Exploring Non-invariance in Classroom Behavior Trajectories using Growth Mixture Modeling


This project was supported by a grant from the Institute for Education Sciences (IES), U.S. Department of Education (R324B060014). For additional information, please direct all correspondence to Sandra Chafouleas at sandra.chafouleas@uconn.edu
Methodological studies using empirical data can improve our understanding of both method and applied research questions.

- Growth mixture models (GMMs) and latent class growth models (LCGMs) represent person centered techniques useful in studying heterogeneity in the characteristics of change.
- Current best practices for model specification emphasize the importance of the parameterization, class enumeration, and testing invariance assumptions.
- The impact of varying assumptions on the parameters and latent class composition has not been addressed in applied research with multilevel GMMs.
- As a case study, changes in student classroom behavior using Direct Behavior Rating Single Item Scales (DBR-SIS) over one year were analyzed using multilevel GMMs, reflecting students nested within classrooms.
- The approach used in this study reflects what Marsh and Hau (2007) referred to as a “methodological-substantive synergy” (p. 552) in which methodological insights are gained through the application of modeling techniques using real data to gain answers to questions of substantive interest.
Conceptual Framework

MLGMM

DBR-SIS

Number of classes

Non-invariance / Random Effects

More complex models = more complex modeling problems

Behavior Typologies

Single-Case Design
Direct Behavior Rating Single Item Scale (DBR-SIS)

Participants and Measures
- 593 students in grade 7 and 8, with 10 DBR-SIS measures for each of three data collection periods: fall, winter, and spring.
- DBR-SIS (aggregate AE, RS, and DB – reverse scored).
- BESS (once per data collection period)
  - Provided Risk identification (>61)
- Demographic Variables:
  - Race, gender, disability status, behavioral supports, office disciplinary referrals, suspensions and expulsions.

The current study utilizes data from a larger validation study of the Direct Behavior Rating Single Item Scales (DBR-SIS; Chafouleas, 2011). Measures range from 0-30 with 30 representing optimal behavior.
Purpose of the applied research question is to define characteristics of variation in classroom behavior over one year.

Raw data plot of average DBR-SIS composite.

Histogram of DBR-SIS Composite from 1 year
Longitudinal path model was selected to align with data collection procedure.
Prior research using latent growth curve analysis found that intercepts and slopes vary by gender, race, risk, and special ed. status.

F=female, M=Male, NR=non-risk, R=risk, NS=non-special education, and S=special education.
The impact of non-invariance has the potential to affect ML GMM in many ways.

- Number of classes
- Proportion of students within each class and class assignment
- Intercepts and slopes of each class
- Random effects of the intercepts and slopes at the within teacher and between teacher level
- Residual variances
Although the number of GMM studies has grown, only two have included a thorough analysis of tests of invariance.

- Morin et al. (2011) investigated adolescent anxiety trajectories, including testing invariance assumptions.
- Kreuter & Muthen (2008) investigated criminological data using GMM, and also tested invariance assumptions.
- Neither of these studies were multilevel.
Multilevel GMM layers on additional dimensions of complexity with corresponding opportunity to error in selecting models.

- In a simulation study to investigate the effect of ignoring a level of nesting in GMM, Chen, Kwok, Luo, and Willson (2010) found that ignoring the nesting of the data resulted in reduced accuracy in classification, biased variance estimates, and biased standard errors.

- Morin, Maïano, Marsh, Nagengast, & Janosz (2013) investigated adolescent trajectories of measures related to student school life using nested data, but used single level GM modeling due to the small number of clusters, which in this case were schools.
The current study uses multilevel GMM in which model parameters were initially set as invariant and then systematically relaxed.

- Intercepts and slopes can vary between latent classes.
- Intercepts and slopes can vary within latent classes within teacher (random effects).
- Intercept and slope variances can vary across latent class within teacher and between teacher.
- Residual variances can vary across time.
- Residual variances can vary across latent classes.
MODEL 1: Latent Class Growth Model (LCGM), Fixed Effects, Class Invariant Residual Variances.

<table>
<thead>
<tr>
<th>Model</th>
<th>Classes</th>
<th>AIC</th>
<th>BIC</th>
<th>SABIC</th>
<th>Entropy</th>
<th>LMR p-value</th>
<th>BLRT p-value</th>
<th>Best LL rep?</th>
<th>Convergence or matrix problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>41,467</td>
<td>41,559</td>
<td>41,492</td>
<td></td>
<td></td>
<td></td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>2</td>
<td>37,687</td>
<td></td>
<td></td>
<td></td>
<td>0.967</td>
<td>0.14</td>
<td>&lt;.001</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>3</td>
<td>36,605</td>
<td></td>
<td></td>
<td></td>
<td>0.971</td>
<td>0.28</td>
<td>&lt;.001</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>4</td>
<td>36,079</td>
<td></td>
<td></td>
<td></td>
<td>0.928</td>
<td>0.68</td>
<td>1.00</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

Class distribution: 76% 21% 3%

<table>
<thead>
<tr>
<th>Model</th>
<th>Classes</th>
<th>AIC</th>
<th>BIC</th>
<th>SABIC</th>
<th>Entropy</th>
<th>LMR p-value</th>
<th>BLRT p-value</th>
<th>Best LL rep?</th>
<th>Convergence or matrix problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td>34,378</td>
<td>34,585</td>
<td>34,435</td>
<td>0.967</td>
<td>0.41</td>
<td>&lt;.001</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>34,250</td>
<td>34,486</td>
<td>34,315</td>
<td>0.955</td>
<td>&lt;.001</td>
<td>yes</td>
<td>no</td>
<td></td>
</tr>
</tbody>
</table>

Class distribution: 92% 2% 6%
MODEL 3: GMM, Intercept Random Effects that vary by Class, Slope Fixed Effects, and Class Invariant Residual Variances.

<table>
<thead>
<tr>
<th>Model</th>
<th>Classes</th>
<th>AIC</th>
<th>BIC</th>
<th>SABIC</th>
<th>Entropy</th>
<th>LMR p-value</th>
<th>BLRT p-value</th>
<th>Best LL rep?</th>
<th>Convergence or matrix problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3</td>
<td>33,362</td>
<td>33,594</td>
<td>33,426</td>
<td>0.837</td>
<td>0.209</td>
<td>&lt;.001</td>
<td>yes</td>
<td>Yes; NPD PSI</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>33,251</td>
<td>33,527</td>
<td>33,327</td>
<td>0.795</td>
<td>0.763</td>
<td>&lt;.001</td>
<td>yes</td>
<td>Yes; NPD PSI &amp; NPD FOD</td>
</tr>
</tbody>
</table>

Class distribution: 44% 34% 22%
MODEL 4: GMM, Intercept Random Effects that vary by class, Slope Fixed Effects, and Class Non-invariant Residual Variances.

<table>
<thead>
<tr>
<th>Model</th>
<th>Classes</th>
<th>AIC</th>
<th>BIC</th>
<th>SABIC</th>
<th>Entropy</th>
<th>LMR p-value</th>
<th>BLRT p-value</th>
<th>Best LL rep?</th>
<th>Convergence or matrix problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>3</td>
<td>28,677</td>
<td>29,041</td>
<td>28,778</td>
<td>.941</td>
<td>&lt;.001</td>
<td></td>
<td>yes</td>
<td>Yes, NPD FOD*</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Did not converge</td>
<td></td>
</tr>
</tbody>
</table>

Class distribution: 24% 41% 35%

*Non-positive definite first order derivative matrix.
Simulation study using these parameters, but with more clusters, provided evidence of over-parameterization.
MODEL 5 (restricted model 4): Intercepts non-invariant, slopes fixed; residual variances vary by class but are invariant across time. *

<table>
<thead>
<tr>
<th>Model</th>
<th>Classes</th>
<th>AIC</th>
<th>BIC</th>
<th>SABIC</th>
<th>Entropy</th>
<th>LMR p-value</th>
<th>BLRT p-value</th>
<th>Best LL rep?</th>
<th>Convergence or matrix problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td>28,799</td>
<td>28,953</td>
<td>28,841</td>
<td>.947</td>
<td>.065</td>
<td>&lt;.001</td>
<td>yes</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Did not converge</td>
</tr>
</tbody>
</table>

Class distribution: 18% 44% 38%

*Parameterization was simplified to eliminate problems due to empirical under identification and non positive definite matrices.
Using Auxiliary E option in Mplus, differences in Risk, Office Discipline Referrals, and Special Education Status support Model 5.

<table>
<thead>
<tr>
<th>Distal Outcome</th>
<th>Class</th>
<th>Mean 1</th>
<th>Mean 2</th>
<th>Mean 3</th>
<th>Chi-Square Test 1 vs. 2</th>
<th>p 1 vs. 2</th>
<th>Chi-Square Test 1 vs. 3</th>
<th>p 1 vs. 3</th>
<th>Chi-Square Test 2 vs. 3</th>
<th>p 2 vs. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk</td>
<td>1</td>
<td>0.15</td>
<td>0.25</td>
<td>0.00</td>
<td>1 vs. 2</td>
<td>&lt;.01</td>
<td>1 vs. 3</td>
<td>&lt;.01</td>
<td>1 vs. 2</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.75</td>
<td>0.59</td>
<td>0.15</td>
<td>1 vs. 3</td>
<td>&lt;.01</td>
<td>1 vs. 3</td>
<td>&lt;.01</td>
<td>1 vs. 3</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.00</td>
<td>0.91</td>
<td>0.61</td>
<td>2 vs. 3</td>
<td>&lt;.01</td>
<td>2 vs. 3</td>
<td>0.05</td>
<td>2 vs. 3</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Overall</td>
<td></td>
<td>Overall</td>
<td></td>
</tr>
<tr>
<td>ODR</td>
<td>1</td>
<td>0.24</td>
<td>0.31</td>
<td>0.10</td>
<td>1 vs. 2</td>
<td>&lt;.01</td>
<td>1 vs. 2</td>
<td>&lt;.01</td>
<td>1 vs. 2</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.69</td>
<td>0.85</td>
<td>0.25</td>
<td>1 vs. 3</td>
<td>&lt;.01</td>
<td>1 vs. 3</td>
<td>&lt;.01</td>
<td>1 vs. 3</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.93</td>
<td>0.83</td>
<td>0.60</td>
<td>2 vs. 3</td>
<td>0.01</td>
<td>2 vs. 3</td>
<td>0.86</td>
<td>2 vs. 3</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Overall</td>
<td></td>
<td>Overall</td>
<td></td>
</tr>
<tr>
<td>SpecEd</td>
<td>1</td>
<td>0.19</td>
<td>0.18</td>
<td>0.03</td>
<td>1 vs. 2</td>
<td>0.16</td>
<td>1 vs. 2</td>
<td>0.74</td>
<td>1 vs. 2</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.22</td>
<td>0.14</td>
<td>0.20</td>
<td>1 vs. 3</td>
<td>0.78</td>
<td>1 vs. 3</td>
<td>0.69</td>
<td>1 vs. 3</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.16</td>
<td>0.21</td>
<td>0.22</td>
<td>2 vs. 3</td>
<td>0.76</td>
<td>2 vs. 3</td>
<td>0.61</td>
<td>2 vs. 3</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td>0.38</td>
<td>Overall</td>
<td>0.92</td>
<td>Overall</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

Comparison of Means using Posterior Probability based Multiple Imputation

Equality test of Means using Posterior Probability based Multiple Imputation
Discussion and Limitations

Discussion

• As the variances are allowed to vary across latent classes, the growth curves and class compositions change dramatically.
• The latent class containing students at risk changes from 3% in Model 1 to 38% in Model 4.
• Modeling non-invariance uncovered variability as the distinguishing feature of the heterogeneous latent subgroups of student classroom behavior.
• The plots displaying variability provide evidence that the classes and students are best characterized by their variability rather than mean levels.
• Next step will include an analysis of students who are movers versus stayers as model parameterization changes.

Limitations

• The GMM is an exploratory technique. With real data it is unknown whether the results are reflective of true latent subpopulations.
• The DBR-SIS lack normality; this may cause over extraction of latent classes.
• Because we used real data, methodological findings may not be generalizable.
Questions? Thank you!

Janice.Kooker@uconn.edu
Key References


