Single-peaked functions and the theory of preference. *Psychological Review*, 84, 216–230.
- Coombs, C. H., & Smith, J. E. K. (1973). On the detection of structure in attitudes and developmental processes. *Psychological Review*, 80, 337–351.
- David, H. A. (1963). *The method of paired comparisons*. New York: Hafner.
- Davidson, R. R., & Farquhar, P. H. (1976). A bibliography on the method of paired comparisons. *Biometrics*, 32, 241–252.
- Davison, M. L. (1977). On a metric, unidimensional unfolding model for attitudinal and developmental data. *Psychometrika*, 42, 523–548.
- Davison, M. L., Robbins, S., & Swanson, D. B. (1978). Stage structure in objective moral judgment. *Developmental Psychology*, 14, 137–146.
- Douglas, G. A. (1982). Conditional inference in a generic Rasch model. In D. Spearritt (Ed.), *The improvement of measurement in education and psychology* (pp. 129–157). Melbourne: Australian Council for Educational Research.
- Fischer, G. H. (1973). The linear logistic test model as an instrument in educational research. *Acta Psychologica*, 37, 359–374.
- Gulliksen, H. (1950). *Theory of mental tests*. New York: Wiley.
- Guttman, L. (1950). The basis for scalogram analysis. In S. A. Stouffer, L. Guttman, E. A. Suchman, P. F. Lazarsfeld, S. A. Star, & J. A. Clausen, *Measurement and prediction (Studies in social psychology in World War II, Vol. 4)*. Princeton NJ: Princeton University Press.
- Guttman, L. (1954). The principal components of scalable attitudes. In P. F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences*. Glencoe IL: The Free Press.
- Jensema, C. J. (1974). An application of latent trait mental test theory. *British Journal of Mathematical and Statistical Psychology*, 27, 29–48.
- Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In B. A. Goslin (Ed.), *Handbook of socialization theory and research*. San Francisco: Rand-McNally.
- Leik, R. K., & Matthews, M. (1968). A scale for developmental processes. *American Sociological Review*, 33, 62–75.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, No. 140.
- Lord, F. M. (1968). An analysis of the Verbal Scholastic Aptitude Test using Birnbaum's three parameter logistic model. *Educational and Psychological Measurement*, 28, 989–1020.
- Lumsden, J. (1976). Test theory. *Annual Review of Psychology*, 27, 251–280.
- Lumsden, J. (1977). Person reliability. *Applied Psychological Measurement*, 1, 477–482.
- Mostellar, F. (1949). *A theory of scalogram analysis using non-cumulative types of items* (Report No. 9). Boston: Harvard University, Laboratory of Social Relations.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research. [Expanded edition, University of Chicago Press, 1980.]
- Rasch, G. (1961). On general laws and the meaning of measurement in psychology. In J. Neyman (Ed.), *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, IV*. Berkeley CA: University of California Press.
- Rasch, G. (1968). *A mathematical theory of objectivity and its consequence for model construction*. Paper presented at the European Meeting on Statistics, Econometrics and Management Science, Amsterdam.
- Sixtl, F. (1973). Probabilistic unfolding. *Psychometrika*, 38, 235–248.

- Thissen, D. (1982). Marginal maximum likelihood estimation for the one-parameter logistic model. *Psychometrika*, 47, 175–186.
- Thurstone, L. L. (1927a). A law of comparative judgment. *Psychological Review*, 34, 278–286.
- Thurstone, L. L. (1927b). The method of paired comparisons for social values. *Journal of Abnormal and Social Psychology*, 21, 384–400.
- Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, 33, 529–554.
- Thurstone, L. L. (1931). Rank order as a psychophysical method. *Journal of Experimental Psychology*, 14, 182–201.
- Traub, R. E., & Wolfe, R. G. (1981). Latent trait theories and the assessment of educational achievement. *Review of Research in Education*, 4, 377–435.
- Wohlwill, J. F. (1963). The measurement of scalability for non-cumulative items. *Educational and Psychological Measurement*, 23, 543–555.
- Wright, B. D., & Panchapakesan, N. (1969). A procedure for sample-free item analysis. *Educational and Psychological Measurement*, 29, 23–48.
- Wright, B. D., & Douglas, G. A. (1977). Conditional versus unconditional procedures for sample-free item analysis. *Educational and Psychological Measurement*, 37, 47–60.
- Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago: Mesa Press.
- Zinnes, J. L., & Griggs, R. A. (1974). Probabilistic, multidimensional unfolding analysis. *Psychometrika*, 39, 327–350.

#### Acknowledgments

Alan Lyne wrote the computer program for the simulation and data analyses.

#### Author's Address

Send requests for reprints or further information to David Andrich, School of Education, Murdoch University, Murdoch, Western Australia 6150, Australia.