Chapter 15

Intrinsic and Extrinsic Work Values as a Single Unfolding Scale

Wijbrandt H. van Schuur

Department of Sociology, University of Groningen

1. Introduction

Students of the sociology of labor have found it useful to describe work values as either intrinsic or extrinsic. *Intrinsic* values reflect such factors as whether the work is interesting, pleasing, and challenging, and whether the worker can take responsibility for his labor, whereas *extrinsic* values reflect job benefits that are unrelated to the worker's tasks, e.g., whether the job pays well, has a good pension plan, and provides for generous holidays.

Researchers have differed in their view of the relationship between intrinsic and extrinsic work values. According to Lenski (1963), intrinsic values are positive and extrinsic values are negative instantiations of the Weberian notion of the Protestant ethic; for him, then, the two sets of values are bipolar opposites, with people differing in the importance they ascribe to them (cf. Lindseth and Listhaug, 1994). Herzberg (1966), in contrast, considered intrinsic and extrinsic work values to be two unrelated concepts; he termed them „motivators“ and „hygiene factors“, respectively.

Competing theoretical interpretations are sometimes due to differences in the outcomes of alternative methods of data analysis. In this chapter I hypothesize that Lenski's bipolar model is the right one, and that Herzberg's finding of two independent factors is the result of his use of factor analysis on data for which it is ill-suited. If intrinsic and extrinsic work values are bipolar opposites, and subjects seek a balance between the two, it is most appropriate to postulate a (bipolar intrinsic-extrinsic) latent trait on which both subjects and items can be assigned a position. Subjects are assumed to stress the importance of work values with positions close to their own. This measurement model is called the *unfolding model* (Coombs, 1964). The unfolding model is best known in connection with data collected as full rank orders. For dichotomous data, Coombs also used the term 'parallellogram analysis' (cf. Cliff et al., 1988) (see also section 1.7 in chapter 1).

If the unfolding model is the correct one for a particular data set, then the measurement model implied in factor analysis is not: the factor analysis of data that conform to the unfolding model will show an extra, artificial factor (Ross and Cliff, 1964). Such a factor solution is often rotated to simple structure so that the two factors (from data that in fact form a unidimensional unfolding scale) are interpreted as two unrelated concepts (see e.g., Davison 1977, and Van Schuur and Kiers, 1994). If intrinsic and extrinsic work values can be shown to form an unfolding scale, this will explain why Herzberg found two unrelated factors.

To test whether work values form an unfolding scale and to illustrate the unfolding model, we will examine work values data from the World Values Study 1990 (Díez Nicolás and Inglehart 1994). The World Values Study 1990 is a cross-national comparative investigation of values in a variety of domains, including work. It succeeds the European Values Study 1981, which was restricted to Western Europe. This chapter uses the German data as a gesture to our hosts at the Sankelmark conference, but similar results are found using data from the other countries. From this data set all subjects were deleted who mentioned none, only one, or all eight of the items under investigation (i.e., 83, 114, and 104 subjects, respectively). This leaves 1805 subjects.

Subjects were asked the following question: „Here are some aspects of a job that some people have said are important. Please look at them and tell me which ones you personally think are important in a job“. Eight of the original fifteen job values are used in the following analyses; they are listed here, along with the proportion of subjects who found each value important:

A: Not too much pressure (0.27);
B: Generous holidays (0.34);
C: Good hours (0.50);
D: Good job security (0.77);
E: A job that meets one's abilities (0.75);
F: A job where you can achieve something (0.66);
G: A responsible job (0.56); and
H: A useful job for society (0.34).

When a factor analysis is carried out on these data, the first four eigenvalues are 1.88, 1.31, 0.97, and 0.95. The first two of these are substantially larger than the remaining ones, and the first two eigenvectors can be interpreted as capturing extrinsic (items A, B, C, and D) and intrinsic (E, F, G, and H) work values as two independent factors. Let us see whether this interpretation is correct for this data set.

2. Unidimensional unfolding models

Unidimensional unfolding models for dichotomous data are Item-Response-Theory models with single-peaked item characteristic curves (ICCs) (see section 1.7. in chapter 1). In our application, selecting a job value (item) is interpreted as the positive response, and not selecting it is the negative response. Both items and subjects are represented along the latent dimension. The highest probability of a positive response to an item is found among subjects with the same scale value as the item, and the probability of a positive response decreases monotonically with increasing distance between subject and item.

The *deterministic* unidimensional unfolding model for dichotomous data, as developed by Coombs (1964), is applied to a data matrix in which cell (v,i) gets the value 1 when the row element v (generally a subject) exceeds the degree of proximity to the column element i (e.g., an item, a statement, or an object) that is necessary in order to select or „prefer“ it, and the value 0 if this is not the case. The ICCs of the deterministic unidimensional unfolding model are double-step functions: subjects have probability 1 of selecting an item within a closed area on the latent trait around the scale value of the item and probability 0 outside that interval on either side.

Parametric unidimensional unfolding models (e.g., Hoijtink, 1990; Andrich and Luo, 1993 see eq. (25) in chapter 1) have a clear relationship to the Rasch model (Andrich and

Luo even call their program RASCHFOLD). In contrast, the *nonparametric* model to be discussed here - MUDFOLD (for **M**ultiple **U**ni**D**imensional un**FOLD**ing, Van Schuur and Post, 1991) - has a clear relationship to the nonparametric model for cumulative scale analysis developed by Mokken (1971) and extended by Sijtsma, Debets and Molenaar (1990) (see section 1.6 in chapter 1). MUDFOLD has also been generalized to multicategory data (Van Schuur, 1993).

3. The MUDFOLD procedure for finding an unfoldable order of items: Item parameter estimation

The MUDFOLD procedure resembles the Mokken scaling procedure in the way (ordinal) parameter values are estimated and goodness-of-fit is evaluated. Both procedures use Loevinger's criterion of homogeneity to find a maximal subset of scalable items. This criterion compares the amount of violation of the deterministic model (O) in a data set to a standard of statistical independence (E). The criterion, defined as H = 1 – O/E, is applied not only to the scale as a whole, H, but also to each item, H(i), and to each of the smallest possible scales - H(ij) for each pair of items (i, j) in a Mokken scale and H(ijk) for each triple of items (i, j, k) in a MUDFOLD scale. Like Mokken scaling, MUDFOLD uses a „bottom-up“ strategy to successively *add* items to an embryonic scale until the optimal scale is reached; this differs from the standard „top-down“ procedure used in most Rasch analysis procedures, in which the worst-fitting items are successively *deleted*. Once a maximal subset of scalable items has been found, other procedures are used to diagnose whether the data conform well enough to a probabilistic (cumulative or unfolding) model for the scale to be accepted.

The smallest possible unfolding scale is an ordered set of three items. The embryonic unfolding scale that MUDFOLD uses to start the „bottom-up“ selection process is defined as the „best triple“, i.e., the triple with a high enough coefficient of homogeneity and the largest number of informative responses (i.e., the 110, 011, and 111 responses). The other responses (000, 100, 010, 001) are not informative about the structure of an unfolding scale. MUDFOLD allows starting the „bottom-up“ process with any user-defined set of items. In order to find the „best triple“, all n(n–1)(n–2)/6 triples (56, for our 8 items) are considered in each of their three different permutations (each stimulus can be the middle one; reflexions are irrelevant). In the „best triple“ the response patterns 110 and 011 may occur frequently, but the 101–pattern should not occur, since of the eight possible combinations of 1s and 0s for three items, it is the only pattern that violates the deterministic model. If i, j, and k form an unfolding scale in this order, then H(ijk) should be positive and H(jik) and H(ikj) should be negative. A triple for which only one of the three homogeneity values is positive is called a „unique“ triple because the items can be interpreted as an unfolding scale in only one order; the best triple will be a unique triple. By default, values of H(i) > 0.30 are required for each item i to be accepted, and of H > 0.30 for the whole set of items to be accepted as an unfolding scale.

The information necessary for scale construction - number of errors observed O, number of errors expected under statistical independence E, and homogeneity coefficient - is shown in Table 1 for four of the 56 triples (the triples consisting of job values ABC, ABD, AEF, and FGH). Three sets of information are shown for each triple: first (to the left) the information for the ordering of the items with the alphabetically first letter in the middle (e.g., BAC), then (in the center) for the ordering with the second letter in the middle (ABC), and finally (to the right) for the ordering with the third letter in the middle (ACB). (The response pattern that counts as an error is always 101 for the items in any particular order, so for the first triple, ABC, 0 is first a response to A, then to B, then to C). Additional information about the probability that the population has a zero-homogeneity given the H-values and the sample size is omitted here (cf. Post, 1992).

Among all the ordered triples that can be formed from our set of eight items, the following are „unique triples“: ADF, ADG, ADH, AEF (shown in table), AEG, AEH, BDH, BEH, and DEH. AEF is the „best triple“, because its H(ijk)–value (0.40) is well beyond the default value of 0.30, and it has the highest sum of the frequencies of the acceptable patterns 011 and 110 (756 + 139). Triple ABD is a „positive triple“; i.e., its three H-coefficients are all positive. Such triples will not contribute to the falsification of an unfolding scale. Triples ABC and FGH are „dual triples“: two of their three H-coefficients are positive. This data set did not contain any „negative triples“: triples for which all three H-coefficients are negative. Such a triple violates the unfolding model in all three of its orders.

	jik			*ijk*			*ikj*		
	O	E	H	O	E	H	O	E	H
ABC	267	226.9	-0.18	110	157.1	0.30	40	83.1	0.52
ABD	283	348.3	0.19	174	241.2	0.28	34	39.0	0.13
AEF	756	650.9	-0.16	47	78.8	0.40	139	124.3	-0.12
FGH	82	118.2	0.31	130	173.0	0.25	463	443.7	-0.04

Table 1: Information about triples of items in their three distinct permutations

Starting with AEF as the „best triple“, the procedure adds new items one by one in any position to the scale as long as the criteria for an unfolding scale are met. The most important criterion is that all ordered triples in the new scale have a positive H(ijk)-value. The next most important criterion is that each new item, as well as the scale as a whole, has a homogeneity value higher than some lower boundary, set by default at 0.30. And the third criterion is that each new item can be added to the existing scale in only one position. The item to be added at each step is the one that results in the highest possible homogeneity value for the new scale. For our scale, the remaining items - C, G, B, H, and D - are now added one by one, each in a unique position.

Table 2 shows (from top to bottom) the order in which the items form an unfolding scale, and (from left to right) the items that took part in the scale as each successive item was added. As it happens, the order of the items reflects their single-peaked popularity order: the first item is the least popular (i.e., it has the smallest percentage of positive responses), and the popularity of each successive item increases to item D and then decreases again. All item-homogeneity coefficients are higher than the default lower boundary 0.30.

Number of items in the scale		*3*	*4*	*5*	*6*	*7*	*8*
	p(i)	H(i)	H(i)	H(i)	H(i)	H(i)	H(i)
A. No Pressure	0.27	0.40	0.49	0.46	0.44	0.41	0.41
B. Generous Holidays	0.34	---	---	---	0.49	0.48	0.46
C. Good Hours	0.50	---	0.46	0.44	0.49	0.47	0.44
D. Good Job Security	0.76	---	---	---	---	---	0.33
E. Job Meets Abilities	0.75	0.40	0.41	0.39	0.43	0.40	0.36
F. Can Achieve Something	0.66	0.40	0.45	0.40	0.45	0.42	0.39
G. Responsible Job	0.57	---	---	0.40	0.43	0.38	0.36
H. Useful for Society	0.34	---	---	---	---	0.35	0.32
H-value of the scale		0.40	0.45	0.42	0.46	0.42	0.39

Table 2: Result of search procedure to find a maximal set of unfoldable items.
p(i): Percentage of respondents who considered value i important for their job;
H(i): Homogeneity value for item i

In this procedure we have used the homogeneity coefficients to establish which items form part of an unfolding scale, and in what order. To illustrate a stricter search procedure we will check whether the items form an unfolding scale for each of the subgroups composed of subjects who picked the same number (between 2 and n-2) of items as important. Such „pick k/n“ analyses are marginally different from the previous „pick any/n“ analysis in the way the number of model violations expected under statistical independence, (E), is calculated (cf. Van Schuur 1989). Table 3 gives the results for $2 \leq k \leq 5$ analyses. The homogeneity coefficients of the scales remain around 0.30 for these analyses, but for the „pick 2/8“ analysis H is 0.25 and for the „pick 6/8“ analysis (not shown) it drops to 0.16. The H(i)-coefficients for the individual items are generally lower than in the „pick any/n“ analysis, but they hover around 0.30 except for item H in the „pick 3/8“ analysis and almost all the items in the „pick 2/8“ analysis. In all these analyses, however, there are several negative H(ijk)-values (see the last line of the table), and these include triples with the most popular items, D and E, which have the lowest homogeneity coefficients. In many cases, however, these negative H(ijk)-values are close to zero so overall the H-values corroborate our interpretation of the data as forming a unidimensional unfolding scale.

	Pick 2/8 N=255		*Pick 3/8* N=403		*Pick 4/8* N=426		*Pick 5/8* N=353	
	p(i)	H(i)	p(i)	H(i)	p(i)	H(i)	p(i)	H(i)
A. No pressure	.12	.28	.13	.36	.22	.39	.34	.31
B. Holidays	.07	.16	.16	.48	.24	.42	.41	.35
C. Good hours	.16	.30	.28	.39	.44	.38	.62	.32
D. Security	.52	.16	.62	.28	.78	.30	.90	.26
E. Abilities	.40	.27	.65	.31	.78	.32	.87	.27
F. Achieve	.32	.34	.55	.38	.65	.36	.74	.31
G. Responsible	.25	.28	.41	.36	.61	.36	.66	.29
H. Useful society	.16	.19	.19	.17	.28	.31	.47	.33
H-value of the scale		.25		.34		.36		.30
Number of triples in the scale with negative H(ijk)'s:		7		7		1		9

Table 3: Results of alternative 'pick k/n' unfolding analyses.
p(i) is the percentage of subjects who give the positive response to an item;
H(i) is the item homogeneity coefficient

4. Subject parameter estimation in a unidimensional unfolding analysis

In nonparametric unfolding, the scale values of subjects (the subject parameters) are estimated with a procedure similar to that used in Mokken's nonparametric cumulative scaling. Molenaar (1982) conceived a subject's scale score as „the number of item steps passed". For a Mokken scale, the number of item steps passed is simply the sum of the subject's scores on the individual items that make up the scale. In nonparametric unfolding, the notion of item steps is a little more complicated. Each item is represented along the latent continuum by a closed interval, bounded by two item steps: the left-sided i_{01} and the right-sided i_{10} item step. Subjects located between these item steps are close to the item and give it the positive response (1). Going from left to right along the latent continuum, they have „passed" the left-sided item step i_{01} but not the right-sided item step i_{10}. The number of item steps they have passed for this item, therefore, is 1. Subjects who give the negative response (0) to the item are represented either to the left of item step i_{01}, in which case they have passed 0 item steps, or to the right of item step i_{10}, in which case they passed 2 item steps. Which of these two representations applies to a specific situation can be determined from the context: if the nonpreferred item is positioned to the left of the preferred item(s) in the scale, then a negative response implies that 2 item steps have been passed for this item, but if it is represented to the right of the preferred item(s), a negative response implies that 0 item steps have been passed. For imperfect response patterns (i.e., response patterns that contain at least one „101" response to a triple of items), the decision about whether a negative response means 0 or 2 item steps passed is based on the position of the item with respect to *most* of the preferred items. A negative response sandwiched between an equal number of positive responses on either side is interpreted as 1 item step passed. In a response pattern that contains at least one positive response, the total number of item steps passed is defined simply as the number of item steps passed, summing over all items.

5. Goodness-of-fit tests for a nonparametric unfolding scale

Like a Mokken scale, an unfolding scale is initially hypothesized on the basis of its homogeneity coefficient, but it is not yet evaluated. To assess the extent to which the data conform to a probabilistic unfolding model, we need other diagnostics. Post (1992) showed that the correlation matrix of items in their unfoldable order should have two sign changes (-+-) at most going from left to right. Davison (1977) suggested that the correlation matrix should have a simplex pattern in which the correlations are highest along the diagonal and decrease both rowwise and columnwise away from the diagonal. The correlation matrix of our 8 items (not shown) conforms well to this pattern.

A second diagnostic, developed by Post and Snijders (1993), is the conditional adjacency matrix. This matrix contains the percentages of subjects who give the positive response to row item i from among all the subjects who give the positive response to column item j. If the items are sequenced both rowwise and columnwise in their unfoldable order, these percentage will be highest for a certain column item k, which is generally close to j. This column item should not be to the left of the column item that received the highest percentage of positive responses to item i-1 in the previous (upper) row. Stated more simply, the highest percentages rowwise should shift from top left to bottom right. Table 4 shows the conditional adjacency matrix. The numbers printed in **boldface** are the highest rowwise. Only row D shows a marginal deviation from the expected pattern.

	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*
A) No Pressure	-	**43**	37	29	25	21	19	19
B) Generous Holidays	**55**	-	**55**	37	35	32	32	29
C) Good Hours	70	**79**	-	53	49	47	46	47
D) Good Job Security	**84**	83	82	-	77	74	74	75
E) Job Meets Abilities	71	76	74	75	-	**81**	80	76
F) Can Achieve something	52	61	62	64	71	-	**76**	73
G) Responsible Job	41	52	52	55	60	**66**	-	**66**
H) Useful for Society	24	28	32	33	34	38	**39**	-

Table 4: Conditional adjacency matrix. Cell (i,j) contains the percentage of respondents who mentioned row value i as important in their job from among all respondents who mentioned column value j.

A third diagnostic of goodness-of-fit evaluates whether subjects with different scale values on the unfolding scale differ in the expected way in their response to the items in the scale (Table 5). Subjects with low scale scores should give the positive response most often to the leftmost items, and decreasingly often to successive items to the right. Subjects with high scale scores should show the reverse pattern. For subject groups with intermediate scale scores, the percentages should first increase rowwise and later decrease (i.e., they should form a single-peaked order). The highest percentages for each subject group, shown in **boldface** in Table 5, conform rowwise to the expected pattern. Since items and subjects have the same status in an unfolding scale, similar monotonically increasing and decreasing patterns should be found for the first and last columns (i.e., items) of the same matrix, and single-peaked patterns for the intermediate columns. The highest percentages columnwise are shown in *italics*. They too conform to the expected pattern, with one small exception each - of only 6% - in columns E and H. (Note that some cells contain the highest percentage in both their row and their column; the values are then shown in both **boldface** and *italics*.

Group	*scale from*	*scores to*	*Freq*	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
1	2	6	336	*70*	*75*	***85***	80	46	21	13	6
2	7	7	236	55	61	83	**89**	81	56	43	19
3	8	8	196	19	41	61	***97***	***97***	58	51	13
4	9	9	381	15	28	50	**90**	86	82	60	41
5	10	10	269	4	12	30	73	**92**	***92***	81	38
6	11	14	387	2	2	7	43	63	79	***86***	*66*

Table 5: Percentages of subjects in each scale score group who mention stimulus i. In **boldface**: highest percentage per row; in *italics*: highest percentage per column.

6. Discussion

The latent trait model demonstrated in this chapter differs from almost all other latent trait models described in this book (except chapter 26) in the form of its item characteristic curves: its ICCs do not increase monotonically, but are single-peaked instead. Such models are often useful when measurement is applied to subjects' preferences rather than to their abilities. In the measurement of ability, the latent trait can be considered as unipolar, going from zero to plus infinity. But in the measurement of preferences, the latent trait can be seen as bipolar, going from minus infinity through zero to plus infinity. In the measurement of ability, subjects are represented between the items they dominate and the items that dominate them. In the measurement of preferences, in contrast, subjects are represented in the center of the items they endorse. The items they do not endorse are too extreme for them toward either the left or the right pole of the scale.

In the case of work values, our success in applying the unfolding model suggests that intrinsic and extrinsic work values can be regarded as poles of a unidimensional continuum. Most subjects adopt a balance between intrinsic and extrinsic values, but their equilibrium point may differ. For reasons of parsimony, Lenski's unidimensional interpretation of intrinsic and extrinsic work values as bipolar opposites should, therefore, be favored over Herzberg's interpretation of these two sets of values as independent.

Acknowledgement

I would like to thank Melissa Bowermann for her editorial help.

References

Andrich, D. and G. Luo (1993). A hyperbolic cosine latent trait model for unfolding dichotomous single-stimulus responses. *Applied Psychological Measurement*, 17, 253-276.

Cliff, N., Collins, L.M., Zatkin, J., Gallipeau, D. and McCormick, D.J. (1988). An ordinal scaling method for questionnaires and other ordinal data. *Applied Psychological Measurement*. 12, 83-97.

Coombs, C.H. (1964). *A Theory of Data*. New York: Wiley.

Davison, M.L. (1977). On a metric, unidimensional unfolding model for attitudinal and developmental data. *Psychometrika*, 42, 523-548.

Díez Nicolás, J. and Inglehart, R. (eds.) (1994). *Tendencias mundiales de cambio en los valores sociales y políticos*. Madrid: Los libros de Fundesco.

Herzberg, F. (1966). *Work and the nature of Man*. New York: World Publishing Company.

Hoijtink, H. (1990). A latent trait model for dichotomous choice items. *Psychometrika*, 55, 641-656.

Lenski, G. (1963). *The religious factor. A sociological study of religion's impact on politics, economics, and family life.* New York: Doubleday, Anchor Books.

Lindseth, O.H. and Listhaug, O. (1994). Religíon y valores del trabajo en los noventa: un estudio comparativo de Europa Occidental y Estados Unidos. In: Díez Nicolás, J. and Inglehart, R. (eds.). *Tendencias mundiales de cambio en los valores sociales y políticos.* Madrid: Los libros de Fundesco, 469-481.

Mokken, R.J. (1971). *A theory and procedure of scale analysis.* New York/Berlin: De Gruyter (Mouton).

Molenaar, I.W. (1982). Mokken scaling revisited. *Kwantitatieve Methoden*, 3, 145-164.

Post, W.J. (1992). *Nonparametric unfolding models: a latent structure approach.* Leiden: DSWO Press.

Post, W.J. and Snijders, T.A.B. (1993). Nonparametric unfolding models for dichotomous data. *Methodika*, 7, 130-156.

Ross, J. and N. Cliff (1964). A generalization of the interpoint distance model. *Psychometrika*, 29, 167-176.

Sijtsma, K., P. Debets and I.W. Molenaar (1990). Mokken scale analysis for polychotomous items: theory, a computer program, and an empirical application. *Quality and Quantity*, 24, 173-188.

Van Schuur, W.H. (1989). Unfolding the German political parties: a description and application of multiple unidimensional unfolding. In: G. de Soete, H. Feger and K.C. Klauer (eds.). New Developments in Psychological Choice Modeling. New York: Elsevier, 259-290.

Van Schuur, W.H. (1993). Nonparametric unidimensional unfolding for multicategory data. *Political Analysis*, 4, 41-74.

Van Schuur, W.H. and W. Post (1991). *User's manual MUDFOLD, a program for multiple unidimensional unfolding.* Groningen, i.e.c. ProGamma, Grote Rozenstraat 15, 9712 GT Groningen, The Netherlands.

Van Schuur, W.H. and Kiers H.A.L. (1994). Why factor analysis is often the wrong model for analyzing bipolar concepts and what model to use instead. *Applid Psychological Measurement*, 18, 97-110.

Van Schuur, W.H. and Kruijtbosch, M. (1995). Measuring subjective well-being: unfolding the Bradburn affect balance scale. *Social Indicators Research*, 36, 49-74.