Chapter 28

Self Monitoring - A Class Variable?

Matthias von Davier and Jürgen Rost

IPN - Institute for Science Education, Kiel

1. Trait or types: On the nature of self-monitoring

The self monitoring (SM) construct was introduced by Snyder (1974) as a trait that describes and explains individual differences in the (self-) control of expressive behavior:

„There are, however, striking and important differences in the extent to which individuals can and do monitor their self-presentation, expressive behavior and non-verbal affective display" (Snyder, 1974, p. 526).

The SM-questionnaire developed by Snyder (1974) covers self-report items like 'I'm not always the person I appear to be' that usually were analyzed in a quantitative fashion, i.e., by summing the item responses after coding all items in the same direction. In hundreds of empirical studies, however, a median-split was applied to these sum scores in order to differentiate between two groups of people, the *high* self monitorer and the *low* self monitorer. The results of those investigations strongly depend on how well two possibly different kinds of individuals, the high SM and low SM, are represented by such a fifty-fifty splitting of the sample.

In the further development, the theoretical construct of SM seemed to adopt to this way of data analysis and SM explicitly changed to a type construct:

„...individuals high and low in self-monitoring are guided by different meta-controls of behavior in social contexts.

... high self monitoring people ... consider the self performances and strategic appearences they generate.

... low self monitoring individuals ... consider the consistency of the self's deeds with ones life world of beliefs (Snyder and Campbell, 1982, p. 323).

The quotations above not only express different kinds of individual differences, i.e., differences *in degree* vs. differences *in type* (see Meehl, 1992), but they also differ in the level of description. The 1974 statement defines self monitoring by means of manifest behavioral differences, whereas in the 1982 statement different meta-controls resulting in two distinct behavioral patterns are postulated.

In order to establish the type-nature of the SM construct, Gangestad & Snyder (1985) showed that by applying taxonometric procedures (Meehl & Golden, 1982) a main part of the variance of a certain SM subscale can be explained by a dichotomous (or binary) latent variable. In accordance with their expectations, the two groups of individuals established by the latent binary variable showed conditional response probabilities on this SM subscale that can be identified as high- and low- self monitoring behavioral patterns.

Miller & Thayer (1989) criticized that the results of Gangestad & Snyder, among other things, were caused by the calculation of correlations for dichotomous variables. The latent *binary* variable, therefore, seemingly explains a greater part of variance than a latent *continuous* variable. This critique of the results of Gangestad & Snyder mainly points out problems which arise by applying methods for continuous variables to categorical data.

In the present paper analyses with a family of models designed for multivariate, *categorical* data are reported. These models are derived by extending the latent class analysis (LCA, Lazarsfeld & Henry, 1968, see also chapter 1 section 2.1) and the Mixed Rasch model (MRM, Rost 1990, 1991) to a kind of „hybrid" model.

The aim of these analyses is to compare the fit of different models with discrete or continuous latent variables, respectively, in order to check whether Gangestad & Snyder's postulate of a binary latent variable is appropriate. The results of these analyses are presented in section 3, after a short description of the model.

2. Hybrid Rasch-LC models

As stated in chapter 1, section 2.5, it is possible to assume different models within different latent classes. In the hybrid model (Yamamoto, 1989), for instance, it is assumed that the population consists of two latent classes (subpopulations) where in class 1 the Rasch model holds and in class 2 a LCA structure, i.e., independence of the manifest variables is assumed.

With regard to the self monitoring construct it is reasonable to extend the 2-class hypothesis in the direction that the high self monitoring class is a Rasch-type class, where individual differences in the tendency to control expressive behavior exist. In contrast, in the low self montoring class, no quantitative differences with regard to the control of expressive bahavior are expected to be found. From this perspective, the self monitoring questionnaire only measures people that *do* monitor their expressive behavior. Low self monitoring people would be *unscalables* in the light of that questionnaire, as these people are fitted by a model without individual differences.

The hybrid Rasch-LC models in their general form allow for free mixtures of Rasch models and latent class models for dichotomous and polytomous data (cf. v. Davier, 1995, see also equation (50) in chapter 1, section 2.5). The probability of a response pattern $\mathbf{x}_v$ in the case of a hybrid Rasch-LC model is

$$P(\mathbf{x}_v) = \sum_{g=1}^{G} \pi_g \, \pi_{rg} \frac{\exp\left(-\sum_{i=1}^{k} \sigma_{ixg}\right)}{\gamma_{rg}} + \sum_{h=1}^{H} \pi_h \prod_{i=1}^{k} \pi_{ixh} \,, \tag{1}$$

where G is the number of Rasch model classes, and H is the number of ordinary LCA classes assuming independence of the manifest variables. In the Rasch model classes, the conditional score probabilities π_{rg} of all raw scores $r \in \{0,\ldots,mk\}$ in classes $g \in \{1,\ldots,G\}$ have to be estimated (m is the number of response categories minus one). To avoid the estimation of such a large number of nuisance parameters, a discrete two parameter distribution can be assumed for the conditional score probabilities in any of the Rasch classes. The parameters π_{rg} are approximated by the distributional assumption

$$\pi_{rg} = \frac{\exp\left(\frac{r}{mk}\tau_g + 4\frac{r(mk-r)}{(mk)^2}\delta_g\right)}{\sum_{s=0}^{mk}\exp\left(\frac{s}{mk}\tau_g + 4\frac{s(mk-s)}{(mk)^2}\delta_g\right)} \tag{2}$$

with location parameter τ_g and dispersion parameter δ_g (Rost & v. Davier, 1995). In contrast to mk parameters in the unrestricted case, this logistic restriction needs only 2 parameters to approximate a latent score distribution.

The general model structure of free mixtures of Rasch models and latent class models allows for testing hypotheses about subpopulations by comparing the fit of different models. Within the framework of these hybrid Rasch-LC models it can be checked, for instance, for which of the latent classes a trait variable has to be assumed.

In addition, in the case of polytomous data, it is possible to impose different restrictions on the class specific item parameters (e.g., different restrictions for the threshold parameters in mixtures of polytomous Rasch models or latent class models for ordinal data (see eq. (17), (18) and (39), (40) in chapter 1)). The program system WINMIRA (von Davier, 1994) used for the analyses is a stand-alone Microsoft Windows 3.1x application, which is capable of estimating various hybrid Rasch-LC models for dichotomous and polytomous data.

3. The analysis of SM data

The analyses are based on the original data set by Gangestad & Snyder (1985) covering responses of 1919 individuals on 25 items of the self monitoring scale. Snyder & Gangestad (1985) revised the self monitoring scale, removed 7 items which „discriminated poorly“, and proposed a new 18-item measure. The revised scale is listed in Table 1. In this table, the 8 items used by Gangestad & Snyder (1985) for their taxonometric analysis are printed in *italics*.

We re-analyzed the 8 item selection in order to be comparable with the results of Gangestad & Snyder (section 3.1). In a second step, we applied a parametric bootstrap procedure for goodnes of fit tests in order to confirm the validity of the selected model for the data (section 3.2).

3.1 Re-analysis of the 8 item self monitoring scale

In a first step of analyses, the models were compared and selected by means of an information criterion, the CAIC which is an extension of the AIC (Akaike, 1973; Bozdogan, 1987). The CAIC is defined as

$$\mathrm{CAIC}_M = -2\log L_M + N_{par}\left(1+\log(N)\right) . \tag{3}$$

where N is the sample size, N_{par} is the number of model parameters and $\log L_M$ is the log-likelihood of model M. According to the minimum CAIC criterion, the model M with the lowest value on this criterion is assumed to fit the data best.

1. I find it hard to imitate the behavior of other people. (F)(.39)
2. At parties and social gatherings, I do not attempt to do or say things that others will like. (F)(.20)
3. *I can only argue for ideas which I already believe. (F)(.24)*
4. I can make impromptu speeches even on topics about which I have almost no information. (T)(.39)
5. *I guess I put on a show to impress or entertain others. (T)(.48)*
6. *I would probably make a good actor. (T)(.59)*
7. *In a group of people I am rarely the center of attention. (F)(.45)*
8. *In different situations and with different people, I often act like very different persons. (T)(.25)*
9. I am not particularly good at making other people like me. (F)(.28)
10. I'm not always the person I appear to be. (T)(.22)
11. I would not change my opinions (or the way I do things) in order to please someone or win their favor. (F)(.17)
12. *I have considered being an entertainer. (T)(.41)*
13. I have never been good at games like charades or improvisational acting. (F)(.49)
14. I have trouble changing my behavior to suit different people and different situations. (F)(.34)
15. *At a party I let others keep the jokes and stories going. (F)(.45)*
16. I feel a bit awkward in public and do not show up quite as well as I should. (F)(.31)
17. I can look anyone in the eye and tell a lie with a straight face (if for a right end). (T)(.30)
18. I may deceive people by being friendly when I reallv dislike them. (T)(.18)

Table 1: Eighteen-item scale of self-monitoring. The keys T (true) or F (false) indicate the direction of item formulation. Item loadings on the first unrotated factor are given in parentheses.

Table 2 shows the log Likelihood and CAIC values as well as the number of estimated parameters for the one- and two-classes models.

	logL	Npar	CAIC
1LC	-10072.9	8	20214.4
RM	-9691.5	9	19477.2
2LC	-9611.2	17	19367.9
1RM1LC	-9585.6	18	19325.4
2RM	-9571.1	19	19304.6

Table 2: Goodness of fit statistics for 1- and 2-class solutions of the latent class, the hybrid Rasch-LC and the mixed Rasch models.

The „non mixture", i.e., one-class models are the Rasch model (RM) and the 1-class latent class analysis (1LC), i.e., the model of multinomial independence in the whole population. In these and the following analyses, the two parameter restriction of the score probabilities as described in section 2 was used when estimating the parameters of Rasch model classes. The two-classes models are the 2-class LCA, the hybrid model (1RM1LC) (one Rasch homogeneous class, one LCA class), as well as the 2-class mixed Rasch model (2RM).

According to Gangestad & Snyder the model with two homogeneous subpopulations, i.e., the 2-class latent class analysis (2LC), should fit the data better than the ordinary unidimensional Rasch model (RM). The CAIC-values of Table 2 show that this expectation is justified, since the CAIC of 2LC is smaller than that of the RM.

This means that the assumption of two homogeneous subpopulations gives a better representation of the data than the assumption of just one population with quantitative individual differences according to the Rasch model.

Figure 1 shows the conditional response probabilities of the two subpopulations LCA (2LC).

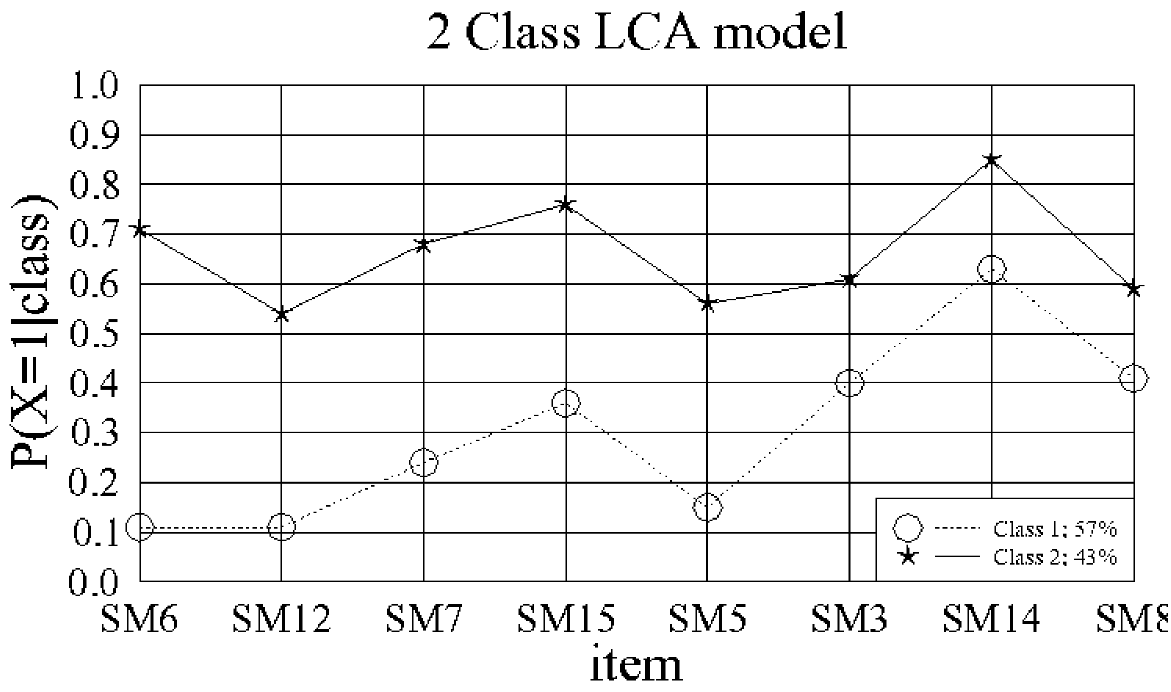

Figure 1: Conditional response probabilities of the 2-classes LCA model for the 8 SM-items ordered according to the differences between the classes.

A large difference in degree between the classes can be found, that is, class 1 has low response probabilities for all items whereas class 2 shows quite high probabilities of positive responses (the response patterns have been coded according to the key given in Table 1, i.e., high response probabilities mean a high tendency to answer in the way self monitoring individuals would do).

The shape of the profiles in Figure 1 support the hypothesis of the the typological nature of self monitoring, as the differences between the class-specific item difficulties vary among the items, so that a simple latent trait model cannot fit the data. As mentioned before, the CAIC values show that one population with quantitative variation does not suffice to describe the data.

	logL	Npar	CAIC
3LC	-9560.4	26	19343.3
1RM2LC	-9535.9	27	19302.1
2RM1LC	-9510.6	28	19260.8
3RM	-9507.4	29	19263.1
4LC	-9517.8	35	19335.3
4RM	-9479.9	39	19293.7

Table 3: Goodness of fit statistics for the 3- and 4-classes models.

In contrast to the assumption of Gangestad & Snyder, however, our analyses indicate the need for a more complex model, since the 2-classes hybrid model as well as the mixed Rasch model have a smaller CAIC than the 2LC model. In order to check whether even more classes have to be assumed, additional analyses were carried out with mixtures of three and four classes (see Table 3).

Two hybrid Rasch-LC models were estimated in addition to the 3- and 4-classes latent class and mixed Rasch models.

According to the CAIC-values, the three-classes hybrid-model with two Rasch model classes (2RM1LC) fits the data best. Figure 2 shows the response profiles of the 2RM1LC model's subpopulations by means of their (mean) response probabilities.

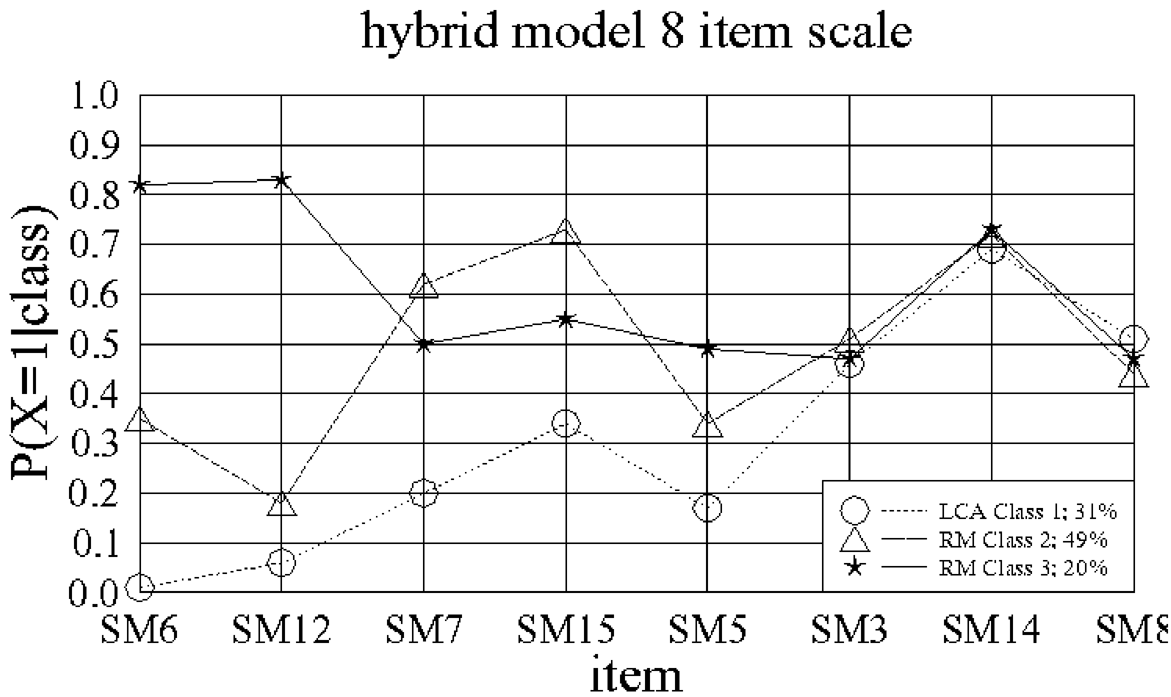

Figure 2: Mean response probabilities of the 3-classes Hybrid Rasch-LC model for the 8 SM items.

By comparing Figures 1 and 2 it can be seen, that the profile of the ordinary latent class has not changed much. This class (marked with O in Figure 2) has very low probabilities of a positive response on 5 items, SM6, SM12, SM7, SM15, and SM5. This class seems to consist of people who do *not* tend to monitor their expressive behavior.

Item SM3: *„I can only argue for ideas which I already believe“* shows almost identical difficulties in all classes. Likewise, only small differences among the classes exist for the items SM8 and SM14 (see Table 1 for the item text). Therefore, these items do not contribute to the interpretation of the class differences.

The Rasch-homogeneous subpopulations (class 2 and 3) are characterized by different item parameter vectors, or mean response probabilities, respectively. For these classes, there are large differences between response probabilities for item SM6: *„I would probably make a good actor“* and item SM12: *„I have considered being an entertainer“*. The probabilities of an agreeing response are very high for these items (SM6: p = 0.82 and SM12: p = 0.85) in class 3 as compared to class 2. Both items emphasize the professional side of behavioral control. In class 2, positive answers on these items are less likely, i.e., p = 0.32 for SM6 and p = 0.15 for item SM12. A similar (but smaller) difference exists for item SM5 *„I guess I put on a show to impress or entertain others“* between the two Rasch homogeneous classes.

Obviously, there are two different kinds of controlling expressive behaviour, one kind employing actor-like skills and techniques of behavioral control, and another kind without playing such roles. According to this result, self monitorer differ in both, the *extend* to which they monitor their self presentations and the *kind* of doing it, i.e., like an actor or not.

Whereas both kinds of self monitorer vary in their degree of self-control (they form Rasch classes), low self-monitorer do not vary in degree, i.e., they form an ordinary LCA-class. It may be regarded as a confirmation of the interpretation of the two high SM classes that the low SM class has the lowest response probabilities for the „actor-type“ items SM5, SM6 and SM12 (compare Figures 1 and 2).

3.2 Goodness of fit tests by means of the parametric bootstrap

The parametric bootstrap method is a resimulation technique where a number of b data sets are simulated on the basis of the parameter estimates of a real data set (Efron and Tibshirani, 1993). The „empirical“ distribution of a goodness of fit statistic, like the Pearson chi-square X^2, obtained for this sample of simulated data sets, serves as the basis for statistical inference. A common decision rule would be that the model under consideration is accepted, i.e., is assumed to fit the data, if the fit statistic of the real data set is smaller than at least 5% of the simulated values of that fit statistic.

The results of a Monte Carlo study on the parametric bootstrap method for extremely sparse tables by von Davier (1997) provide evidence that the Pearson X^2 and the Cressie-Read statistic (CR), see Cressie and Read (1984), are suitable fit statistics for the bootstrap in the case of tables like the 2^8 table of the self monitoring scale. However, the bootstrap criterion turned out to be much stronger than information criteria like the CAIC and, in our case of the 8-item scale, the bootstrap even rejects the 3-classes hybrid model presented above. In particular, none of b = 40 bootstrap samples showed a fit statistic as high as the fit value of the real data set when all possible 2- and 3-classes hybrid models were examined.

In order to avoid the assumption of an even higher number of classes we tried to keep the 3-classes hybrid model for the self monitoring scale by replacing odd items by possibly better ones drawn from the remaining 10 items of the original 18-item pool. The criterion for selecting poorly fitting items was the item-Q index by Rost and von Davier (1994). According to this item fit measure, the three items SM3, SM8, and SM14 (see Table 1) were removed, because they showed a significant misfit.

Classes	Model	X^2	CR
2	2LC	.000	.000
2	1RM1LC	.000	.000
2	2RM	.000	.000
3	3LC	.000	.000
3	1RM2LC	.025	.000
3	2RM1LC	.125	.150
3	3RM	.400	.400

Table 4: Bootstrap results for different 2- and 3-class models: LC stands for an ordinary latent class (multinomial independence), RM for a Rasch class.

They were replaced by those two items of the remaining 10 items that had highest factor loadings in the one-factor solution of the entire set of 18 items, i.e., items SM1 and SM4. Table 4 presents the bootstrap results for this new 7-item scale showing that the 3-classes hybrid model with 2 Rasch classes is the most parsimonious model that fits these data.

In this table, a p-value of, e.g., p=0.125 means that 12.5 % of the simulated data sets show a higher (i.e., worse) fit value than the real data set. In this sense, the p-values represent the level of significance under which the null-hypothesis of model fit can be retained.

Although the 3-class mixed Rasch model has an even better fit, we prefer the more parsimonious hybrid model, because the assumption of a low self monitoring class *without* any variation in degree is theoretically more appealing. In fact, the item profiles of the hybrid model for this new 7-item scale confirm the interpretation given for the 8-item selection: there is a low self monitoring class covering 27 % of the population and showing no variation in their reponse behavior. The two Rasch classes, again, show substantial differences on the „actor“ and „entertainment“ items SM6 and SM12.

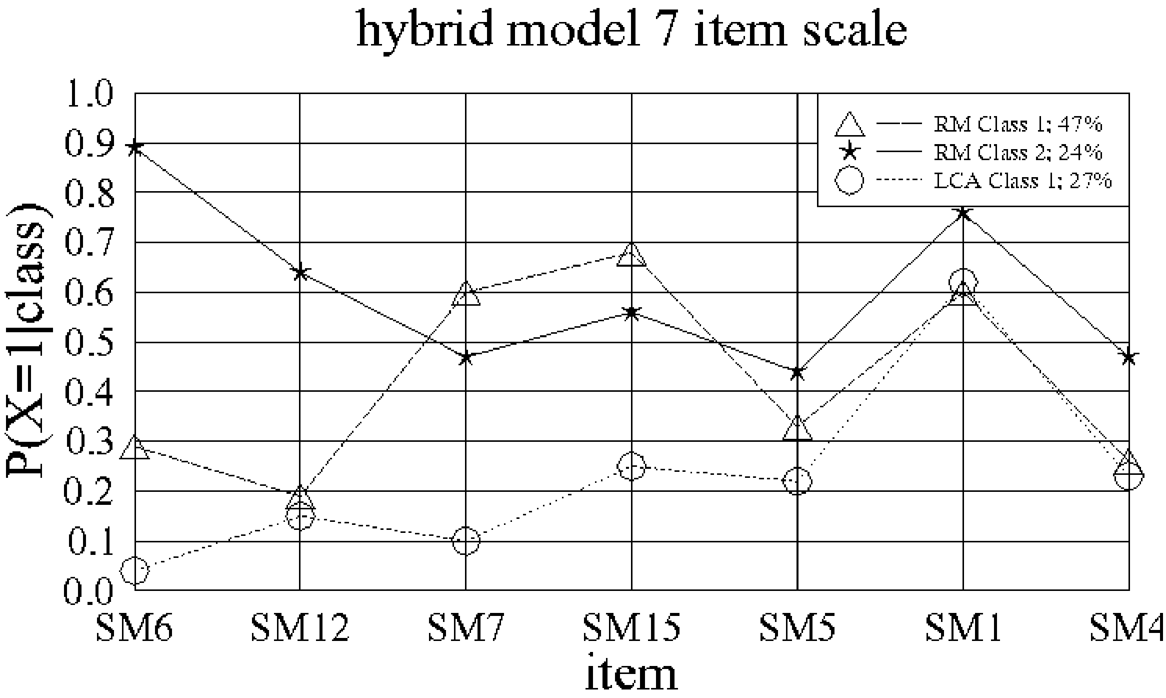

Figure 3: Conditional response probabilities of the remaining 10 item for the two subsamples gained from the 8 items analysis.

As a result, there is a high stability of the 3-class solution even when some items were replaced, and we now have a 7-item short version of the SM-scale that is fitted by a typological psychometric model.

4. Discussion

The analyses of the self monitoring scale showed that there are different types or classes of people *differing in kind* with respect to controlling expressive behavior. Gangestad & Snyder's hypothesis is partly supported, since a homogeneous group of low self monitoring people has been identified in the analysis of the 8-item scale.

In contrast to their hypothesis, the group of high self monitoring people was found not to be homogeneous. According to our results, it is more appropriate to allow for both qualitative and quantitative differences within the group of high self monitorers. This group not only shows individual *differences in degree* of controlling expressive behavior, what low self-monitoring people do not do, but they can also be subdivided in two Rasch homogeneous classes with different item parameters. The main difference between the two classes of high SM individuals can be found with respect to items with a content addressing almost professional behavioral control like *„I would probably make a good actor"*. In one Rasch class (class size 20 percent), these items have very high probabilities for a positive response. In the other Rasch class (49 percent), these items are not that easy.

The analyses with discrete mixture distribtion models showed that the self monitoring data should be described by a hybrid Rasch-LC model allowing for quantitative differences in some subpopulations (differences in degree) as well as qualitative differences (differences in kind) between the classes. We employed very different methods for analyzing the self monitoring data, but, in the end, we can join Gangestad & Snyder (1985) in pointing out that *"... discreteness in personality can be shown to exist"*, though the structure might be more complex than expected.

Acknowledgements

We are indebted to Mark Snyder and Steve Gangestad for allowing us to re-analyse their data. This research was part of the DFG-Project: Development and validation of psychometric mixture distribution models.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In: *Petrov, B.N. & Csaki, F.: 2nd international symposium on information theory (Akademiai Kiado, Budapest).*

Bozdogan, H. (1987). Model selection and Akaikes information criterion (AIC): The general theory and its analytical extensions. *Psychometrika, 52,* 3, 345-370.

Davier, von, M. (1994). WINMIRA: A program system for analyses with the Rasch model, with the latent class analysis and with the mixed Rasch model. Kiel: *Institute for Science Education* (IPN).

Davier, von, M. (1995). Mixtures of polytomous Rasch models and latent class models for ordinal variables. In F. Faulbaum & W. Bandilla (eds.), *Softstat 95 - Advances in Statistical Sofware 5*. Stuttgart: Lucius & Lucius.

Gangestad, S. & Snyder, M. (1985). „To carve nature at its joints“: On the existence of discrete classes in Personality. *Psychological Review, 92,* 3, 317-349.

Lazarsfeld, P.F. & Henry, N.W. (1968). *Latent structure analysis,* Boston: Houghton Mifflin.

Meehl, P.E. & Golden, R.R. (1982). Taxometric methods. In J. N. Butcher & P.C. Kendall (Eds.*) The handbook of research methods in clinical psychology,* (pp.127-181). New York: Wiley.

Meehl, P.E. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality, 60,* 1, 117-174.

Miller, M.L. & Thayer, J.F. (1989) On the existence of discrete classes in personality: Is self monitoring the correct joint to carve? *Journal of Personality and Social Psychology, 57,* 1, 143-155.

Rost, J. (1990). Rasch models in latent classes: An integration of two approaches to item analysis. *Applied Psychological Measurement, 3,* 271-282.

Rost, J. (1991). A logistic mixture distribution model for polychotomous item responses. *The British Journal for Mathematical and Statistical Psychology, 44,* 75-92.

Rost, J. & von Davier, M. (1995) Mixture distribution Rasch models, in: Fischer, G. & Molenaar, I. (eds.): Rasch models: *Foundations, recent developments and applications*. New York: Springer.

Snyder, M. (1974). Self Monitoring of expressive behavior. *Journal of Personality and Social Psychology, 30,* 526-537.

Snyder, M. & Gangestad, S. (1986). On the nature of self-monitoring: Matters of assessment, matters of validity. *Journal of Personality and Social Psychology, 51,* 1, 125-139.

Yamamoto, K. (1989). A Hybrid model of IRT and latent class models. ETS *Research Report* (RR-89-41), Princeton, NJ: Educational Testing Service.