Beyond Multilevel Regression Modeling: Multilevel Analysis in a General Latent Variable Framework

Bengt Muthén & Tihomir Asparouhov[*]

To appear in The Handbook of Advanced Multilevel Analysis.
J. Hox & J.K Roberts (eds), Taylor and Francis

January 23, 2009

[*]This paper builds on a presentation by the first author at the AERA HLM SIG, San Francisco, April 8, 2006. The research of the first author was supported by grant R21 AA10948-01A1 from the NIAAA, by NIMH under grant No. MH40859, and by grant P30 MH066247 from the NIDA and the NIMH. We thank Kristopher Preacher for helpful comments.

Abstract

Multilevel modeling is often treated as if it concerns only regression analysis and growth modeling. Multilevel modeling, however, is relevant for nested data not only with regression and growth analysis but with all types of statistical analyses. This chapter has two aims. First, it shows that already in the traditional multilevel analysis areas of regression and growth there are several new modeling opportunities that should be considered. Second, it gives an overview with examples of multilevel modeling for path analysis, factor analysis, structural equation modeling, and growth mixture modeling. Examples include two extensions of two-level regression analysis with measurement error in the level 2 covariate and a level 1 mixture; two-level path analysis and structural equation modeling; two-level exploratory factor analysis of classroom misbehavior; two-level growth modeling using a two-part model for heavy drinking development; an unconventional approach to three-level growth modeling of math achievement; and multilevel latent class mediation of high school dropout using multilevel growth mixture modeling of math achievement development.

1 Introduction

Multilevel modeling is often treated as if it concerns only regression analysis and growth modeling (Raudenbush & Bryk, 2002; Snijders & Bosker, 1999). Furthermore, growth modeling is merely seen as a variation on the regression theme, regressing the outcome on a time-related covariate. Multilevel modeling, however, is relevant for nested data not only with regression analysis but with all types of statistical analyses, including

- Regression analysis
- Path analysis
- Factor analysis
- Structural equation modeling
- Growth modeling
- Survival analysis
- Latent class analysis
- Latent transition analysis
- Growth mixture modeling

This chapter has two aims. First, it shows that already in the traditional multilevel analysis areas of regression and growth there are several new modeling opportunities that should be considered. Second, it gives an overview with examples of multilevel modeling for path analysis, factor analysis, structural equation modeling, and growth mixture modeling. Due to lack of space, survival, latent class, and latent transition analysis are not covered. All of these topics, however, are covered within the latent variable framework of the Mplus software, which is the basis for this chapter. A technical description of this framework including not only multilevel features but also finite mixtures is given in Muthén and Asparouhov (2008). Survival mixture analysis is discussed in Asparouhov, Masyn and Muthén (2006). See also examples in the Mplus User's Guide (Muthén & Muthén, 2008). The User's Guide is available online at www.statmodel.com.

The outline of the chapter is as follows. Section 2 discusses two extensions of two-level regression analysis, Section 3 discusses two-level path analysis and structural equation modeling, Section 4 presents an example of two-level exploratory factor analysis, Section 5 discusses two-level growth modeling using a two-part model, Section 6 discusses an unconventional approach to three-level growth modeling, and Section 7 presents an example of multilevel growth mixture modeling.

2 Two-level regression

One may ask if there really is anything new that can be said about multilevel regression. The answer, surprisingly, is yes. Two extensions of conventional two-level regression analysis will be discussed here, taking into account measurement error in covariates and taking into account unobserved heterogeneity among level 1 subjects.

2.1 Measurement error in covariates

It is well known that measurement error in covariates creates biased regression slopes. In multilevel regression a particularly critical covariate is the level 2 covariate $\bar{x}_{.j}$, drawing on information from individuals within clusters to reflect cluster characteristics, as for example with students rating the school environment. Based on relatively few students such covariates may contain a considerable amount of measurement error, but this fact seems to not have gained widespread recognition in multilevel regression modeling. The following discussion draws on Asparouhov and Muthén (2006) and Ludtke et al (2008). The topic seems to be rediscovered every two decades given earlier contributions by Schmidt (1969) and Muthén (1989).

Raudenbush and Bryk (2002; p. 140, Table 5.11) considered the two-level, random intercept, group-centered regression model

$$y_{ij} = \beta_{0j} + \beta_{1j}\,(x_{ij} - \bar{x}_{.j}) + r_{ij}, \quad (1)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\,\bar{x}_{.j} + u_j, \quad (2)$$

$$\beta_{1j} = \gamma_{10}, \quad (3)$$

defining the "contextual effect" as

$$\beta_c = \gamma_{01} - \gamma_{10}. \quad (4)$$

Often, $\bar{x}_{.j}$ can be seen as an estimate of a level 2 construct which has not been directly measured. In fact, the covariates $(x_{ij} - \bar{x}_{.j})$ and $\bar{x}_{.j}$ may be seen as proxies for latent covariates (cf Asparouhov & Muthén, 2006),

$$x_{ij} - \bar{x}_{.j} \approx x_{ijw}, \quad (5)$$

$$\bar{x}_{.j} \approx x_{jb}, \quad (6)$$

where the latent covariates are obtained in line with the nested, random effects ANOVA decomposition into uncorrelated components of variation,

$$x_{ij} = x_{jb} + x_{ijw}. \tag{7}$$

Using the latent covariate approach, a two-level regression model may be written as

$$\begin{aligned} y_{ij} &= y_{jb} + y_{ijw} & (8)\\ &= \alpha + \beta_b\, x_{jb} + \epsilon_j & (9)\\ &\quad + \beta_w\, x_{ijw} + \epsilon_{ij}, & (10) \end{aligned}$$

defining the contextual effect as

$$\beta_c = \beta_b - \beta_w. \tag{11}$$

The latent covariate approach of (9) and (10) can be compared to the observed covariate approach (1) - (3). Assuming the model of the latent covariate approach of (9) and (10), Asparouhov and Muthén (2006) and Ludtke et al (2008) show that the observed covariate approach introduces a bias in the estimation of the level 2 slope γ_{01} in (3),

$$E(\hat{\gamma}_{01}) - \beta_b = \frac{(\beta_w - \beta_b)\psi_w/c}{\psi_b + \psi_w/c} = (\beta_w - \beta_b)\,\frac{1}{c}\,\frac{1 - icc}{icc + (1 - icc)/c}, \tag{12}$$

where c is the common cluster size and icc is the covariate intraclass correlation $(\psi_b/(\psi_b + \psi_w))$. In contrast, there is no bias in the level 1 slope estimate $\hat{\gamma}_{10}$. It is clear from (12) that the between slope bias increases for decreasing cluster size c and for decreasing icc. For example, with $c = 15$, $icc = 0.20$, and $\beta_w - \beta_b = 1.0$, the bias is 0.21.

Similarly, it can be shown that the contextual effect for the observed covariate approach $\hat{\gamma}_{01} - \hat{\gamma}_{10}$ is a biased estimate of $\beta_b - \beta_w$ from the latent covariate approach. For a detailed discussion, see Ludtke et al (2008), where the magnitudes of the biases are studied under different conditions.

As a simple example, consider data from the German Third International Mathematics and Science Study (TIMSS). Here there are $n = 1,980$ students in 98 schools with average cluster (school) size $= 20$. The dependent variable is a math test score in grade 8 and the covariate is student-reported disruptiveness level in the school.

The intraclass correlation for disruptiveness is 0.21. Using maximum-likelihood (ML) estimation for the latent covariate approach to two-level regression with a random intercept in line with (9) and (10) results in $\hat{\beta}_b = -1.35$ (SE= 0.36), $\hat{\beta}_w = -0.098$ (SE= 0.03), and contextual effect $\hat{\beta}_c = -1.25$ (SE = 0.36). The observed covariate approach results in the corresponding estimates $\hat{\gamma}_{01} = -1.18$ (SE= 0.29), $\hat{\gamma}_{10} = -0.097$ (SE= 0.03), and contextual effect $\hat{\beta}_c = -1.08$ (SE = 0.30).

Using the latent covariate approach in Mplus, the observed covariate *disrupt* is automatically decomposed as $disrupt_{ij} = x_{jb} + x_{ijw}$. The use of Mplus to analyze models under the latent covariate approach is described in Chapter 9 of the User's Guide (Muthén & Muthén, 2008).

2.2 Unobserved heterogeneity among level 1 subjects

This section reanalyzes the classic High School & Beyond (HSB) data used as a key illustration in Raudenbush and Bryk (2002; RB from now on). HSB is a nationally representative survey of U.S. public and Catholic high schools. The data used in RB are a subsample with 7,185 students from 160 schools, 90 public and 70 Catholic. The RB model presented on pages 80-83 is considered here for individual i in cluster (school) j:

$$y_{ij} = \beta_{0j} + \beta_{1j}\ (ses_{ij} - mean\_ses_j) + r_{ij}, \tag{13}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\ sector_j + \gamma_{02}\ mean\_ses_j + u_{0j}, \tag{14}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\ sector_j + \gamma_{12}\ mean\_ses_j + u_{1j}, \tag{15}$$

where *mean_ses* is the school-averaged student *ses* and *sector* is a 0/1 dummy variable with 0 for public and 1 for Catholic schools. The estimates are shown in Table 1. The results show for example that, holding *mean_ses* constant, Catholic schools have significantly higher mean math achievement than public schools (see the γ_{02} estimate) and that Catholic schools have significantly lower *ses* slope than public schools (see the γ_{12} estimate).

What is overlooked in the above modeling is that a potentially large source of unobserved heterogeneity resides in variation of the regression coefficients between groups of individuals sharing similar but unobserved background characteristics. It seems possible that this

Table 1: High School & Beyond two-level regression estimates from Raudenbush & Bryk (2002)

Loglikelihood	-23,248
Number of Parameters	10
BIC	46,585

Parameter	Estimate	S.E	Est./S.E.	Two-Tailed P-Value
Within level				
Residual variance				
math	36.720	0.721	50.944	0.000
Between level				
math (β_{0j}) ON				
sector (γ_{01})	1.227	0.308	3.982	0.000
mean_ses (γ_{02})	5.332	0.336	15.871	0.000
s_ses (β_{1j}) ON				
sector (γ_{11})	-1.640	0.238	-6.905	0.000
mean_ses (γ_{12})	1.033	0.333	3.100	0.002
math WITH				
s_ses	0.200	0.192	1.041	0.298
Intercepts				
math (γ_{00})	12.096	0.174	69.669	0.000
s_ses (γ_{10})	2.938	0.147	19.986	0.000
Residual variances				
math	2.316	0.414	5.591	0.000
s_ses	0.071	0.201	0.352	0.725

phenomenon is quite common due to heterogeneous sub-populations in general population surveys. Such heterogeneity is captured by level 1 latent classes. Drawing on Muthén and Asparouhov (2009), these ideas can be formalized as follows.

Consider a two-level regression mixture model where the random intercept and slope of a linear regression of a continuous variable y on a covariate x for individual i in cluster j vary across the latent classes of an individual-level latent class variable C with K categories labelled $c = 1, 2, \ldots, K$,

$$y_{ij|C_{ij}=c} = \beta_{0cj} + \beta_{1cj}\, x_{ij} + r_{ij}, \tag{16}$$

where the residual $r_{ij} \sim N(0, \theta_c)$ and a single covariate is used for simplicity. The probability of latent class membership varies as a two-level multinomial logistic regression function of a covariate z,

$$P(C_{ij} = c \mid z_{ij}) = \frac{e^{a_{cj} + b_c\, z_{ij}}}{\sum_{s=1}^{K} e^{a_{sj} + b_s\, z_{ij}}}. \tag{17}$$

The corresponding level-2 equations are

$$\beta_{0cj} = \gamma_{00c} + \gamma_{01c}\, w_{0j} + u_{0j}, \tag{18}$$

$$\beta_{1cj} = \gamma_{10c} + \gamma_{11c}\, w_{1j} + u_{1j}, \tag{19}$$

$$a_{cj} = \gamma_{20c} + \gamma_{21c}\, w_{2j} + u_{2cj}. \tag{20}$$

With K categories for the latent class variable there are $K - 1$ equations (20). Here, w_{0j}, w_{1j}, and w_{2j} are level-2 covariates and the residuals u_{0j}, u_{1j}, and u_{2cj} are (2+K-1)-variate normally distributed with means zero and covariance matrix Θ_2 and are independent of r_{ij}. In many cases $z = x$ in (17). Also, the level 2 covariates in (18) - (20) may be the same as is the case in the High School & Beyond example considered below, where there is a common $w_j = w_{0j} = w_{1j} = w_{2j}$. To reduce the dimensionality, a continuous factor f will represent the random intercept variation of (20) in line with Muthén and Asparouhov (2009).

Figure 1 shows a diagram of a two-level regression mixture model applied to the High School & Beyond data. A four-class model is chosen and obtains a loglikelihood value of $22,812$ with 30 parameters, and BIC $= 45,891$. This BIC value is considerably better than the conventional two-level regression BIC value of $46,585$ reported in Table 1 and the mixture model is therefore preferable. The mixture model and its ML estimates can be interpreted as follows. Because this type of model is new to readers, Figure 1 will be used to understand the estimates rather than reporting a table of the parameter estimates for (16) - (20).

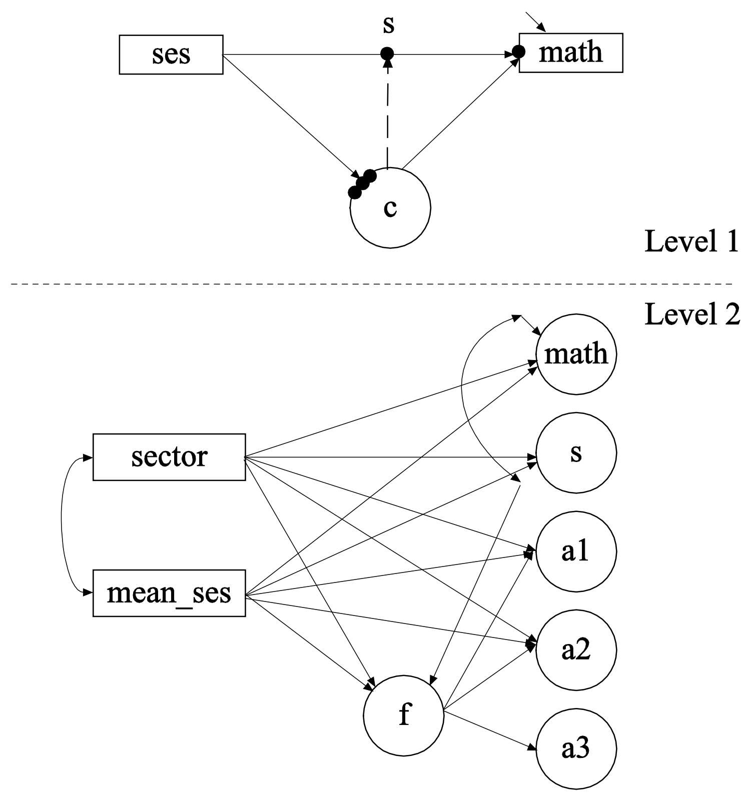
The latent class variable c in the level 1 part of Figure 1 has four classes. As indicated by the arrows from c, the four classes are characterized by having different intercepts for *math* and different slopes for *math* regressed on *ses*. In particular, the *math* mean changes significantly across the classes. An increasing value of the *ses* covariate gives an increasing odds of being in the highest math class which contains 31% of the students. For three classes with lowest *math* intercept, *ses* does not have a further, direct influence on *math*: the mean of the random slope s is only significant in the class with the highest *math* intercept, where it is positive.

The random intercepts of c, marked with filled circles on the circle for c on level 1, are continuous latent variables on level 2, denoted $a1 - a3$ (four classes gives three intercepts because the last one is standardized to zero). The (co-)variation of the random intercepts is for simplicity represented via a factor f. These random effects carry information about the influence of the school context on the probability of a student's latent class membership. For example, the influence of the level 2 covariate *sector* (public=0, Catholic = 1) is such that Catholic schools are less likely to contribute to students being in the lower *math* intercept classes relative to the highest *math* intercept class. Similarly, a high value of the level 2 covariate *mean_ses* causes students to be less likely to be in the lower *math* intercept classes relative to the highest *math* intercept class.

The influence of the level 2 covariates on the random slope s is such that Catholic schools have lower values and higher *mean_ses* schools have higher values. The influence of the level 2 covariates on the random intercept *math* is insignificant for *sector* while positive significant for *mean_ses*. The insignificant effect of *sector* does not mean, however, that *sector* is unimportant to math performance given that *sector* had a significant influence on the random effects of the latent class variable c.

It is interesting to compare the mixture results to those of the conventional two-level regression in Table 1. The key results for the conventional analysis is that (1) Catholic schools show less influence of *ses* on *math*, and (2) Catholic schools have higher mean math achievement. Neither of these results are contradicted by the mix-

Figure 1: Model diagram for two-level regression mixture analysis.



ture analysis. But using a model that has considerably better BIC, the mixture model explains these results by a mediating latent class variable on level 1. In other words, students' latent class membership is what influences math performance and latent class membership is predicted by both student-level ses and school characteristics. The Catholic school effect on math performance is not direct as an effect on the level 2 math intercept (this path is insignificant), but indirect via the student's latent class membership. For more details on two-level regression mixture modeling and a math achievement example focusing on gender differences, see Muthén and Asparouhov (2009).

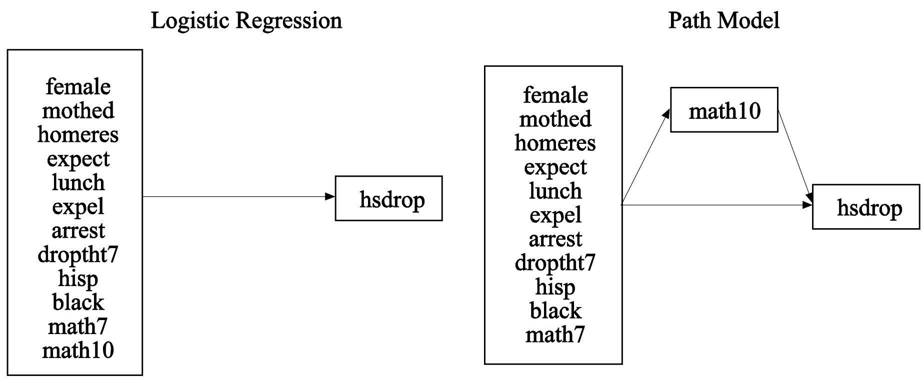
3 Two-level path analysis and structural equation modeling

Regression analysis is often only a small part of a researcher's modeling agenda. Frequently a system of regression equations is specified as in path analysis and structural equation modeling (SEM). There have been recent developments for path analysis and SEM in multilevel data and a brief overview of new kinds of models will be presented in this section. No data analysis is done, but focus is instead on modeling ideas.

Consider the left part of Figure 2 where the binary dependent variable *hsdrop*, representing dropping out by Grade 12, is related to a set of covariates using logistic regression. A complication in this analysis is that many of those who drop out by Grade 12 have missing data on $math10$, the mathematics score in Grade 10, where the missingness is not completely at random. Missingness among covariates can be handled by adding a distributional assumption for the covariates, either by multiple imputation or by not treating them as exogenous. Either way, this complicates the analysis without learning more about the relationships among the variables in the model. The right part of Figure 2 shows an alternative approach using a path model that acknowledges the temporal position of $math10$ as an intervening variable that is predicted by the remaining covariates measured earlier. In this path model, "missing at random" (MAR; Little & Rubin, 2002) is reasonable in that the covariate may well predict the missingness in *math*. The resulting path model has a combination of a linear regression for a continuous dependent variable and a logistic regression for a binary dependent variable.

Figure 3 shows a two-level counterpart to the path model. The top part of the figure shows the within-level part of the model for the student relationships. Here, the filled circles at the end of the arrows indicate random intercepts. On the between level these random intercepts are continuous latent variables varying across schools. The two random intercepts are not treated symmetrically, but it is hypothesized that increasing $math10$ intercept decreases the *hsdrop* intercept in that schools with good mean math performance in Grade 10 tend to have an environment less conducive to dropping out. Two school-level covariates are used as predictors of the random intercepts, *lunch* which is a dummy variable used as a poverty proxy and *mstrat*, measuring math teacher workload as the ratio of students to full-time

Figure 2: Model diagram for logistic regression path analysis



math teachers.

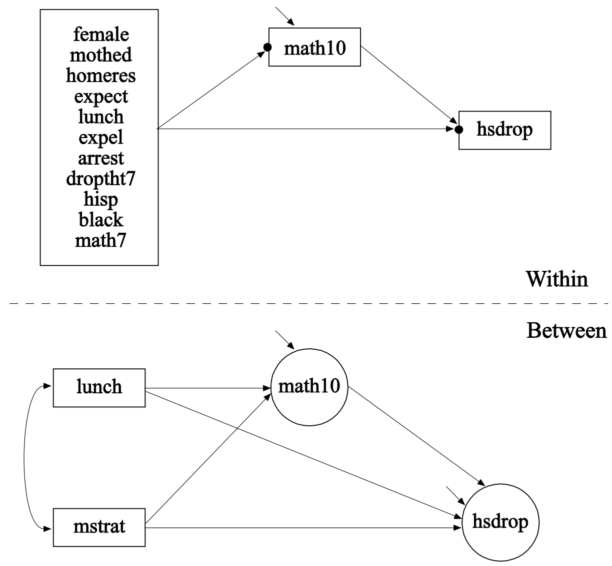
Another path analysis example is shown in Figure 4. Here, u is again a categorical dependent variable and both u and the continuous variable y have random intercepts. Figure 4 further illustrates the flexibility of current two-level path analysis by adding an observed between-level dependent variable z which intervenes between the between-level covariate w and the random intercept of u. Between-level variables that play a role as dependent variables are not used in conventional multilevel modeling.

Figure 5 shows a path analysis example with random slopes a_j, b_j, and c'_j. This illustrates a two-level mediational model. As described in e.g. Bauer, Preacher and Gil (2006), the indirect effect is here $\alpha \times \beta + Cov(a_j, b_j)$, where α and β are the means of the corresponding random slopes a_j and b_j.

Figure 6 specifies a MIMIC model with two factors $fw1$ and $fw2$ for students on the within level. The filled circles at the binary indicators $u1 - u6$ indicate random intercepts that are continuous latent variables on the between level. The between level has a single factor fb describing the variation and covariation among the random intercepts. The between level has the unique feature of also adding between-level indicators $y1 - y4$ for a between-level factor f, another example of between-level dependent variables. Two-level factor analysis will be discussed in more detail in Section 4.

Figure 7 shows a structural equation model with an exogeneous

Figure 3: Model diagram for two-level logistic regression path analysis



and an endogenous factor that has both within-level and between-level variation. The special feature here is that the structural slope s is random. The slope s is regressed on a between-level covariate x.

4 Two-level exploratory factor analysis

A recent multilevel development concerns a practical alternative to ML estimation in situations that would lead to heavy ML computations (cf Asparouhov & Muthén, 2007). Heavy ML computations occur when numerical integration is needed, as for instance with categorical outcomes. Many models, including factor analysis models, involve many random effects, each one of which adds a dimension of integration. The new estimator uses limited information from first- and second-order moments to formulate a weighted least squares approach that reduces multidimensional integration into a series of one- and two-dimensional integrations for the uni- and bivariate moments. This weighted least squares approach is particularly useful in exploratory factor analysis (EFA) where there are typically many random effects due to having

Figure 4: Model diagram for path analysis with between-level dependent variable

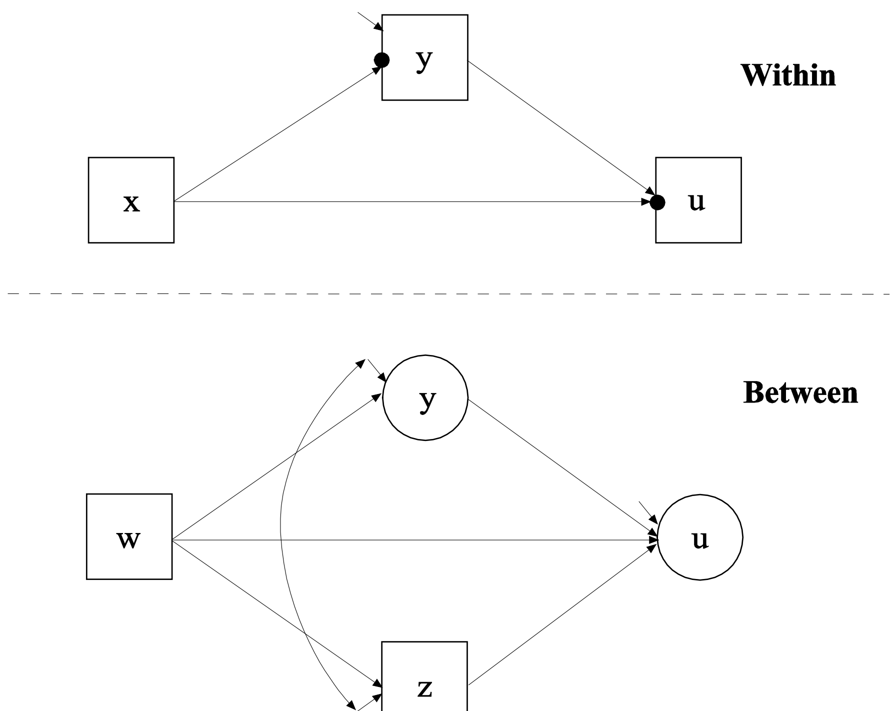


Figure 5: Model diagram for path analysis with mediation and random slopes

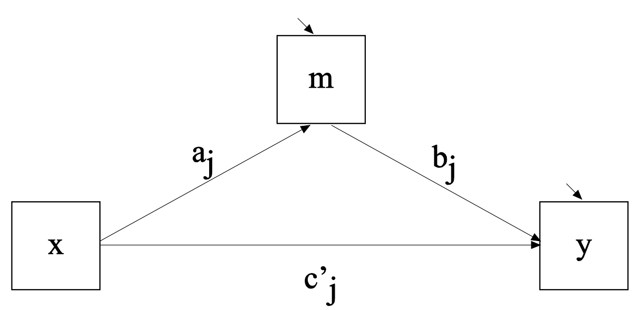
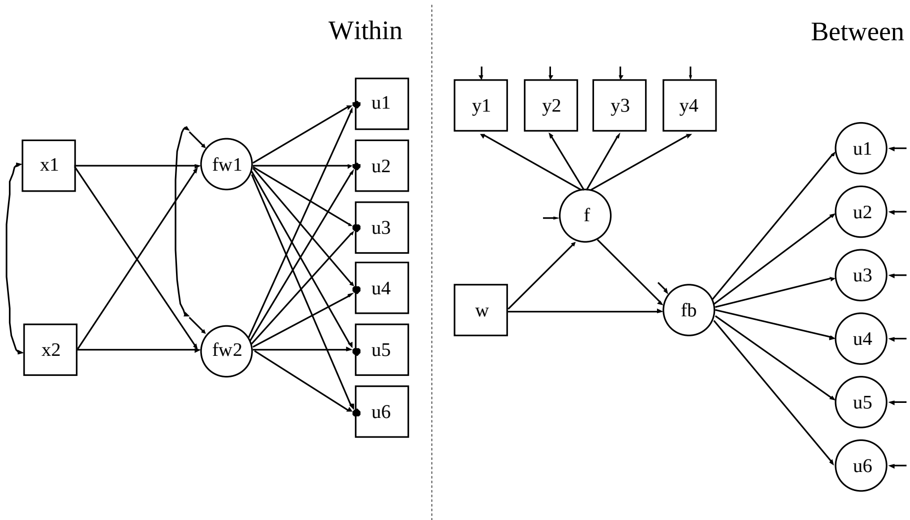


Figure 6: Model diagram for two-level SEM



many variables and many factors.

Consider the following EFA example. Table 2 shows the item distribution for a set of 13 items measuring aggressive-disruptive behavior in the classroom among 363 boys in 27 classrooms in Baltimore public schools. It is clear that the variables have very skewed distributions with a strong floor effects so that $40\% - 80\%$ are at the lowest value. If treated as continuous outcomes, even non-normality robust standard errors and χ^2 tests of model fit would not give correct results in that a linear model is not suitable for data with such strong floor effects. The variables will instead be treated as ordered polytomous (ordinal). The 13-item instrument is hypothesized to capture three aspects of aggressive-disruptive behavior: property, verbal, and person. Figure 8 shows a model diagram with notation analogous to two-level regression. On the within (student) level the three hypothesized factors are denoted $fw1 - fw3$. The filled circles at the observed items indicate random measurement intercepts. On the between level these random intercepts are continuous latent variables varying over classrooms, where the variation and covariation is represented by the classroom-level factors $fb1 - fb3$. The meaning of the student-level factors $fw1 - fw3$ is in line with regular factor analysis. In con-

Figure 7: Model diagram for two-level SEM with a random structural slope

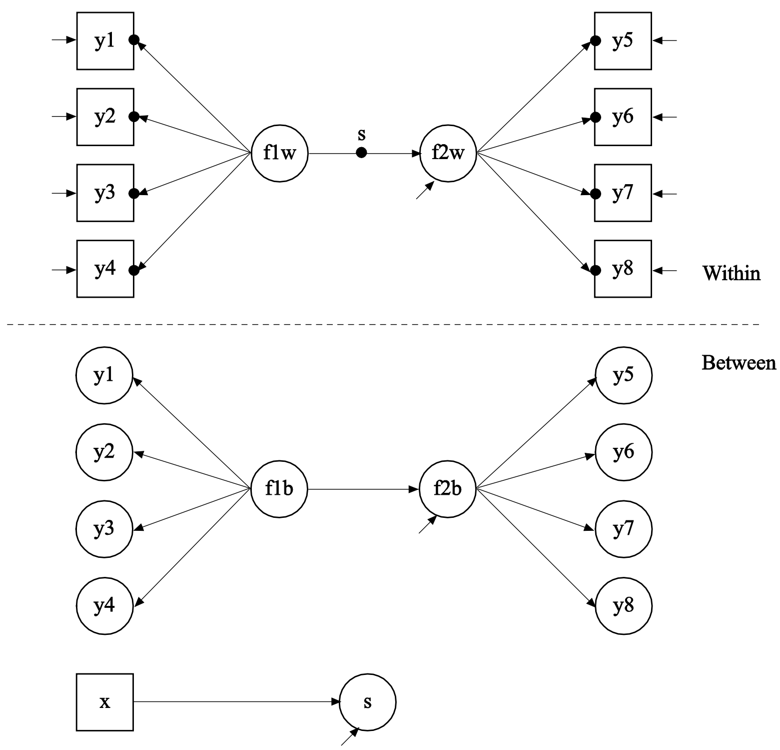


Figure 8: Two-level factor analysis model

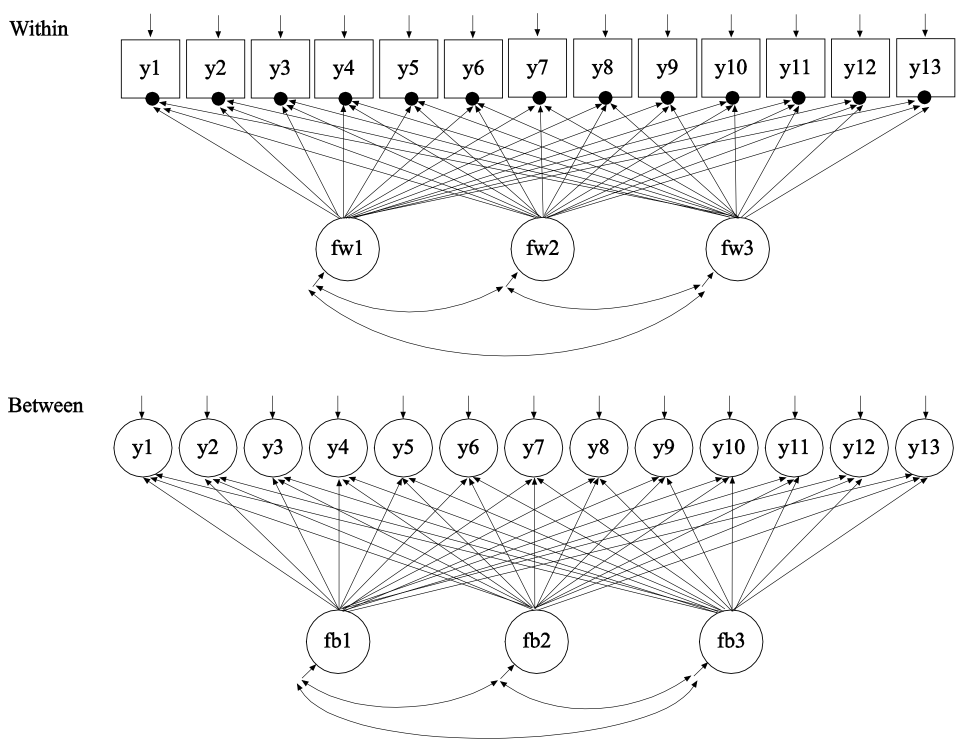


Table 2: Distributions for aggressive-disruptive items

Aggression Items	Almost Never (scored as 1)	Rarely (scored as 2)	Sometimes (scored as 3)	Often (scored as 4)	Very Often (scored as 5)	Almost Always (scored as 6)
stubborn	42.5	21.3	18.5	7.2	6.4	4.1
breaks rules	37.6	16.0	22.7	7.5	8.3	8.0
harms others and property	69.3	12.4	9.40	3.9	2.5	2.5
breaks things	79.8	6.60	5.20	3.9	3.6	0.8
yells at others	61.9	14.1	11.9	5.8	4.1	2.2
takes others' property	72.9	9.70	10.8	2.5	2.2	1.9
fights	60.5	13.8	13.5	5.5	3.0	3.6
harms property	74.9	9.90	9.10	2.8	2.8	0.6
lies	72.4	12.4	8.00	2.8	3.3	1.1
talks back to adults	79.6	9.70	7.80	1.4	0.8	1.4
teases classmates	55.0	14.4	17.7	7.2	4.4	1.4
fights with classmates	67.4	12.4	10.2	5.0	3.3	1.7
loses temper	61.6	15.5	13.8	4.7	3.0	1.4

trast, the classroom-level factors $fb1 - fb3$ represent classroom-level phenomena for which a researcher typically has less understanding. These factors require new kinds of considerations as follows. If the same set of three within-level factors (property, verbal, and person) are to explain the (co-)variation on the between level, classroom teachers must vary in their skills to manage their classrooms with respect to all three of these aspects. That is, some teachers are good at controlling property-oriented aggressive-disruptive behavior and some are not, some teachers are good at controlling verbally-oriented aggressive-disruptive behavior and some are not, etc. This is not very likely and it is more likely that teachers simply vary in their ability to manage their classrooms in all three respects fairly equally. This would lead to a single factor fb on the between level instead of three factors.

As shown in Figure 8, ML estimation would require 19 dimensions of numerical integration, which is currently an impossible task. A reduction is possible if the between-level, variable-specific residuals are zero, which is often a good approximation. This makes for a reduction to 6 dimensions of integration which is still a very difficult task. The Asparouhov and Muthén (2007) weighted least squares approach is suitable for such a situation and will be used here. The approach assumes that the factors are normally distributed and uses an ordered

Table 3: Two-level EFA model test result for aggressive-disruptive items

Within-level factors	Between-level factors	Df	Chi-square	CFI	RMSEA
unrestricted	1	65	66 (p = 0.43)	1.000	0.007
1	1	130	670	0.991	0.107
2	1	118	430	0.995	0.084
3	1	107	258	0.997	0.062
4*	1	97	193	0.998	0.052

*4th factor has no significant loadings

probit link function for the item probabilities as functions of the factors. This amounts to assuming multivariate normality for continuous latent response variables underlying the items in line with using polychoric correlations in single-level analysis. Rotation of loadings on both levels is provided along with standard errors for rotated loadings and resulting factor correlations.

Table 3 shows a series of analyses varying the number of factors on the within and between levels. To better understand how many factors are needed on a certain level, an unrestricted correlation model can be used on the other level. Using an unrestricted within-level model it is clear that a single between-level factor is sufficient. Adding within-level factors shows an improvement in fit going up to 4 factors. The 4-factor solution, however, has no significant loadings for the additional, fourth factor. Also, the 3-factor solution captures the three hypothesized factors. The factor solution is shown in Table 4 using Geomin rotation (Asparouhov & Muthén, 2008) for the within level. Factor loadings with asterisks represent loadings significant on the 5% level, while bolded loadings are the more substantial ones. The loadings for the single between-level factor are fairly homogeneous supporting the idea that there is a single classroom management dimension.

Table 4: Two-level EFA of aggressive-disruptive items using WLSM and Geomin rotation

Aggression Items	Within-Level Loadings			Between-Level Loadings
	Property	Verbal	Person	General
stubborn	0.00	**0.78***	0.01	**0.65***
breaks rules	0.31*	0.25*	0.32*	**0.61***
harms others and property	**0.64***	0.12	0.25*	**0.68***
breaks things	**0.98***	0.08	-0.12*	**0.98***
yells at others	0.11	**0.67***	0.10	**0.93***
takes others' property	**0.73***	**-0.15***	0.31*	**0.80***
fights	0.10	0.03	**0.86***	**0.79***
harms property	**0.81***	0.12	0.05	**0.86***
lies	**0.60***	0.25*	0.10	**0.86***
talks back to adults	0.09	**0.78***	0.05	**0.81***
teases classmates	0.12	0.16*	**0.59***	**0.83***
fights with classmates	-0.02	0.13	**0.88***	**0.84***
loses temper	-0.02	**0.85***	0.05	**0.87***

5 Growth modeling (two-level analysis)

Growth modeling concerns repeated measurement data nested within individuals and possibly also within higher-order units (clusters such as schools). This will be referred to as two- and three-level growth analysis, respectively. Often, two-level growth analysis can be performed in a multivariate, wide data format fashion, letting the level 1 repeated measurement on y over T time points be represented by a multivariate outcome vector $y = (y_1, y_2, \ldots, y_T)'$, reducing the two levels to one. This reduction by one level is typically used in the latent variable framework of Mplus. More common, however, is to view growth modeling as a two-level model with features analogous to those of two-level regression (see, e.g., Raudenbush & Bryk, 2002). In this case, data are arranged in a univariate, long format.

Following is a simple example with linear growth, for simplicity using the notation of Raudenbush and Bryk (2002). For time point t and individual i, consider

y_{ti} : individual-level, outcome variable
a_{ti} : individual-level, time-related variable (age, grade)
x_i : individual-level, time-invariant covariate

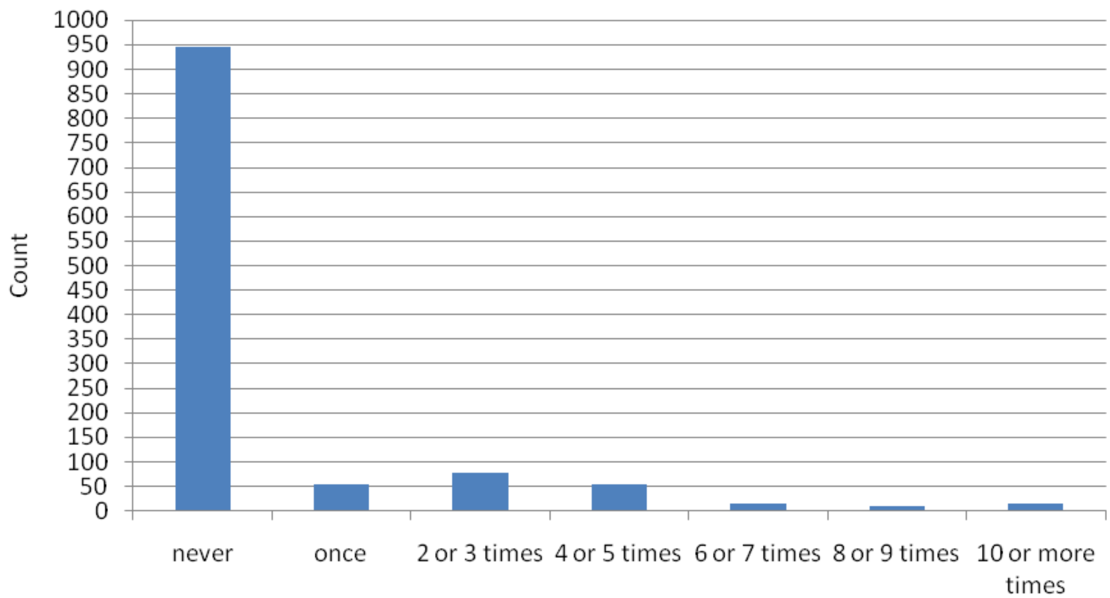
and the 2-level growth model

$$\textit{Level } 1: \; y_{ti} = \pi_{0i} + \pi_{1i}\, a_{ti} + e_{ti}, \tag{21}$$

$$\textit{Level } 2: \; \begin{cases} \pi_{0i} = \gamma_{00} + \gamma_{01}\, x_i + r_{0i}, \\ \pi_{1i} = \gamma_{10} + \gamma_{11}\, x_i + r_{1i}, \end{cases} \tag{22}$$

where π_0 is a random intercept and π_1 is a random slope. One may ask if there really is anything new that can be said about (two-level) growth analysis. The answer, surprisingly, is again yes. Following is a discussion of a relatively recent and still underutilized extension to situations with very skewed outcomes similar to those studied in the above EFA. Here, the example concerns frequency of heavy drinking in the last 30 days from the National Longitudinal Survey of Youth (NLSY), a U.S. national survey. The distribution of the outcome at age 24 is shown in Figure 9, where a majority of individuals did not engage in heavy drinking in the last 30 days. Olsen and Schafer (2001) proposed a two-part or semicontinuous growth model for data of this type, treating the outcome as continuous but adding a special modeling feature to take into account the strong floor effect.

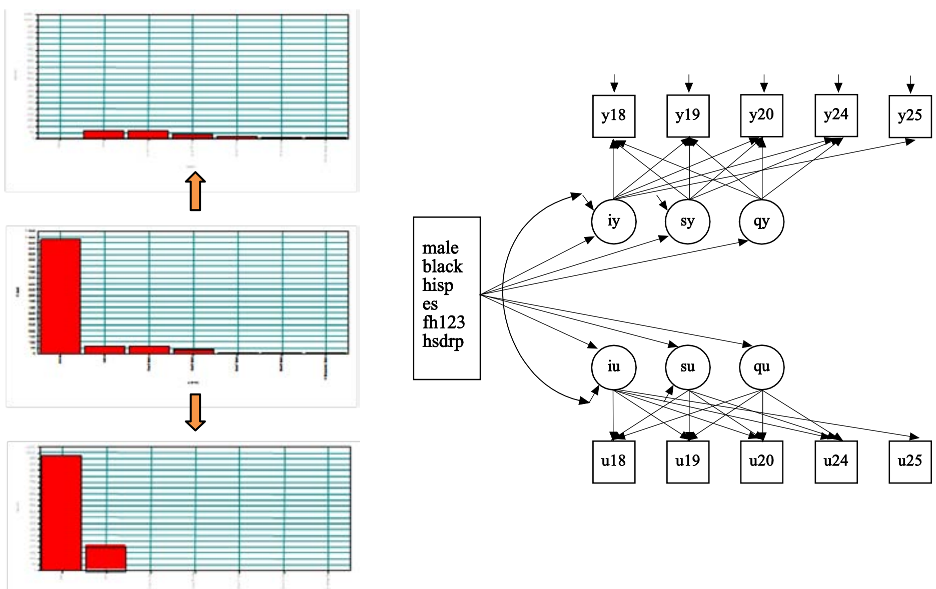
Figure 9: Histogram for heavy drinking at age 24



The two-part growth modeling idea is shown in Figure 10, where the outcome is split into two parts, a binary part and a continuous part. Here, iy and iu represent random intercepts π_0, whereas sy and su represent random linear slopes π_1. In addition, the model has random quadratic slopes qy and qu. The binary part is a growth model describing for each time point the probability of an individual experiencing the event, whereas for those who experienced it the continuous part describes the amount, in this case the number of heavy drinking occasions in the last 30 days. For an individual who does not experience the event, the continuous part is recorded as missing. A joint growth model for the binary and the continuous process scored in this way represents the likelihood given by Olsen and Schafer (2001).

Non-normally distributed outcomes can often be handled by ML using a non-normality robust standard error approach, but this is not sufficient for outcomes such as shown in Figure 9 given that a linear model is unlikely to hold. To show the difference in results as compared two-part growth modeling, Table 5 shows the Mplus output for

Figure 10: Two-part growth model for heavy drinking



the estimated growth model for frequency of heavy drinking ages 18 - 25. The results focus on the regression of the random intercept i on the time-invariant covariates in the model. The time scores are centered at age 25 so that the random intercept refers to the systematic part of the growth curve at age 25. It is seen that the regular growth modeling finds all but the last two covariates significant. In contrast, the two-level modeling finds several of the covariates insignificant in one part or the other (the two parts are labeled *iy ON* for the continuous part and *iu ON* for the binary part. Consider as an example, the covariate *black*. As is typically found being black has a significant negative influence in the regular growth modeling, lowering the frequency of heavy drinking. In the two-part modeling this covariate is insignificant for the continuous part and significant only for the binary part. This implies that, holding other covariates constant, being black significantly lowers the risk of engaging in heavy drinking, but among blacks who are engaging in heavy drinking there is no difference in amount compared to other ethnic groups. These two paths of influence are confounded in the regular growth modeling.

As shown in Figure 11, a distal outcome can also be added to the growth model. In this example, the distal outcome is a DSM-based classification into alcohol dependence or not by age 30. The distal outcome is predicted by the age 25 random intercept using a logistic regression model part. Table 6 shows that the distal outcome is significantly influenced only by the age 25-defined random intercept *iu* for the binary part, not by the random intercept for the continuous part. In other words, if the probability of engaging in heavy drinking at age 25 is high the probability of alcohol dependence by age 30 is high. But the alcohol dependence probability is not significantly influenced by the frequency of heavy drinking at age 25. The results also show that controlling for age 25 heavy drinking behavior, none of the covariates has a significant influence on the distal outcome.

6 Growth modeling (three-level analysis)

This section considers growth modeling of individual- and cluster-level data. A typical example is repeated measures over grades for students nested within schools. One may again ask if there really is anything new that can be said about growth modeling in cluster data. The

Table 5: Two-part growth modeling of frequency of heavy drinking ages 18-25

Parameter	Estimate	S.E.	Est./S.E.	Std	StdYX
Regular growth modeling, treating outcome as continuous. Non-normality robust ML (MLR)					
i ON					
male	0.769	0.076	10.066	0.653	0.326
black	-0.336	0.083	-4.034	-0.286	-0.127
hisp	-0.227	0.103	-2.208	-0.193	-0.071
es	0.291	0.128	2.283	0.247	0.088
fh123	0.286	0.137	2.089	0.243	0.075
hsdrp	-0.024	0.104	-0.232	-0.020	-0.008
coll	-0.131	0.086	-1.527	-0.111	-0.052
Two-part growth modeling					
iy ON					
male	0.262	0.052	5.065	0.610	0.305
black	-0.096	0.059	-1.619	-0.223	-0.099
hisp	-0.130	0.066	-1.963	-0.301	-0.111
es	0.082	0.062	1.333	0.191	0.068
fh123	0.213	0.076	2.815	0.495	0.152
hsdrp	0.084	0.065	1.289	0.195	0.078
coll	-0.015	0.053	-0.280	-0.035	-0.016
iu ON					
male	2.041	0.176	11.594	0.949	0.474
black	-1.072	0.203	-5.286	-0.499	-0.222
hisp	-0.0545	0.234	-2.331	-0.254	-0.093
es	0.364	0.234	1.560	0.169	0.060
fh123	0.562	0.275	2.045	0.262	0.080
hsdrp	-0.238	0.216	-1.103	-0.111	-0.044
coll	-0.259	0.196	-1.317	-0.120	-0.056

Figure 11: Two-part growth model for heavy drinking and a distal outcome

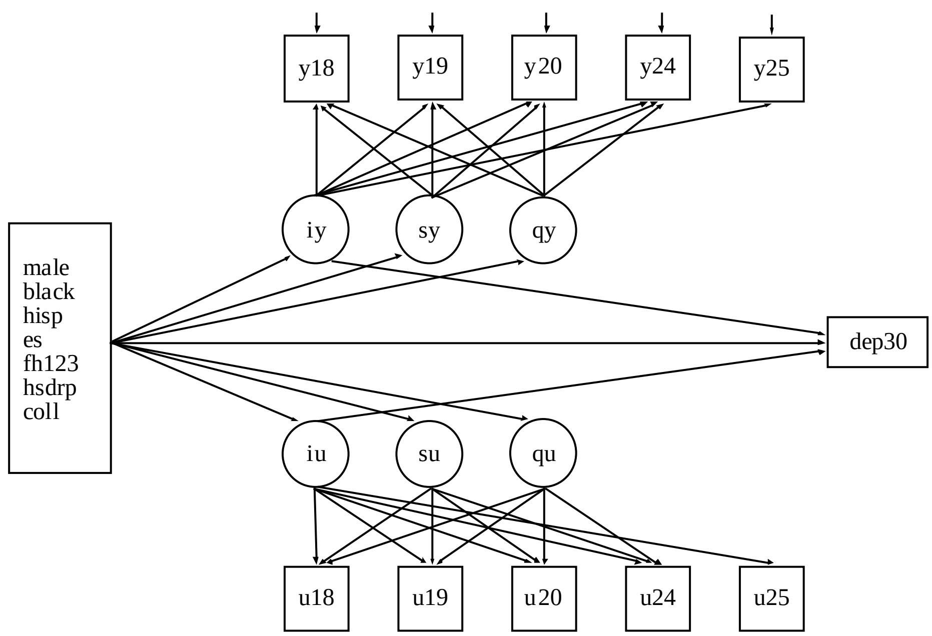


Table 6: Two-part growth modeling of frequency of heavy drinking ages 18-25 with a distal outcome

Parameter	Estimate	S.E.	Est./S.E.	Std	StdYX
dep30 ON					
iu	0.440	0.141	3.120	0.949	0.427
iy	0.874	0.736	1.187	0.373	0.168
dep30 ON					
male	-0.098	0.291	-0.337	-0.098	-0.022
black	0.415	0.294	1.414	0.415	0.083
hisp	0.025	0.326	0.075	0.025	0.004
es	0.237	0.286	0.830	0.237	0.038
fh123	0.498	0.325	1.532	0.498	0.069
hsdrp	0.565	0.312	1.812	0.545	0.101
coll	-0.384	0.276	-1.390	-0.384	-0.081

answer, surprisingly, is once again yes. An important extension to the conventional 3-level analysis becomes clear when viewed from a general latent variable modeling perspective.

For simplicity, the notation will be chosen to coincide with that of Raudenbush and Bryk (2002). Consider the observed variables for time point t, individual i, and cluster j,

y_{tij}	:	individual-level, outcome variable
a_{1tij}	:	individual-level, time-related variable (age, grade)
a_{2tij}	:	individual-level, time-varying covariate
x_{ij}	:	individual-level, time-invariant covariate
w_j	:	cluster-level covariate

and the 3-level growth model

$$Level\ 1: \ y_{tij} = \pi_{0ij} + \pi_{1ij}\ a_{1tij} + \pi_{2tij}\ a_{2tij} + e_{tij}, \tag{23}$$

$$Level\ 2: \begin{cases} \pi_{0ij} = \beta_{00j} + \beta_{01j}\ x_{ij} + r_{0ij}, \\ \pi_{1ij} = \beta_{10j} + \beta_{11j}\ x_{ij} + r_{1ij}, \\ \pi_{2tij} = \beta_{20tj} + \beta_{21tj}\ x_{ij} + r_{2tij}, \end{cases} \tag{24}$$

$$Level\ 3: \begin{cases} \beta_{00j} = \gamma_{000} + \gamma_{001}\ w_j + u_{00j}, \\ \beta_{10j} = \gamma_{100} + \gamma_{101}\ w_j + u_{10j}, \\ \beta_{20tj} = \gamma_{200t} + \gamma_{201t}\ w_j + u_{20tj}, \\ \beta_{01j} = \gamma_{010} + \gamma_{011}\ w_j + u_{01j}, \\ \beta_{11j} = \gamma_{110} + \gamma_{111}\ w_j + u_{11j}, \\ \beta_{21tj} = \gamma_{21t0} + \gamma_{21t1}\ w_j + u_{21tj}. \end{cases} \tag{25}$$

Here, the πs are random intercepts and slopes varying across individuals and clusters, and the βs are random intercepts and slopes varying across clusters. The residuals e, r and u are assumed normally distributed with zero means, uncorrelated with respective right-hand side covariates, and uncorrelated across levels.

In Mplus, growth modeling in cluster data is represented in a similar, but slightly different way that offers further modeling flexibility. As mentioned in Section 5 the first difference arises from the level 1 repeated measurement on y over time being represented by a multivariate outcome vector $y = (y_1, y_2, \ldots, y_T)'$ so that the number of levels is reduced from three to two. The second difference is that each variable, with the exception of variables multiplied by random slopes, is decomposed into uncorrelated within- and between-cluster components. Using subscripts w and b to represent within- and between-cluster variation, one may write the variables in (23) as

$$y_{tij} = y_{btj} + y_{wtij}, \tag{26}$$

$$\pi_{0ij} = \pi_{0bj} + \pi_{0wij}, \tag{27}$$

$$\pi_{1ij} = \pi_{1bj} + \pi_{1wij}, \tag{28}$$

$$\pi_{2tij} = \pi_{2tbj} + \pi_{2twij}, \tag{29}$$

$$e_{tij} = e_{btj} + e_{wtij}, \tag{30}$$

so that the level 1 equation (23) can be expressed as

$$y_{tij} = \pi_{0bj} + \pi_{0wij} + (\pi_{1bj} + \pi_{1wij})\ a_{1tij} + (\pi_{2btj} + \pi_{2wtij})\ a_{2tij} + e_{btj} + e_{wtij}. \tag{31}$$

The 3-level model of (23) - (25) can then be rewritten as a 2-level model with levels corresponding to within- and between-cluster variation,

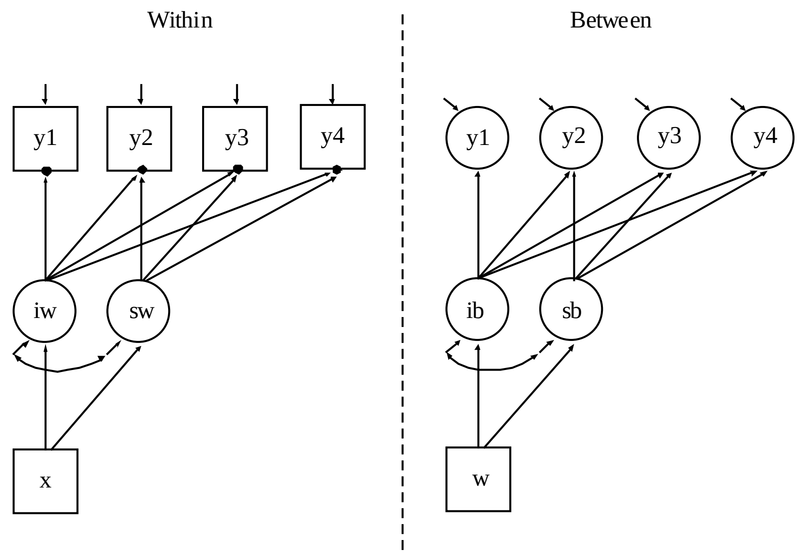
$$Within: \begin{cases} y_{wtij} = \pi_{0wij} + \pi_{1wij}\ a_{1tij} + \pi_{2wtij}\ a_{2tij} + e_{wtij}, \\ \pi_{0wij} = \beta_{01j}\ x_{ij} + r_{0ij}, \\ \pi_{1wij} = \beta_{11j}\ x_{ij} + r_{1ij}, \\ \pi_{2wtij} = \beta_{21tj}\ x_{ij} + r_{2tij}, \end{cases} \tag{32}$$

$$Between: \begin{cases} y_{btj} = \pi_{0bj} + \pi_{1bj}\ a_{1tij} + \pi_{2btj}\ a_{2tij} + e_{btj}, \\ \pi_{0bj} = \beta_{00j} = \gamma_{000} + \gamma_{001}\ w_j + u_{00j}, \\ \pi_{1bj} = \beta_{10j} = \gamma_{100} + \gamma_{101}\ w_j + u_{10j}, \\ \pi_{2btj} = \beta_{20tj} = \gamma_{200t} + \gamma_{201t}\ w_j + u_{20tj}, \\ \beta_{01j} = \gamma_{010} + \gamma_{011}\ w_j + u_{01j}, \\ \beta_{11j} = \gamma_{110} + \gamma_{111}\ w_j + u_{11j}, \\ \beta_{21tj} = \gamma_{21t0} + \gamma_{21t1}\ w_j + u_{21tj}. \end{cases} \tag{33}$$

From the latent variable perspective taken in Mplus, the first line of the within level (32) and the first line of the between level (33) is the measurement part of the model with growth factors π_0, π_1 measured by multiple indicators y_t. The next lines of each level contain the structural part of the model. As is highlighted in (31), the rearrangement of the 3-level model as (32), (33) shows that the 3-level model typically assumes that the measurement part of the model is invariant across within and between in that the same time scores a_{1tij} are used on both levels.

As seen in (32), (33) the decomposition into within and between components also occurs for the residual $e_{tij} = e_{wtij} + e_{btj}$. The e_{btj} term is typically fixed at zero in conventional multilevel modeling, but this is an important restriction. This restriction is not clear from the way the model is written in (23). Time-specific, between-level variance parameters for the residuals e_{btj} are often needed to represent across-cluster variation in time-specific residuals.

Figure 12: A two-level growth model (3-level analysis)



Consider a simple example with no time-varying covariates and where the time scores do not vary across individuals or clusters, $a_{1tij} = a_{1t}$. To simplify notation in the actual Mplus analyses, and dropping the ij and j subscripts, let $iw = \pi_{0w}$, $sw = \pi_{1w}$, $ib = \pi_{0b}$, and $sb = \pi_{1b}$ be the within-level and between-level growth factors, respectively. Figure 12 shows the model diagram for four time points using the within-level covariate x and the between-level covariate w. The model diagram may be seen as analogous to the two-level factor analysis model, adding covariates. The between-level part of the model is drawn with residual arrows pointing to the time-specific latent variables $y1 - y4$. These are the residuals e_{btj} which conventional growth analysis assumes are zero.

The model of Figure 12 is analyzed with and without the zero residual restriction using mathematics scores in Grades 7 - 10 from the Longitudinal Survey of American Youth (LSAY). Two between-level covariates are added, *lunch* (a poverty index) and *mstrat* (math teacher workload). The between-level Mplus ML results from the two analyses are shown in Table 7. The χ^2 model test of fit results show a big improvement when adding the residual variances to the model. The sb regression on *mstrat* also shows large differences between the two approaches with a smaller and insignificant effect in the conventional approach. Given that the *sb* residual variance estimate is larger for the conventional approach, it appears that the conventional model tries to absorb the residual variances into the slope growth factor variance. The residual variance for Grade 10 has a negative insignificant value which could be fixed at zero but does not change other results much.

6.1 Further 3-level growth modeling extensions

Figure 13 shows a student-level regression of the random slope *sw* regressed on the random intercept *iw*. With *iw* defined at the first time point, the study investigates to which extent the initial status influences the growth rate. The regression of the growth rate on the initial status has a random slope *s* that varies across clusters. For example, a researcher may be interested in how schools vary in their ability to reduce the influence of initial status on growth rate. Seltzer, Choi, and Thum (2002) studied this topic using Bayesian MCMC estimation, but ML can be used in Mplus. Figure 13 shows how the school variation in *s* can be explained by a school-level covariate *w*. The rest of the school-level model is specified as in the previous section.

Figure 14 shows an example of a multiple-indicator, multilevel growth model. In this case the growth model simply uses a random intercept. The data have four levels in that the observations are indicators nested within time points, time points nested within individuals, and individuals nested within twin pairs. The model diagram, however, shows how this case can be expressed as a single-level model. This is accomplished using a triply multivariate representation where the indicators (two in this case), time points (five in this case), and twins (two) create a 20-variate observation vector. With categorical outcomes, ML estimation needs numerical integration which is prohibitive given that there are 10 dimensions of integration, but weighted least squares estimation is straightforward.

Figure 15 shows an alternative, two-level approach. The data vector is arranged as doubly multivariate with indicators and twins creating 4 outcomes. The two levels are time and person. This approach assumes time-invariant measurement parameters and constant time-specific factor variances. These assumptions can be tested using the single-level approach in Figure 14 with weighted least squares esti-

Table 7: Two-level growth modeling (three-level modeling) of LSAY math achievement, Grades 7-10

Parameter	Estimate	S.E.	Est./S.E.	Std	StdYX
Conventional growth modeling: Chi-square (32) = 179.58. Between-level estimates and SEs:					
sb ON					
lunch	-1.271	0.402	-3.160	-1.919	-0.397
mstrat	1.724	1.022	1.688	2.605	0.185
Residual variances					
math7	0.000	0.000	0.000	0.000	0.000
math8	0.000	0.000	0.000	0.000	0.000
math9	0.000	0.000	0.000	0.000	0.000
math10	0.000	0.000	0.000	0.000	0.000
ib	5.866	1.401	4.186	0.736	0.736
sb	0.354	0.138	2.564	0.809	0.809
Allowing time-specific level 3 residual variances: Chi-square (28) = 83.69. Between-level estimates and SEs:					
sb ON					
lunch	-1.312	0.367	-3.576	-2.495	-0.516
mstrat	2.281	0.771	2.957	4.338	0.308
Residual variances					
math7	1.396	0.749	1.863	1.396	0.159
math8	1.414	0.480	2.946	1.414	0.154
math9	0.382	0.381	1.002	0.382	0.042
math10	-0.121	0.518	-0.234	-0.121	-0.012
ib	5.211	1.410	3.694	0.704	0.704
sb	0.177	0.155	1.143	0.640	0.640

Figure 13: Multilevel modeling of a random slope regressing growth rate on initial status

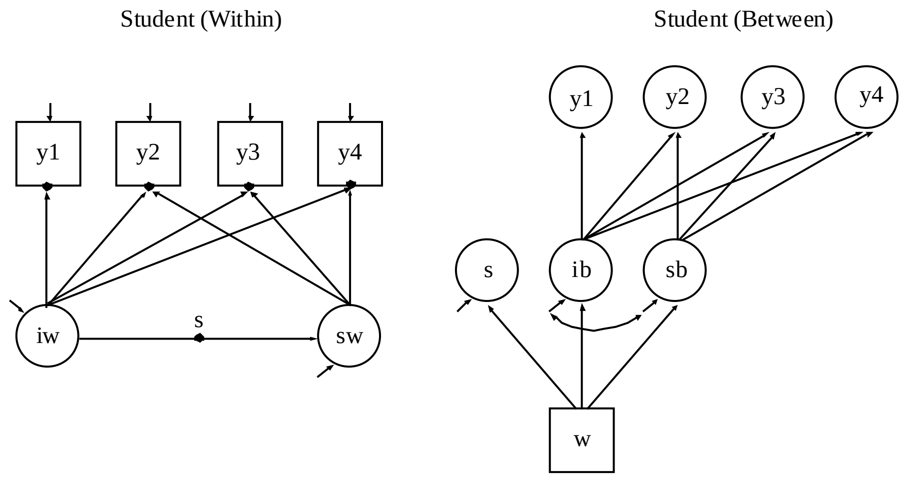


Figure 14: Multiple indicator multilevel growth

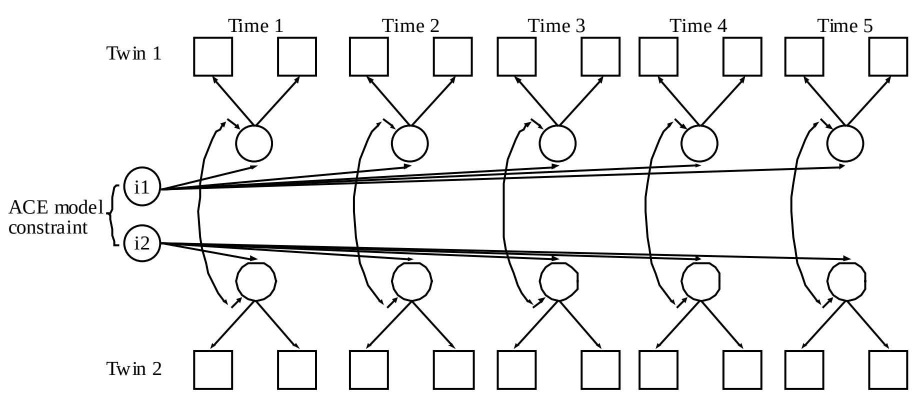
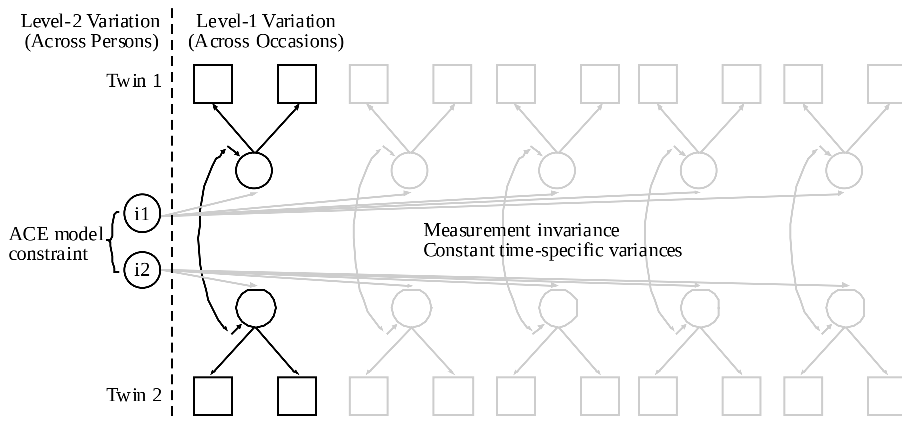


Figure 15: Multiple indicator multilevel growth in long form



mation. With categorical outcomes, the two-level formulation of Figure 15 leads to 4 dimensions of integration with ML, which is possible but still quite heavy. A simple alternative is provided by the new two-level weighted least squares approach discussed for multilevel exploratory factor analysis in Section 4.

7 Multilevel growth mixture modeling

The growth model of Section 5 assumes that all individuals come from one and the same population. This is seen in (22) where there is only one set of γ parameters. Similar to the two-level regression mixture example of Section 2.2, however, there may be unobserved heterogeneity in the data corresponding to different types of trajectories. This type of heterogeneity is captured by latent classes, i.e. finite mixture modeling.

Consider the following example which was briefly discussed in Muthén (2004), but is more fully presented here. Figure 16 shows the results of growth mixture modeling (GMM) for mathematics achievement in Grades 7 - 10 from the LSAY data. The analysis provides a sorting of the observed trajectories into three latent classes. The leftmost class with poor development also shows a significantly higher percentage of students who drop out of high school, suggesting predictive validity for the classification.

Figure 16: Growth mixture modeling with a distal outcome

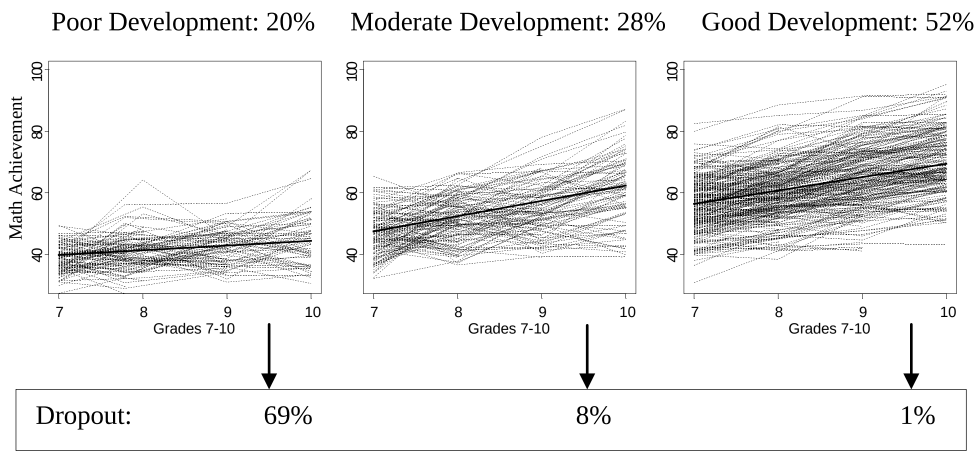
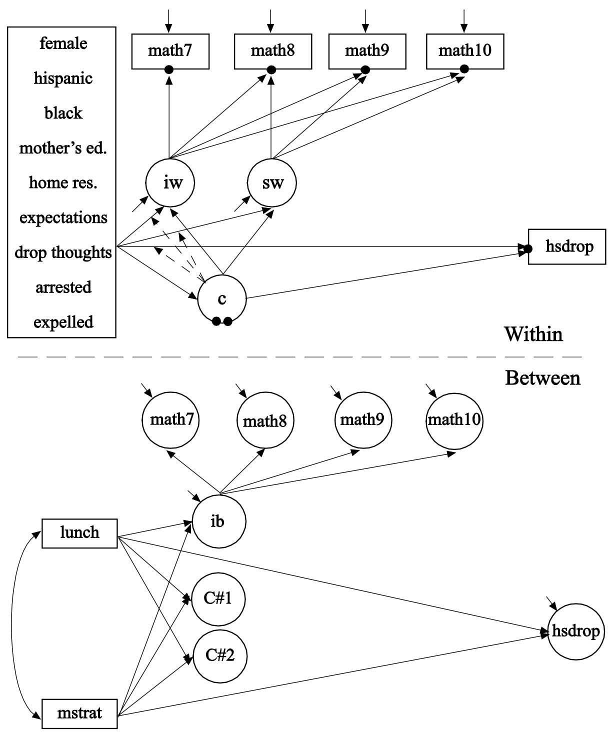


Figure 17 shows the model diagram for the two-level GMM for the LSAY example. In the within (student-level) part of the model, the latent class variable c is seen to influence the growth factors iw and sw, as well as the binary distal outcome $hsdrop$. The broken arrows from c to the arrows from the set of covariates to the growth factors indicate that the covariate influence may also differ across the latent classes. The filled circles for the dependent variables $math7$ - $math10$, $hsdrop$, and c indicate random intercepts. These random intercepts are continuous latent variables which are modeled in the between (school-level) part of the model. For the between part of the growth model only the intercept is random, not the slope. In other words, the slope varies only over students, not schools. Because there are 3 latent classes, there are 2 random intercepts for c, labeled $c\#1$ and $c\#2$. On between there are two covariates discussed in earlier examples, $lunch$ (a poverty index) and $mstrat$ (math teacher workload).

Table 8 gives the estimates for the multinomial logistic regression of c on the covariates. On the within level (student level), the estimates are logistic regression slopes, whereas on the between level (school level), the estimates are linear regression slopes. The within level results show that the odds of membership in class 1, the poorly developing class, relative to the well developing reference class 3 are

Figure 17: Two-level growth mixture modeling with a distal outcome



significantly increased by being male, black, having dropout thoughts in Grade 7, and having been expelled or arrested by Grade 7. The odds are decreased by having high schooling expectations in Grade 7. The between level results pertain to how the school environment influences the student's latent class membership. The probability of membership in the poorly developing class is significantly increased by *lunch*, that is being in the poverty category, whereas *mstrat* has no influence on this probability.

The top part of Table 9 shows the within-level logistic regression results for the binary distal outcome *hsdrop*. It is seen that the probability of dropping out of high school is significantly increased by being female, having dropout thoughts in Grade 7, and having been expelled by Grade 7. The dropout probability is significantly decreased by having high mother's education and having high schooling expectations in Grade 7.

The bottom part of Table 9 pertains to the between level and gives results for the random intercept *ib* of the growth model and the random intercept of the *hsdrop* logistic regression. These results concern the influence of the school environment on the level of math performance and on dropping out. For *ib* it is seen that increasing *mstrat* (math teacher workload) lowers the school average math performance. For *hsdrop* it is seen that poverty status increases the probability that a student drops out of high school. The two random intercepts are negatively correlated so that lower math performance in a school is associated with a higher dropout probability.

It is interesting to study the effects of the school level poverty index covariate *lunch*. The model says that poverty has both direct and indirect effects on dropping out of high school. The direct, school-level effect was just discussed in connection with the bottom part of Table 9. The indirect effect can be seen by poverty increasing the probability of being in the poorly developing math trajectory class as shown in the between-level results of Table 8. As seen in Figure 16 and also in the top part of the model diagram of Figure 17, the latent class variable c influences the probability of dropping out on the student level. In other words, poverty has an indirect, multilevel effect mediated by the within-level latent class variable. This illustrates the richness of detail that a multilevel growth mixture model can extract from the data.

Table 8: Two-level GMM for LSAY math achievement: Latent class regression results

Parameter	Estimate	S.E.	Est./S.E.
Within level			
c#1 ON			
female	-0.751	0.188	-3.998
hisp	0.094	0.705	0.133
black	0.900	0.385	2.339
mothed	-0.003	0.106	-0.028
homeres	-0.060	0.069	0.864
expect	-0.251	0.074	-3.406
droptht7	1.616	0.451	3.583
expel	0.698	0.337	2.068
arrest	1.093	0.384	2.842
Between level			
c#1 ON			
lunch	2.265	0.706	3.208
mstrat	-2.876	2.909	-0.988
c#2 ON			
lunch	-0.088	1.343	-0.065
mstrat	-0.608	2.324	-0.262

Table 9: Two-level GMM for LSAY math achievement: Distal outcome and school-level random intercept results

Parameter	Estimate	S.E.	Est./S.E.	Std	StdYX
Within level					
hsdrop ON					
female	0.521	0.232	2.251		
hisp	0.208	0.322	0.647		
black	-0.242	0.256	-0.944		
mothed	-0.434	0.121	-3.583		
homeres	-0.089	0.052	-1.716		
expect	-0.333	0.052	-6.417		
droptht7	0.629	0.320	1.968		
expel	1.212	0.195	6.225		
arrest	0.157	0.263	0.597		
Between level					
ib ON					
lunch	-1.805	1.310	-1.378	-0.851	-0.176
mstrat	-13.365	3.086	-4.331	-6.299	-0.448
hsdrop ON					
lunch	1.087	0.543	2.004	1.087	0.290
mstrat	-0.178	1.478	-0.120	-0.178	-0.016
ib WITH					
hsdrop	-0.416	0.328	-1.267	-0.196	-0.253

8 Conclusions

This chapter has given an overview of latent variable techniques for multilevel modeling that are more general than those commonly described in text books. Most if not all of the models cannot be handled by conventional multilevel modeling or software. If space permitted, many more examples could have been given. For example, using combinations of model types, one may formulate a two-part growth model with individuals nested within clusters, or a two-part growth mixture model. Several multilevel models such as latent class analysis, latent transition analysis, and discrete- and continuous-time survival analysis can also be combined with the models discussed. All these model types fit into the general latent variable modeling framework available in the Mplus program.

References

[1] Asparouhov, T., Masyn, K. & Muthen, B. (2006). Continuous time survival in latent variable models. Proceedings of the Joint Statistical Meeting in Seattle, August 2006. ASA section on Biometrics, 180-187.

[2] Asparouhov, T., & Muthén, B. (2006). Constructing covariates in multilevel regression. Mplus Web Notes: No. 11.

[3] Asparouhov, T. & Muthén, B. (2007). Computationally efficient estimation of multilevel high-dimensional latent variable models. Proceedings of the 2007 JSM meeting in Salt Lake City, Utah, Section on Statistics in Epidemiology.

[4] Asparouhov, T. & Muthén, B. (2008). Exploratory structural equation modeling. Forthcoming in Structural Equation Modeling.

[5] Bauer, Preacher & Gil (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. Psychological Methods, 11, 142-163.

[6] Little, R.J., & Rubin, D.B. (2002). Statistical analysis with missing data. Second edition. New York: John Wiley & Sons.

[7] Ludtke, O., Marsh, H.W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. Psychological Methods, 13, 203-229.

[8] Muthén, B. (1989). Latent variable modeling in heterogeneous populations. Psychometrika, 54, 557-585.

[9] Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (ed.), Handbook of quantitative methodology for the social sciences (pp. 345-368). Newbury Park, CA: Sage Publications.

[10] Muthén, B. & Asparouhov, T. (2008). Growth mixture modeling: Analysis with non-Gaussian random effects. In Fitzmaurice, G., Davidian, M., Verbeke, G. & Molenberghs, G. (eds.), Longitudinal Data Analysis, pp. 143-165. Boca Raton: Chapman & Hall/CRC Press.

[11] Muthén, B. & Asparouhov, T. (2009). Multilevel regression mixture analysis. Forthcoming in Journal of the Royal Statistical Society, Series A.

[12] Muthén, L. & Muthén, B. (2008). Mplus User's Guide. Los Angeles, CA: Muthén & Muthén.

[13] Olsen, M.K. & Schafer, J.L. (2001). A two-part random effects model for semicontinuous longitudinal data. Journal of the American Statistical Association, 96, 730-745.

[14] Raudenbush, S.W. & Bryk, A.S. (2002). Hierarchical linear models: Applications and data analysis methods. Second edition. Newbury Park, CA: Sage Publications.

[15] Seltzer, M., Choi, K., Thum, Y.M. (2002). Examining relationships between where students start and how rapidly they progress: Implications for conducting analyses that help illuminate the distribution of achievement within schools. CSE Technical Report 560. CRESST, University of California, Los Angeles.

[16] Schmidt, W. H. (1969). Covariance structure analysis of the multivariate random effects model. Unpublished doctoral dissertation, University of Chicago.

[17] Snijders, T. & Bosker, R. (1999). Multilevel analysis. An introduction to basic and advanced multilevel modeling. Thousand Oakes, CA: Sage Publications.