KM 44 (1993)
pg. 9-20

Determining Scale Values for Subjects in MUDFOLD[1]

Wouter van der Brug[2]

Abstract

The method for determining scale values for subjects in a multicategorical unfolding scale, proposed by Van Schuur (1993b), treats large proportions of the response patterns as 'missing'. In this article an alternative method is proposed which determines scale values for almost all response patterns. The advantage of the alternative method is illustrated with an empirical analysis of a multicategorical unfolding scale.

Preface

MUDFOLD -a nonparametric unfolding model developed by Van Schuur (1984, 1988, 1993a)- has recently been extended to be applicable to multicategorical data, such as Likert-type rating items (Van Schuur, 1993b). Van Schuur (1993a-b) proposes a method for determining the scale values of subjects, which has the disadvantage that no scale value can be determined for respondents who do not give the highest rating to at least one of the items. The scale value of these respondents is coded as a 'missing value'. This is a problem because, depending on the number of response categories and the type of question, a substantial proportion of the respondents will not use the highest response-category (Saris, c.s., 1987). The method proposed by Van Schuur (1993a-b) will be referred to here as the 'step counting method'.

With an alternative method, based on a weighted average procedure, it is possible to determine scale values for almost all response patterns. Moreover, it may often give a more realistic representation of the respondents position on the latent continuum. This method will be referred to as the 'averaging method'. In this article I will first explain both procedures for determining the scale values of subjects for scales with dichotomous data. Then I will present a general formula for the 'averaging method', which applies to dichotomous and multicategorical data.

The purpose of this article is to compare these two alternative methods for determining scale values for subjects in a MUDFOLD-scale. This is obviously only one of the problems involved with unfolding analyses of rating data. It is explicitly not my intention to confront other methodological issues here, which do not relate directly to the issue of determining scale values for respondents.

Although both the 'step counting method' and the 'averaging method' give good ordinal representations of the respondents, I will conclude that the 'averaging method' has some definite advantages over the 'step counting method'. This will be illustrated with an empirical analysis of a multicategorical unfolding scale.

---

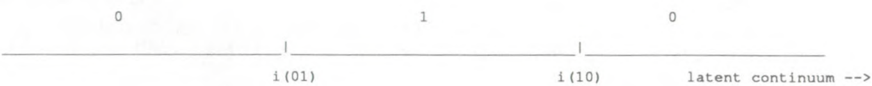
[1] The author expresses his gratitude to Cees van der Eijk, Wijbrandt van Schuur and Norman Verhelst for their usefull comments on prior versions of this paper.

[2] Vakgroep Algemene Politicologie. Universiteit van Amsterdam. Oudezijds Achterburgwal 237. 1012 DL Amsterdam. Tel: 020-5252765

1. MUDFOLD with dichotomous items

MUDFOLD is an unfolding procedure that determines whether a set of items classified by Coombs (1964) as proximity data, form a unidimensional unfolding scale (Van Schuur, 1993a). "In this procedure MUDFOLD evaluates and selects among candidate unfolding orders of items by comparing the number of times the dataset violates these orders under the deterministic model of perfect unfoldability with the number of violations under the null-model of statistical independence."[3] An important goodness-of-fit diagnostic is the H-value. If the items are ordered on the latent continuum in such a way that the dataset does not violate the deterministic model, the H-value will be 1. The H-value will be approximately 0, if the items are statistically independent. According to Van Schuur, the H-value must be at least .30 for the items to form an unfolding scale. Additional goodness-of-fit diagnostics are also provided (Van Schuur, 1984-1993b). Note that MUDFOLD is a procedure for non-parametric unfolding. This means that the procedure determines the ordering of the items, not their scale values.

figure 1: unfoldable dichotomous item i with two itemsteps[4]



Proximity data differ fundamentally from dominance data in this respect that respondents are expected to respond positively to items adjacent to them, and negatively to items that are further away. As a consequence, when respondents and items are represented on the same latent dimension, the respondent should be placed in the middle of the items that (s)he responds positively to.

A weighted average procedure, like the one proposed in this paper, is an obvious way to determine the scale values for respondents. The items are given a value based on their order on the latent continuum. The respondent is assigned a scale value that is the average score of the items picked. So, if the items A, B, C, D, and E form a MUDFOLD scale in the same order, they are assigned the scale values 1, 2, 3, 4, and 5. If a respondent would pick the items B and C (scale value 2 and 3), his/her scale value would be the average of 2 and 3, i.e., 2.5. The method developed originally by Van Schuur (1984) to compute scale values for respondents, was also based on a weighted average procedure.

Although this method gives a good ordinal representation of the subjects, Van Schuur developed an alternative procedure for computing scale values that is analogous to the procedure followed in the cumulative Mokkenscale (Mokken, 1970). This method is

---

[3] Van Schuur (1993a: 40).

[4] This figure was copied from Van Schuur (1993).

based on the principle of counting the itemsteps the respondent passed. Figure 1 shows that each dichotomous unfolding item contains two itemsteps, dividing the latent continuum in three areas. If this would be a deterministic model, then only the respondents in the middle of the continuum would respond positively to item i. Respondents positioned to the left of itemstep i(01) or to the right of itemstep i(10) would respond negatively to it.

Determining how many itemsteps a respondent passed, presents a problem. If the respondent answers positively to an item (s)he passed 1 itemstep, but a negative answer could mean that either 0 or 2 itemsteps were passed. In other words, the respondent could be situated on the far left, or on the far right area of the continuum. Whether the respondent passed 0 or 2 itemsteps is decided by observing the whole response pattern. If the items that the respondent answers positively to are situated on the right of the item responded negatively to, then it will be interpreted to indicate that 2 itemsteps were passed. If they are mostly to the left of it, it will be assumed that 0 itemsteps have been passed. If there are the same number of positive responses on either side of this item, it will be interpreted to indicate that 1 itemstep was passed.

Table 1: examples of response patterns in a dichotomous MUDFOLD scale, and scale values assigned by both the 'step counting method'(SCM) and by the 'averaging method'(AM)

	1	3	5	7	9	scale values	
item	A	B	C	D	E	SCM	AM
respondent 1	1	1	1	0	0	3	3
respondent 2	1	1	0	0	1	3	4.3
respondent 3	1	1	0	1	1	5	5
respondent 4	0	1	0	1	1	7	6.3
respondent 5	0	0	0	0	1	9	9

The hypothetical response patterns in table 1 can be used to explain this procedure. Respondent 1 gives positive scores to the left items A, B, and C. Positive scores are always interpreted to mean that one itemstep has been passed. The negative scores for the items right of the items that (s)he responds positively to are interpreted to indicate that none of these itemsteps have been passed. Respondent 1 has therefore passed 1+1+1+0+0+0=3 itemsteps. Respondent 5, on the other hand, only responds positively to the far right item. The negative responses to the other four items are therefore interpreted to indicate that both itemsteps of these items have been passed. So, the conclusion is that respondent 5 has passed 2+2+2+2+1=9 itemsteps.

Another way to compute the scale values for respondents according to the 'step counting procedure', is the following. After establishing that a group of items forms an unfolding scale in a certain order, the items can be given a value corresponding to the order in which they are ranked. Using only the odd numbers 1, 3, 5, 7, etc., the scale value turns out to be the *median* value of the items picked. Note that the earlier procedure determined the scale value of the respondent by computing the *average* value of the items picked.

The hypothetical response patterns of table 1 can now be used to compare the scale values determined by both the 'step counting method' and by the 'averaging method'. The first thing to be noticed is that for 'perfect' response patterns the mean and the median are the same. So, for these response patterns both methods are indifferent for *scales with dichotomous data*. Assigning scale values to respondents with 'imperfect' response patterns is more arbitrary. As mentioned earlier, a respondent should be represented on the latent continuum in the middle of the items picked. Respondent 2 should therefore be represented more to the middle of the scale then respondent 1, since his/her preference for the left items is more unsettled. In my view, the scale value assigned to the respondents should take this into account.

In practice however, the two procedures are virtually indifferent for *dichotomous* MUDFOLD scales. As a matter of experiment I tried both methods on a dichotomous unfolding scale that I constructed with one of my own datasets.[5] The scales created by both procedures had a Pearson-correlation of .9912. For *multicategorical* scales though, the 'averaging method' has some definite advantages, which I will elaborate on in the next paragraph.

2. MUDFOLD with multicategorical data

Figure 2 shows an unfolding item i with four response categories (0, 1, 2 and 3), and 6 itemsteps. In general we can say that each item with r response categories has 2(r-1) measuring points, or so called 'itemsteps', on the latent continuum. If this was a deterministic unfolding item, a respondent positioned on the latent continuum between itemstep i(01) and i(12) would give item i a rating of '1'.[6] A respondent positioned between itemstep i(23) and i(32) would give a rating of '3', etc. Since MUDFOLD is a probabilistic scaling procedure, we cannot be certain what rating a respondent will give. Since the item i presented here is a hypothetical example, we could however define a response function so that the ratings presented in figure 2, are the 'most likely' responses, given each position on the latent continuum.[7]

The procedure to determine whether a set of multicategorical scales form an unfolding scale is very similar to the procedure for dichotomous items. Deviations from perfect scalability are defined by the number of times a respondent gives a lower rating to

---

[5] These data came from a survey held in Februari 1991 among a sample of 423 students from the Charles University in Prague. The respondents were asked to give ratings on a 7-point scale for seven Czech politicians. Five of these items formed an unfolding scale in a MUDFOLD procedure, after the responses had been dichotomized.

[6] The notation for the 'itemstep' which divides between an interval on the left on which a respondent will give a deterministic item i a rating of '1', and an interval on the right on which a respondent will give item i a rating of '2', is i(12).

[7] In this example all response categories are the 'most likely' ratings at certain positions of the latent continuum. This does not have to be necessarily so. It is possible that a certain rating category is not the most likely response on any position of the latent continuum.

an item 'in between' two items that get a higher rating. The assumption is that each category discriminates identical for each stimuli, irrespective of the items rated. Therefore the response pattern 303 violates the unfolding model to a higher extent then the response pattern 323. MUDFOLD does not assume identical intervals between the ratings (Van Schuur, 1993b).

The 'step counting procedure' for determining the scale value of respondents for rating data, is analogous to the one developed for dichotomous data. Determining how many itemsteps have been passed, we are faced with the same problem as in the case of dichotomous items. A rating of '2' for item i could mean that the respondent passed either two or four itemsteps. We determine whether the respondent passed two or four itemsteps, by observing the whole response pattern. If the highest ranking is given to an item situated to the left of item i, we will interpret this to indicate that only two itemsteps of item i have been passed. If, on the other hand, the highest ranking is given to an item to the right of item i, it will be interpreted to mean that four itemsteps from item i have been passed. If the same number of highest rankings occur on both sides of item i, we will assume that three itemsteps have been passed. It is important to note that the lowest rating category is 0. If the rating categories in the original dataset range, for instance, from 1 to 5, they should be recoded so that they range from 0 to 4. This should also be done when using the 'averaging method'.

Figure 2: unfoldable item i with four response categories (0, 1, 2 and 3) and 6 itemsteps

```
           0                     1    2     3     2    1        0
_________________________________|____|_____|_____|____|____|__________________________
                               i(01) i(12) i(23)  i(32) i(21) i(10) latent continuum -->
```

The step counting method will be illustrated with an example. The response pattern 03100 will be interpreted to indicate that all six itemsteps of the far left, three itemsteps of the second item and one itemstep of the third item, have been passed. The scale value assigned will be 6+3+1+0+0=10. If we are willing to accept the general idea from figure 2 that an item has 2(r-1) itemsteps on the latent continuum, and that the presented responses are the most likely responses on those positions on the latent continuum, then we can do the same with a hypothetical scale that contains three such items. Figure 3 shows a hypothetical unfolding scale, which contains three items i, j, and k, each having four response categories (0, 1, 2 and 3). Also, the most likely response pattern for each position on the latent continuum is given.[8]

---

[8] The concept of the most likely response pattern is based on the principle of 'local independence'. This principle says that all systemetic variation in the response of subjects to items is solely due to the position of subjects on the latent continuum. In a single point on the latent continuum all variation is random, i.e., responses to items are statistically independent.

Figure 3: Hypothetical unfolding scale with three items i, j and k, each having four response categories (0, 1, 2 and 3) and 'most likely' response patterns[9]

```
000        200         200   110   020       020    011    002          002          000
____|_____|______|_______|___|__|__|__|___|________|__|___|_|__|__|__|_______|_______|_____|_____
      100          300         210   120      030        021   012       003           001
                                                                    latent continuum -->
```

In figure 3 the intervals between the response categories are identical for each item, i.e., the i(12)-i(23)=j(12)-j(23). The intervals between rating categories are not identical, i.e., i(01)-i(12) is not equal to i(12)-i(23). In figure 3 one can see that response patterns that do not contain the highest possible rating, like for instance the response pattern 020, can show up on different positions on the latent continuum. The 'step counting method' regards these response patterns therefore as 'missing data'. With the 'averaging method' that is proposed in this paper, we do not have to do that. The averaging method is presented in a formal way by formula 1, which can be applied to MUDFOLD scales with multicategorical data. Note that the method 'works' for all response patterns, except for the response pattern that contains only zero's.

Formula 1:

$$\theta_j = \frac{\sum_{i=1}^{k} (R_{ij} V_i)}{\sum_{i=1}^{k} R_{ij}}$$

Where:
k= number of items
θ_j= scale value from subject j
R_{ij}= rating subject j gives item i
V_i= scale value of item i (the odd numbers 1, 3, 5, etc.)

Table 2 presents the scale values for the 'most likely' response patterns from figure 3, assigned with the 'step counting' and with the averaging' method. The example presented in figure 3 and table 2 suggests that there may be large intervals on the latent continuum on which respondents are unlikely to give the highest rating to any of the items offered. Accordingly, the 'step counting method' will treat most answers given by respondents on these positions as missing data. The most important advantage of formula 1 is that scale values can be assigned to respondents on almost every position on the latent continuum. The intervals on which respondents are unlikely to give the highest rating to any of the items, may not generally be as large as in this hypothetical example. But, if a scale consists of items with many rating categories, it may be expected that a large proportion of the responses will be treated as 'missing data'.

---

[9] Figure 3 only presents the most likely response patterns. The itemsteps in their corresponding order on the latent continuum are : i(01), i(12), i(23), i(32), j(01), i(21), j(12), i(10), j(23), j(32), k(01), j(21), k(12), j(10), k(23), k(32), k(21), k(10).

Table 2: The most likely response patterns from figure 3, and their scale values, computed with the 'step counting method' (SCM) and with the averaging method (AM)

Most likely response pattern	SCM	AM*	most likely response pattern	SCM	AM*
000	m	m	030	9	9
100	m	3	020	m	9
200	m	3	021	m	11
300	3	3	011	m	12
200	m	3	012	m	13
210	m	5	002	m	15
110	m	6	003	15	15
120	m	7	002	m	15
020	m	9	001	m	15

* The values assigned by formula 1, have been multiplied by a factor 3 (the number of response categories minus one), to make them better comparable to the values assigned by the 'step counting method'.
** m= missing

Formula 1 computes scale values for respondents according to a weighted average procedure. Formula 2 computes the standard deviation S_j around this weighted average score, for each respondent separately. The standard deviation could be used as an indicator for the 'region of acceptance'. One should however be careful with the interpretation of this measure. First, MUDFOLD is not a parametric scale, so the intervals between the scale values of the items (the V_i's) are not estimated by MUDFOLD. Instead, we fix the scale values with equal intervals between each consecutive item. Second, even if the intervals were equal, the 'region of acceptance' depends to a large extent on the respondent's position on the continuum. This means that a wide 'region of acceptance' in terms of the stimuli does not indicate a wide 'region of acceptance' in terms of the underlying dimension. Therefore, the standard deviation S_j should in my view only be used to compare *within* groups, i.e., between respondents who have (approximately) the same scale value on the continuum.

Formula 2:

$$S_j^2 = \frac{\sum_{i=1}^{k} R_{ij}[V_i - \theta_j]^2}{\sum_{i=1}^{k} R_{ij}}$$

The response patterns presented in figure 3 are the ones that are expected most likely to occur, given the assumptions of the model. Since MUDFOLD is a probabilistic and not a deterministic model, a proportion of the response patterns will violate the deterministic model of perfect unfoldability. The performance of both methods for such response patterns cannot be illustrated with figure 2 and 3, since these are not the 'most likely response patterns'. To still get an idea of the performance of both methods, some hypothetical 'perfect' and 'imperfect' response patterns are presented in table 3.

Table 3: Hypothetical response patterns for a MUDFOLD scale with five items, each having six response categories (0, 1, 2, 3, 4, and 5). Scale values assigned by both the 'step counting method'(SCM) and by the 'averaging method'(AM)

	1	3	5	7	9	scale values	
item	A	B	C	D	E	SCM	AM
respondent 1	0	2	3	0	0	missing	21
respondent 2	0	2	4	0	0	missing	21.3
respondent 3	0	2	5	0	0	23	22.1
respondent 4	1	2	5	2	0	24	23
respondent 5	0	0	5	3	0	28	28.8
respondent 6	0	0	0	5	1	36	36.7
respondent 7	4	0	0	5	0	31	21.6
respondent 8	5	0	0	4	0	9	18.3

The respondents' scale values computed by formula 1 have been multiplied by 5 (which is r-1), to make them better comparable to the scores computed by the 'step counting method'. Both methods now assign scores in the range of 5 to 45.[10] If we compare the scores assigned to the 'perfect response patterns' from the respondents 3, 4, 5, and 6, both methods seem to be virtually indifferent. Note that this is also true for the scale values in table 2.

The 'missing data' problem becomes apparent again, when the scale values of the first three respondents are compared. In my view all three respondents provide valuable information, and their response patterns are very similar. They should therefore not be treated as 'missing', but should be assigned highly similar scale values. Formula 1 seems to perform quite well in this respect.

Another thing that becomes apparent when looking at the respondents 6, 7, and 8, is that the 'step counting method' can be very sensitive to small differences between response patterns. To show this, I have presented an example that is a bit 'extreme', in the

[10] If formula 1 is multiplied by (r-1) the scale values will have the same range as when they are computed with the 'step counting method'. If r is the number of rating categories and k is the number of items, then there are 2k(r-1) itemsteps on the latent continuum. We have noted that on both extreme ends of the scale (r-1) itemsteps cannot be counted. Hence, the 'step counting method' has a theoretical minimum of (r-1) and a theoretical maximum of 2k(r-1)-(r-1)=(r-1)(2k-1). If we fix the scale values of the items at the odd numbers 1, 3, 5, etc., then formula 1 has a theoretical minimum of 1 and a theoretical maximum of (2k-1). The reader can now verify that the range of scores will be identical for both methods if formula 1 is multiplied by (r-1).

sense that the response patterns 7 and 8 fit the requirements of the unfolding scale very badly. Although, extreme cases like this are unlikely to occur in practice, the example serves its purpose to illustrate the sensitivity of the 'step counting method' for small changes in response patterns.

In my view the responses of the respondents 7 and 8 are very similar, and both very different from the response pattern of respondent 6. The scale values assigned with formula 1 takes this into account. The 'step counting method', however, grants so much importance to the item with the highest rating, that respondent 6 and 7 are positioned close together, very far from respondent 8. I consider it a disadvantage of the 'step counting method' that it can be so sensitive for very small changes in the response pattern. The 'averaging method' seems to be more robust in this respect.

3. An empirical application

In order to compare both methods I have applied them both to a multicategorical unfolding scale constructed with a dataset that was collected by Tillie (1993). The dataset is a survey of a representative sample of the Dutch electorate (N=704), conducted in september 1992. The respondents were asked to indicate on a 1 to 10 rating scale the probability that they would ever vote for any of seven Dutch political parties. The resulting unfolding scale, after listwise deletion of missing data (N=624), is presented in table 4.

Table 4: Unfolding scale of seven Dutch political parties

Stimuli	H-value
C.D. (Extreme Right, Racist Party)	.45
G.P.V (Christian Fundamentalists)	.56
V.V.D (Liberal Conservatives)	.52
C.D.A (Christian Democrats)	.51
D66 (Liberal Democrats)	.59
P.v.d.A (Labour)	.62
Groen Links (Green party)	.71

The seven stimuli form a strong unfolding scale (H=.57).[11] Moreover, the stimuli are ranked in the same order as their position on a left-right scale[12], which is another indication for the validity of the unfolding scale. The respondents were given a scale value with both procedures. Only 266 respondents assigned the maximum score of 10 to at least one stimulus. Therefore, only 266 (42.6%) of the 624 respondents who responded to all the seven items of the scale, could be assigned a scale value with the 'step counting

[11] According to Van Schuur (1984-1993b), if H>.50 the items form a strong unfolding scale. This means that there are little violations of the deterministic model of perfect unfoldability.

[12] The respondents were also asked to assign a position to political parties on a 10 point left-right scale, where a 1 means extremely left-wing and a 10 means extremely right-wing. If the average score is computed for each party, they are ranked in the same order, except for the C.D. The dispute about the position of this party in the Dutch ideological space continues.

method'. Using formula 1, 605 respondents (95.3% of the remaining sample of 624) were assigned a scale value.

In order to compare both methods, the correlation was computed between the scale values determined by both methods. Here only the 266 respondents who were assigned a scale value with the 'step counting method' could be used. The Pearson correlation between the scale values assigned with both methods is .97. This indicates that in instances where a scale value can be assigned, both methods are almost indifferent.

In this survey the respondents were also asked to indicate their own position in terms of left and right on a scale ranging from 1 to 100. Theoretically the variable left-right selfplacement can be used as a criterion to validate the unfolding scale, since the respondents position on the left-right scale is the best predictor of his/her party preference (v.d.Eijk and Niemöller, 1983).

From the 605 respondents who were assigned a scale value with formula 1, 551 respondents answered the left-right selfplacement question. The correlation of the party-preference unfolding scale and the variable left-right selfplacement, is .63 for this group of 551 respondents. From the 266 respondents who were assigned a scale value with the 'step counting method', 242 respondents also answered the left-right selfrating question. For this group the Pearson correlation between the variable left-right selfplacement and the unfolding scale, was .68 if the 'step counting method' was used and .67 if formula 1 was used. So, if the variable left-right selfplacement is used as a criterion to validate both scaling procedures, the conclusion should be that formula 1 performs just as well as the 'step counting method'. Moreover, respondents who cannot be assigned a scale value with the 'step counting method', get a valid scale value with formula 1. These conclusions stem from the fact that correlations of .68, .67 and .63 are virtually indifferent.

Table 5: Scale values of respondents computed with formula 1 for subgroups

group	N	Minimum	Maximum	mean	S.D.
all respondents	605	27	117	74.94	17.00
included group	266	27	108	76.06	17.87
excluded group[13]	339	27	117	74.06	16.26

In table 5 some basic statistics of the scale values of the respondents computed with formula 1, are presented. The statistics have been computed separately for the whole group of 605 respondents for whom a scale score could be computed with formula 1, for the 266 respondents for whom a scale score could be computed with the 'step counting method', and for the 'excluded' group of 339 respondents. The results presented in table 4 confirm that the respondents who cannot be assigned a scale value with the 'step counting method', does not differ fundamentally from the other respondents. The mean and standard deviation of the two groups of respondents are virtually indifferent.

---

[13] The excluded group is the group of respondents for whom the 'step counting method' cannot determine a scale value.

4. Conclusions

After comparing formula 1 and the 'step counting method', we can arrive at the following conclusions. Both methods can be used in order to give an ordinal representation of the positions of the respondents. Also, both methods assign virtually the same scale values to 'perfect response patterns', i.e., response patterns that do not violate the deterministic model. Because large deviations from 'perfect' response patterns will not occur often, otherwise the items would not form a scale, the two methods are to a large extent indifferent. Application of both methods on an empirical dataset, confirms this. The scale values assigned with both methods have a Pearson correlation of .97.

One of the reasons why Van Schuur prefers the 'step counting method' is that subjects will be more evenly distributed on the latent continuum, if this method is applied. This results from the fact that there is a limited amount of categories in which respondents will be scaled. Using the 'averaging method', 'imperfect' response patterns may result in categories that fall in between others, and contain only few cases. This presents a problem when MUDFOLD creates a table in which the average itemscore is presented for every score group of respondents. Since this is a very 'rough' diagnostic, I would suggest to round of the fractional numbers that result from applying formula 1. The fractional numbers do not present a theoretical problem as long as the ordinal representation remains intact. Moreover, distributions of ordinal scales are difficult to judge, since the relative length of the intervals cannot be estimated.

A major drawback of the 'step counting method' is that in unfolding scales with multicategorical data, the scale value for response patterns that do not contain the highest rating for any of the items, are treated as 'missing data'. The empirical application of both methods highlighted this drawback. Using the 'step counting method', 57.4% of the sample had to be coded 'missing', whereas using formula 1 this was only 4.7%. Van Schuur (1993) acknowledges that the disadvantage of the 'step counting method' is that a large proportion of the response patterns has to be coded 'missing'. The solution he offers is to recode the highest response categories, so that every respondent gives the highest score to at least one item. This solution is not satisfactory, because the consequence is that we loose some of the information that we originally saved by using multicategorical instead of dichotomous data. Other ways of recoding the response categories may be developed, but all together these solutions seem rather artificial. The straightforward way in which the 'averaging method' can determine scale values for almost all respondents, may be considered a real advantage.

Moreover, scale values assigned to respondents who give 'imperfect' response patterns, may be more realistic using formula 1. Therefore, it is my position that if the researcher does not object to fractional numbers (s)he should use the 'averaging method'. Formula 1 has the practical advantage of being easily applicable in standard software such as SPSS, so no additional software is required.

References

Coombs, C.H. (1964). *A Theory of Data*. John Wiley and Sons, New york/ London. (2nd ed. 1976).

Eijk, C. van der, and B. Niemöller (1983). *Electoral Change in the Netherlands. Empirical results and methods of measurement*. CT Press, Amsterdam.

Mokken, R.J. (1971). *A Theory and Procedure of Scale Analysis. With applications in political research.* New York/ Berlin: Walter de Gruyter (Mouton).

Saris, W.E., Putte, B. van de, Maas, K., Seib, H. (1987). De vergelijkbaarheid van antwoorden: De noodzaak van meerdere referentiepunten. In Zouwen, J. van de, Jong-Gierveld, J. de (ed.). *De Vragenlijst in het Sociaal Onderzoek. Een Confrontatie van Onderzoekspraktijk en -Methodiek.* Van Loghum Slaterus B.V., Deventer.

Schuur, W.H. van (1984). *Structure in Political Beliefs. A new unfolding model with application to European party activists.* CT Press, Amsterdam.

_______________ (1988). Stochastic Unfolding. In Saris, W.E. en Gallhofer, I.N. (eds.). *Sociometric Research, Vol. I: Data Collection and Scaling.* MacMillan, London.

_______________ (1993a). MUDFOLD. *Kwantitatieve Methoden* 42:39-54.

_______________ (1993b, forthcomming). Nonparametric Unidimensional Unfolding for Multicategory data. Offered for publication to *Political Analysis.*

Tillie, J. (1993, forthcomming). *Dutch Electorate Study 1992.*

ontvangen 30- 3 -1993
geaccepteerd 23- 6 -1993