Chapter 34

Detecting Artifacts in Panel Studies by Latent Class Analysis

Herbert Matschinger and Matthias C. Angermeyer

Department of Psychiatry, University of Leipzig

1. Introduction

Measuring change for different kinds of domains by employing Items of the rating-scale type for at least two points in time several artifacts may lead to heavy distortions of the instrument. Besides other kinds of well- known sources for bias in panel studies the necessarily invariant application of both the items and the categories may result in increasing inattentiveness for all following repetitions. At least for the particular example presented in this paper we have to assume that a „latent change“ into a certain direction may *cause* the interviewer to neglect this subject rather then to record the manifestations of the change. This may result in a very simple structure of the reduced space, dominated by missing-values or „non-recording“. Additionally, tangible effects of the mere layout of the questionnaire may be observed during later waves. The aim of this paper is less to focus on the change itself or to test a particular restrictive latent class model, but rather to explore the structure of a particular instrument with respect to non-recording during later applications. This will be done also in order to detect and to describe subsets of observations suffering from systematic non-recording for at least part of the questionnaire at T_2.

2. Data

To evaluate the efficacy of an antidepressant drug a set of 15 items indicating 15 different symptoms of depression has been applied before and eight weeks after treatment. The list of symptoms was presented in two columns: the first 8 symptoms were located on the left side, the latter 7 symptoms on the right side of the page. A total of 8927 patients where included in the study. The physicians were asked to evaluate each of the 15 symptoms (listed in Table 1) on a 4-point scale ranging from

„not present“ (1) „weak“ (2) „average“ (3) to „heavy“ (4).

1)	depressive irritation	9)	trouble concentrating/thinking
2)	anhedonia	10)	repeated thoughts of death
3)	change of weight	11)	anxiety
4)	sleep disturbances	12)	suicidal tendency
5)	restlessness	13)	cardiovascular symptoms
6)	psychomotor inhibition	14)	gastrointenstinal symptoms
7)	anergy	15)	headache
8)	guilt		

Table 1: Order of symptoms as listed in the questionnaire

Changes in patient's symptoms between T_1 and T_2 are considered to provide an indication of the success of the treatment. Unfortunately a considerable loss of information was observed (at least one symptom was unrecorded in the case of more then 15% of the patients). This kind of missings can not likely to be missing at random (MAR) or missing at random within classes (MARC), and we would expect heavy biases of the remaining sample (Little & Rubin 1987, 1990).

The main goal of the following analysis is not to impute the „non-recording“, but rather to describe and partially explain a very specific mechanism of „panel death“ resulting both from an interaction of individual characteristics of the patients and from the particular lay out of the questionnaire at T_2. It is possible to evaluate the structure by special forms of non-linear PCA in order to represent the data in a low-dimensional space regarding the variables as nominal and using non-recording as a category. This approach takes into account all bivariate crosstabs (Burt-table) and not all possible patterns, as is done by latent-class analysis as a special form of log-linear analysis (comp. Van Buuren & Van Rijckevorsel, 1992; Hagenaars, 1993; Vermunt, 1993).This will be carried out by

1. evaluating the structure of the symptoms at T_2 by latent-class analysis (Formann, 1984; Langeheine, 1988; Rost, 1988) including the „non-recordings“ as a single category for each of the 15 symptoms. In order to reduce the number of cells the items have been dichotomized {1,2} {3,4} resulting in 3-point items;
2. applying a multinomial logit model to explain the probability of falling into one of the latent-classes at T_2.

3. Interpretation of the latent-class analysis

Since we had no strict assumptions about the structure at T_2 we calculated several solutions from 2 to 6 classes imposing no restrictions on the responseprobabilities at all. Since we have to deal with 3^{15} possible but only 1084 observed pattern we only look at the BIC and at the distribution of the membership probabilities in order to choose a particular solution as the very one explaining the data sufficiently and most parsimoneously. All calculations were carried out using model 9 of the program LACORD (Rost, 1990).

The 4-class solution shows the best fit with respect to the BIC but to facilitate the interpretation of this solution both the 3- and the 4-class solution and the cross-tabulation between the membershipfiles will be presented.

3 classes = 57244 4 classes = 55254 5 classes = 58912 6 classes = 58948

3.1 Three-class solution

The response probabilities of the 3-class solution exhibit a very clear and simple structure for the second measure if non-recording is taken into account for each symptom:

	Var.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	Kat.															
1.Class	1	0.306	0.309	0.308	0.299	0.306	0.293	0.278	0.286	0.031	0.017	0.028	0.023	0.016	0.022	0.023
0.036	2	0.036	0.023	0.023	0.023	0.023	0.023	0.022	0.033	0.013	0.006	0.013	0.007	0.010	0.007	0.010
	3	0.568	0.668	0.670	0.678	0.671	0.684	0.700	0.681	0.955	0.976	0.959	0.970	0.974	0.971	0.967
2.Class	1	0.300	0.379	0.896	0.495	0.615	0.693	0.691	0.679	0.568	0.827	0.600	0.940	0.790	0.826	0.792
0.174	2	0.696	0.616	0.104	0.502	0.383	0.305	0.301	0.317	0.428	0.168	0.391	0.054	0.200	0.170	0.199
	3	0.004	0.005	0.000	0.003	0.002	0.001	0.008	0.004	0.003	0.005	0.009	0.006	0.011	0.003	0.009
3.Class	1	0.963	0.980	0.981	0.959	0.984	0.993	0.993	0.991	0.973	0.997	0.987	0.996	0.982	0.988	0.984
0.790	2	0.035	0.018	0.016	0.040	0.013	0.006	0.004	0.007	0.026	0.002	0.011	0.002	0.015	0.010	0.013
	3	0.002	0.002	0.003	0.001	0.003	0.001	0.003	0.001	0.001	0.001	0.002	0.002	0.003	0.002	0.002

Table 2: Response probabilities for the 15 symptoms at T_2
(1= not present/weak, 2 = average/heavy, 3 = not observed)

Class 1: The first and most interesting class with respect to the degeneration of the symptom structure at T_2 represents about 4% of the population. It is dominated by a high probability for *non-recording* for all items and a much lower but still considerable probability for category 1 (*not present/weak*). The latter holds true only for the first 8 symptoms, written down on the left side of the page (Table 1). For item 9-15 the probability of non-recording is almost 1. This class represents both an artifact from the particular lay-out of the questionnaire and those patients who showed up at T_2 but were not recorded for their symptoms. Responseprobabilities show that the lack of a symptom and non-recording has something in common at least for some of the patients.

Class 2: The second class characterizes a subpopulation for which there is still a considerable probability to show one or more of the symptoms. The probability of a symptom not to be recorded is almost zero. This class represents about 17.5% of the total population. It is the class of patients not yet totally recovered from depression.

Class 3: The third class representing almost 80% of the population can be addressed as a class of patients which exhibit no more symptoms of depression after 8 weeks of treatment.

3.2 Four-class solution

To facilitate the interpretation of a 4-class solution we first cross tabulate the memberships for both the 3 and 4 class solutions. This table shows that the class 1 („non- recording“ class) splits into class 3 and 4 in the 4-class solution (Table 3).

	class 1	*class 2*	*class 3*	*class 4*
class 1	0	0	96	207
class 2	1451	0	0	0
class 3	0	6713	0	0

Table 3: Cross tabulation 3-class solution by 4-class solution

Class 1 and 2: The first class represents the same type of respondents as the second class of solution described earlier. It is the class of patients still exhibiting all or part of the symptoms. Class 2 coincides without exceptions with class 3 of the 3-class solution. It is, just as before the class of patients exhibiting no symptoms after 8 weeks of treatment with the antidepressive drug.

Class 3 and 4: Class 3 is the tiny class of patients who showed no symptoms for the first half of the questionnaire and for whom the second part of the symptom list was (therefore ?) not recorded. For class 4 the probability of „not- recorded“ is almost 1 for all 15 symptoms.

	Var.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	Kat.															
1.Class	1	0.300	0.380	0.895	0.495	0.615	0.694	0.690	0.679	0.569	0.828	0.600	0.940	0.790	0.826	0.792
0.174	2	0.695	0.615	0.104	0.501	0.383	0.305	0.301	0.317	0.428	0.168	0.391	0.054	0.199	0.170	0.199
	3	0.004	0.005	0.001	0.003	0.002	0.001	0.008	0.004	0.004	0.005	0.009	0.005	0.011	0.003	0.009
2.Class	1	0.963	0.980	0.982	0.959	0.984	0.993	0.993	0.991	0.973	0.998	0.987	0.997	0.982	0.988	0.984
0.790	2	0.035	0.018	0.016	0.040	0.013	0.006	0.004	0.007	0.026	0.002	0.011	0.000	0.015	0.010	0.013
	3	0.002	0.002	0.003	0.001	0.003	0.001	0.003	0.001	0.001	0.001	0.002	0.003	0.003	0.002	0.003
3.Class	1	0.886	0.917	0.969	0.917	0.927	0.917	0.887	0.897	0.061	0.019	0.040	0.019	0.011	0.021	0.021
0.011	2	0.114	0.073	0.021	0.072	0.063	0.062	0.062	0.062	0.010	0.000	0.000	0.000	0.000	0.000	0.010
	3	0.000	0.010	0.010	0.010	0.010	0.021	0.051	0.041	0.929	0.981	0.960	0.981	0.989	0.979	0.968
4.Class	1	0.036	0.029	0.012	0.014	0.019	0.005	0.000	0.005	0.022	0.022	0.027	0.027	0.017	0.022	0.022
0.024	2	0.000	0.000	0.000	0.000	0.005	0.005	0.000	0.019	0.015	0.005	0.019	0.010	0.014	0.010	0.010
	3	0.964	0.971	0.988	0.986	0.976	0.990	1.000	0.976	0.964	0.973	0.954	0.964	0.969	0.969	0.969

Table 4: Response probabilities for the 15 symptoms at T_2
(1= not present/weak, 2 = average/heavy, 3 = not observed)

The histograms for the probability of membership for each class show that it is particularly easy to allocate patients to class 3 and 4 and fairly precise for class 1 and 2. Particularly for class 3 and 4 the probability to belong to one of else classes is 1 for virtually all observations.

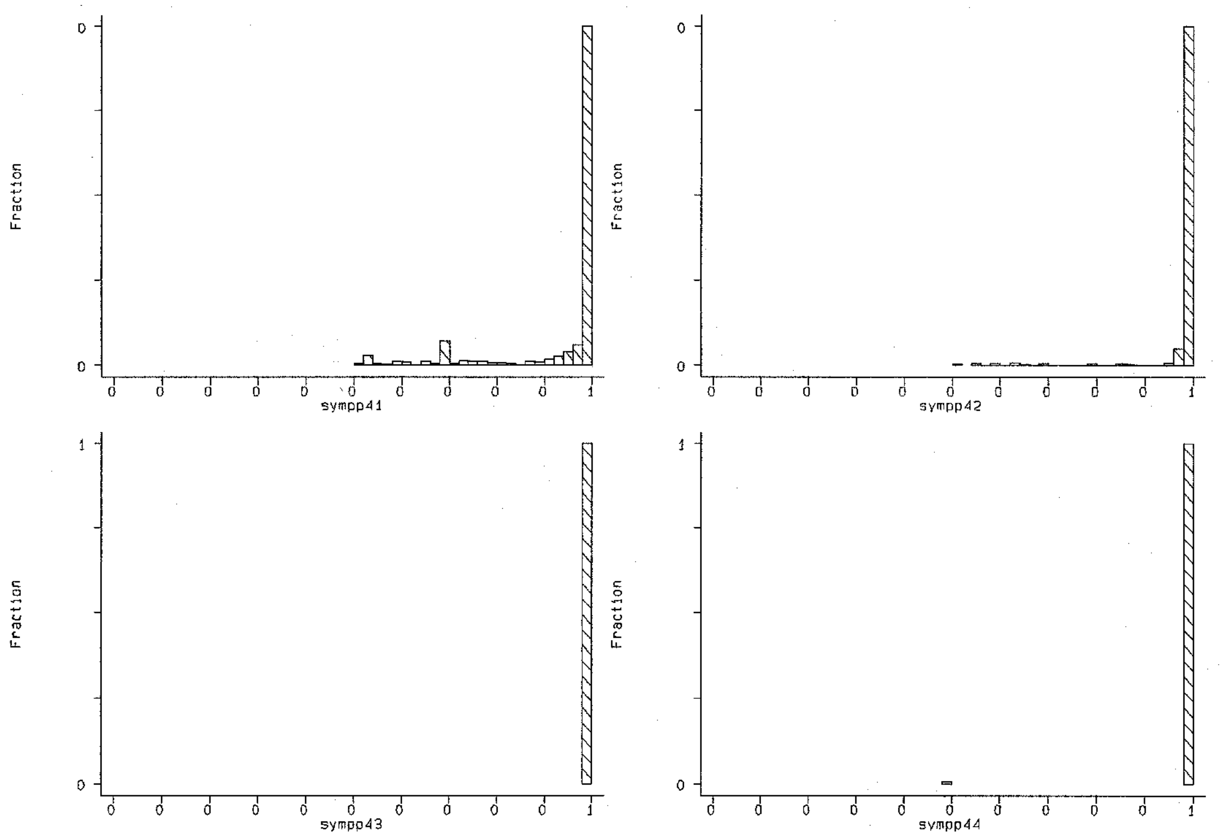

Figure 1: Membershipprobabilities for the 4-class solution

3.3 Four-class solution with equality constraints

The straightforward interpretation of particularly class 2,3 and 4 was very much facilitated by the similarity of the probabilities for certain categories. Therefore, it is of great interest to compare the fit of a solution where certain probabilities are constrained to be equal, with the 4-class solution just mentioned above. The bold figures in Table 5 indicate the probabilities constrained to be equal with respect to one particular category in one class. The BIC - value for this solution is 55861.578 which is still better then the BIC for the unconstrained 3-class solution *and* better then the unconstrained 5-class solution. The precision of allocating respondents to the classes is almost the same as for the unconstrained 4-class solution, therefore the histograms will not be presented here. The solution just discussed is furthermore an indication to prefer the 4-class solution, but still needs a correct statistical proof using another dataset, since the constraints employed result from a sound inspection of the unconstrained solution and are not derived from any elaborated theory.

	Var.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	Kat.															
1.Class 0.194	1	0.324	0.380	0.898	0.501	0.641	0.725	0.722	0.708	0.587	0.847	0.633	0.947	0.799	0.839	0.804
	2	0.670	0.573	0.100	0.495	0.357	0.273	0.270	0.287	0.410	0.149	0.359	0.049	0.190	0.158	0.187
	3	0.006	0.005	0.002	0.004	0.003	0.002	0.008	0.005	0.004	0.003	0.008	0.004	0.011	0.004	0.009
2.Class 0.770	1	**0.986**	**0.986**	**0.986**	**0.986**	**0.986**	**0.986**	**0.986**	**0.986**	**0.986**	**0.986**	**0.986**	**0.986**	**0.986**	**0.986**	**0.986**
	2	0.013	0.012	0.011	0.013	0.010	0.011	0.008	0.011	0.013	0.009	0.011	0.000	0.011	0.012	0.012
	3	0.001	0.002	0.002	0.000	0.003	0.002	0.005	0.002	0.001	0.004	0.003	0.014	0.002	0.002	0.002
3.Class 0.012	1	**0.898**	**0.898**	**0.898**	**0.898**	**0.898**	**0.898**	**0.898**	**0.898**	0.045	0.049	0.049	0.049	0.049	0.035	0.027
	2	0.093	0.072	0.040	0.079	0.087	0.055	0.044	0.055	0.004	0.000	0.000	0.000	0.000	0.014	0.022
	3	0.008	0.029	0.061	0.023	0.014	0.046	0.058	0.047	**0.951**	**0.951**	**0.951**	**0.951**	**0.951**	**0.951**	**0.951**
4.Class 0.024	1	0.016	0.016	0.016	0.008	0.008	0.008	0.008	0.004	0.000	0.012	0.008	0.009	0.008	0.012	0.012
	2	0.000	0.000	0.000	0.008	0.008	0.008	0.008	0.012	0.016	0.004	0.008	0.006	0.008	0.004	0.004
	3	**0.984**	**0.984**	**0.984**	**0.984**	**0.984**	**0.984**	**0.984**	**0.984**	**0.984**	**0.984**	**0.984**	**0.984**	**0.984**	**0.984**	**0.984**

Table 5: Response probabilities for the 15 symptoms at T_2 with equality-constraints
(1= not present/weak, 2 = average/heavy, 3 = not observed)

4. Explanation of membership by multinomial logit models

There is no essential difference between the restricted and two unrestricted solutions, because classes 3 and 4 are basically similar. Therefore only the membership probabilities for the 4-class solution (unconstrained) will be explained by means of the following exogenous characteristics.

Age
Gender of the patients. (0= male 1=female)

Severity of the depression at first screening (4-point scale)
0 - slight 1 - moderate 2 - heavy 3 - extremely heavy

Compliance measured (3-point scale)
0 - good 1 - moderate 2 - poor

Efficacy of treatment (5-point scale).
0 - total remission of symptoms
1 - partial remission of symptoms
2 - further treatment is necessary
3 - no change in symptoms
4 - symptoms become worse

To include an overall estimate of severity seems to be redundant at first glance, but is very informative because this variable is strongly associated with symptoms recorded at T_1. Table 6 shows the parameters exp(p) of the multinomial logistic regression (RRR = relative risk ratio). The parameter for severity shows that this variable increases the odd of being in class 1 (the relative risk ratio is 3.41) or 4 (the relative risk ratio is 2.2) instead of falling into class 2. The more severe the depression is classified at the first screening, the more it is likely to exhibit symptoms at T_2 *or that symptoms are not recorded at* T_2.

The same holds true for the compliance of the patients and the efficacy of the treatment. Bad compliance is strongly associated with class 1 or class 4, so bad compliance enhances the probability of not recorded symptoms at T_2. Minor effectiveness of the antidepressant drug is strongly associated with class 1 (RRR = 6.00) and 4 (RRR = 4.48) either, thus enhancing the probability of exhibiting symptoms or not to be recorded at the second point in time. Of course, the overall estimate of efficacy explains most of the dependent variable (30%) whereas severity and compliance explain both about 4%. Age and gender of the patients are of less importance, which seems to be surprising, as bias due to non-recording is expected to be dependent at least from the age of the patients.

Number of obs = 8211 chi2(15) = 2844.25
Prob > chi2 = 0.0000 Pseudo R2 = 0.3593

sympkl4	*RRR*	*Std. Err.*	*z*	*P>|z|*
1				
gender	1.088366	.1023874	0.900	0.368
age	1.011694	.0030866	3.811	0.000
severity	3.413889	.2590507	16.181	0.000
compl.	2.072039	.1684454	8.962	0.000
efficacy	6.001707	.3418422	31.463	0.000
3				
gender	1.148822	.2914684	0.547	0.584
age	1.005222	.0083705	0.626	0.532
severity	1.134248	.2169665	0.659	0.510
compl.	0.884225	.3373408	-0.323	0.747
efficacy	1.086939	.1985779	0.456	0.648
4				
gender	0.943701	.3020966	-0.181	0.856
age	1.005704	.0102384	0.559	0.576
severity	2.205047	.5380364	3.241	0.001
compl.	2.935095	.6303643	5.014	0.000
efficacy	4.477691	.7433052	9.031	0.000

(Outcome sympkl4 = 2 is the comparison group)
Table 6: Multinomial logit model without interaction effects

So far we have considered only main effects in the models for latent class membership. We now extend the model to include interaction effects for severity*efficacy, compliance*efficacy and severity*compliance. Only two of them show a considerable impact and will be interpreted. The conditional effects of both severity, compliance are now the effects of one of these variables conditional on the value 0 of the other variable, e.g., the effect of compliance conditional to only slight depression at T_1, or the effect of severity conditional to a good compliance. These effects on class 4 vanished statistically, as the effect of compliance obviously depends on the severity estimated at T_1.

Conditional effects (derived from interaction parameters) can be calculated just as is done in linear models using the parameters of the model instead of the RRR = exp(p) simply be taking the ln(RRR). To calculate the effect of compliance conditioning for the four categories of severity (0,1,2,3) the following expression is used:

$$\ln(0.5533723) + \ln(2.063339)*\text{severity} \quad (1)$$

This can be rewritten as

$$-0.59172 + 0.72432*\text{severity} \quad (2)$$

Taking:

$$\exp(-0.59172 * 0.72432*\text{severity})$$

provides the RRR of compliance conditioned on a particular value of severity, i.e., 0,1,2 or 3. If the severity-code for depression is 0, the expression simplifies to exp(–0.59172) = 0.5533723, the very parameter for the „simple effect“ of compliance. This parameter is printed out in bold face in Table 7. Of course we don't know anything about the statistical significance for the effect of compliance conditioning on severity code 1,2 or 3, but we are able to compare the size of the effects, comparing the relativ risk ratios (RRR) which becomes 2.3559 for heavy depression (code 2) and 4.861 for extremly heavy depression (code 3).

The interaction between compliance and efficacy shows an effect on the probability of being a member of class 3. The effect of compliance becomes considerably large only for the categories 2-4 of efficacy. Only little efficacy *and* poor compliance result in an increased probability of falling into this class. Nevertheless, it must be emphasized that the Pseudo-R^2 is not increased substantially by adding interaction effects of any kind to the model.

Number of obs = 8211 chi2(24) = 2881.78
Prob > chi2 = 0.0000 Pseudo R2 = 0.3641

sympkl4	*RRR*	*Std. Err.*	*z*	*P>|z|*
1				
gender	1.094008	.1027656	0.956	0.339
age	1.011173	.0030779	3.650	0.000
severity	4.394044	.654082	9.944	0.000
compl.	5.970739	2.023021	5.274	0.000
efficacy	8.532867	1.947426	9.394	0.000
se	.9010482	.0839519	-1.118	0.263
ce	.7202218	.0649128	-3.641	0.000
sc	.8046095	.096302	-1.816	0.069
3				
gender	1.148427	.2914838	0.545	0.586
age	1.006321	.0083843	0.756	0.449
severity	.9539857	.2282138	-0.197	0.844
compl.	.3681148	.5289198	-0.696	0.487
efficacy	.4382994	.3130899	-1.155	0.248
se	1.439419	.4419584	1.186	0.236
ce	2.152986	.7284263	2.267	0.023
sc	.8486928	.5031226	-0.277	0.782
4				
gender	.9454203	.3038156	-0.175	0.861
age	1.006218	.0102995	0.606	0.545
severity	.7282141	.3464267	-0.667	0.505
compl.	.5533723	.5499147	-0.595	0.552
efficacy	2.535159	1.514341	1.557	0.119
se	1.288204	.3417375	0.955	0.340
ce	.9106015	.203747	-0.419	0.676
sc	2.063339	.7481438	1.998	0.046

(Outcome sympkl4 = 2 is the comparison group)
Table 7: Multinomial logit model with interaction effects
se = severity * efficacy, ce = complianc * efficacy, sc = severity * compliance

5. Discussion

The results presented show that to record a particular symptom is not independent from specific characteristics of the observations and the layout of the questionnaire. It is not surprising that both the severity of the depression recorded at T_1 and the efficacy of the treatment recorded at T_2 increase the probability of being a member of class 1 (still exhibiting symptoms after eight weeks of treatment. The effect on the probability of class 4 sheds light on the mechanism of „non-recording“, indicating that this kind of drop-out is not random at all. Exploratory latent- class analysis turned out to be the very tool to identify subgroups of subjects prone to this mechanism. In this example it was shown that only a small subgroup of patients (303) was not or only partially recorded after treatment and should be excluded from further analysis. The overall simplicity of the structure of symptoms shows that there was only little differential weighting of the symptoms at T_2 at all. Regarding class 3 it was possible to identify another subgroup, where only the first part (8 symptoms on the left half of the symptomlist) was recorded. The probability of being a member of this class hardly depends on any of those characteristics used in the regression model. It might be of great interest to explore in detail the characteristics of the subgroups mentioned above in order to gain a deeper insight into the mechanisms of non-recording. But this would go far beyond the scope of this study.

References

Formann, A.K. (1984). *Die Latent- Class Analyse*. Weinheim: Beltz.

Hagenaars, J. (1993). *Loglinear models with latent variables*. Newbury Park: Sage.

Langeheine, R. (1988). New developments in latent class theory. In: Langeheine R & Rost J (eds.), *Latent trait and latent class models*. New York: Plenum.

Little, R.J.A. & Rubin, D.B. (1987). *Statistical analysis with missing data*. New York: Wiley.

Little, R.J.A. & Rubin, D.B. (1990). The analysis of social science data with missing values. *Sociological Methods and Research*, 18, 292-326.

Rost, J. (1988). *Quantitative und qualitative probabilistische Testtheorie*. Bern: Huber.

Rost, J. (1990). *LACORD. Latent class analysis for ordinal variables. A FORTRAN Program. 2nd edition.* Kiel: IPN.

Van Buuren, S. & Van Rijckevorsel, J.L.A. (1992). Imputation of missing categorical data by maximizing internal consistency. *Psychometrika*, 57, 567-580.

Vermunt, K.J. (1993). *Log-linear and event history analysis with missing data using the EM algorithm.* Tilburg: Work and Organizations Research Centre, WORC PAPER 93.09.015/7