# 12

# Rasch Models for Longitudinal Data

Thorsten Meiser

Friedrich-Schiller-Universität Jena

## 12.1 Introduction

The chapter gives an overview of Rasch models for the measurement of change across repeated observations of the same individuals and items. The models described herein include extensions of the original Rasch model that allow one to analyze multidimensional latent constructs and to incorporate heterogeneity of change across individuals. In particular, the use of mixture-distribution Rasch models in longitudinal research allows one to model quantitative interindividual differences in a latent trait at each occasion, together with qualitative interindividual differences in the course of development. A mover–stayer mixed-Rasch model can be specified as a special case that reflects the assumption that change over time occurs for some latent subpopulation but not for another. An empirical example illustrates that the mover–stayer mixed-Rasch model can provide a parsimonious and viable account of observed heterogeneity of change.

## 12.2 Rasch Models for Repeated Observations

The Rasch model (RM, Rasch, 1968, 1980) is usually applied to the responses of individuals to items observed at one point in time. However, the RM can also be used in situations in which a set of items is repeatedly administered to the same sample of individuals. In those longitudinal designs, the RM specifies the probability that item $i$, $i = 1, \dots, I$, is solved by person $v$, $v = 1, \dots, N$, at occasion $t$, $t = 1, \dots, T$:

$$P(X_{vit} = 1 | \theta_{vt}, \beta_{it}) = \frac{\exp(\theta_{vt} - \beta_{it})}{1 + \exp(\theta_{vt} - \beta_{it})}. \tag{12.1}$$

The parameter $\theta_{vt}$ denotes person $v$'s latent ability at occasion $t$, and $\beta_{it}$ denotes item $i$'s difficulty at occasion $t$.

Modeling changes in ability or item difficulty over time is the aim of extensions of the RM and other IRT models, such as those by Fischer (1983, 1995d), Wilson (1989), and Embretson (1991). This chapter gives an overview of RMs for modeling change and presents an application of some of the presented models to a longitudinal data set. Additional applications of this class of models can be found in the chapters by Draney and Wilson, and Glück and Spiel in this volume.

### 12.2.1 Modeling Homogeneous and Person-Specific Change

Aside from specifying the probability of solving an item at a particular measurement occasion $t$ in terms of Equation 12.1, the RM and its extensions allow one to measure change from one occasion to another and to test hypotheses about the latent course of development. In the simplest case, one may assume that the person and item parameters are invariant over time, that is, $\theta_{vt} = \theta_v$ and $\beta_{it} = \beta_i$ for all occasions $t$, which means that no change occurs at all. Alternatively, one can specify the hypothesis that all individuals exhibit the same amount of change on the latent continuum by introducing a change parameter $\lambda_t$ that is constant across individuals and items:

$$P(X_{vit} = 1|\theta_v, \beta_i, \lambda_t) = \frac{\exp(\theta_v + \lambda_t - \beta_i)}{1 + \exp(\theta_v + \lambda_t - \beta_i)}. \qquad (12.2)$$

The model in Equation 12.2 represents a linear logistic test model (LLTM; Fischer, 1983, 1995d,b; Spada & McGaw, 1985) that decomposes the item parameter $\beta_{it}$ into basic parameters that capture the item's initial difficulty $\beta_i$ and the change $\lambda_t$ that has occurred until occasion $t$, $\beta_{it} = \beta_i - \lambda_t$ with $\lambda_1 = 0$. Technically, the person parameter $\theta_v$ is considered constant over time in the linear logistic model, so that the relative position of person $v$ is preserved across the measurement occasions. Because change in overall item difficulty is equivalent to global change in latent-person ability, however, the model reflects the assumptions that change may occur and that change is homogeneous across persons. Accordingly, $\lambda_t$ can be interpreted as the average increase (or decrease) in ability from the first measurement occasion until occasion $t$ for all individuals. Due to the additive decomposition of $\beta_{it}$ into item difficulty $\beta_i$ and the change effect $\lambda_t$, the relative positions of the items are also maintained over time. The latter assumption can be dropped by allowing for time-specific item parameters.

The linear logistic RM in Equation 12.2 is based on a unidimensional latent space. That is, the position of person $v$ on the one latent trait $\theta$ underlies her or his responses to all items $i$ at all occasions $t$, although the person's absolute position on the latent continuum may shift from one occasion to another.

In many applications, however, it may be plausible to assume that different items measure different latent constructs, such as distinct aspects of a syndrome in clinical research or specific cognitive abilities in educational assessment. To measure time or treatment effects across repeated observations in such cases, the linear logistic test model with relaxed assumptions (LLRA; Fischer, 1983, 1995c) accommodates multidimensionality of items. This model allows for item-specific latent traits by specifying interactions between persons and items, $\theta_{iv}$. To measure time or treatment effects, the model contains change parameters that are considered constant across the items and their latent dimensions. Although such generalizations of the RM to incorporate multidimensionality in an item set are suitable specifications in many instances, the remainder of this chapter will largely focus on the issue of homogeneity versus heterogeneity of change across individuals, which can also be addressed by modeling and testing for particular types of multidimensionality in longitudinal RMs.

The RM in Equation 12.2 contains the rather restrictive assumption that change is homogeneous across persons, that is, that the amount of change $\lambda_t$ is supposed to be the same for all persons $v$. This restrictive assumption can be dropped by specifying a multidimensional latent space that contains one latent-trait continuum for each measurement occasion (e.g., Andersen, 1985). The resulting model reflects the concept of person-specific change that can be written as

$$P(X_{vit} = 1|\theta_{vt}, \beta_i) = \frac{\exp(\theta_{vt} - \beta_i)}{1 + \exp(\theta_{vt} - \beta_i)}. \tag{12.3}$$

Formally, the parameter $\theta_{vt}$ represents an interaction between person $v$ and measurement occasion $t$, which implies that change in the latent ability $\theta$ may be person-specific rather than homogeneous, whereas item difficulty $\beta_i$ is assumed to remain constant over time. In contrast to the linear logistic RM (12.2), the relative position of person $v$ may therefore change from one occasion to another in Equation 12.3, so that the amount of change cannot be measured by means of a global change parameter $\lambda_t$. The model of person-specific change in Equation 12.3 is appropriate in longitudinal research designs in which the items form a unidimensional scale at each occasion with stationary item parameters, and in which the speed or direction of development may vary between persons, for example because individuals profit to different degrees from training or intervention programs.

The models of homogeneous change in Equation 12.2 and of person-specific change in Equation 12.3 result from particular restrictions of the person parameters $\theta_{vt}$ and the item parameters $\beta_{it}$ in the general RM for repeated observations as defined in Equation 12.1. Alternative restrictions are also possible, including the decomposition of the item parameters $\beta_{it}$ into linear combinations of specific treatment effects and general trends (e.g., Fischer, 1995c).

### 12.2.2 Loglinear Rasch Models for Measuring Change

The loglinear representation of RMs (Cressie & Holland, 1983; Kelderman, 1984; see also the chapter by Kelderman in this volume) forms a suitable framework for the specification and test of hypotheses about change in longitudinal data. In the loglinear notation of the conditional RM, the expected probabilities of response vectors are reparameterized as linear combinations of item parameters and of parameters representing the total scores of the response vectors. This notation facilitates the specification of theoretical assumptions concerning latent change and affords straightforward statistical tests especially for small item sets (Meiser, 1996; Meiser et al., 1998).

To illustrate, let a set of $I$ items be administered to a sample of individuals at $T = 2$ measurement occasions. Because the total score $R_v = \sum_t \sum_i x_{vit}$ forms the sufficient statistic for person parameter $\theta_v$ under the unidimensional RM of homogeneous change in Equation 12.2, the probability of a given response vector $\boldsymbol{x} = (x_{11}, \ldots, x_{I1}, x_{12}, \ldots, x_{I2})$ with total score $R$ can be expressed without latent-person parameter $\theta_v$. In the loglinear reparameterization of the RM of homogeneous change, the logarithm of the expected probability of response vector $\boldsymbol{x}$ can therefore be written as

$$\ln P(\boldsymbol{x} = (x_{11}, \ldots, x_{I1}, x_{12}, \ldots, x_{I2})) = u - \sum_{t=1}^{2} \sum_{i=1}^{I} x_{it}\beta_i + \sum_{i=1}^{I} x_{i2}\lambda_2 + u_R. \tag{12.4}$$

Likewise, the sufficient statistic of the two-dimensional latent-ability vector $(\theta_{v1}, \theta_{v2})$ in the model of person-specific change (12.3), for two occasions is given by the pair of total scores $(R_{v1}, R_{v2})$, with $R_{v1} = \sum_i x_{i1}$ being the total score at the first occasion and $R_{v2} = \sum_i x_{i2}$ being the total score at the second occasion. The conditional RM of person-specific change can therefore be specified by the following loglinear model for the expected probability of response vector $\boldsymbol{x}$ with the two total scores $R_1$ and $R_2$:

$$\ln P(\boldsymbol{x} = (x_{11}, \ldots, x_{I1}, x_{12}, \ldots, x_{I2})) = u - \sum_{t=1}^{2} \sum_{i=1}^{I} x_{it}\beta_i + u_{(R_1,R_2)} \tag{12.5}$$

To achieve identifiability of the loglinear RMs in (12.4) and (12.5), some parameter restrictions need to be imposed. The usual restrictions include the constraints that the set of item parameters and the set of total score parameters sum to zero, that is, $\sum_i \beta_i = 0$, $\sum_R u_R = 0$, and $\sum_{R_1} \sum_{R_2} u_{(R_1,R_2)} = 0$.

The loglinear framework facilitates straightforward specifications and tests of hypotheses about change, such as stationarity of the change parameters or invariance of item parameters across measurement occasions. Stationarity of latent change is reflected by the constraints $\lambda_t = \lambda$ for all $t > 1$ in Equation 12.4. Invariance of the item parameters can be tested by introducing time-specific difficulty parameters $\beta_{it}$ in Equations 12.4 and 12.5 and by comparing the resulting more-general model variants with the models assuming constant item parameters.

For model-testing purposes, it is of particular importance to note that the loglinear model of homogeneous change in (12.4) can be derived as a special case from the loglinear model of person-specific change in Equation 12.5. That is, imposing the restrictions $u_{(R_1,R_2)} = u_{(R_1+R_2)} + R_2\lambda_2$ in Equation 12.5 yields Equation 12.4. This hierarchical relation between the loglinear models (12.4) and (12.5) allows one to test for homogeneity of change across persons by means of a statistical model comparison.

### 12.2.3 Extensions to Multiple Response Categories and Multidimensional Latent Traits

Linear logistic RMs for the measurement of change have been extended from the analysis of dichotomous items to the analysis of items with several response categories (Fischer & Parzer, 1991; Fischer & Ponocny, 1994, 1995). For that purpose, the item and category parameters in the rating-scale model (Andrich, 1978) or the partial-credit model (Masters, 1982) for polytomous items are specified with regard to different points in time, analogous to the item parameter in Equation 12.1. The item and category parameters are then decomposed into basic parameters that reflect item and category difficulty on the one hand and change or treatment effects over time on the other hand.

Moreover, longitudinal RMs can be extended to include more than one latent trait at each occasion, as in the linear logistic model with relaxed assumptions. Generalizing the concepts of homogeneous and person-specific change from unidimensional RMs, change can be specified to be homogeneous across persons or person-specific within each of the latent traits of a multidimensional latent-trait model.

In very general terms, the probability to observe response category $x$ in item $i$ with $m+1$ response categories $0,\ldots,m$ at occasion $t$ can be specified by a longitudinal RM with $D$ latent traits at each measurement occasion:

$$P(X_{vit} = x|\boldsymbol{\theta}_{\mathbf{vt}}, \boldsymbol{\tau}_{\mathbf{ist}}) = \frac{\exp(\sum_{d=1}^{D}\sum_{s=1}^{x} w_{isd}\theta_{vtd} - \sum_{d=1}^{D}\sum_{s=1}^{x} w_{isd}\tau_{istd})}{\sum_{y=0}^{m}\exp(\sum_{d=1}^{D}\sum_{s=1}^{y} w_{isd}\theta_{vtd} - \sum_{d=1}^{D}\sum_{s=1}^{y} w_{isd}\tau_{istd})}. \tag{12.6}$$

In Equation 12.6, $\theta_{vtd}$ denotes the latent-ability parameter of person $v$ at occasion $t$ on the latent dimension $d$, and $\tau_{istd}$ represents the difficulty of the threshold between categories $s-1$ and $s$ for item $i$ at occasion $t$ on dimension $d$. The values $w_{isd}$ are weights that reflect the degree to which the latent-ability dimensions are involved in reaching the various response categories.

These weights are determined a priori by the researcher and are thus part of the model specification. Usually, the weights are restricted to the binary values of zero and one, indicating whether a particular trait is involved in reaching a category or not (see Meiser, 1996; Meiser et al., 1998).

The polytomous multidimensional latent-trait model for longitudinal data in Equation 12.6 serves as a superordinate framework, or metastructure, to derive more specific models by sets of parameter constraints. For example, a multidimensional model of homogeneous change within each latent trait can be specified by setting $\theta_{vtd} = \theta_{vd}$ and by decomposing the threshold parameters into their initial difficulty and change parameters, $\tau_{istd} = \tau_{isd} - \lambda_{td}$. The resulting model reflects global change with persisting relative positions of the persons and items on each of the $D$ latent traits (Meiser, 1996; Meiser & Rudinger, 1997). Together with a theory-based and parsimonious selection of the weights $w_{isd}$, such parameter constraints are often necessary to yield identifiable submodels of the superordinate framework in Equation 12.6 for a given data set.

## 12.3 Mixture-Distribution Rasch Models for the Analysis of Change

The aforementioned distinction between change that is completely homogeneous across persons versus change that is purely person-specific marks two extremes of homogeneity and heterogeneity, respectively. In a given population, a limited number of latent-developmental trajectories may coexist, so that change is neither completely homogeneous nor completely person-specific. Instead, the direction and the amount of change may be rather homogeneous within each of several subpopulations, whereas the course of change may differ between the subpopulations.

In some cases, relevant subpopulations can be defined by manifest extraneous grouping variables, such as gender, socioeconomic status, or treatment group, so that differences in the developmental trajectories can be analyzed by parameter comparisons between groups (e.g., Fischer, 1983, 1995d). In other cases, either extraneous grouping variables may either not be available, or they may not account for observed heterogeneity of change in the population (e.g., Wilson, 1989). Then, the subpopulations are latent and have to be identified by statistical modeling techniques in order to separate the developmental patterns that are mixed in the total population.

The goal to identify latent subpopulations and to measure change within each subpopulation can be pursued by means of finite mixture-distribution models (McLachlan & Peel, 2000; Titterington et al., 1985). Finite mixture-distribution models characterize the probabilities of events in terms of a weighted sum of component distributions. Each component distribution is specified to hold within a subpopulation $c$, $c = 1, \ldots, C$, and the weights correspond to the proportions of the subpopulations in the entire population, $\pi_c$.

### 12.3.1 Class-Specific Homogeneous Change

Applying the notion of finite mixture-distribution models to longitudinal RMs, one may assume that a population consists of $C$ latent subpopulations and that change is homogeneous within each subpopulation. This assumption can be specified by a mixed RM (Rost, 1990, 1991; von Davier & Rost, 1995; see also the chapter by von Davier & Yamamoto in this volume) of the form

$$P(X_{vit} = 1) = \sum_{c=1}^{C} \pi_c P(X_{vit} = 1 \mid c)$$

with component probabilities

$$P(X_{vit} = 1|c, \theta_{v|c}, \beta_{i|c}) = \frac{\exp(\theta_{v|c} + \lambda_{t|c} - \beta_{i|c})}{1 + \exp(\theta_{v|c} + \lambda_{t|c} - \beta_{i|c})}. \tag{12.7}$$

The mixed RM for longitudinal data in Equation 12.7 combines the unidimensional RM, which allows for quantitative differences among persons while implying homogeneity of change, and the latent-class approach, which allows for qualitatively distinct patterns of change (see Meiser et al., 1995, for a discussion of Rasch and latent-class models in longitudinal research).

More specifically, in contrast to the model of homogeneous change in Equation 12.2, the parameters of the mixed RM in Equation 12.7 are specified conditional on latent class $c$ that contains a proportion $\pi_c$ of the entire population. By introducing class-specific change parameters $\lambda_{t|c}$, the model accounts for qualitative differences in change. In contrast to usual latent-class models, however, the mixed RM also allows for quantitative differences between individuals of the same subpopulation in terms of the person parameter $\theta_{v|c}$. Together, the mixed longitudinal RM (12.7) integrates interindividual differences and homogeneous change within each latent subpopulation with qualitative differences in change between subpopulations.

### 12.3.2 A Mover–Stayer Mixed-Rasch Model

With appropriate parameter restrictions, mixed RMs can be used to disentangle latent subpopulations of "movers" and "stayers" within a latent-trait framework that incorporates quantitative interindividual differences as well as differences in change over time.

The distinction between a latent subpopulation that exhibits change over time, the "movers," and a latent subpopulation that shows invariant response behavior over time, the "stayers," has been incorporated into mixed Markov chain models (e.g., Langeheine & van de Pol, 1994; van de Pol & Langeheine, 1990) to express the idea that observed heterogeneity of change may reflect the coexistence of two simple mechanisms in a given population: change and no change. The distinction between a latent class of movers and a latent class of stayers can easily be transferred to the mixed longitudinal RM in Equation 12.7 by setting $C = 2$ and imposing the restriction $\lambda_{t|2} = 0$ for all $t$. Thereby, the change parameters for the first subpopulation $\lambda_{t|1}$ are free to differ from zero, which means that the latent class $c = 1$ may exhibit global change in the latent ability across the measurement occasions. Thus, class 1 represents a subpopulation of movers. By restricting $\lambda_{t|2}$ to zero, the latent ability is constrained to be invariant over time in class $c = 2$. Thereby, class 2 forms a subpopulation of stayers. Extending mover–stayer models in the framework of mixed Markov models, the mover–stayer mixed RM admits persisting interindividual differences in latent ability within both subpopulations of movers and stayers.

Mover–stayer mixed RMs were successfully applied to longitudinal data concerning the development of observed activity in childhood (Meiser & Rudinger, 1997) and concerning the development of mathematical problem-solving skills in primary school (Meiser et al., 1998). The latter analysis is briefly summarized in the following section in order to illustrate the various RMs for the measurement of change that were discussed throughout this chapter. For further applications of RMs to longitudinal data, see the chapters by Draney and Wilson and by Glück and Spiel in this volume.

## 12.4 An Empirical Illustration

In an analysis of mathematical problem-solving skills in primary-school children, Meiser et al. (1998) applied a series of longitudinal RMs to investigate the course of latent development. The empirical data were taken from a large-scale longitudinal study on the cognitive abilities and achievements of school children in Germany (Weinert & Helmke, 1997). The selected items encompassed three arithmetic word problems that were administered to a sample of 1030 children in the second and third grades. The series of models was specified as conditional RMs in their loglinear representation, and the analyses were run with the software LEM (Vermunt, 1997a). This software facilitates loglinear model specification in terms of design matrices (see Meiser, 2005; Rindskopf, 1990) and allows the inclusion of latent-class variables in the loglinear modeling framework.

In a first step, we applied the conditional RM of homogeneous change in its loglinear representation (see Equation 12.4) to the three items at the two occasions. This model was rejected on grounds of a poor overall goodness of fit, as revealed by the likelihood ratio statistic of $G^2(54) = 72.53$, $p = .047$. The loglinear RM of person-specific change (see Equation 12.5), in contrast, showed a satisfactory goodness of fit with $G^2(46) = 54.84$, $p = .175$. As delineated above, the two loglinear models are hierarchically related. A model comparison by means of the conditional likelihood ratio statistic therefore yields a focused test of homogeneity of change across persons. The model comparison showed a significant difference in model fit, $\Delta G^2(8) = 17.69$, $p = .024$, which indicated that the homogeneity assumption was violated for the given data set.

To analyze the structure of developmental heterogeneity further, we specified a mover–stayer mixed RM that follows from Equation 12.7 with the specification of two latent subpopulations and with the restriction $\lambda_{t|2} = 0$ for the stayer class $c = 2$. In addition, we imposed equality restrictions on the item parameters across the latent classes $c = 1$ and $c = 2$, $\beta_{i|1} = \beta_{i|2}$, which reflect the assumption that the items form an invariant scale not only across time but also across the different latent subpopulations. The resulting model provided an acceptable overall goodness of fit to the data, $G^2(49) = 64.79$, $p = .065$. The latent subpopulation of movers comprised an estimated proportion of $\hat{\pi}_1 = .43$ of the children, and the latent subpopulation of stayers comprised the complementary estimated proportion of $\hat{\pi}_2 = .57$.

The mover–stayer mixed RM cannot be compared with the models of global change and person-specific change by means of a conditional likelihood ratio test using the chi-square distribution. This is due to the fact that the mover–stayer mixed RM is not hierarchically related to the other models, so that the regularity conditions for a statistical model comparison are not met. Therefore, a descriptive comparison between the model of person-specific change and the mover–stayer mixed RM was conducted with the information criterion CAIC (Burnham & Anderson, 2002). This model comparison demonstrated that the mover–stayer mixed RM provided a better balance between model fit and model parsimony, CAIC=7740.27, than did the model of person-specific change, CAIC=7442.36.

Together, the empirical results of the Rasch analysis of the given data set on arithmetic problem-solving highlight that the mover–stayer mixed RM may offer a parsimonious account of observed heterogeneity in change by specifying the two simple underlying mechanisms of change and no change in a given population. In fact, the subpopulation of movers, $c = 1$, showed an estimated change parameter of $\hat{\lambda}_{2|1} = 1.19$ that was significantly larger than zero, as indicated by a $z$-value of 4.01. In terms of the expected probabilities to solve the arithmetic problems, the movers improved their chances to provide the correct responses to the three items from an average of .47, .35, and .43 at second grade to an average of .72, .61, and .69 at third grade. Because the change parameter was fixed to zero for the subpopulation of stayers, $c = 2$ with $\lambda_{2|2} = 0$, the expected probabilities of successful item solution did not differ between the two assessment occasions for stayers. Children in this latent subpopulation had average chances of .41, .34, and .38 to solve the three items at both second grade and third grade.

The mover–stayer mixed RM allows for differences in item difficulty and person ability at each measurement occasion, and it separates qualitatively different patterns of development. In the school data analyzed by Meiser et al. (1998), a latent subpopulation of children who improved performance from one grade to the next could be distinguished from another latent subpopulation of children whose performance remained unchanged. The separation of latent subpopulations with different developmental trajectories by mixed RMs can also be used to investigate possible associations between qualitative patterns of development and external variables such as gender and socioeconomic indices (e.g., Meiser et al., 1995). The combination of RMs for measuring change in a (sub)population and finite-mixture models for analyzing heterogeneity between latent subpopulations thus provides a flexible framework for specifying and testing hypotheses about change in longitudinal data.