Chapter 10

Modeling longitudinal changes in old age: From covariance structures to dynamic systems

Ulman Lindenberger

Paolo Ghisletta

Methods in cognitive aging

A major goal in aging research is to identify the structural dynamics and causal mechanisms of senescent changes in behavior. This goal is pursued within and across psychometric, cognitive-experimental, and cognitive-neuroscience research traditions, among others (Lindenberger and Baltes, 1994a; Li and Lindenberger, 1999). Each of these traditions has developed modal research strategies. For instance, the psychometric tradition often applies variance-partitioning procedures to cross-sectional data sets; the cognitive-experimental tradition produces predominantly ordinal age-by-treatment interactions; and aging researchers within the emerging field of cognitive neuroscience try to establish links between age differences in behavior and age differences in patterns of brain activation, which are generally obtained by subtracting patterns of baseline activation from patterns of activation observed under experimental conditions.

Each of these methods has its strengths and weaknesses. Variance-partitioning techniques applied to age-heterogeneous cross-sectional data are a useful heuristic for identifying variables that share large amounts of variance with other variables including age. However, such variables may function remarkably well as 'mediators' (alleged proximal causes) of age differences in other variables for reasons that are unrelated to the structure and mechanisms of age-based changes in cognition (Hertzog, 1996; Lindenberger and Pötter, 1998). Ordinal age-by-condition interactions revealed by the cognitive-experimental research tradition, such as greater negative age differences for the more difficult of two task conditions, cannot be taken as firm evidence that the mechanisms underlying performance in different task conditions age independently or at different rates (cf. Dunn and Kirsner, 1988; Kliegl, Mayr, and Krampe, 1994). Rather, such interactions need to be interpreted with reference to parameterized theoretical models that predict over- or underadditivity of age and condition effects for specific scalings of the dependent variable. Finally, the existence of vast differences in functional connectivity between young and older adults revealed by recent neuroimaging studies (Schreurs *et al.*, 2001) adds complexity to the functional interpretation of age-based activation differences in specific areas of the brain (cf. Cabeza, 2001, 2002).

Latent growth models and multilevel models

In this chapter, we discuss yet another set of methods aimed at discerning structural dynamics and causal mechanisms of ontogenetic changes: the multivariate analysis of longitudinal changes by means of latent growth models (LGM) and multilevel models (MLM). Similar to other methods, LGM and MLM have specific strengths, limitations, and problems. The aim of the present chapter is to highlight central characteristics, using recent examples from our own research for illustration (Ghisletta and Lindenberger, in press, 2002; Lindenberger and Ghisletta, 2002).

LGM evolved in psychometrics, biometrics, and behavior genetics, and can be considered as variants of structural equation modeling aimed at longitudinal data analysis. MLM evolved independently of LGM, and are extensively used in educational research (Laird and Ware, 1982; Bryk and Raudenbush, 1987). When applied to the analysis of longitudinal data, LGM and MLM are closely related and sometimes identical (Chou, Bentler, and Pentz, 1998; McArdle and Hamagami, 1996; MacCallum *et al.*, 1997; Little, Schnabel, and Baumert, 2000). However, as will be shown later on, some research questions are more easily addressed with MLM, whereas others can only be formalized with LGM.

LGM and MLM enable researchers to explore the structure of change, and to test hypotheses about the underlying dynamics that drive this structure. In the context of research on behavioral aging, an important feature of LGM and MLM is to further examine, and potentially disambiguate, the functional status of correlations observed in age-heterogeneous cross-sectional data. Take, for instance, the strong correlations between sensory and intellectual functioning in old and very old age observed in various cross-sectional data sets (Anstey, Lord, and Williams, 1997; Lindenberger and Baltes, 1994b). On the one hand, these correlations may indicate a functional connection between sensory and intellectual domains in advanced old age. On the other, they may result from the superimposition of functionally independent age-linked processes (Bäckman *et al.*, 2000: 505; for a formal treatment, see Lindenberger and Pötter, 1998). Here, LGM and MLM help to discern the relative importance of the two alternatives by examining whether, and to what degree and in which manner, changes are correlated (coupled) across domains.

If the analyses, perhaps due to the nature of available data (e.g., many individuals, few occasions), focus on interindividual differences, the investigation of correlated change corresponds to asking the question whether individuals who show more decline in one domain also show more decline in others. Of course, this question only makes sense if rates of change differ reliably across individuals; if individuals change at similar rates in the domains involved, these domains may still be linked *within* subjects, but between-subject methods are unable to inform us about it. Conversely, the existence of a link between two variables at the level of interindividual differences in change does not necessarily imply a functional coupling within each individual, especially if crucial statistical assumptions are violated (e.g., sample homogeneity, ergodicity; cf. Molenaar, Huizenga, and Nesselroade, 2003). Despite these complications, longitudinal data generally provide more valid information about the functional relation between different age changes than cross-sectional data.

LGM and MLM are general and adaptive tools for testing hypotheses about dynamic structural relations. In contrast to alternative standard data-analytic procedures, LGM and MLM impose fewer restrictions on the shape and variability of the change phenomena under study. As an example, take empirical situations in which variances or covariances change over time. Such situations violate the variance–covariance homogeneity assumption underlying repeated measures ANOVA. That is, changes in variance and covariance cannot be examined using repeated measures ANOVA because use of this method assumes their constancy. In contrast, LGM and MLM permit time-dependent changes in variances and covariances, and empower researchers to formally specify and test hypotheses about their magnitude, direction, and dynamics (LGM), in conjunction with or independent of hypotheses about changes in level. In addition, shapes, variances, and covariances of change are controlled for error, according to the error structure specified by the user.

In sum, LGM and MLM allow researchers to represent a wide range of change processes in a flexible and reliable manner. To make optimal use of these benefits, researchers need to provide a formal expression of guiding hypotheses and plausible competitors, and to choose an appropriate representation of available data. That is, they need to specify which questions to ask to the data, and to arrange the available data in a manner that maximizes the chances of obtaining a meaningful answer. Depending upon the questions asked and the data at hand, this process is either facilitated by LGM or MLM. In the following, we first focus on notation and commonalities between LGM and MLM, and then list their discriminating features.

Latent growth and multilevel models: Commonalities

In the simplest longitudinal case, LGM and MLM aim at estimating the parameter values of a specified function representing the longitudinal change of the sample analyzed. The parameters ought to minimize the discrepancies between the shape of the assumed change function and observed individual trajectories. All individuals' change trajectories are assumed to follow the same curve. Individual differences are expressed as deviations from average (population) parameters.

A popular expression for the change function is a linear curve, defined as the sum of a time-independent intercept and a time-dependent slope:

$$Y_{i,j} = \beta_{0,i} + \beta_{1,i} \cdot T_{i,j} + E_{i,j} \qquad (1)$$

The measurement $Y_{i,j}$ for individual i at time j is the sum of an intercept $\beta_{0,i}$ plus the linear slope coefficient $\beta_{1,i}$ multiplied by time plus the error component $E_{i,j}$. Note however that both $\beta_{0,i}$ and $\beta_{1,i}$ and $E_{i,j}$ have subscripts i. This indicates that each individual's change function is represented as a deviation from the population. The error component and the time indicator are both time-dependent. In the most common application the change function in equation (1) will lead to the estimation of six parameters:

1 The average intercept (the average of $\beta_{0,i}$);
2 The average slope (the average $\beta_{1,i}$);
3 The variance of the intercept (the variance of $\beta_{0,i}$);
4 The variance of the slope (the variance of $\beta_{0,i}$);
5 The covariance between intercept and slope (the covariance between $\beta_{0,i}$ and $\beta_{1,i}$); and
6 The error variance (the variance of $E_{i,j}$).

More parameters may be estimated by modeling the error structure. In the most usual application the error is assumed constant across time and uncorrelated (e.g., no autocorrelations over time).

Note that these six parameters are named differently in LGM and MLM terminology. In LGM, the six parameters are called in the order of presentation:

1 Average of the level factor,
2 Average of the change factor,
3 Variance of the level factor,
4 Variance of the change factor,
5 Covariance between level and change factors, and
6 Variance of the error factors.

In MLM the respective terms are:

1 Fixed intercept effect,
2 Fixed slope effect,
3 Random intercept effect,
4 Random slope effect,
5 Intercept-slope covariance (or level 2 covariance); and
6 Residual variance (or level 1 variance).

The change function of equation (1) can be used to evaluate basic hypotheses about change. Often, the covariance between level and change is of theoretical interest (do individuals with higher levels of performance show more positive change?). In interpreting the magnitude of this correlation, however, two things need to be kept in mind. First, this covariance is limited by the corresponding variances of level and change. Specifically, the covariance between level and change is not defined if level, change, or both do not display individual differences. Second, the correlation between level and change is dependent upon the point in time at which time equals zero, that is, the placement of the intercept. Caution is thus required when substantive interpretations of the correlation between level and change are being made.

With LGM and MLM, time can be represented in any possible way such as time in study, age from birth, or time to onset of disease. For each of these representations, the intercept can be shifted to the beginning, the center, or the end of the observation period, or to any other point in time. For most applications of MLM and LGM, it is advantageous to center the intercept at the middle of the observation period (Kreft, de Leeuw, and Aiken, 1995; Mehta and West, 2000; Wainer, 2000; see Figure 10.1 for an example).

Another generally overlooked property of change functions of the type in equation (1) is that several statistical variations are feasible. Both level and change may display one of four combinations of effects: only a fixed, average effect different from zero, only a random, variance effect different from zero, neither effect different from zero, or both effects different from zero. Many research applications of equation (1) simply assume that both level and change have a fixed, average effect different from zero, *and* a random, variance effect different from zero. However, while many permutations of level and change effects are possible ($4^2 = 16$), only those allowing random, variance effects for both level and change can define the covariance effect between level and change. Hence, only four of the possible 16 permutations define the covariance effect! To simply assume that both level and change have both fixed, average effect and random, variance effect may lead to flawed conclusions. Specifically, it is not meaningful to interpret covariances based on variances that do not reliably differ from zero.

Latent growth and multilevel models: Differences

Despite their overall similarity, researchers sometimes may prefer MLM over LGM and vice versa. Both LGM and MLM can accommodate incomplete data and unbalanced data structures (under missing at random assumptions). LGM require extra efforts for this adjustment because time is defined as a series of discrete variables, so that each measurement point has to be specified. If the data structure is highly unbalanced, if many data points are missing, or both, this results in a large number of missing observations (for an application in which each individual contributed at most three measurements to a curve covering 26 possible measurement occasions, see Ghisletta and McArdle, [2001]). To reduce missing data points, the time variable needs to be rounded into discrete bins (e.g., days, months, or years of age). With MLM, this problem does not exist because time is treated as a single continuous variable (Hedeker and Gibbons, 1997).

In most applications of LGM and MLM, the shape of average change is specified through the researcher. However, the relation between time and the variable of interest Y, which is linear in equation (1), can also be optimized in a data-driven manner. This exercise is straightforward with LGM because the loadings defining the change factor do not have to be specified a priori but can be estimated, as in ordinary factor analysis. Technically, two loadings still need to be specified (e.g., fixed) for identification purposes (see Ghisletta and McArdle, 2001). In MLM, this exercise is cumbersome but possible (see McArdle and Woodcock, in preparation).

At times, researchers may wish to specify the error structure. In equation (1) it is assumed that the error is constant in time and uncorrelated. This need not be the case. Substantive considerations may dictate an in-depth analysis of the error structure as well as nested comparisons between alternative structures. Depending upon the structures in question, this is more easily accomplished with LGM or with MLM. However, only LGM provides all possible degrees of freedom in the specification of the error structure.

In multivariate applications, relations among different variables are generally of greatest interest. In MLM, relations among time-varying variables are typically confined to covariances. For instance, the level of one variable may be found to correlate significantly with the change of another variable, or various changes may be found to correlate. With LGM, as with any other structural equation model, relations among time-varying variables can be freely specified. The level of one variable may thus be specified to exert a unidirectional, regression-type effect onto the change of another variable (as compared to a bidirectional covariance relation). With more advanced models and more complex theories (see Examples 2 and 3 below), greater flexibility currently mandates an LGM approach.

Time-invariant covariates are often included in the study of change. In the multilevel modeling tradition, which is strictly hierarchical (top-down) in orientation, time-invariant covariates are conceived as predictors or antecedents of time-varying variables. In contrast, LGM does not prescribe the relation between the two sets of variables. For instance, instead of assuming that SES predicts interindividual differences in cognitive change, the researcher may 'turn the arrow around' and posit that interindividual differences in cognitive change predict SES.

In summary, if there are many incomplete data and the design results in highly unbalanced data, if the change function is rather common (e.g., linear, polynomial, logistic, exponential), if the theory does not call for unusual error structures, if the multivariate application focuses on covariances among the variables' change features, and if time-invariant covariates are included as predictors, then MLM may be the preferred method of analysis. Note that these conditions cover a wide range of applications in the psychological literature. However, if the analyzer is able to accommodate unbalanced and incomplete data structure by a series of discrete variables without producing too many missing data points, if the shape of the change function is to be estimated by the data, if the error structure is of primary interest, if directional relations among the variables' change features are posited, and if time-invariant covariates are also considered as predicted variables or correlates, and not only as predictors, then LGM may be the preferred. Both methods are rather easily implemented thanks to new software (see Zhou, Perkins, and Hui, 1999). However, advanced applications of LGM, such as Examples 2 and 3 reported below, still require more than common familiarity with structural equation modeling software.

Research examples: Participants, observations, and variables

The following three illustrations all concern various aspects of the longitudinal cognitive and sensory test batteries of the Berlin Aging Study (BASE) (Baltes and

Mayer, 1999). BASE is a multidisciplinary study of aging, involving psychology, sociology and social policy, internal medicine and geriatrics, and psychiatry. The sample originated from a random draw of addresses from the city registry of West Berlin. On the first occasion, it comprised 516 individuals and was stratified by age and gender (mean age = 85 years; age range = 70 – 103 years) (cf. Lindenberger *et al.*, 1999). For a summary of research on intellectual functioning in the context of BASE, see Lövdén, Ghisletta, and Lindenberger (in press).

Started in 1989, the BASE longitudinal design currently consists of five measurement occasions. Data collection and data entry has been completed up to the fourth occasion. The first, third, and fourth measurement occasions include a reduced multidisciplinary measurement protocol (Intake assessment) followed by an in-depth, discipline-specific follow-up (Intensive protocol); the second occasion was limited to the Intake assessment. Table 10.1 displays the longitudinal design up to the fourth measurement occasion. For each occasion, the relevant variables and the effective sample sizes are specified. The average distance in time to the first occasion is also provided. As is generally true for longitudinal studies of aging populations, the effective sample decreased considerably over occasions (see also Lindenberger, Singer, and Baltes, 2002 and Singer, *et al.*, 2003).

Four intellectual abilities were included in the longitudinal test battery, represented by two variables each: Perceptual speed was marked by Digit Letter

Table 10.1 Description of cognitive and sensory variables from the BASE longitudinal design

	IA_1	IP_1	IA_2	IA_3	IP_3	IA_4	IP_4
				Occasion of measurement			
Variables							
Perceptual speed	DL	DL, IP	DL	DL	DL, IP	DL	DL, IP
Memory	PA, MT			PA, MT			PA, MT
Fluency		CA, WB	CA	CA	CA, WB	CA	CA, WB
Knowledge		VO, SW			VO, SW		VO, SW
Vision		CV, DV	CV,DV		CV, DV		CV, DV
Hearing		H	H		H		H
Time in study	0.00	0.13	1.95	3.76	3.99	5.55	6.03
N	516	516	361	244	208	164	132

Note. IA = Intake Assessment; IP_n = Intensive Protocol on the n-tn longitudinal measurement occasion; DL = Digit Letter; IP = Identical Pictures; PA = Paired Associates; MT = Memory for Text; CA = Categories; WB = Word Beginnings; VO = Vocabulary; SW = Spot-a-Word; CV = close visual acuity; DV = distance visual acuity; H = hearing. Time in study is calculated as average in years.

(DL) and Identical Pictures (IP), verbal knowledge by Vocabulary (VO) and Spot-a-Word (SW), episodic memory by Memory for Text (MT) and Paired Associates (PA), and fluency by Categories (CA) and Word Beginnings (WB) (for details, see Lindenberger, Mayr, and Kliegl, 1993). The sensory variables included close visual acuity (CV) and distant visual acuity (DV), assessed with regular Snellen reading tables, as well as auditory acuity, measured by pure-tone thresholds across four frequencies in both ears (H; for details see Marsiske *et al.*, 1999).

Research example number 1

Covariance structures of level and change analyzed by means of MLM followed by exploratory factor analysis

In this first example, we will describe an analysis of interrelations of change among cognitive and sensory variables of BASE (Lindenberger and Ghisletta, 2002). The focus of the analysis was on exploring the amount of shared variance among cognitive and sensory variables, especially with respect to change. The eight cognitive and two vision variables in Table 10.1 were analyzed; auditory measurements were aggregated into a single hearing indicator. We first wished to obtain the ideal specification of the change function for each variable according to equation (1). Then we wanted to calculate the amount of shared random, variance effects among the variables' changes. We assumed, for each variable, that error variances are constant and uncorrelated. As mentioned above, dropout caused the data to be incomplete, and the fact that participants were not measured at identical time intervals implied unbalanced data. Hence, the MLM approach proved ideal. Change was defined as the passing of time in years since the beginning of the study. *We thus applied the time = time in study definition.* To provide unbiased estimates of covariances between level and change, the data were linearly transformed for each individual such that the time in study summed over all available data points was zero (centering; see Figure 10.1).

At first we investigated the functional expression of each variable's change, in particular testing for random, variance effects of change. This implied 11 separate univariate growth models. All 11 variables displayed reliable fixed, average as well as random, variance level effects. Nine (that is, all but VO and HE) displayed reliable fixed, average change effects while only six (DL, IP, PA, CA, CV, and DV) displayed significant random, variance effects of change over the six years of observations. Hence, only six variables could define their correlation between level and change and could potentially share change variance among each other.

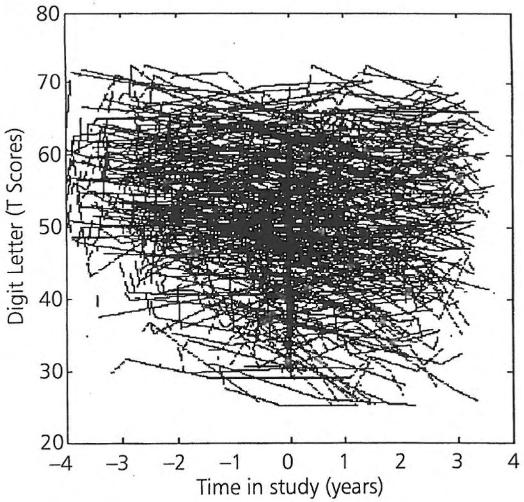

Fig. 10.1 Performance on the Digit Letter, a test of perceptual speed, as a function of time in study. For each individual, data are centered in time. Individuals with incomplete data are included in the analysis.

In a second step, we specified one multivariate growth model with 11 variables, which defined 17 random, variance effects (11 for the level and six for the change components). Besides estimating the analogous parameters of the 11 separate univariate models, this model also contained the relations among the 11 variables. In particular, it estimated a 17 by 17 covariance matrix expressing the relations among all non-zero level and change components. The previous univariate analyses were restricted to the diagonal of the random, variance effects matrix, that is, the variance of the random, variance effects, plus the covariances between level and change within each variable. The additional 130 ($= 17^{*}\ 18/2 - 17 - 6$) elements can only be estimated through multivariate analysis. The parameters analogous to the 11 univariate models of the previous step were very closely approximated by the estimates of the multivariate model.

The resulting multivariate covariance matrix was finally analyzed in a third step by exploratory factor analysis. The two factors with eigenvalues greater than one were extracted and obliquely rotated. Generally, the level variables had positive, moderate to high loadings on the first factor, whereas the six change variables showed positive, moderate to high loadings on the second factor. The two factors were positively correlated ($r = 0.41$), suggesting that individuals with higher levels of performance showed less negative change. Taken together, the two factors explained 61 per cent of the total variance across the 17 measures. Thus, this analysis suggests that a sizeable portion of the variance in level and change is shared across intellectual and visual domains.

Research example number 2

Testing the dedifferentiation hypothesis of old-age cognition

The second research example refers to a formal test of the dedifferentiation of cognitive abilities hypothesis in old and very old-age (Ghisletta and Lindenberger, in press). According to this hypothesis (Baltes *et al.*, 1998; Lindenberger, 2001), behavioral aging is marked by a contraction of 'functional cerebral space' (Kinsbourne and Hicks, 1978), or by decreasing specificity and increasing overlap among representations, induced by biological changes in the aging brain. With respect to the distinction between broad fluid and broad crystallized abilities (Horn, 1989), or the mechanics and pragmatics of cognition (Baltes, 1987), the dedifferentiation hypothesis of old-age cognition predicts that changes in the mechanics *drive* (i.e., temporally precede and causally influence) age changes in the pragmatics.

Thus far, evidence regarding dedifferentiation is cross-sectional and indirect. In contrast to cross-sectional data, longitudinal data allow for a direct test of age-based dedifferentiation dynamics. We addressed the dedifferentiation hypothesis by concentrating on perceptual speed as a marker of the mechanics, and knowledge as a marker of the pragmatics of cognition. Each of the two abilities was represented by two variables, DL and IP for perceptual speed, Vo and SW for knowledge.

Unfortunately, standard LGM or MLM do not allow for the specification of lead – lag relations because relations among time-dependent processes are modeled symmetrically. However, a recently developed variant of LGM, called the Dual Change Score Model (DCM) (McArdle, 1986, 2001a, 2001b; McArdle, Hamagami, Meredith *et al.*, 2000; McArdle and Hamagami, 2001), can be used for this purpose. In addition to modeling the time-invariant (stationary and often linear) change process, DCM has a 'dynamic' parameter, expressing the effect of a variable's true score at t-1 time on the true change of that variable from t-1 to t. This extra parameter is accommodated by respecification and extension of the standard LGM approach; the MLM approach cannot be adapted correspondingly.

The dedifferentiation hypothesis is clearly longitudinal in nature and is usually formulated in relation to chronological age; the process of dedifferentiation is assumed to unfold with advancing age. Therefore, in the present case, and in

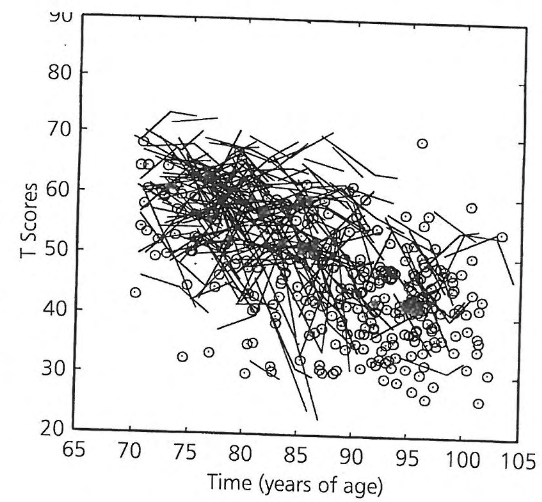

Fig. 10.2 Performance on perceptual speed, a unit-weighted composite of Digit Letter and Identical Pictures, as a function of chronological age. Individuals with incomplete data are included in the analysis. The resulting age trajectory is influenced by both cross-sectional and longitudinal data components.

contrast to the previous example, time was defined in terms of chronological age (see Figure 10.2; cf. Raudenbush and Chan, 1992). Within the DCM, the expression of the variable Y at any time point is directly dependent upon its expression at the previous time point:

$$Y_{i,j} = (1 + \beta) Y_{i,j-1} + \bar{S} \tag{2}$$

where $Y_{i,j}$ is the true value Y of individual i at time j, β is the dynamical, auto-regressive parameter, $Y_{i,j-1}$ is the true value Y of individual i at time $j - 1$, and $\bar{S}$ is the average of the slope factor (McArdle, 2001b). The multivariate extension of the DCM easily adds a similar dynamic parameter, say γ, quantifying the effect of another variable for individual i at time $j - 1$, say $Z_{i,j-i}$ on the value of Y at time t. Hence equation (2) is expanded to

$$Y_{i,j} = (1 + \beta) Y_{i,j-1} + \gamma \cdot Z_{i,j-1} + \bar{S} \tag{3}$$

Equation 3 estimates the strength of the auto-proportion (β) as well as that of the proportion of the second variable (γ). In such a system of two equations we can compare the effect sizes of β and γ and draw conclusions on the relative importance of each variable upon its change and upon the change of the other variable. Hence, within the system analyzed, possible 'leading' and 'lagging' variables may be statistically identified.

The bivariate DCM applied to perceptual speed and knowledge clearly indicated that perceptual speed is the leader of the system and knowledge the lagger (for statistical tests, see Ghisletta and Lindenberger, in press). As a consequence, the variance of knowledge is increasingly saturated by variance in perceptual speed with advancing age. Thus, this application provides strong longitudinal evidence for the hypothesis of age-based directional dedifferentiation.

Research example number 3

Time-based structural dynamics between intellectual and visual changes in old age

The final example refers to the longitudinal exploration of relations between intellectual and sensory abilities (Ghisletta and Lindenberger, 2002). Specifically, we examined three current hypotheses about the link between intellectual and sensory domains in old age: the mediation hypothesis (Anstey, 1999), the common cause hypothesis (Lindenberger and Baltes, 1994b), and the cascade hypothesis (Birren, 1964). Two intellectual abilities, perceptual speed and knowledge, and two sensory abilities, close and distance visual acuity, were included in the analysis to yield a quadrivariate DCM. In contrast to the second research example, the basic time dimension was defined in terms of measurement occasion instead of chronological age. This focus on measurement occasions was in part motivated by practical considerations; given the large amount of missing data and the maximum longitudinal observation period of six years spread out over 34 years of age, a model with age as the basic time dimension would have run the risk of empirical under-identification. Instead, age was specified as a covariate of the level factors and as a possible direct influence on latent changes of each of the four variables. In this manner, we were able to examine whether a variable is changing, whether its change is affected by itself, by one of the other three variables, and by chronological age.

The results indicated that individuals changed on all four variables across occasions. The average, group changes were more marked for close vision and perceptual speed than for distance vision and knowledge. Occasion-based changes in perceptual speed were reliably affected by chronological age and close vision, changes in knowledge by chronological age and itself, changes in close vision by itself and distance vision, and changes in distance vision by perceptual speed and knowledge. Intellectual and sensory abilities were intimately connected in time, but neither one was dominating the other. Thus, contrary to the previous age-based specification of the DCM for perceptual speed and knowledge, no single leader of change emerged. As argued in detail elsewhere (Ghisletta and Lindenberger, 2002), this finding provides limited support for the common-cause hypothesis, and no support for the mediation and cascade hypotheses.

Concluding remarks

The study of age changes in behavior requires the concerted use of different statistical tools and research designs (Baltes, Reese, and Nesselroade, 1988; Hertzog, 1996). By structuring longitudinal change, LGM and MLM make a unique contributions towards a better understanding of the structural dynamics and causal mechanisms of behavioral aging. However, as is true for other methods, LGM and MLM are not without problems and limitations. In the context of longitudinal investigations of behavioral aging, four issues seem especially relevant.

First, LGM and MLM generally assume sample homogeneity (cf. Molenaar *et al.*, 2003). Each individual's change is described as the deviation from a single average curve. This procedure makes only limited sense if the shape of the average curve differs from the shape of individual curves, or if the sample consists of a mixture of subsamples with different change trajectories (for an early treatment of this issue in the gerontological literature, see Baltes and Labouvie, 1973: 146; cf. Wohlwill, 1970). For instance, modal shapes of change appear to differ between individuals with and without a dementing illness before this illness has reached a clinical threshold (Bäckman, Small, and Fratiglioni, 2001; Sliwinski *et al.*, 1996). Uncritical use of the full-information approach, in which each individual is included in the estimation of the overall curve regardless of how much data are missing, may hide sample heterogeneity and distort shapes of change. Thus, an important goal in longitudinal research is to identify early markers of subgroup membership, and to examine group differences in change trajectories. A useful means to this end is to separately model change in the total sample and in subgroups to examine the existence of systematic differences in shape of change (Singer *et al.*, 2003).

Second, a more general but related assumption of LGM and MLM is ergodicity, or the assumption that the structure of interindividual differences is governed by the same regularities as the structure of within-person changes. If the ergodicity assumption does not hold, the analysis of interindividual differences may yield an invalid picture of within-person changes. Unfortunately, lack of ergodicity is not detected easily (Molenaar *et al.*, 2003). Innovative formal analyses and quantitative simulations are needed to determine whether certain violations of the ergodicity assumption are more detrimental to the validity of between-person differences in the study of ontogenetic change than others (Lövdén and Lindenberger, in press).

Third, both LGM and MLM assume that data are missing at random (MAR). In longitudinal studies of aging, this assumption is often violated. In dealing with this fundamental problem, several strategies seem productive. For example, one may consider, in each particular case, whether violation of the MAR assumption precludes an unbiased and meaningful test of the relevant guiding hypothesis. Also, attempts should be made to explicitly model sample selectivity by including key markers such as onset of disease or distance to death. This may require multiple representations of time within the same model. For instance, time in study may serve as the basic time dimension, and chronological age, distance to death, or distance to onset of disease may be specified as antecedents or consequents of change.

Finally, with typical longitudinal data, variances and covariances of change are generally small in relation to variances and covariances in level. As argued above, variances that differ reliably from zero are a necessary condition for investigating structural hypotheses about covariances. However, the sampling distributions of covariances may be overly wide with small variances, even if they differ significantly from zero. Again, formal analyses and statistical simulations are needed to examine the robustness of LGM and MLM under such conditions.

To a large degree, the problems listed above refer to fundamental aspects of behavioral aging and ontogenetic change in general. Thus, working on these problems in the context of LGM and MLM represents not only a methodological but also a conceptual challenge. Both seem well worth the effort.

Author note

Preparation of this chapter was financially supported by a grant from the Deutsche Forschungsgemeinschaft to the Ulman Lindenberger; SFB 378, project: Aging, Resources, and Cognition (ARC).

References

Anstey, K. J. (1999). Construct overlap in resource theories of memory aging. *Gerontology,* **45**, 348–350.

Anstey, K. J., Lord, S. R., and Williams, P. (1997). Strength in the lower limbs, visual contrast sensitivity and simple reaction time predict cognition in older women. *Psychology and Aging,* **12**, 137–144.

Bäckman, L., Small, B. J., Wahlin, A., and Larsson, M. (2000). Cognitive functioning in very old age. In F. I. M. Craik and T. A. Salthouse (eds), *The Handbook of Aging and Cognition* (2nd ed., pp. 499–558). Mahwah, NJ: Lawrence Erlbaum Associates.

Bäckman, L., Small, B. J., and Fratiglioni, L. (2001). Stability of the preclinical episodic memory deficit in Alzheimer's disease. *Brain,* **124**, 96–102.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology,* **23**, 611–626.

Baltes, P. B., and Labouvie, G. V. (1973). Adult development of intellectual performance: Description, explanation, modification. In C. Eisdrofer and M. P. Lawton (eds), *The Psychology of Adult Development and Aging.* Washington, DC: Americal Psychological Association.

Baltes, P. B., Lindenberger, U., and Staudinger, U. M. (1998). Life-span theory in developmental psychology. In R. M. Lerner (ed.), *Theoretical Models of Human Development. Vol. 1 of the Handbook of Child Psychology* (pp. 1029–1143). New York: Wiley.

Baltes, P. B., and Mayer, K. U. (1999). *The Berlin Aging Study: Aging from 70 to 100.* New York: Cambridge University Press.

Baltes, P. B., Reese, H. W., and Nesselroade, J. R. (1988). *Life-span Developmental Psychology: Introduction to research methods* (reprint of 1977 ed.). Hillsdale, NJ: Erlbaum.

Birren, J. E. (1964). *The Psychology of Aging.* Englewood Cliffs, N.J.: Prentice-Hall.

Bryk, A. S., and Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin,* **101**, 147–158.

Cabeza, R. (2001). Functional neuroimaging of cognitive aging. In R. Cabeza and A. Kingstone (eds), *Handbook of Functional Neuroimaging of Cognition* (pp. 331–377). Cambridge, MA: MIT Press.

Cabeza, R. (2002). Hemispheric asymmetry reduction in old adults: The HAROLD Model. *Psychology and Aging,* **17**, 85–100.

Chou, C.-P., Benter, P. M., and Pentz, M. A. (1998). Comparisons of two statistical approaches to study growth curves: The multilevel model and latent curve analysis. *Structural Equation Modeling,* **5**, 247–266.

Dunn, J. C., and Kirsner, K. (1988). Discovering functionally independent mental processes: The principle of reversed association. *Psychological Review,* **95**, 91–101.

Ghisletta, P., and Lindenberger, U. (in press). Age-based structural dynamics between perceptual speed and knowledge in the Berlin Aging Study: Direct evidence for ability dedifferentiation in old age. *Psychology and Aging.*

Ghisletta, P., and Lindenberger, U. (2002). Exploring the structural dynamics of the link between sensory and cognitive functioning in old age: Longitudinal evidence from the Berlin Aging Study. Manuscript in preparation.

Ghisletta, P., and McArdle, J. J. (2001). Latent growth curve analyses of the development of height. *Structural Equation Modeling,* **8**, 531–555.

Hedeker, D., and Gibbons, R. D. (1997). Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological Methods,* **2**, 64–78.

Hertzog, C. A. (1996). Research design in studies of aging and cognition. In J. E. Birren and K. W. Schaie (eds), *Handbook of the Psychology of Aging* (pp. 24–37). San Diego, CA: Academic Press.

Horn, J. L. (1989). Models of intelligence. In R. L. Linn (ed.), *Intelligence: Measurement, theory, and public policy* (pp. 29–73). Urbana, IL: University of Illinois Press.

Kinsbourne, M., and Hicks, R. E. (1978). Functional cerebral space: A model for overflow, transfer and interference effects in human performance: A tutorial review. In J. Requin (ed.), *Attention and Performance VII.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Kliegl, R., Mayr, U., and Krampe, R. T. (1994). Time-accuracy functions for determining process and person differences: An application to cognitive aging. *Cognitive Psychology,* **26**, 134–164.

Kreft, I. G. G., de Leeuw, J., and Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research,* **30**, 1–21.

Laird, N. M., and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics,* **38**, 963–974.

Li, S.-C., and Lindenberger, U. (1999). Cross-level unification: A computational exploration of the link between deterioration of neurotransmitter systems and dedifferentiation of cognitive abilities in old age. In L.-G. Nilsson and H. Markowitsch (eds), *Cognitive Neuroscience of Memory* (pp. 104–146). Toronto: Hogrefe and Huber.

Lindenberger, U. (2001). Lifespan theories of cognitive development. In N. J. Smelser and P. B. Baltes (eds), *International Encyclopedia of the Social and Behavioral Sciences.* Oxford, UK: Elsevier.

Lindenberger, U., and Baltes, P. B. (1994a). Aging and intelligence. In R. J. Sternberg (ed.), *Encyclopedia of Human Intelligence* (Vol. 1, pp. 52–66). New York: Macmillan.

Lindenberger, U. and Baltes, P. B. (1994b). Sensory fuctioning and intelligence in old age: A strong connection. *Psychology and Aging,* **9**, 339–355.

Lindenberger, U., Gilberg, R., Nuthmann, R., Pötter, U., Little, T. L., and Baltes, P. B. (1999). Sample selectivity and generalizability of results from the Berlin Aging Study. In P. B. Baltes and K. U. Mayer (eds), *The Berlin Aging Study: Aging from 70 to 100* (pp. 56– 82). New York: Cambridge University Press.

Lindenberger, U., and Ghisletta, P. (2002). *Cognitive and sensory changes in the Berlin Aging Study: Longitudinal analyses of the total sample.* Paper presented at the Cognitive Aging Conference, Atlanta, Georgia.

Lindenberger, U., Mayr, U., and Kliegl, R. (1993). Speed and intelligence in old age. *Psychology and Aging,* **8**, 207–220.

Lindenberger, U., and Pötter, U. (1998). The complex nature of unique and shared effects in hierarchical linear regression: Implications for developmental psychology. *Psychological Methods,* **3**, 218–230.

Lindenberger, U., Singer, T., and Baltes, P. B. (2002). Longitudinal selectivity in aging populations: Separating mortality-associated versus experimental components in the Berlin Aging Study (BASE). *Journal of Gerontology: Psychological Science.*

Little, T. D., Schnabel, K. U., and Baumert, J. (eds) (2000). *Modeling Longitudinal and Multilevel Data: Practical issues, applied approaches, and specific examples.* Mahwah, NJ: Lawrence Erlbaum Associates.

Lövdén, M., Ghisletta, P., and Lindenberger, U. (in press). Cognition in the Berlin Aging Study: The first ten years. *Aging, Neuropsychology, and Cognition.*

Lövdén, M., and Lindenberger, U. (in press). Development of intellectual abilities in old age: From age gradients to individuals. In O. Wilhelm and R. W. Engle (Eds.), *Understanding and measuring intelligence.* Thousand Oaks, CA: Sage.

McArdle, J. J. (1986). Latent growth within behavior genetic models. *Behavior Genetics,* **16**, 163–200.

McArdle, J. J. (2001a). Growth curve analysis. In N. J. Smelser and P. B. Baltes (eds), *The International Encyclopedia of the Behavioral and Social Sciences.* New York, NY: Pergamon Press.

McArdle, J. J. (2001b). A latent difference score approach to longitudinal dynamic structural analyses. In R. Cudeck, S. duToit, and D. Sorbom (eds), *Structural Equation Modeling: Present and future* (pp. 342–380). Lincolnwood, IL: Scientific Software International.

McArdle, J. J. and Hamagami, F. (1996). Multilevel models from a multiple group structural equation perspective. In G. A. Marcoulides and R. E. Schumaker (eds), *Advanced Structural Equation Modeling. Issues and techniques* (pp. 89–124). Mahwah, NJ: Lawrence Erlbaum Associates.

McArdle, J. J. and Hamagami, F. (2001). Linear dynamic analyses of incomplete longitudinal data. In L. M. Collins and M. Sayer (eds), *New Methods for the Analysis of Change.* Washington, DC: American Psychological Association.

McArdle, J. J., Hamagami, F., Meredith, W., and Bradway, K. P. (2000). Modeling the dynamic hypotheses of Gf-Gc theory using longitudinal life-span data. *Learning and Individual Differences,* **12**, 53–79.

McArdle, J. J., and Woodcock, R. W. (in preparation). Longitudinal multilevel analyses of test-retest data of the growth and decline of cognitive abilities.

MacCallum, R. C., Kim, C., Malarkey, W. B., and Kiecolt-Glaser, J. K. (1997). Studying multivariate change using multilevel models and latent curve models. *Multivariate Behavioral Research,* **32**, 215–253.

Marsiske, M., Delius, J., Maas, I., Lindenberger, U., Scherer, H., and Tesch-Römer, C. (1999). Sensory systems in old age. In P. B. Baltes and K. U. Mayer (eds), *The Berlin Aging Study: Aging from 70 to 100.* London, UK: Cambridge University Press.

Mehta, P. D., and West, S. G. (2000). Putting the individual back into individual growth curves. *Psychological Methods,* **5**, 23–43.

Molenaar, P. C. M., Huizenga, H. M., and Nesselroade, J. R. (2003). The relationship between the structure of inter-individual and intra-individual variability: A theoretical and empirical vindication of developmental systems theory. In U. M. Staudinger and U. Lindenberger (eds), *Understanding Human Development: Lifespan psychology in exchange with other disciplines* (pp. 339–360). Dordrecht, NL: Kluwer Academic Publishers.

Raudenbush, S. W., and Chan, W. (1992). Growth curve analysis in accelerated longitudinal designs. *Journal of Consulting and Clinical Psychology,* **61**, 941–951.

Schreurs, B. G., Bahro, M., Molchan, S. E., Sunderland, T., and McIntosh, A. R. (2001). Interactions of prefrontal cortex during exeblink conditioning as a function of age. *Neurobiology of Aging,* **22**, 237–246.

Singer, T., Verhaeghen, P., Ghisletta, P., Lindenberger, U., and Baltes, P. B. (2003). The fate of cognition in very old age: Six-year longitudinal findings in the Berlin Aging Study (BASE). *Psychology and Aging,* **18**, 318–331.

Sliwinski, M., Lipton, R. B., Buschke, H., and Stewart, W. (1996). The effects of preclinical dementia on estimates of normal cognitive functioning in aging. *Journal of Gerontology: Psychological Sciences,* **50B**, P162–P170.

Wainer, H. (2000). The centercept: An estimable and meaningful regression parameter. *Psychological Science,* **11**, 434–436.

Wohlwill, J. F. (1970). The age variable in psychological research. *Psychological Review,* **77**, 49–64.

Zhou, X.-H., Perkins, A. J., and Hui, S. L. (1999). Comparisons of software packages for generalized linear multilevel models. *The American Statistician,* **53**, 282–290.