

Modelling Attitude Measurement: The Unfolding Mechanism and its Relationship with the Cumulative Mechanism

Guanzhong Luo

Research Institute of Education
South China Normal University

Abstract

It is not uncommon that the models designed for the measurement of attainment, in which the response mechanism is cumulative, are used in the measurement of attitude. However, there is a more appropriate mechanism called unfolding for the measurement of attitude. This paper presents a model for attitude measurement which relates the unfolding mechanism to the cumulative mechanism. In relating them, each of the mechanisms is clarified. An example of measuring the attitude towards capital punishment is presented.

Acknowledgment: This project was supported in part by the Australian Development Cooperation Scholarship Scheme (ADCSS) and the Australian Research Council (ARC), and was carried out while the author was on study leave at the School of Education, Murdoch University, Western Australia.

Modelling Attitude Measurement: The Unfolding Mechanism and its Relationship with the Cumulative Mechanism

Guanzhong Luo

Research Institute of Education
South China Normal University

Introduction

In his *The measurement of Values*, Thurstone (1959) classifies his

works into two chapters: Subjective Measurement and Attitude measurement. Under the title Subjective Measurement, he deals with psychological measurement mainly under the basic Assumption that satisfaction increases with increase in the amount of the commodity possessed by an individual (Thurstone, 1959; p. 124). This assumption in fact applies the cumulative mechanism. In general modelling theory, items and persons can be located on the same continuum. And if the cumulative mechanism applies, then the probability that a person gives a positive response to an item increases as the location of the person relative to the location of the item gets greater, and if the persons and items are ordered with their location, the data matrix will have basically a triangle structure.

On the other hand, under the title Attitude Measurement, Thurstone devotes the main effort to the scaling of attitude

variables. He pointed that "A person who usually talks in favor of preparedness, for example, would be represented somewhere to the right of the neutral zone. A person who is more interested in disarmament would be represented somewhere to the left of the neutral zone." (Thurstone, 1959; p. 220)

Thurstone realised the difference between the attitude measurement and the other types of measurement he dealt with under the title of subjective measurement. He described the type of the scales he used in attitude measurement as the maximum probability type (Thurstone, 1959; p. 214), which in fact revealed the embryo of the single-peaked (Unfolding) mechanism for attitude measurement which is developed explicitly by his student, C. H. Coombs.

Coombs and Avrunin (1977) clarify the occurrence of Single-peaked functions in attitude measurement, using the example of vacation time taking, they show that when the vacation time increases, though the 'utility function' for both good and bad are monotonous, the total preference function, which is the combination of these two utility functions, is single peaked. With further investigation, they come to the conclusion that single-peakedness is inevitable if there is only one component (dimension).

As Coombs and Avrunin (1977) pointed out, single-peaked functions are the foundation underlying unfolding theory, which was introduced even earlier than when the significance of single-peaked functions was realised. Coombs (1950) writes the following to explain the principle of the unfolding theory:

Let us assume now that we have asked each of a group of

individuals to place a set of stimuli in rank order with respect to the relative degree to which he would prefer to endorse them. Our understanding of the results that would follow will be clearer if we build a mechanical model which has the appropriate properties. This is very simply done by imagining a hinge located on the J scale at the C_i value of the individual and folding the left side of the J scale over and merging it with the right side. The stimuli on the two sides of the individual will mesh in such a way that quantity $|C_i - Q_i|$ will be in progressively ascending magnitude from left to right. The order of the stimuli on the folded J scale is the I scale for the individual whose C_i value coincides with the hinge.

Coombs(1964) also identifies the ideal structure of the data governed by this mechanism: if the persons and items are ordered by their locations, the data matrix will have basically the structure of parallelogram. Then the task of measurement is to unfold the individual scale, called the I scale, so that the J scale, the joint continuum on which stimuli (items or statements) and individuals (persons) are located simultaneously, is recovered.

Now we can see that there are two classes of modelling theories : one based on the cumulative mechanism and the other based on the single-peaked mechanism. During the last three decades, modelling theory on the cumulative mechanism has been well developed and great progress has been made on its application. Among the models confirming the cumulative mechanism, the Rasch model, which was introduced by Rasch(1960) and developed further by his followers, has shown its significance in both theoretical and practical aspects. The Rasch model explicitly specifies that the probability function for a positive response to be

where b is the location of the person (ability) and d is the location of the item (difficulty).

As the fundamental measurement in physics, the measurement underlined by the Rasch model have measurement unit and origin. From the viewpoint of statistics, the Rasch model belongs to the exponential family of distributions and provides sufficient statistics for parameters (person ability and item difficulty). Therefore, this model is widely used in the measurement of types of educational and psychological variables. In fact, the Rasch model is the kernel of the cumulative scaling mechanism. Any other cumulative model can be understood as a generalisation, modification or approximation of the Rasch model in some sense. Their favourable properties are just the reflection of those implied in the Rasch model. In this sense, we say that the Rasch

model is representative of the cumulative class of scaling models and the framework underlying the cumulative scaling models is based on the Rasch model.

Compared with the modelling theory on cumulative mechanism, the unfolding theory has not achieved as much progress, and its real application is limited. From the theoretical point of view, this mechanism is still poorly understood; from the practical point of view, feasible programs according to the unfolding theory are not available. In fact, it is not uncommon that the models designed for the measurement of attainment, in which the response mechanism is cumulative, are used in the measurement of attitude. It is not surprising this kind of misuse often leads to some confusing or misleading results.

To clarify the unfolding mechanism clearly and provide practitioners with efficient procedures for the measurement of attitude, this paper presents a model called the Hyperbolic Cosine Model (HCM). Instead of presenting a descriptive model, we derive this model from the well known Rasch model for ordered response categories). In this way, the relationship between these two types of response mechanisms introduced above is revealed. Some other recent results on this model are also briefly reported. At the end of this paper, an example of measuring the attitude towards capital punishment is presented.

Folded data

Many attributes have two extreme directions. If the attribute of an object goes to infinity at either direction, it may tend to be judged "bad" by any person. In general, after some quantification, such an attribute can be treated as a variable with one or more dimensions. Hereafter we focus on the situation with one dimension. Then for every person, there is an ideal value of the attribute that he/she would mostly ask for if given the choice. It is called the ideal point of the person. The ideal point may be different from person to person. We are then able to locate persons on the continuum according to their ideal point. Imagining that we could ask a group of people to judge whether the attribute value of a object is

0 --- below his/her ideal point;

1 ---- close enough to or exactly at his/her ideal point; or

2 ---- over his/her ideal point.

then the possible response outcome are (0, 1, 2). We could make up a questionnaire about a set of objects in this way and get a response matrix in which every cell was 0, 1 or 2. The Generalised Rasch Model (GRM) could be used then without difficulty to analyse this matrix to estimate the attribute value of objects and the location of every person. When number of

categories $m+1=3$, according to GRM, the probabilistic function is

Usually, attitude measurement does not perform in this way. The most common method in attitude measurement is to ask the respondent whether he/she agrees or disagrees to the statement (item) presented, and the possible outcomes can be numerically recorded as 0 (disagree) or 1 (agree). It is not surprising that if the Rasch model is used to analyse a data set collected with this method, the results will be confusing and misleading. The Real example given at the end of this paper will show this matter.

From a psychological point of view, it is interesting to explore whether a respondent will judge his/her relative location to the statement(item) before making a response. It is believed that a respondent will do that consciously or subconsciously. If so, consciously or subconsciously again, the respondent has to fold down the judgement to make the required response. Response outcome 2 is illegal in this situation and must be folded down to 0:

1--- agree ----- the location of the item is close enough to or exactly at his/her ideal point (location of person) ;

0 --- disagree -- otherwise , that is, the location of the item is below or over his/her ideal point.

As the result of this “folding while responding”, we get “folded data” to be analysed. The purpose of the analysis on the folded data is to

- 1) “unfold” the data set as the reverse action of “folding while responding”, and
- 2) estimate the location of statements and persons.

As we will see, the model and technique developed in this paper are the realisation of this general thinking.

Hyperbolic Cosine model

We begin with the GRM (1). As shown in Fig 1, while the probability for $(x_{ni} = 2)$ is cumulative, the probability for $(x_{ni} = 1)$ is single-peaked: when the person location increases, this probability first climbs up to its maximum value, and then decreases. Folding the probability of $(x_{ni} = 0)$ and $(x_{ni} = 2)$, we have

Here we invoke hyperbolic cosine function

which is a symmetric function with minimum value $\cosh(0) = 1$. It is increasing monotonically when $x > 0$.

DEFINITION 1. If the response outcome $\{ x_{ni} , n=1, \dots , N; i=1, \dots , I \}$ obeys the following probabilistic distribution (2)

where

b_n ----- person parameter;
 d_i , q_i --- item parameters;
 N ----- number of persons;
 I ----- number of items (statements).

then we say $\{ x_{ni} \}$ obeys the Hyperbolic Cosine Model (HCM). The probabilistic function of the HCM can be expressed uniformly

as the following.

(3)

The Structure of the HCM

After deriving the HCM by folding the response 0 and 2 within the GRM to 0, what we are eager to see is whether the HCM actually implies the unfolding mechanism. In fact, the probability $P(x=1)$, which has a single-peaked shape within the GRM, is intact during "folding while responding". The probability $P(x=0)$ and $P(x=2)$ in the GRM has been summed, forming an inverted single-peaked curve in Fig 2. In HCM, therefore, the probability $P(x=1)$ has the maximum value when $b-d=0$ and the probability decrease as the absolute distance between b and d gets greater. In this sense, we conclude that HCM belongs to the class of unfolding models. The HCM is distinguished from any other type of unfolding model by the fact that the HCM is the inevitable logical result of "folding while responding" within the GRM. No ad hoc or artificial "design" is needed. We conceivably expect that the HCM has the potency to reveal the natural structure of unfolding models. And furthermore, now that its "mother model" -- Rasch Model (the SRM and the GRM) has many excellent properties from both the statistical and measuremental point of view, we also expect that the HCM has some advantages over other unfolding models. After "folding while responding", nevertheless, the mechanism of the Rasch model is corrupted. Its

properties certainly are destroyed or damaged to some extent. The problem arises: what is retained? First of all, and most unfortunately, is the corruption of sufficiency. The advantage of the Rasch model over other cumulative models is greatly based on the existence of sufficient statistics. In practice, the great utility of sufficient statistics is brought into full play in reducing raw data without any loss of information about parameters to be estimated. Once sufficient statistics have been found, all the feasible estimation methods would work on the sufficient statistics rather than the raw data.

To observe the corruption of sufficiency within the HCM, we consider the likelihood function

By the factorisation theorem, no sufficient statistics for $\{(b_n), (d_i), (q_i)\}$ are found (except raw data $\{x_{ni}\}$). If we treat $\{(b_n), (d_i)\}$ as known coefficients, however, the statistic (s_i) , which is independent of $\{(b_n), (d_i)\}$, is sufficient for (q_i) .

Due to the corruption of sufficiency, another property of the Rasch model -- objectivity, which is the corollary of sufficiency, is also invalid. Of the four basic properties of the Rasch model, therefore, only one, unidimensionality, is retained within HCM. We should point out, however, that in addition to item location d , which is of one dimension, q has a great effect on response behaviour.

The synthesis of Unfolding models

Stressing the importance of the single-peaked function for unfolding theory, Coombs (1977) wrote:

'Single-peaked functions are the foundation underlying unfolding theory which provides an algorithm for the measurement of the variable(s) underlying preferential choice. ...There are, of course, an infinite variety of theoretical systems that can be constructed to yield single-peaked preference functions. One might assume that the elemental component utility functions are

single or doubly inflected, for example. Then an appropriate combination rule for a preference function and suitable conditions on the options could be constructed that would ensure single-peakedness of the preference function.'

Here Coombs predicts the direction of constructing the general form of unfolding models. Using Coombs' words, this general form should "provide general and plausible conditions that can account for the ambiguity of this phenomenon -- the single peakedness of preference and hedonic tone".

According to the single-peakedness principle of unfolding models, the general form of the probabilistic function for unfolding models is

or in a combined form

For detailed derivation, see Luo(1993). Furthermore, any of the unfolding probabilistic functions can be taken as a special case of (2.6.16). For example:

I) In (6), if

then we obtain HCM (2) again.

II) if we let

$a_1=1$, $a_k = 0$ for all $k > 1$; and

then this leads to the Simple Square Logistic Model (Andrich, 1988).

(See Luo(1993) for detail and further discussion).

In the derivation of the general form of unfolding models, it is an essential condition that $f(0)$ is finite. Otherwise the Taylor's series of the function $f(x)$ at $x_0 = 0$ is not available and the deduction above is no longer valid. This assumption is equivalent to requiring that the probabilities (both for positive and negative responses) are always smaller than 1, even when $b_n - d_i = 0$. It means that random effect always exist. Recall the situation within the Rasch model: no matter how high an ability a person has, the probability of a positive response is always smaller than 1. In this sense we say this assumption is reasonable for both cumulative and unfolding models. It is noted here that though the probability when $b_n - d_i = 0$ is always smaller than 1, if we let item parameters other than location, that is, the item unit, vary, this value can approximate to 1 to any given extent.

The coefficients for the serial expansion of $\log \cosh(b_n - d_i)$ are yet to be specified. Using the Taylor's expansion method, we have

Comparing the HCM with other unfolding models, we see that the probabilistic functions of them all involve a quadratic function of $(b_n - d_i)$. In fact, other unfolding models can be considered as different approximate forms of the HCM(ignoring a constant multiplier of exponential power). In this sense, we say that the HCM catches the essence of the unfolding mechanism and is the representative of unfolding models.

Now we see the relationship of the two main mechanisms in the

modelling theory for measurement. The dichotomous models for attainment measurement employ cumulative mechanism. When these models are generalised for the multi-category situation, the single-peaked curves inevitably appear for the middle category, though the response mechanism is still cumulative. Then the models which employ unfolding mechanism can be brought out by

folding the categories pairwise and symmetrically about the middle category.

Parameter Estimation: Example

Once a model is established, from the viewpoint of practitioners, the most important issue is to find a feasible and efficient procedure for estimating the parameters of the model. Though a variety of approaches can be used to derive such a procedure, the maximum likelihood principle generally leads to simple and elegant solution equations and therefore is widely applied. The estimation obtained by the maximum likelihood principle is called the Maximum Likelihood Estimation (MLE). To maximise the likelihood function (7), partial differentiation are taken with respect to each parameter respectively on the log L, setting the differential results equal to 0, we get the maximum log-likelihood equations:

By the Newton-Raphson method, equation (17) can be solved iteratively with a necessary constraint

which deletes the indeterminacy in (17). For more detailed discussion see Andrich and Luo (1991,1993). The computation program is available from the authors. Other alternative procedures for the item parameter estimation are also developed (Luo, 1993).

Real example: Attitude toward capital punishment. Andrich(1988) lists the following statement on capital punishment:

1. Capital punishment is one of the most hideous practice of our time.
2. the state cannot teach the sacredness of human life by destroy it.
3. Capital punishment is not an effective deterrent to crime.
4. I don't believe in capital punishment. but I am not sure it isn't necessary.
5. I think capital punishment is necessary but I wish it were not.
6. Until we find a more civilised way to prevent crime we must have capital punishment.
7. Capital punishment is justified because it does act as a

deterrent to crime.

8. Capital punishment gives the criminal what he deserves.

The response patterns and their observed frequencies (Andrich, 1988) are shown in the following:

Table 1. The response patterns and their
observed frequencies

Frequencies	Response Pattern	Observed
4	01100000	
10	11100000	
3	01110000	
8	11110000	
1	01111000	
1	01110010	
2	11111100	
1	10111100	
1	01101010	
1	01111110	
1	10011011	
1	00111101	
2	01011111	
1	01001110	
2	00101101	
2	01001101	
1	10001111	
2	00011111	
3		

1 00010011
 5 00001111
 1 00001100
 1 00000111

The following is the result of analysis using HCM , compared with that using the Rasch model. It is common practice when using the Rasch model that the responses on “negative” statements should be reversed. In this example, item 1 to 4 are reversed in using the Rasch model.

Table 2(a) Item Parameter

Estimation

Item unit qi by HCM reversed)	Statement	Item location di	
	Number	HCM	Rasch model
		(item 1-4 reversed)	
1	-4.973	-0.894	5.020
2	-6.229	0.515	8.168
3	-6.099	0.770	8.279
4	-1.155	-0.529	4.051
5	3.582	-0.599	4.056
6	4.525	-0.282	4.076
7	5.171	0.477	3.378
8	5.178	0.543	3.382

Table 2(b) Personal Parameter

Estimation

Personal location by Rasch model reversed)	Response
	Pattern
	(item 1-4 reversed)
	HCM

-1.115	01100000	-11.810
-0.478	11100000	-7.597
-0.478	01110000	-1.041
0.079	11110000	-3.324
0.079	01111000	-0.044
0.079	01110010	-0.043
0.079	11111100	-0.045
1.259	10111100	0.704
0.632	01101010	1.358
0.078	01111110	1.364
1.259	10011011	2.013
0.632	00111101	2.007
0.632	01011111	2.702
1.259	01001110	2.712
0.079	00101101	2.712
0.079	01001101	2.712
0.079	10001111	3.618
0.632	00011111	3.618
0.632	00010011	7.234
-0.478	00001111	5.387
0.079	00001100	8.347
-1.115	00000111	7.245
-0.478		

As shown in Table 2, the estimated parameter values of both items and persons are quite different when different models are used. Furthermore, when the persons and items are ordered in accordance with the locations estimated by the HCM, as shown in

Table 2, the data matrix is basically a parallelogram(with some random error). But if we order the persons and items in accordance with the locations estimated by the Rasch model, the data matrix has no clear triangle structure. This reveals that the response patterns have a single-peaked structure and the Rasch model is misused if we use it in this situation. The example shows then that it is crucial to choose the appropriate model to process the data. The question then arises: how to choose the appropriate model for a given set of data? In most situations, the data structure can be predicted by carefully

analysing the items and statements comprising the questionnaire. For detailed discussion, see Luo(1993).

Reference

Andrich, D.(1982) An extension of the Rasch model for ratings providing both location and dispersion parameters. *Psychometrika*, Vol. 47, 105-113.

Andrich, D. (1988) the application of an unfolding model of the PIRT type to the measurement of attitude. *Applied Psychological Measurement*. Vol. 12, 33-51.

Andrich, D. and Luo, G.(1991) A latent trait model for unfolding based on a Rasch model for ordered response categories. Paper presented at the International Psychological Measurement Symposium, Nanjing, P. R. China, December.

Andrich, D. and Luo, G.(1992) Hyperbolic Cosine Model for unfolding direct responses. *Applied Psychological Measurement* (in press).

Luo, G.(1993) Hyperbolic Cosine Models: Implications and Applications. Ph. D thesis(in completion), Murdoch University.

Coombs, C. H.(1950) Psychological scaling without a unit of measurement. *Psychological Review*, 1950, 57, 145-158.

Coombs, C. H.(1964) A theory of data.. New York: Wiley.

Desarbo, W. S.(1986) Simple and weighted unfolding threshold models for the spatial representation of binary choice data. *Applied Psychological Measurement*. 10, 247-264.

Rasch, G.(1960) Probabilistic models for some intelligence and attainment test. Copenhagen: Danish Institute for Educational

Research.