Determining subjects' scale scores in MUDFOLD for multicategory data. A research note in reaction to Van der Brug.

Wijbrandt H. van Schuur[1]

Van der Brug (1993) gave an excellent account of the procedure used in the nonparametric unidimensional unfolding procedure MUDFOLD for determining scale values for subjects. He called this procedure 'step counting', because each item with r response categories divided the latent continuum in 2(r-1)+1 areas, bounded by 2(r-1) item steps (see Figure 1). For instance: An item with four response categories (prefer very much=3; prefer somewhat=2; prefer a little=1; prefer not at all=0) has 2(4-1)=6 item steps, which divide the latent continuum in 7 areas (0 to 7). See Figure 1.

```
response:                 0     1     2     3     2     1     0
nr of steps passed:       0     1     2     3     4     5     6
                        ----|-------|-----|-----|-----|-----|---
item step:                i(01) i(12) i(23) i(32) i(21) i(10)
                                         latent continuum ->
```

The highest response (3) is unequivocally represented by having 3 item steps passed. But for lower responses two representations are possible: to the left or to the right of the highest possible response. Which of these two representations is chosen depends on the context, i.e., whether the highest response was given to one or more other stimuli represented to the left or to the right of it.

So in the response ABCDE,12321, to five items that form an unfolding scale in this order, in which item A is given response value 1, item B the value 2, etcetera, the highest value (3) is given to item C. This means that the lower values for A and B to the left of C should be interpreted as having passed more than 3 item steps: 5 for item A and 4 for item B, whereas the values for D and E should be interpreted as having passed less than three item steps: 2 for item D and 1 for item E.

The justification for this is given with the assumption in the model that for all unfoldable items the order of the left-sided item steps (i.e., the item steps i(01) ... i(r-1,r)) is the same as the order of the right-sided item steps (i.e., the item steps i(r,r-1)... i(10)) (Van Schuur, 1993a). This assumption is sufficient for the unequivocal determination of the order of the items along the underlying dimension.

Van der Brug cites two problems with this approach. The most serious problem is that for response patterns that violate the unfolding model relatively minor differences in

---

[1]Full address Vakgroep Sociologie, Grote Rozenstraat 31, 9712 TG Groningen. Tel.: 050-636436; Fax: 050-636226; e-mail: H.VAN.SCHUUR@PPSW.RUG.NL.

responses may have large consequences for the scale values. His example: the response pattern ABCDE,40050 to five items with six response categories each gets the scale value (6+10+10+5+0=) 31, whereas the response pattern ABCDE,50040 gets scale value (5+0+0+4+0=) 9.

The second problem he mentions deals with the question how MUDFOLD handles response patterns that do not contain the highest response value for any of the items. This has been a problem to which different strategies have been tried in the past. Van der Brug correctly criticizes the procedure in which all such response patterns are deleted as missing values. Indeed, in many empirical applications this would lead to a deletion of more than half of the subjects. He is also correct in pointing out that recoding the response categories such that this problem does not occur anymore implies throwing out a lot of important information.

The strategy implemented in MUDFOLD now is to interpret any subject's highest response as the maximal possible response. This means that, for instance, the response patterns 02300, 02400, and 02500 all receive the same scale value (10+8+5+0+0=) 23.

Van der Brug's solution to overcome both problems is to go back to an earlier proposal (Van Schuur, 1993b, which had not appeared at the time vdB wrote his paper) of calculating a weighted average of the responses. The responses are then weighted by the rank number of the items in the unfolding scale (or a linear function thereof). For instance, the items ABCDE get weights 1,2,3,4,5, so the response pattern 12343 would get the scale value (1*1 + 2*2 +3*3 + 4*4 + 5*3)/(1+2+3+4+5) = 3.2.

It is true that using such a procedure would create smaller differences in scale values to similar imperfect response patterns than the step counting method. It is also true that the fact that the averaging method gives broken numbers rather than integers is not a fundamental problem. Nevertheless two counter arguments should be made. First: the averaging method is not an ordinal method, in that it assumes equal distance between the items. The step counting method remains fully ordinal. Second, the problem that imperfect response patterns that are largely similar may get widely different scale values should be approached from the perspective of person fit: the problem increases to the extent that the response patterns violate the model more.

Van der Brug showed that scale values determined by the step counting method and the averaging method for dichotomous data correlated .99 and for multicategory data, in which response patterns without any maximum response value were deleted, correlated .97. Using the same data that Van der Brug used (624 respondents and 7 unfolding items) (see Tillie, 1994) and using the adapted version of the step counting approach , in which each respondent's highest response value is recoded to the maximum possible value, we now find a correlation of .94 between the step counting method and the averaging method.

Two remarks should be made about this finding. First, among the 624 respondents only 43% used the maximum score of 10 at least once; 20% had 9 as their highest score; 19% had 8, 8% had 7, and 11% of the subjects had 6 or less as their highest score). Only 19 subjects (3%) gave the same value (in this case, the lowest value: 1) to all 7 items. Leaving out the 5% subjects with 4 or less as their highest score hardly changes the correlation (from .937 to .942). This suggests that the solution of increasing the highest response value to the maximal response value is less problematic then might be feared.

Second, 51% of the 624 subjects made no errors against the deterministic model, 10% made between 1 and 9 errors, another 15% made between 10 and 22 errors, the next 20% made between 23 and 62 errors, and the last 5% made between 64 and 154 errors. It turns out that the correlation between the two scale scoring methods hardly changes if the 5% subjects with the largest number of errors are deleted (from .937 to .944). Here again in this example Van der Brug's fears about large differences in scale values for small changes in imperfect response patterns does not seem to be warranted. Obviously, more (systematic simulation) study is needed to make stronger claims.

In the present version of MUDFOLD scale values and number of model violations for each response patterns can be saved on a separate ASCII file, that can then later be added to the original file. For purposes of further comparison of both methods to calculate scale values, the averaging method will also be incorporated in the next update of MUDFOLD that is presently in progress

References

Tillie, J. (forthcoming). Dutch Electorate Study 1992. Dissertation, University of Amster dam.

Van der Brug, W. (1993). Determining scale values for subjects in MUDFOLD. Kwantita tieve Methoden, 44, 9-20

Van Schuur, W.H. (1993a). Nonparametric undimensional unfolding for multicategory data. Political analysis, 4, 43-72

Van Schuur, W.H. (1993b). Nonparametric undimensional unfolding for rating data. In: R. Steyer, K.F.. Wender, and K.F. Widaman (eds.), Psychometric methodology. Proceedings of the 7th European Meeting of the Psychometric Society in Trier. Stuttgart: Gustav Fischer, 469-474

Ontvangen: 3-12-1993

Geaccepteerd: 7-1-1994