Chapter 38

Modeling Stability and Regularity of Change: Latent Structure Analysis of Longitudinal Discrete Data

Thorsten Meiser and Georg Rudinger

Institute of Psychology, University of Bonn

The decisive step in the statistical analysis of longitudinal data is to „translate“ substantive developmental hypotheses into a formal model that accounts for the presumed data generating process. This modeling approach aims at testing a statistical model which represents the scientific model of development, which may be a psychological, sociological or educational model. In many fields of longitudinal research, quantitative or qualitative latent variables are assumed to affect the observations at each of the measurement occasions (cf. Henning & Rudinger, 1985). In this case, individual developmental sequences are defined as series of *locations* on a *latent continuum* or as series of *stages* referring to a *discrete latent variable*. Hypotheses concerning the direction, the speed and the universality of change on a latent continuum or concerning the order of passing stages and the duration time in the categories of a discrete latent variable are part of the scientific concept of development. The major advantage of latent structure models is that they facilitate the specification of such developmental assumptions directly in terms of latent variables and appropriate measurement models (Erdfelder, 1990).

In this chapter, two applications of latent structure analysis to longitudinal discrete data are presented. The first one focusses on models with discrete latent variables for each measurement occasion. It is demonstrated that *regularity of change* can be modeled by both restrictions of transitions between latent stages as well as constraints on the variability of parameters that are part of the measurement models. In the second application, we present models comprising quantitative latent dimensions. In this context different aspects of *stability* are considered, including different restrictions on stability for different latent dimensions or for different subpopulations.

1. Development of Deductive Reasoning

Here, a reanalysis of data on the development of deductive reasoning is described. The data and the original analysis are published by Schröder, Edelstein and Hoppe-Graff (1991, pp. 181ff.). The sample consists of $N = 101$ children from a longitudinal study in Iceland. Data are taken from three measurement occasions, i.e., from the assessments at the children's ages of 9, 12 and 15 years. From each point in time, two types of reasoning tasks are used for the analysis. The first type is related to experiences each child presumably had before (premise: „If there is a fire drill at school, the school bell rings“). The second type is characterized by decontextualization, that means it does not refer to real-life experiences (premise: „If I travel to A, I pass B“). Each task was considered as „solved“, if four questions involving deductive reasoning were answered correctly, and as „not solved“ otherwise. Consequently, two dichotomous indicators result per measurement occasion, i.e reasoning related to experience (RE) and reasoning given an abstract premise (RA). It is assumed that at each occasion a discrete latent variable θ affects the observed reasoning performance. The resulting latent structure model is displayed in Figure 1.

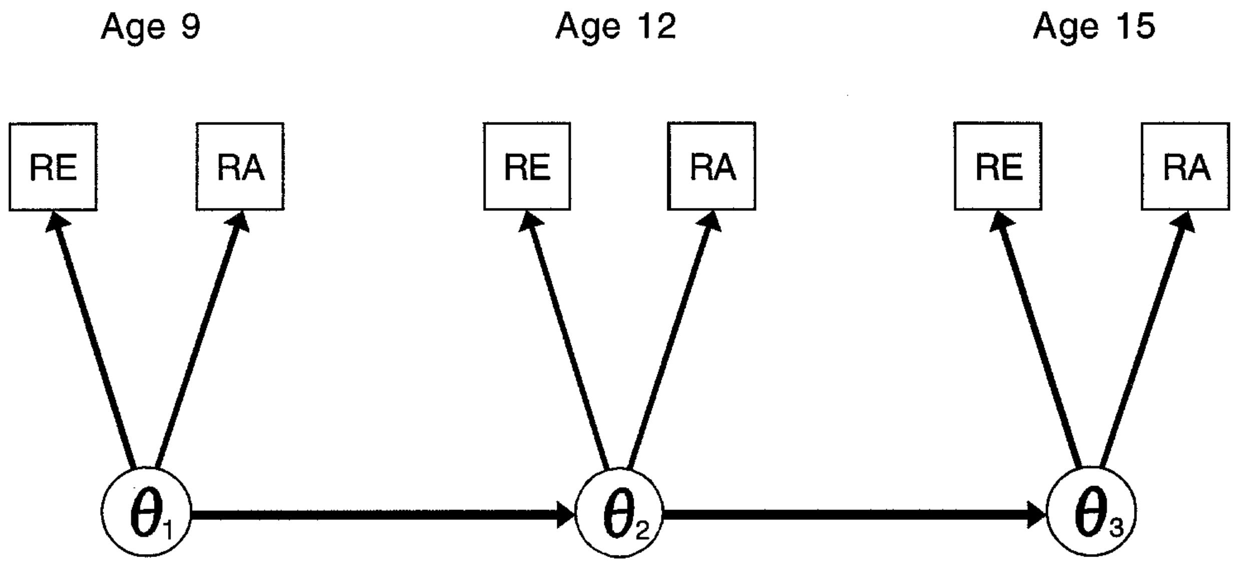

Figure 1: Dynamic latent structure model for the deductive reasoning data

1.1 Stage-Sequential Assumptions and Model Specification

The developmental model is based on two assumptions: First, the process of change is confined to a process of *cumulative growth*, since - according to Piagetian tradition - the construction of cognitive abilities is regarded as irreversible. Second, although both tasks require the same cognitive operations of handling conditional implications, it is assumed that mastering the task related to experience, RE, precedes mastering the abstract task, RA. This hypothesis corresponds to Piaget's concept of *décalage*, stating that mastery of logically equivalent tasks may be acquired with a time-lag due to task characteristics. Hence, the developmental model combines the assumptions of cumulative growth and of a *universal sequence of generalization* in deductive reasoning ability, namely from mastering contextual tasks to mastering abstract tasks.

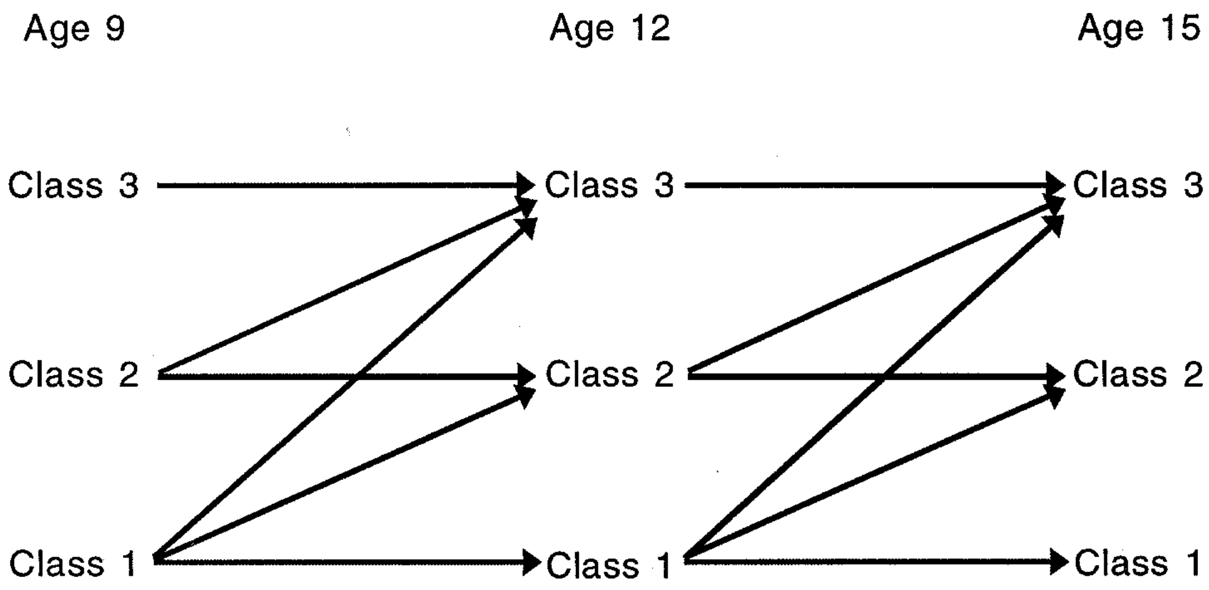

Figure 2: Latent transitions admitted by the stage-sequential assumptions

The structural model which results from these developmental hypotheses includes the specification of the cross-sectional latent variables in Figure 1 as well as the transitions between the latent variables: Cross-sectionally, we specify three latent classes, i.e., the subpopulations of 'nonmasters' failing in RE and in RA (Class 1), 'partial masters' solving only RE (Class 2) and 'masters' solving both tasks (Class 3). Note that due to the presumed universal sequence of generalization there is no subpopulation of subjects who solve the abstract task without solving the other. Longitudinally, only those latent transitions are admitted that maintain the previous latent stage or lead to a higher stage of reasoning ability (see Figure 2). Because of the assumption of cumulative growth, transitions to lower latent classes are not included.

In the present analysis, the measurement models are specified in terms of *intrusion-omission* latent class models (Dayton & Macready, 1980). The intrusion and omission rates are constrained to be constant over time, yielding two nonredundant parameters for the measurement models.

There are different approaches to testing the developmental model outlined so far: that is the usual latent class approach (LCA) on the one hand and the latent Markov chain approach for multiple indicators on the other (see chapter 1). In the LCA approach, each trajectory across the three occasions displayed in Figure 2 is regarded as a latent class of its own. In this case, the latent structure model is based on *one* discrete latent variable which is defined by the combination of three latent variables (cf. Goodman, 1974). As can be seen from Figure 2, the LCA model includes ten latent classes, resulting in nine non-redundant proportion parameters for the structural model. In the latent Markov chain model, the latent variables are not combined to a joint latent variable, rather they are linked by transition rates (Langeheine, 1994; Langeheine & van de Pol, 1990; van de Pol & Langeheine, 1990). Since the transition rates from the second measurement occasion to the third do not depend on initial class membership, there are two non-redundant initial class proportions and three non-redundant transition rates for each pair of adjacent measurement occasions (see Figure 2). Therefore, the latent Markov model yields only eight parameters in the structural model. The latent Markov model is more restrictive than the LCA model, because it specifies a first order autoregressive process among the latent variables, whereas in the latent class approach a saturated model is specified for the non-zero cells in the latent contingency table (one parameter is included for each non-redundant combination of the latent variables).

1.2 Empirical Results

The empirical results of both the LCA model, which was originally proposed by Schröder et al. (1991), and the latent Markov model are displayed in Table 1. The results were obtained using the program PANMARK (van de Pol, Langeheine & de Jong, 1991).

	LCA Model (Cumulative Growth)	*Latent Markov Model (Cumulative Growth)*	*Latent Markov Model (Free Transitions)*
G^2	48.55, 52 df	49.43, 53 df	44.17, 47 df
BIC	612.94	609.20	631.63
AIC	584.17	583.05	589.79
Omission Rate	0.180	0.198	0.218
Intrusion Rate	0.073	0.072	0.033

Table 1: Results of the LCA and latent Markov chain models

Empirically, the latent Markov chain model hardly differs from the LCA model. Because of the sparse data situation (only 27 of the 64 cells in the contingency table have non-zero entries), we have to be cautious of the interpretation of the likelihood ratio test statistics G^2. Therefore, we take the information criteria BIC and AIC for model comparison which both indicate that the more parsimonious latent Markov model may be preferred. Although the small sample size prevents any definite model selection, here and in the following the information criteria are taken as guides for tentative conclusions.

In order to further investigate the cumulative growth hypothesis a latent Markov chain model was applied in which transitions between all latent classes are admitted, including transitions to lower developmental stages. As can be seen in Table 1 (right column), this model shows worse values in the information criteria than the restricted model. So we tentatively conclude that a first order autoregressive process of latent cumulative growth is appropriate for the empirical data.

1.3 A Modified Developmental Model

Considering the structural models discussed so far, it is striking that subjects who have reached the highest latent class are necessarily subjected to an artificial *ceiling effect* in regards to latent transitions. That is, these subjects are forced to be *latent stayers* in Class 3, because latent growth is not provided for after having acquired the ability to master conditional implications and having generalized this ability to abstract premises. However, it may be hypothesized that subjects in this class still improve in deductive reasoning, i.e., that they make fewer errors after a period of *consolidation*. In the process of consolidation, already acquired cognitive strategies may be affirmed and confusion among cognitive strategies may decrease. To integrate consolidation as an aspect of development additional to cumulative growth and generalization means to specify assumptions on the omission rate in the above latent Markov chain model. In particular, the omission rate is regarded as a dynamic, rather than time-invariant, parameter for subjects in Class 3, and it is expected to decrease monotonically with duration in this developmental stage. Formally, this is accomplished by additional subscripts to the ρ parameters in equation (44) of chapter 1.

Three versions of a latent Markov model with dynamic omission rates were used. In *Version A*, a consolidation process is assumed only for those subjects who are in the highest class throughout the whole assessment period. For the first occasion, the omission rate in this subpopulation was set equal to the omission rate in the other classes (OM1), the omission rate at the second occasion (OM2) was not constrained, and the omission rate at the third occasion (OM3) was fixed to zero in one run and set free in another run. Since in the latter case OM3 was bounded at zero, we regard OM3 as fixed in the sequel. In *Version B*, also subjects who reach latent Class 3 at the second measurement occasion are considered to have a decreasing omission rate at the subsequent assessment. For this subpopulation, the omission rate at the age of twelve was constrained to be equal to the initial omission rate (OM1), whereas the rate at the age of fifteen is equated with OM2. So, OM2 denotes the omission rate referring to the second occurence of individual membership in Class 3, and OM3 refers to the third occurence. Again, OM3 is fixed to zero. Finally, *Version C* results from Version B by also restricting OM2 to zero. The results of the three models are given in Table 2.

Again we use BIC and AIC for model comparison, because of the sparse data situation and because not all of the compared models are hierarchically ordered. Looking at Tables 1 and 2 we see that for the given data situation the models with dynamic error rates for the class of masters are superior to the latent Markov model with time-invariant error rates. In addition, the estimated omission rates satisfy the required monotonic relation. Both outcomes support the concept of consolidation in terms of decreasing omission rates. Table 2 also reveals that the favorite Version B provides a somewhat different explanation of change from the models considered in Table 1: Since OM1 is as high as 0.576, the focus of development is on consolidation, i.e., on affirming and training cognitive strategies, rather than on „developmental jumps“ from failure to success.

	Version A	*Version B*	*Version C*
G^2	41.88, 52 df	39.66, 52 df	46.85, 53 df
BIC	606.26	604.04	606.62
AIC	577.50	575.27	580.46
Omission Rates OM1	0.365	0.576	0.415
OM2	0.186	0.162	0.0 (fixed)
OM3	0.0 (fixed)	0.0 (fixed)	0.0 (fixed)
Intrusion Rate	0.047	0.050	0.086

Table 2: Results of the latent Markov chain models with dynamic omission rates for subjects in latent Class 3

2. Development of the Personality Dimension(s) of Activity

In this second example we use repeated ratings of activity gathered in the *German Postwar Generation Study* (cf. Thomae, 1965). The sample of the analysis consists of $N = 545$ children born in 1945/46. At each measurement occasion the children were rated concerning their behavior during several test situations and an interview. For the analysis, three-point ratings of the children's activity are used from three measurement occasions, namely 1955, 1956, and 1957. The categories of the ratings are „low“, „medium“, and „high“ level of activity (for a more detailed description of the data, including the empirical cell frequencies, and analyses, see Meiser, in press; Meiser, Hein-Eggers, Rompe & Rudinger, 1995). A basic assumption of the models discussed here is that the manifest ratings reflect the children's locations in a *latent space of activity* in a probabilistic fashion. The goal is to investigate „movements“ in this latent space, i.e., *intra-individual change*, with respect to *inter-individual differences* in intra-individual change as well as *intra-individual differences* in intra-individual change (for a taxonomy of change see Buss, 1974).

2.1 Analyzing Homogeneity of Change Using Unidimensional Rasch Models

Let X be a polytomous random variable with m+1 ordered response categories. If the variable is observed repeatedly, X_i denotes the variable observed at occasion i. The general unidimensional Rasch model for polytomous indicators measuring a latent trait θ, i.e., the *partial credit model* (Masters, 1982), specifies the category probabilities of X_i in terms of equation (14) of chapter 1. If all threshold parameters τ_{is} of indicator X_i differ from $\tau_{(i-1)s}$ of indicator $X_{(i-1)}$ by a constant λ_i ($i \geq 2$), that is if the constraint

$$\tau_{is} = \tau_{(i-1)s} - \lambda_i \ , \ s \in \{1, \ldots, m\}, \ i \geq 2 \tag{1}$$

holds true, this constant equals the 'amount of intra-individual change' in the latent trait θ from measurement occasion i-1 to i. This can easily be seen by inserting constraint (1) into equation (14) of the introductory chapter:

$$p(X_{vi} = x) = \frac{\exp\left(x(\theta_v + \lambda_i) - \sum_{s=1}^{x} \tau_{(i-1)s} \right)}{\sum_{y=0}^{m} \exp\left(y(\theta_v + \lambda_i) - \sum_{s=1}^{y} \tau_{(i-1)s} \right)} \ . \tag{2}$$

Since λ_i is independent from θ, all subjects are assumed to have the same direction and speed of latent change independent from their initial level on the latent continuum, i.e., change is constrained to be *homogeneous over subjects*. Therefore, inter-individual differences in θ are preserved over time, which corresponds to the concept of *relative stability*. Without restriction (1), threshold-dependent change may occur which can hardly be interpreted as unidimensional global change of θ. Note that the polytomous Rasch model with restriction (1) is equivalent to the well-known *rating scale model* proposed by Andrich (1978, see also equation (17) in chapter 1).

For the given trichotomous indicators of activity the rating scale model is rejected with a likelihood ratio statistic of G^2=42.63, 17 df. (estimation by conditional maximum likelihood - CML). Threshold-dependent variability does not sufficiently account for this misfit, since the partial credit model is also rejected with G^2=30.92, 15 df Therefore it is reasonable to investigate violations of the developmental model of unidimensional homogeneous change by admitting either intra-individual differences or inter-individual differences in intra-individual change.

2.2 Specification of Intra-Individual Differences by Two-Dimensional Rasch Models

As an alternative to the unidimensional models, we now consider models in which homogeneity of change refers to several latent dimensions and distinct developmental trajectories may occur for different latent dimensions. For the given ratings it may be presumed that passing the threshold from „low“ to „medium“ level of activity involves another latent characteristic than passing from „medium“ to „high“. This assumption is based on the fact that the categories „low“ and „high“ enclose the extremes of „extremely low degree of activity“ and „hyperactive“ (Thomae, 1965, p. 89), respectively. Hence, passing the first threshold may depend on a general trait of activity in social or achievement situations, whereas the second threshold may involve a kind of 'hyperactivity trait' which is qualitatively different from the former. The resulting model is a *two-dimensional partial credit model* in which each threshold refers to a latent dimension of its own. Within each dimension, the homogeneity of change hypothesis is maintained. However, there may be intra-individual differences between the initial levels and the trajectories of development in the latent dimensions (for a more detailed account, see Meiser, in press).

Maximum likelihood estimation of the two-dimensional partial credit model was carried out using the loglinear representation of conditional multidimensional Rasch models (cf. Kelderman & Rijkes, 1994; Meiser, in press). The results are displayed in Table 3 (*Model A*). Empirically, the model is not rejected. As can be seen from the estimates of λ, the amount of change appears to be larger in the second dimension (hyperactivity trait) than in the first one (general activity trait). This difference can be further examined using a restricted two-dimensional partial credit model: In *Model B*, equality constraints are imposed on the λ parameters of both dimensions. As can be seen in Table 3, this model of „equal change“ is rejected and differs significantly from the unconstrained Model A. We may conclude that development in the general activity trait is quantitatively different from development in the hyperactivity trait. Since the amount of latent change is actually much smaller in the first dimension, the processes of development might also be qualitatively different. That is, while *relative stability* applies to the second dimension, *absolute stability* may hold for the first. In order to test this assumption, in *Model C* the λ parameters of the general activity trait are fixed to zero. By this, no change is permitted in this dimension. Model C is not rejected and does not differ significantly from the unconstrained Model A (see Table 3). So, we may prefer this more parsimonious model representing absolute stability in the first and relative stability in the second latent dimension.

		Model A	*Model B*	*Model C*
G^2		18.61, 13 df	32.18, 15 df	21.72, 15 df
BIC		2731.90	2732.87	2722.42
AIC		2675.99	2685.56	2675.11
Dimension 1	λ_2	0.37	1.06	0.0 (fixed)
	λ_3	-0.35	-0.99	0.0 (fixed)
Dimension 2	λ_2	1.45	1.06	1.49
	λ_3	-1.33	-0.99	-1.36

Table 3: Results of the two-dimensional partial credit models

2.3 Investigation of Change in Subpopulations Using Mixed Rasch Models

Another way to account for the misfit of the unidimensional rating scale model is to introduce inter-individual differences in intra-individual change. Therefore, we now focus on the investigation of different developmental trajectories for different latent subpopulations. For this purpose we used a two-class mixed rating scale model (cf. Rost, 1991; see also section 2.3 of chapter 1). This model implies homogeneity of change in one dimension *within* the classes, whereas the patterns of development may be different *between* the classes (for a more detailed discussion, see Meiser, Hein-Eggers, Rompe & Rudinger, 1995).

The analysis was carried out using the program LEM (Vermunt, 1993), which facilitates loglinear model specifications in latent class models. As shown in Table 4, the two-class rating scale model is not rejected (*Model D*), and the class-specific developmental pathways are clearly distinguished as indicated by the estimates of λ parameters. Fixing the λ parameters to zero for latent Class 1 (*Model E*) yields only a negligible decrease of the likelihood ratio statistic (see Table 4). We may conclude that there is one latent subpopulation of absolute stability (Class 1) and another one of homogeneous change, i.e., relative stability (Class 2).

		Model D	*Model E*
	G^2	6.25, 9 df	6.37, 11 df
	BIC	2744.75	2732.27
	AIC	2671.63	2667.76
Class 1	Proportion[a)]	0.285	0.290
	λ_2	0.01	0.0 (fixed)
	λ_3	0.06	0.0 (fixed)
Class 2	Proportion[a)]	0.608	0.603
	λ_2	3.69	3.77
	λ_3	-3.61	-3.66

a) Note that the class proportions do not sum to one, because response patterns with minimum or maximum scores cannot be assigned to latent classes (cf. Rost, 1991).

Table 4: Results of the two-class mixed rating scale models

Tables 3 and 4 provide competing models for the activity data, one representing intra-individual differences (Model C) and the other one representing inter-class differences (Model E) in change. Since the two-dimensional partial credit model and the mixed rating scale model are not hierarchically related, the model comparison is based on BIC and AIC. Unfortunately, these criteria provide contradictory results so that an unequivocal decision between the two models is not available. However, since BIC stresses the methodological aspect of model parsimony, we may cautiously decide in favor of Model C.

3. Conclusions

The series of models outlined for the two empirical data sets are meant to demonstrate the flexibility of latent structure analysis for modeling and testing various hypotheses in longitudinal research. In the first section, the focus was on the *multidimensionality of change*, that is on modeling the processes of cognitive construction, generalization and consolidation in a joint model. By the specification of developmental assumptions on the 'error rates' in latent class models, these rates are no longer regarded as purely technical parameters that account for observed deviations from expected response patterns. Moreover, these parameters are also taken to be part of the psychological model, not only as part of the statistical model. Considering the empirical results, this approach appears to be fruitful.

In the second section the focus was on *change in a multidimensional space*. While the unidimensional Rasch model of relative stability was not appropriate for the given data set, different two-dimensional latent structure models fitted the data satisfactorily: A model including two quantitative latent variables on the one hand and a model including a qualitative and a quantitative latent variable on the other. Though in the latter change is limited to the quantitative latent variable, the qualitative mixture variable determines the shape of development. The empirical results reveal that intra- or inter-individual differences in development can be uncovered by the use of multidimensional latent structure models.

References

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43*, 561-573.

Buss, A.R. (1974). A general developmental model for interindividual differences, intraindividual differences, and intraindividual changes. *Developmental Psychology*, *10*, 70-78.

Dayton, M.C. & Macready, G.B. (1980). A scaling model with response errors and intrinsically unscalable respondents. *Psychometrika*, *45*, 343-356.

Erdfelder, E. (1990). Deterministic developmental hypotheses, probabilistic rules of manifestation, and the analysis of finite mixture distributions. In A. von Eye (Ed.), *Statistical methods in longitudinal research. Vol. II: Time series and categorical longitudinal data* (pp. 471-509). New York: Academic Press.

Goodman, L.A. (1974). The analysis of systems of qualitative variables when some of the variables are unobservable. Part I: A modified latent structure approach. *American Journal of Sociology*, *79*, 1179-1259.

Henning, H.J. & Rudinger, G. (1985). Analysis of qualitative data in developmental psychology. In J.R. Nesselroade & A. von Eye (Eds.), *Individual development and social change: Explanatory analysis* (pp. 295-341). Orlando, FL: Academic Press.

Kelderman, H. & Rijkes, C.P.M. (1994). Loglinear multidimensional IRT models for polytomously scored items. *Psychometrika*, *59*, 149-176.

Langeheine, R. (1994). Latent variables Markov models. In A. von Eye & C.C. Clogg (Eds.), *Latent variables analysis. Applications for developmental research* (pp. 373-395). Thousand Oaks, CA: Sage.

Langeheine, R. & van de Pol, F. (1990). A unifying framework for Markov modeling in discrete space and discrete time. *Sociological Methods and Research*, *18*, 416-441.

Masters, G.N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*, 149-174.

Meiser, T. (1995). Loglinear Rasch models for the analysis of stability and change. *Psychometrika,19*, 377-391.

Meiser, T., Hein-Eggers, M., Rompe, P. & Rudinger, G. (in press). Analyzing homogeneity and heterogeneity of change using Rasch and latent class models: A comparative and integrative approach. *Applied Psychological Measurement*.

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

Schröder, E., Edelstein, W. & Hoppe-Graff, S. (1991). Qualitative analyses of individual differences in intraindividual change: Examples from cognitive development. In D. Magnusson, L. R. Bergman, G. Rudinger & B. Törestad (Eds.), *Problems and methods in longitudinal research: Stability and change* (pp. 166-189). New York: Cambridge University Press.

Thomae, H. (1965). Objective socialization variables and personality development. Findings from a longitudinal study. *Human Development*, *8*, 87-116.

van de Pol, F. & Langeheine, R. (1990). Mixed Markov latent class models. In C. C. Clogg (Ed.), *Sociological Methodology 1990* (pp. 213-247). Oxford: Blackwell.

van de Pol, F., Langeheine, R. & de Jong, W. (1991). *PANMARK user manual. Panel analysis using Markov chains - version 2.2*. Voorburg: Netherlands Central Bureau of Statistics.

Vermunt, J. K. (1993). LEM: Log-linear and event history analysis with missing data using the EM algorithm. Tilburg University (WORC Paper 93.09.015/7).