Infant and Child Development
Inf. Child Dev. **15**: 627–634 (2006)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/icd.483

Commentary

On Growth Curves and Mixture Models

Jan B. Hoeksma* and Henk Kelderman
Vrije Universiteit, Amsterdam, The Netherlands

The multilevel model of change and the latent growth model are flexible means to describe all sorts of population heterogeneity with respect to growth and development, including the presence of sub-populations. The growth mixture model is a natural extension of these models. It comes at hand when information about sub-populations is missing and researchers nevertheless want to retrieve developmental trajectories from sub-populations. We argue that researchers have to make rather strong assumptions about the sub-populations or latent trajectory classes in order to retrieve existing population differences. A simulated example is discussed, showing that a sample of repeated measures drawn from two sub-populations easily leads to the mistaken inference of three sub-populations, when assumptions are not met. The merits of methodological advises on this issue are discussed. It is concluded that growth mixture models should be used with understanding, and offer no free way to growth patterns in unknown sub-populations. Copyright © 2006 John Wiley & Sons, Ltd.

Key words: longitudinal models; mixed models; growth curves; multilevel models

GROWTH MODELS

Wohlwill (1973) argued that chronological age is part and parcel of the dependent variable in developmental psychology. He maintained that developmental psychology is about identifying developmental functions, about what is now known as developmental trajectories. Developmental functions describe how specific characteristics or attributes change with age. Once developmental functions are identified, researchers should try to explain them and find out how

*Correspondence to: Jan B. Hoeksma, Vrije Universiteit, Developmental Psychology, van der Boechorststraat1, 1081 BT Amsterdam, The Netherlands.

they vary. Some years later Baltes and Nesselroade (1979) described the rationales of longitudinal research. They argued that longitudinal research aims at describing and explaining individual differences in individual change.

Most notably the methodological proposals of Wohlwill, and Baltes and Nesselroade appeared in the 1970s, at a time when statisticians were highly skeptical about measuring and analysing individual change. These doubts largely disappeared through the work on longitudinal data analyses of Laird and Ware (1982), Goldstein (1986) and Longford (1987). Although their papers were highly technical, and therefore hardly accessible to developmentalists, the proposed statistical models soon found their way in the developmental research. In the Old and New World, the statistical models became known as the multilevel model (Goldstein, 1987) and the hierarchical linear model (Bryk & Raudenbush, 1992), respectively.

Today, developmental researchers are more and more inclined to think about development in terms of developmental trajectories and developmental curves. The psychological literature offers many fine examples of research adhering to the methodology of Wohlwill, and Baltes and Nesselroade, by means of the multilevel model of change. However, as Singer and Willet (2003) argue, that the number of applications is still small given the generality of the model.

LATENT GROWTH MODELS

Through the work of Meredith and Tisak (1990), it soon became apparent that individual differences in growth and development could also be analysed by means of structural equation models (SEM). Other early contributions showing the flexibility of the structural equations model in analysing change were McArdle (1986a, b, 1989), McArdle and Epstein (1987) and Muthén (1989, 1991, 1992). When viewed from a SEM perspective the model is called latent growth model. An excellent explication on how the multilevel model of change translates into a latent growth model can be found in Singer and Willet (2003).

The multilevel model of change and the latent growth model are equivalent from a mathematical point-of-view (Willett & Sayer, 1994). Statistically they differ on how model parameters are estimated. The algorithms used for that purpose are quite different. The multilevel model of change and the latent growth model thus require different computer programs, although the results obtained will hardly differ from one another.

However, the multilevel model of change and the latent growth model differ in the researchers mind, as models shape the way researcher think about their phenomenon they study (Hoeksma, 1999). Figure 1(a) shows a population of growth curves. Only a small number is given, because the picture would otherwise become cluttered. When looking at Figure 1(a), researchers with a latent growth model in mind will think of two *latent growth factors* accounting for the growth pattern. The factors are thought to refer to underlying constructs. Researchers familiar with the multilevel model of change would think of *individual growth* curves that make up the observed pattern and observe how the variability increases with time or age (Hoeksma, 2005). As recently shown by Curran, Bauer, and Willoughby (2004), these differences may have profound effects on the validity of researcher's interpretations.

Suppose that the population of growth curves consists of two sub-populations having different growth patterns (which, by the way, it does). Figure 1(b) displays the mean growth curves of the *two* sub-populations underlying the growth

DOI: 10.1002/icd

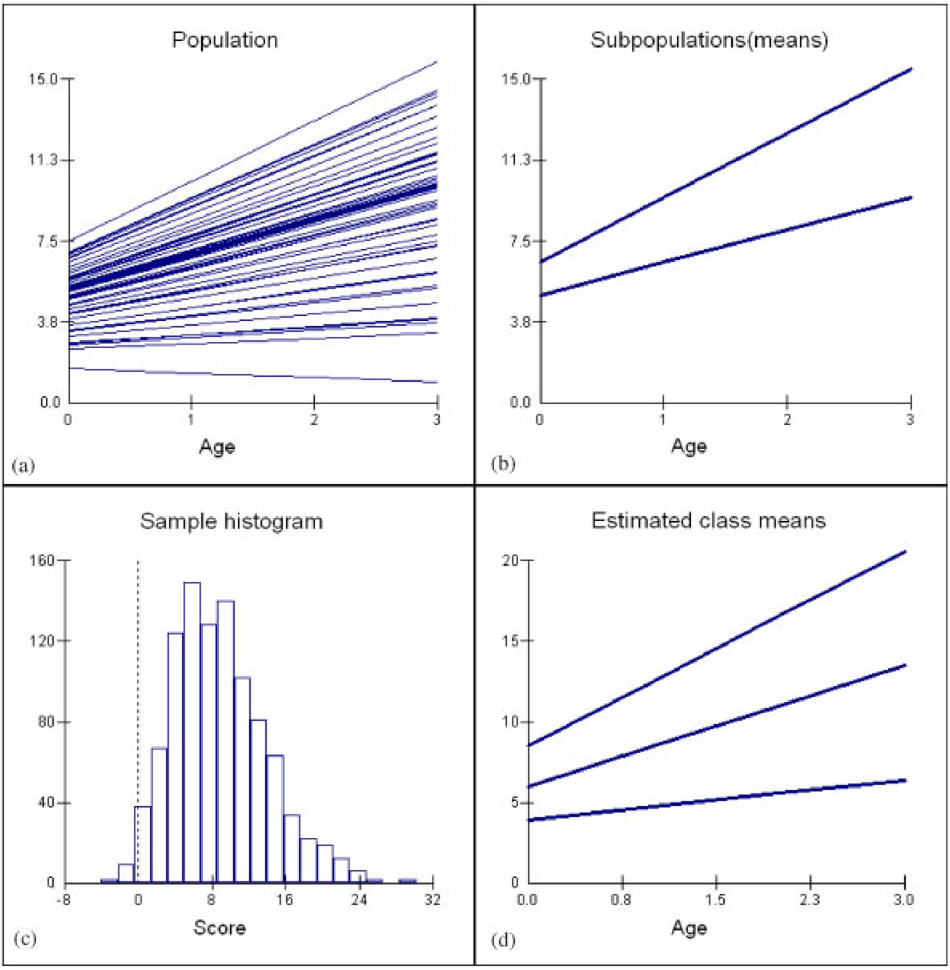

Figure 1. Population of growth curves (a). Mean growth curves in two sub-populations making up the population (b). Sample histogram of 4 equidistant observations of $N = 250$ persons (c). Estimated mean trajectories (mixed growth model) from the sample (1d).

curves in Figure 1(a). A sample of observation from the *two* sub-populations, with measurement error added, is portrayed in Figure 1(c). The histogram shows the distribution of 1000 observations, consisting of four repeated measurements of 250 individuals, randomly sampled from the distribution in Figure 1(a). The distribution is clearly non-normal.

Looking for group differences, researchers using latent growth models will perform a multigroup analysis and try to find out in what respect the *latent growth factors* including their *mean structure* differ between both groups (Muthén & Muthén,1998). In addition, they will look for possible differences of the residuals between both groups. Researchers using the multilevel model of change will try to find out whether the *mean growth patterns* differ between both groups and whether *inter-individual differences in intra-individual change* are homogeneous or non-homogeneous across groups. Moreover, they will try to model possible

DOI: 10.1002/icd

differences in the residual distributions (Goldstein, 2003). Whether the latent growth model or the multilevel model of change is used to analyse the data in Figure 1(c), the population difference displayed in Figure 1(a) and (b) would be largely retrieved, be it within sampling error. The model found would thus account for the distribution in Figure 1(c).

Especially, the multilevel model of change (when analysed using MlwinN, Goldstein *et al.*, 1998) is rather flexible with regard to accommodating non-homogeneity, including qualitatively different distributions in different groups. There is thus no ground for the suggestion made by Nagin (1999) and Nagin and Tremblay (2001), and reiterated by Connel and Frye (2006), that growth models assume 'that there is a single underlying distribution with respect to changes over time about which children are normally distributed' (Connell & Frye, 2006). On the contrary, heterogeneity can be studied quite well within a multilevel framework. But if the data come from a priori unknown sub-populations, growth mixture models can be very helpful.

GROWTH MIXTURE MODELS

Growth mixture models can be used when the number of groups and group membership are *unknown*. That is, when the data are expected to come from *unobserved* sub-populations, and dummies indicating group membership are missing. Growth mixture models (cf. Muthén & Muthén, 2000) and their special case the semi-parametric groups-based method proposed by Nagin (1999) offer suitable means to analyse growth trajectories from unobserved sub-populations. A nice introductory example is given by Connell and Frye (2006). A balanced and technically more adequate example can be found in Muthén and Muthén (2000).

Growth mixture models can be seen as an advanced cluster method. Persons with common growth patterns are 'taken together' to make up a cluster or group (Nagin, 1999). Each resulting cluster or group is characterized by a common growth pattern that differs from the patterns in other groups. An analysis typically results in two to five groups with different growth patterns (cf. Broidy *et al.*, 2003; van Lier, & Crijnen 2005). Resulting groups may include increasers, decreasers, and high and low groups of no change. Akin to other clustering techniques based on latent mixture distributions (cf. Everitt and Dunn 2001), growth mixture models maximize homogeneity or similarity of trajectories within clusters and heterogeneity or dissimilarity of trajectories between clusters (Bauer & Curran, 2003). Besides the estimates of the common trajectories in each group, the analysis results in estimates of each person's probability of belonging to each of the different groups.

Dedicated software such as Mplus (Muthén & Muthén, 1998) or the SAS procedure described by Jones, Nagin, and Roeder (2001) do not find group-trajectories by (re-)distributing growth curves across different groups, but by using so-called mixture distributions. Mixtures are complex distributions consisting of a weighted sum of two or more elementary distributions. Mixtures are generally *non-normal*. The elementary distributions making up the mixture may be normal, censored-normal, Poisson or logistic (Muthén & Muthén, 1998; Nagin, 1999). When a researcher uses Mixed Growth Models, he or she assumes that the repeated measurements within each group follow a similar *known* distribution (e.g. normal distributions with varying means and standard deviations) which, when mixed together, account for the observed non-normal distribution.

DOI: 10.1002/icd

The sample distribution Figure 1(c) is non-normal. We analysed the data of Figure 1(c) by means of a mixed growth model to find (or retrieve) the trajectories in the sub-populations. We estimated so-called one class, two class, three class, and four-class models. The respective values of the Bayesian Information Criterion (BIC) were 5518.8, 5450.5, 5442.3, and 5456.3. The three-class model has the lowest BIC value, and thus fits better than the two and four-class models. It is the model to be preferred. Figure 1(d) portrays the mean growth trajectories of the model found. The results indicate that the sample distribution in Figure 1(b) can be conceived of as a mixture of *three* normal distributions, with age- or time-dependent means as given in Figure 1(d).

It may come as a surprise that the two sub-populations were not retrieved. Why did the growth mixture model extract three groups (classes), whereas the data came from only two sub-populations? The main reason is that the growth mixture model does not search for *existing groups* in the data (how could it), but for *optimal groups* that summarize the data most *parsimoniously.* In statistical terms, the model maximizes the homogeneity (similarity) within groups and heterogeneity (differences) between groups (Bauer & Curran, 2003). The growth mixtures model explains heterogeneity in a *sample* of repeated measurements by creating homogeneous groups. This sample heterogeneity, however, is not *necessarily only caused* by the presence of sub-populations. There are many other potential causes, including heteroscadisticity, sample fluctuations and most importantly non-normality of the original distribution(s) in the population and sub-populations (Bauer & Curran, 2003, 2004). In the present sample, the heterogeneity was the result of both population differences and non-normality in the sub-populations, whereas the latent growth mixture model only assumes population differences.

When information on grouping is missing, there will be a price to be paid to retrieve it. When one uses the growth mixture model, the price is paid by means of assumptions. First, it is assumed that the observed distribution is *non-normal*, and consists of a finite mixture of unobserved non-identical distributions. Second, it assumed that unobserved distributions have a *known* distributional form (e.g. normal) that is *similar* across the groups to be extracted. If either or both these assumptions are not met, the number of groups may be over-extracted. The problem is that non-normality in repeated measures may come about easily and for various reasons. Non-normality may reflect group differences, and also sample fluctuations, or a non-normal distribution in the population (e.g. a lognormal, skewed or peaked population distributions). Also method effects, such as response tendencies, memory effects, bottom and ceiling effects may produce non-normality. As was shown by Bauer and Curran (2003), even mild non-normality in an otherwise homogeneous population results in over-extraction.

VALIDITY OF GROWTH MODELS

If information on group membership has been collected, non-homogeneity can be modelled using the multilevel model of change and the latent growth model. The first model is possibly somewhat less restrictive than the latter. Growth mixture models should be seen as a natural extension of growth models. They are not an alternative. They allow researchers to derive latent classes from repeated measurements. Bauer and Curran (2003) showed however, that multiple trajectory classes can be estimated and appear optimal even if only one group exists in the population. Following a similar line of thought, our example shows

DOI: 10.1002/icd

that if two groups are present in the populations they do not have to be retrieved. In other words, researchers looking for existing population heterogeneity by means of growth mixture models, may easily draw wrong conclusions. This will happen when the sample data are non-normal for any other reason than existing population differences. There are thus as many causes for finding group differences as there are causes for sample non-normality (including the existing group differences in the population). Even design issues, including sample size (Nagin & Tremblay, 2001) and the number of measurement occasions may affect the number of latent classes found (Connell & Frye, 2006). For researchers looking for sub-populations (not for sub-groups in the sample), this is a rather annoying problem.

Unfortunately, the problem cannot be solved by means of replication. Replicating a study in a new sample will give a similar distribution reflecting the same population characteristics that lead to over-extraction of trajectory classes. The example discussed is one of many possible replications. The sample was simulated from given population model. Another replication (simulated sample) from the same model would (within sampling) error result in a similar outcome, i.e. three trajectory classes, because the distribution of observations reflects true non-normality that is not result of differences between sub-populations only.

A better option is to extend the Growth Mixture Model by means of explanatory variables, as suggested by Muthén and Muthén (2000). If one could show, for instance, in the present analysis that the three classes found were related to other well-known phenomenon, this apparently adds validity to the model. We agree, however, with Bauer and Curran (2003) that this comes down to 'affirming the consequence' and does not correspond to a common empirical test of a psychological hypothesis.

A final option is to mitigate the problem by saying that the model found is only a summary of the data and that the trajectory classes found do not have to correspond to psychological meaningful categories. The categories are not carved in stone. Although we think this view of models is legitimate from a statistical point-of-view it disregards both the goal and practices of researchers. Developmental researchers try to understand causal relationships. They think of categories as real or existent. Once their quantitative analyses are finished, they will try to make general statements about the categories or classes found. They will try to generalize their findings and write about these findings as if the categories exist and refer to existing sub-populations.

Should, given these considerations, mixed growth model be discarded from the armoury of statistical techniques? Of course not! In our view any method should be tried to further psychological knowledge (cf. Feyerabend, 1975). However, one should use these methods with an understanding about the assumptions they make, because only than they will lead to valid inferences. Alas, mixed growth models do not offer a free way to growth patterns in unknown sub-populations.

REFERENCES

Baltes, P. B., & Nesselroade, J. F. (1979). History and rationale of longitudinal research. In J. F. Nesselroade, & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 1–59). New York: Academic Press.

Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods*, **8**: 338–363.

DOI: 10.1002/icd

Bauer, D. J., & Curran, P. J. (2004). The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychological Methods*, **9**(1): 3–39.
Broidy, L. M., Tremblay, R. E., Brame, B., Fergusson, D., Horwood, J. L., Laird, R., *et al.* (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site cross-national study. *Developmental psychology*, **39**(2): 222–245.
Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.
Connel, A. M., & Frye, A. A. (2006). Growth mixture modeling in developmental psychology: Overview and demonstration of heterogeneity in developmental trajectories of adolescent antisocial behavior. *Infant and Child Development*, in press.
Curran, P. J., Bauer, D. J., & Willoughby, M. T. (2004). Testing main effects and interactions in latent curve analysis. *Psychological Methods*, **9**(2): 220–237.
Everitt, B. S., & Dunn, G. (2001). *Applied multivariate data analysis*. London: Arnold.
Feyerabend, P. (1975). *Against method*. London: New Left Books.
Goldstein, H. (1986). Multilevel mixed linear models analysis using iterative generalized least squares. *Biometrika*, **73**: 43–56.
Goldstein, H. (1987). *Multilevel models in educational and social research*. London: Oxford University Press.
Goldstein, H., Rashbash, J., Plewis, L., Draper, D., Browne, W., Yang, M., *et al.* (1998). A user's guide to ML Win London: Institute of Education.
Goldstein, H. (2003). *Multilevel statistical models (third edition)*. London: Edward Arnold.
Hoeksma, J. B. (1999). On models and theories. In H. Ader & G. J. Mellenbergh (Eds.), *Research methodology, in the social behavioral & lifesciences* (pp. 335–339). London: Sage.
Hoeksma, J. B. (2005). Multilevel models. In B. Hopkins, R. Barr, G. Michel & P. Rocha, (Eds.), *Cambridge encyclopedia of child development* (pp. 142–146). Cambridge: Cambridge University Press.
Jones, B., Nagin, D. S., & Roeder. (2001). A SAS procedure based on mixed models for estimating developmental trajectories. *Sociological Methods Research*, **229**: 374–393.
Laird, N. M., & Ware, J. H. (1982). Random effects models for longitudinal data. *Biometrics*, **38**: 963–974.
Longford. (1987). A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika*, **74**(4): 812–827.
McArdle, J. J. (1986a). Dynamic but structural equation modeling of repeated measures data. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (Vol. 2, pp. 561–614). New York: Plenum.
McArdle, J. J. (1986b). Latent variable growth within behavior genetic models. *Behavior Genetics*, **16**: 163–200.
McArdle, J. J. (1989). A structural modeling experiment with multiple growth functions. In P. Ackerman, R. Kanfer & R. Cudek (Eds.), *Learning and individual differences: Abilities, motivation, and methodology.* Hilssdale, NJ: Lawrence Erlbaum.
McArdle, J. J., & Epstein, D. (1987). Latent growth curves within developmental structural equation models. *Child Development*, **58**: 110–133.
Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, **55**: 107–122.
Muthén, B. O. (1989). Latent variable modelling in heterogeneous populations. *Psychometrika*, **54**: 557–558.
Muthén, B. O. (1991). Analysis of longitudinal data using latent variable models with varying parameters. In L. M. Collins, & J. L. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 1–17). Washington DC: American Psychological Association.
Muthén, B. O. (1992). Latent variable modeling of growth with missing data and multilevel data. *Paper presented at the Seventh International Conference on Multivariate Analysis*. Barcelona, Spain.
Muthén, B. O., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses. *Alcoholism: Clinical and experimental research*, **24**: 882–891.
Muthén, L. K., & Muthén, B. O. (1998). *Mplus (version 2) [computer software and manual]*. Los Angeles: Muthen & Muthen.
Nagin, D. (1999). Analyzing developmental trajectories: A semi-parametric group based approach. *Psychological Methods*, **4**: 139–157.

DOI: 10.1002/icd

Nagin, D., & Tremblay, R. E. (2001). Analyzing developmental trajectories of distinct but related behaviors: A group-based method. *Psychological Methods*, **6**(1): 18–34.
Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.
van Lier, P. A. C., & Crijnen, A. A. M. (2005). Trajectories of peer-nominated aggression: Risk status, predictors and outcomes. *Journal of Abnormal Child Psychology*, **33**: 99–112.
Willet, J. B., Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change. *Psychological Bulletin* **110**, 268–290.
Wohlwill, J. F. (1973). *The study of behavioral development*. New York: Academic Press.

DOI: 10.1002/icd