Implications of Using Multilevel Latent Class Analyses on School Policy Interventions

Julio C. Cabrera, Stacy R. Karl, Michael C. Rodriguez
University of Minnesota

Maria Elena Oliveri
Educational Testing Services

Minnesota Youth Development Research Group

April, 2014

Paper presented at the annual meeting of the
American Educational Research Association
Philadelphia, PA.

Citation:

Cabrera, J.C., Oliveri, M., Karl, S.R., & Rodriguez, M.C. (2014, April). *Implications of using multilevel latent class analyses on school policy interventions*. Paper presented at the annual meeting of the American Educational Research Association, Philadelphia, PA.

Implications of Using Multilevel Latent Class Analyses on School Policy Interventions

Abstract

In this study, a nonparametric random effects Multilevel Latent Class Analysis (MLCA) model is utilized to analyze individual and contextual-level predictors of bullying behaviors encountered on school property. Surveys administered to 33,612 12th grade students from across 216 schools were used for these analyses. Results show that a four student-level latent class and two school-level latent class model with a school-level common factor for student-level indicators had the best model fit for the data as evidenced by having the lowest values of model fit indices robust to MLCAs (i.e., AIC, AIC3, modified BIC, and sample-adjusted BIC). Multinomial logistic regressions of the four student-level latent classes, depicting bullying encountered in schools, and indicators on both individual- and contextual-level predictors showed that a number of student- and school-level variables were significant predictors of both student- and school-level latent classes and student-level latent class indicators. Results are presented with an emphasis on the effects of the contextual predictors on both the student- and school-level units and their possible impact on school policy and interventions.

INTRODUCTION

Educational researchers often work with datasets that are nested in nature. The ubiquity of nested data has led to a plethora of statistical methods and tools, including hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002), generalized mixed linear modeling, and growth curve modeling, that are capable of modeling the often ignored dependency of data units of observation. HLM has been the primary approach for modeling nested data due to various reasons. HLM has been largely utilized because the majority of the time data is gathered from naturally occurring nesting situations that contain only two levels of hierarchy, for example, data related to student outcomes collected from various different schools (i.e., students nested within schools). HLM offers the most straightforward method to model these data, as well as three-level, longitudinal, and other hierarchical models. However, other modeling techniques (e.g., bootstrap methods that resample clusters of observations, clustered robust standard errors, or Multilevel Latent Class Analysis [MLCA]) have often been ignored despite their potential usefulness in deriving meaningful and robust results, especially when dealing with categorical data (see Arceneaux & Nickerson, 2009). The purpose of the current study was threefold: (a) the main purpose was to demonstrate an application of one of these less-used methods that models nested data–MLCA–to examine survey data administered to 33,605 12th grade students, and model the bullying behaviors encountered by these students across 216 schools from a large Midwestern state of the United States; (b) the secondary purpose was to emphasize the necessity to model between-school differences in order to describe student level relationships, as often as the data deem it possible; and, (c) the third purpose was to highlight the usefulness of the degree of contextualization and specificity that this methodology provides when designing data-driven policy interventions.

Perspective

More often than not, data analysts using nested data ignore the hierarchical structure of the data due to practical reasons, including a lack of a proper sample size at both the individual and group levels (which are required to accurately model the data), a significant degree of data missing on one or more variables of interest, attrition, or low intraclass correlations (ICCs) (de Leeuw, & Kreft, 1995). However, ignoring the nested structure of the data has been shown ad nauseam to bias parameter estimates and standard error estimations of the parameters of interest (e.g., group-level effects) as well as increase Type I error rates (see McCoach, 2010; McCoach & Adelson, 2010), even in the presence of low levels of ICCs (i.e., low ratios of between group variance to total variance). Although a number of hierarchical data analysis methods are available that account for nesting, the precision estimates for some methods are poorly understood, making it difficult for researchers to select the most appropriate method (Arceneaux & Nickerson, 2009). Some of these hierarchical data analysis methods (e.g., multilevel multinomial logistic regression) are able to model both continuous and categorical (binary, polytomous, and count) dependent variables, however, models dealing with multiple categorical dependent variables have yet to be readily implemented outside of structural equation modeling (SEM), without first having to scale the variables into a single continuous factor. In this study, we briefly describe the methodology and usefulness of MLCA when dealing with nested survey data containing multiple indicators of dependent categorical variables.

Multilevel Latent Class Analysis

Latent Class Analysis (LCA) is one of the general latent structure models (i.e., latent class models) used to identify and classify clusters (classes) of units that are closely homogenous with respect to observed variables (see Hagenaars & McCutcheon, 2002). Similar to factor analysis, latent classes are extracted in a manner so that the correlations between observed variables disappear completely within each latent class (Mutz, Bornmann, & Daniel, 2013). LCA is most appropriately used when mutually independent observed indicators are associated because of an underlying unobserved (latent) variable, and when the observed indicator variables are categorical in nature, although observed variables could also be continuous and censored, which allows for a more flexible approach to multilevel analyses (Muthén & Muthén, 2012).

LCA has several assumptions, including homogeneity, local independence, unidimensionality, and monotonicity (Hagenaars & McCutcheon, 2002). The assumption pertinent to the current discussion is the assumption that observations are independent of one another. In LCA models, and other models that assume local independence, this assumption is not often met with many data structures collected through surveys or state assessments, especially when the data include responses from students nested in schools, or students nested in classrooms that are nested in schools, which is often the case in large-scale data collection practices. MLCA, a multilevel version of LCA that accounts for the nested structure of data, allows latent class intercepts to vary across level-2 (e.g., schools) units and examines how these units affect the latent classes at level-1 (Asparouhov & Muthen, 2008; Henry & Muthen, 2010; Vermunt, 2008). By allowing these random intercepts to vary across level-2 units, the probability of membership in a level-1 latent class varies across level-2 units. For example, the probability that a student will belong to one of the latent classes (e.g., high bullying encounters) is likely to vary significantly across schools (e.g., urban versus suburban schools). Additionally, since the dependent variable in MLCA is latent rather than observed, there is an advantage of being able to combine the measurement model with a structural model and analyze additional models such as structural equation modeling (SEM) and growth modeling (Muthen & Muthen, 2012). Thus, covariates at both level-1 and 2 may be added to an MLCA to assess the influence of level-1 covariates in predicting the probability that an individual will belong to a level-1 class while simultaneously assessing the effect of level-2 predictors to both level-1 and level-2 classes (Henry & Muthen, 2010).

MLCA generally relies on various model selection indices to aid in the identification of the model with the best fit (i.e., models with small deviations between the expected values and observed values). Operationally, models with higher number of latent classes usually have better model fits, thus, the trade-off between model complexity and model fit must be a priority in model selection (Zhang, Zhang, Zhang, & Jiao, 2013). Reviews of LCA studies that look at the performance of such indices have identified that traditional indices, such as the Pearson's χ^2 and the likelihood ratio, are vulnerable to estimation bias due to their sensitivity to large sample sizes, thus reviewers have advocated for the use of information evaluation criteria that counter such limitations when selecting models (Zhang et al, 2013). The performance of these information criteria under MLCA modeling has also been reviewed–although its body of literature is miniscule–and three indices have been highlighted as the best model selection tools, including Akaike's information criterion (AIC; when the separation between level-2 classes is low), a slightly modified Bayesian information criterion (BIC; see equation 1), and AIC with a penalty factor of three (AIC3) (Lukociene, & Vermunt, 2010; Lukociene, Varriale, & Vermunt, 2010; Zhang et al., 2013). These information criteria are expressed as:

$$\begin{aligned} \text{AIC} &= -2\ln(L[\hat{\theta}]) + 2k; \\ \text{BIC} &= -2\ln(L[\hat{\theta}]) + \ln(J)\,k; \text{ and,} \\ \text{AIC3} &= -2\ln(L[\hat{\theta}]) + 3k, \end{aligned} \tag{1}$$

where $\ln(L[\hat{\theta}])$ is the maximized log-likelihood value for a model with parameters $\hat{\theta}$, k is the number of the independent (free) parameters, and J is the number of level-2 units (as opposed to using the sample size of level-1 as it is used in LCA; Lukociene et al., 2010). Additionally, these model selection indices have been shown to be sensitive in accuracy of estimation, depending on sample size J and/or level-1 per level-2 units (i.e., n_j) (Zhang et al., 2013).

Methodology

Data

Minnesota Student Survey (MSS). The current study entails a secondary analysis of the 2010 MSS database. Data from the MSS are provided by public school students in Minnesota via local public school districts and managed by the MSS Interagency Team, including the MN Departments of Education, Health, Human Services, Public Safety, and Corrections. The MSS is administered every three years, most recently in 2013. During each administration year, all operating public school districts are invited to participate. In 2010, a total of 130,908 students participated from grades 6, 9, and 12.

Procedure

In this paper, MLCA was used to differentiate students into latent classes of bullying behavior encountered by 36,734 students in 12th grade from 347 schools, where the number of students per school ranged from 1 to 620. The majority of the sample was White (78.5%) and about half of the students were females (50.3%). Taking into consideration the minimal sample size recommendations from Zhang et al. (2013), to accurately estimate adequate values of model selection indices for MLCAs, schools with fewer than 40 students were removed from the dataset and were not included in the analyses. Consequently, the total number of schools included in the analyses was reduced to 216, with the total student sample size being reduced by 8.5%; no major group changes were noted, with ethnic minorities composing 21.9% of the sample and females composing 50.4% of the sample. Individual and contextual-level predictors of the level-1 latent class (i.e., bullying behavior encountered) membership were also considered. Student level predictors were included and consisted of constructs measuring individual, family, and school related beliefs and attitudes (e.g., family support), as well as age, social economic status (SES), student performance (as measured by grade point average [GPA]), and gender. At the contextual-level, the following school predictors were considered: school location (e.g., urban versus rural), student attendance rate, and the SES of the school. For all of these analyses, the software Mplus (Ver.7.1; Muthén & Muthén, 2012) were used, and a similar methodology as that of Henry and Muthen (2010) was followed. The current analyses add to their analyses a couple of important pieces of information: (a) Zhang et al.'s (2013) analysis of how model indices are biased if level-1 and/or level-2 sample sizes do not meet the required sizes, something that Henry and Muthen do not readily present when describing their sample, but it is important to note that the researchers did not have Zhang et al.'s information available at the time of their analysis as evidenced by they're acknowledgement that research of the performance of model indices such as the BIC for MLCA models were still needed; and, (b) we present the results of the nonparametric random model since Henry and Muthen presented results for the parametric random model.

Measures

Latent class membership. In order to differentiate students into latent classes of bullying behavior encountered on school property, a total of eight categorical items (seven dichotomous) were used. The majority of the dichotomous items (0 = No, 1 = Yes) included questions regarding the occurrence of physical and/or sexual forms of bullying (e.g., during the last 12 months has a student: "… pushed, shoved, or grabbed you?" "… touched, grabbed or pinched you in a sexual way?"), another dichotomous item (0 = No, 1 = Yes) inquired about illegal drug solicitation on school property, and one polytomous item was included (0 = 0 times, 1 = 1 times, 2 = 2 or 3 times, 3 = 4 or 5 times, and 4 = 6 or more times) to collect information on the number of times that someone stole or damaged the property of the student (see National

Society for the Prevention of Cruelty to Children, 2004; and U.S. Department of Health and Human Services, 2014; for definitions and types of bullying behaviors).

Individual predictors. As presented above, several student-level predictors were entered in the analyses, including age (*M* = 17.6, *SD* = 0.6), GPA (an average of two self-reported grades; *M* = 3.0, *SD* = 0.7), SES (0 = High, 1 = Low), and gender (0 = Male, 1 = Female). Additional individual constructs were included that measured: (a) mental distress (α = .85), including eight categorical items regarding students' irritability, concentration, sleep troubles, sadness, stress, hopelessness, and nervousness; (b) family support (α = .65), including eight categorical items measuring perceptions on how the parents, relatives, and other adults care about the students, family respect, and parental availability; (c) family sources of information (α = .60), including four categorical items; (d) problems with family (α = .65), including four categorical items measuring family-related alcohol, drug, and/or physical abuse; (e) school safety/climate (α = .81), including five categorical items measuring school safety and climate perceptions of students; and, (f) teacher/community support (α = .80), including six categorical items regarding students' perceptions of how teachers and other adults in the community (e.g., church leaders) feel about them and/or care about them. These constructs have adequate internal consistency and strong psychometric properties, and have been constructed with strong theoretical properties (see Cabrera & Rodriguez, 2011; Rodriguez & Cabrera, 2010; Vue, Stanke, Palma, Cabrera, Bulut, Latterell, & Rodriguez, 2013).

Contextual predictors. Given the flexibility of the MLCA models to model school-level predictors, three variables were included in the analyses, including (a) the SES of the school (as measured by the percentage of students qualified for free or reduced lunch; two dummy variables were created to indicate three levels of SES including High [reference group], Medium, and Low

SES), (b) the location of the school (three dummy variables were created to indicate four locations including city, suburb, town, and rural [reference group]), and (c) average daily student attendance (percentage) at each school.

Results from MLCA will be compared to the individual-level LCA models that do not take into consideration the hierarchical nature of the data, additionally, level-1 and level-2 predictors will be analyzed in order to investigate their associations with the latent class membership of the bullying behavior encountered.

Results

Student-latent class analysis

In the first step of this MLCA, the eight indicators of bullying encountered at school were analyzed with one to five latent classes, while the nested data structure was ignored. LCA results for the five latent classes are presented in Table 1. Out of the five latent class models, the model with four latent classes and 47 free parameters showed the best model fit (lowest values of AIC, AIC3, and BIC) without estimation errors; the model with five latent classes had the best model fit values, but estimation errors were noted and the model was no longer considered. Additionally, the difference in entropy values between the 4-class solution and the 3-class solution was of negligible size (.02), providing further evidence of a 4-class level-1 structure. Posterior probabilities for the 4-class solution were adequate, indicating that this model was able to distinguish between students in all four latent classes. Given this evidence, the model with four latent classes was chosen as the best level-1 model and was used in subsequent analyses.

The four latent classes modeled with LCA represent four groups of students who encountered bullying differently. The largest group (71.8%) represents students that experienced minimal to no bullying encounters, throughout the paper referred to as *Low*, whereas the remaining three groups experienced bullying much differently: (a) the smallest group, termed *High* (4.5%), reported high levels of both physical and sexual bullying; (b) similarly, the second largest group (12.4%), termed *Moderate-Sexual*, reported both physical and sexual bullying encounters, but the majority of bullying behavior encountered tended to be more of a sexual nature (e.g., being touched, grabbed or pinched in a sexual way); and (c) the third largest group 11.4%) , reported encountering both types of physical and sexual bullying as well, but the bullying tended to be more of the physical type (e.g., being pushed, shoved or grabbed), termed *Moderate-Physical*. It is important to note that the difference of one percent in sample size between the two moderate groups may seem negligible; however, the type of bullying the two groups encountered differed substantially, with one group experiencing bullying that was physical in nature whereas the other group encountered more sexual bullying than physical bullying (see Table 2).

Multilevel latent analysis of students within schools

Following the identification of the best-fitted student-level model (i.e., the four latent class model), one to three level-2 latent class models that take into consideration the nested structure of the data were analyzed using a nonparametric approach. Another viable approach could have been to use a parametric MLCA model; however, the random means (i.e., the student-level latent variables allowed to vary across level-2 components) are assumed to be distributed differently between the two approaches (Henry & Muthen, 2010). In the parametric approach, the random means are normally distributed, assumed to be highly correlated, and vary across level-2 units, whereas in the nonparametric approach the random means have different distributions (i.e., multinomial distribution) and vary across level-2 latent classes (Vermunt, 2008). The differences in interpretations of the outcomes mainly lie in the level-2 covariate effects on the level-2 components, more specifically, and for example, in the parametric approach school covariates predict a school's probability that a student will belong to a certain student-level latent class, whereas in the nonparametric approach school covariates predict the probability that a school will belong to a school-level latent class. Given that the parametric model assumes the random means to be normally distributed, computation time for models with more than two level-1 latent classes takes much longer than for nonparametric models. Due to computation time and previous parametric examples shown (e.g., Henry & Muthen, 2010), a nonparametric model was used in the subsequent analyses.

In the nonparametric model, the level-2 model with two latent classes and 51 free parameters showed the best model fit (lowest values of AIC, AIC3, and BIC) over the one latent class and the three latent class models. The differences in entropy values between the three models were small, providing further evidence for a 4-class student-level with a 2-class level-2 structure (see Table 3). These level-2 classes represent two different types of schools. The first school-level latent class represents 46.1% of the total number of students in the sample, with 11.2% of the total number of students (24.4% within the class) experiencing bullying, while the second school-level latent class represents 53.9% of the total number of students with 21.3% (39.5% within the class) experiencing bullying (see Figure 1). The second school-level latent class contains twice as many students from the total sample who experience some type of bullying than the first school-level latent class.

Using the 4-class student-level and 2-class school-level model as the best fitted model, a common school-level factor for the student-level indicators (i.e., the bullying items) was then included in the model to assess the influence of schools on student-level indicators that define the student-level classes. Not only does this common factor allow us to assess indicator-specific school influence, the inclusion of this school-level factor also significantly decreases the amount of computation time by reducing the dimensionality of the intercepts, and it allows the examination of school-level covariates on the student-level latent class indicators (Henry & Muthen, 2010). The improvements in model fits from a fixed-effects model (Table 1), to a nonparametric random-effects model (Table 3), and finally to a nonparametric random effects model with a level-2 common factor for level-1 indicators (Table 3), provide empirical evidence that schools have an influence on student-level latent classes and must be taken into consideration when modeling these data. This influence can be graphically seen as well, by comparing the composition of the student-level latent classes, based on the estimated posterior probabilities of the school-level latent classes, with the composition of the student-level latent classes that ignore the hierarchical nature of the data (see Figure 2).

Student- and School-Level Predictors

Student and school predictors were added to the model described above. Multinomial regressions were used to assess the effects of student and school predictors on student-level latent classes, student-level latent class indicators, and school-level latent classes (see Figure 3 and Table 4). The results in the first three columns of Table 4 are the results from the comparison of students who encountered a high degree of bullying to students who encountered moderate-physical bullying. The odds that students would encounter high bullying (compared to moderate-physical bullying) were significantly higher for students from low SES (versus high SES), those that were male, identified themselves as being of other/multiple races, with higher mental distress and problems with family, as well as with lower perceptions of school/safety climate and teacher/community support. Additionally, the odds of being in the high-bullied group as compared to the moderate-physically bullied group for Asian Americans were about 0.31 times lower than for Whites. The results in the next three columns (i.e., columns 4, 5, and 6) in Table 4 are the results from the comparison of students from the moderate-sexual bullied group and the moderate-physically bullied group. The odds that a student would belong in the moderate-sexually bullied group (compared to the moderate-physically bullied group) were significantly higher for females, with lower perceptions of school/safety climate, and for students with higher mental distress, teacher/community support, and GPA. Whereas the odds that a student would belong to the moderate-physically bullied group (compared to the moderate-sexually bullied group) were significantly higher for males, with lower teacher/community support, and lower GPAs, but with lower mental distress and higher perception of teacher/community support.

The third set of results in Table 4 (columns 7, 8, and 9) present the odds for students to belong to the low bullying group versus the moderate-physical bullying group. The odds that students experience low to no bullying, compared to moderate-physical bullying, are higher for students that are high in SES, female, older in age; for those with higher perceptions of school/safety climate and teacher/community support; for students with high reported GPAs; and for students with lower problems with family. Additionally, the odds of experiencing low to no bullying for Asian American students were 1.90 times higher than the odds for White students. Overall, female students were 6 to 7 times more likely to be in the moderate-sexually bullied group or being in the low bullied group than in the moderate-physically bullied group, and 0.46 times less likely to be in the high bullied group than in the moderate-physically bullied group.

School-level covariates were used to predict the random means for the student-level classes. The first three columns have the comparisons of the high bullied group with the moderate-physical bullied group, which show that suburban schools and school with medium SES (compared with high SES) more likely have students with high bullying encounters than with moderate-physical bullying encounters. Medium SES schools have 21% higher odds of having students with high bullying encounters than with moderate-physical bullying encounters. The following three columns show that there is no between-school differences in the likelihood of having students from moderate-sexual bullying groups or moderate-physical bullying groups, meaning that both are equally likely. The last three columns show the comparisons between students with low bullying encounters and moderate-physical bullying encounters, and indicate that the odds of a student encountering low bullying behaviors are 120% higher if the school is located in a city, 68% higher if the school is located in a town, and 34% higher if the school is located in a rural area, as compared to a school being located in a suburb. Lastly, the odds of a student encountering low bullying (compared to moderate-physical bullying) are lowered by 27% if the school is considered of medium SES (versus high SES).

The last set of analyses looked at the impact of school-level predictors on student-latent class indicators (i.e., the items indicators of bullying encountered). The school-level predictors that had a statistically significant effect on student-level latent class indicators are listed in Table 5. It is seen that the probability of being “kicked, bitten or hit” was 94% higher in rural schools and 37% higher in town schools than in suburban schools, and 17% higher in medium SES schools than in high SES schools. Rural and town schools (in comparison to suburban schools) also showed a higher probability of students being “touched, grabbed, or pinched in a sexual way” (36% and 16% higher, respectively), while medium SES schools (versus high SES schools) showed a marginally significant 8% higher probability. Rural and town schools also show a lower probability than suburban schools to have students that were offered, sold, or given an illegal drug on school property. These significant school-level predictor effects on student-level latent class indicators are clear evidence of the impact that these higher level effects have on lower level components that would be missed when modeling nested data without a multilevel model.

Overall, our findings indicate that a nonparametric random effects MLCA model with a four student-level class and two school-level class structure and a school-level common factor for student-level indicators that takes into consideration the complexity of the nested survey data has strong empirical model properties over individual-level model analysis. This superiority in nested data modeling is evidenced by: (a) better model fit indices (i.e., AIC, AIC3, BIC, and sample-size adjusted BIC), that have taken into account some adjustments needed (e.g., $n_j > 40$), for multilevel models over individual-level models; (b) significant influence of school-level predictors on student-level latent class indicators; and (c) significant influence of school-level predictors on membership of student-level latent classes.

Discussion

Most data analyses used by administrators and teachers, and some educational researchers, to inform their practices are performed at the individual-level (i.e., at the student or classroom or school or district or state level), with only a few analyses performed, aside from multilevel research studies, that take the multilevel dependency into account and provide robust results. In general, data findings used for school decision-making have simply focused on one of the hierarchical levels at a time; however, recent value-added movements have been driving this practice to change even more so than when HLM was introduced. These changes in practice are warranted given that possible negative outcomes could occur when decisions are made using (possibly) biased results. Models that are able to specify how school-level variables influence relationships of student-level variables are highly beneficial since these contextual-level predictors offer an avenue to model what is actually happening in school classrooms (Adcock &

Phillips, 1997). Furthermore, our results have implications on the types of policy interventions that can be implemented based upon results from nested survey questions. To illustrate, the question: “has a student … kicked, bitten, or hit you?” may yield different results depending on whether single- versus multilevel analyses methods are used. Results from single-level analyses may indicate that students do not encounter these behaviors on average whereas results from multilevel analyses–which consider the nature and type of schools from which the data were collected–may indicate that students are being kicked, bitten, or hit more often in certain (e.g., town) school locations. Results derived from the single-level analysis may not prompt any interventions given that students do not seem to encounter these behaviors (on average) whereas results from the multilevel analysis may foster interventions (e.g., instituting a cognitive behavioral intervention). Such degree of specificity and contextualization can be useful when designing school, district, or state policy to generate interventions that are better targeted to the groups or schools in need. Therein lies this study’s key contribution, which is to guide policy and intervention by obtaining more specific and meaningful results from data analyses. Most studies fail to make these types of contextualized analyses, thus leading to broad and general applications and policy interventions, which may be of limited use and generalizability for particular subgroups.

Multilevel models that include contextual-level predictors for both the individual and school levels, like MLCA, offer a clear avenue to what is actually happening in school classrooms (Adcock & Phillips, 1997). In this study, we presented a set of data results that could be potentially useful in directing data-driven policy and/or interventions. Our analysis included both student-level as well as school-level variables that could prove to be highly valuable and informative for policymakers in implementing school, district, or state policies or interventions regarding bullying occurring at their schools. Furthermore, the MLCA results presented in this paper highlight the effects that individual- and school-related variables had on the four groups of students that encountered bullying differently, these results have implications in educational settings at both levels and can inform student practices that can be changed and school interventions that can be introduced. For example, an emphasis on targeting schools in suburban areas could help to strategically specify the allocation of funds for bullying interventions, given that these school are more likely to have students that are highly bullied than schools in rural areas, towns, or cities. At the student-level, these interventions could also specifically target students that are prone to or suffer from mental distress, given that these students are highly likely to experience high levels of bullying. On the flipside, interventions could also specifically work on increasing school safety/climate perceptions since students are 23% to 28% less likely to experience bullying if their perceptions of school safety/climate are high. Lastly, gender differences in the amount of bullying and type of bullying encountered are evident in these results, with females experiencing 6.92 times more low to no bullying (versus physical bullying) than males, but when they did experience moderate bullying, females experience 6.74 times more moderate sexual bullying than moderate physical bullying. Different interventions for these students could be constructed to target the different types of bullying prevalent at their schools, and to schools located in different locations, e.g., female students in schools in rural areas and in towns are 36% and 16% more likely, respectively, than schools in suburban areas, to state that another student touched, grabbed, or pinched them in a sexual matter.

Given that schools in rural areas and in towns were noted as more likely to experience less bullying (34% and 68%, respectively), this kind of specificity of student-level indicators would provide us with information hard to ignore; mainly, that rural and town schools are in need of a specific type of bullying intervention, i.e., an intervention to reduce the amount of touching, grabbing, or pinching in a sexual matter, even though looking at the school level results that rural and town schools are less likely to have students being bullied would indicate otherwise. This degree of specificity and contextualization would be pertinent when designing school, district, or state interventions that would be more targeted to the groups or schools that are in actual need. Even though these types of analyses that obtain more specific and meaningful results from survey data analyses would be great contributions to guide policy and interventions, the current system of data-driven decision- and policy-making is behind the times.

Up until the last few years, considering data in decision making at the district, school, and classroom level has been the exception rather than the rule, with decision makers referring more to educational philosophy and political necessity rather than data results in their decision making processes (Means, Padilla, & Gallagher, 2010). Nowadays, the use of data in school settings is mainly to aid in meeting individual student needs and to determine if existing school goals are being met, while being far less applied in determining the indicators of problems or the decisions about needed systemic changes (Rodriguez, Matuska, Cabrera & Karl, 2010). Moreover, Dunn, Airola, Lo, and Garrison (2013) note major differences in data use between hierarchical units: at the classroom level, data are more often used as a teaching tool to implement differentiated instruction to both the class and students, whereas at the school level, data are used for decision making, including school improvements, curricula decisions, and specialized instructional grouping or placement of students (Means et al., 2010). However, when data findings are used by schools to make policy decisions, the findings often come from data that can be easily compiled and analyzed (i.e., student achievement data [e.g., GPAs, standardized assessment scores], biographical data including ethnicity, free and reduced lunch, special education status, or

English language proficiency status; Means et al., 2010), while other types of data such as behavioral records, survey data, or discipline referrals are often low in priority (Rodriguez et al., 2010) to be considered for further analysis to inform decisions. One of the main barriers in using this latter type of data, such as the data we used in this study, for informing policy or interventions is that the data do not come from only one source but from different sources or warehouses. These data are often not easily linked and available for use in decision-making, which may deter schools from using it, as schools do not have the capabilities to readily do so (Means et al., 2010). However, when administrators and teachers know or are trained to use data effectively and correctly, the benefits are manifold for both schools and students (Messelt, 2004; Rodriguez et al., 2010). It is this clear advantage that schools can gain if school district data management systems and evaluation offices get in sync with the research, evaluation and accountability needs that are fulfilled by multilevel analysis models (Adcock &Phillips, 1997).

Looking at the results of these analyses of bullying encounters, the data-driven changes advocated by Adcock and Phillips are much to be desired, especially given the high number of negative effects that victims of bullying display compared to their non-bullied counterparts, including increased absenteeism, depression, anxiety, loneliness, helplessness, emotional instability, establishment of poor relationships, and/or loss of friendships (see Brown, Aalsma, & Ott, 2013; Morgan, 2012). Additionally, contextual-level variables are able bring to light school-related effects that students are not able to verbalize, as some teachers may notice some of the differences in behavior mentioned above in students but be unaware of the circumstances affecting students in their classrooms. These unseen problems are further exacerbated, as students are simply more likely to admit that they have been bullied to a parental figure rather than to a school official (Brown et al., 2013). If students have problems with parental figures or their family members, to whom do they turn to? Unfortunately, the effect of high number of problems with family members is evident in our results, as these students are 1.61 times more likely to experience high levels of bullying. The lack of student-school communication regarding the reporting of bullying is partly due to the lack of awareness of school administrators to the reality of bullying behaviors taking place at their schools (Dornfeld-Januzzi, 2006). This unawareness of school staff is not surprising, as administrators rank bullying prevention efforts as less important than other school matters, such as staff development, simply because the staff believed that bullying is less problematic at their schools than at other schools. These differences in the number of bullying instances reported by students in this and other surveys (e.g., Youth Risk Behavior Surveillance System, and School Crime Supplement), and the perceptions of school administrators about the lack of bullying behaviors occurring at their schools could be partially based on data summaries, briefs and information gathered from student-level analyses rather than from information that take into account the influence(s) that one level has on the other holistically, thus the importance to conduct multilevel analysis, such as MLCA, whenever possible to examine how school-level and student-level factors interact and influence each other.

Limitations

Several limitations exist in the presentation of these analyses, including the construction of the mental distress, family support, family source of information, problems with family, school safety/climate, and teacher/community support scales prior to these data analyses, which are limited to data collected from a single state in the United States. These limited data may render the results to not being generalizable to other geographical locations. In addition, the data are self-reported, including age, SES, grades, and all other independent and dependent indicators. Although some data were excluded from the analyses due to aberrant response patterns, a more detailed and systematic approach to identify non-responders, fakers, and other data cleaning methods ought to be taken into consideration if the analyses were to be replicated with students' responses from other grades and/or other administration years. Also, data that are incomplete or perceived to be inaccurate at the school-level ought to be examined and verified for accuracy. For example, in terms of enrollment, having a school with one student enrolled in 12th grade seems highly improbable, but may be a function of the voluntary nature of the survey. Lastly, software limitations are noted as well for those that plan to run future similar models, including the large amount of computation time for a single model.

Table 1
Fit Statistics and Criteria by Number of Latent Classes for Level-1 Model Specification

	Number of Level-1 Latent Classes				
Model	1 Class	2 Classes	3 Classes	4 Classes	5 Classes
Fixed effects model					
No. of free parameters	11	23	35	47	59
Log-likelihood	-131825	-113355	-111625	-110824	-110623
AIC	263671	226755	223320	221741	221363
AIC3	263682	226778	223355	221788	221422
BIC	263765	226951	223618	222141	221865
Sample-adjusted BIC	263730	226878	223507	221992	221678
Entropy	1.00	0.84	0.79	0.77	0.79

Note: AIC = Akaike information criterion; AIC3 = Akaike information criterion 3; BIC = Bayesian information criterion

Table 2

Latent Class Proportions for the Four-Class Level-1 Model

		Class 1	Class 2	Class 3	Class 4
		4.5%	11.4%	12.4%	71.8%
		Bullying			
Indicator		High	Moderate Physical	Moderate Sexual	Low
During the last 12 months					
has a student:					
threatened you?	No	0.14	0.55	0.82	0.98
	Yes	0.86	0.45	0.18	0.02
pushed, shoved or grabbed you?	No	0.04	0.10	0.74	0.95
	Yes	0.96	0.90	0.26	0.05
kicked, bitten, or hit you?	No	0.15	0.52	0.96	0.99
	Yes	0.85	0.48	0.04	0.01
stabbed you or fired a gun at you?	No	0.73	0.99	0.97	1.00
	Yes	0.27	0.01	0.03	0.00
touched, grabbed or pinched you in a sexual way?	No	0.09	0.61	0.52	0.98
	Yes	0.91	0.39	0.48	0.02
made unwanted sexual comments, jokes, gestures, or looks towards you?	No	0.11	0.61	0.35	0.94
	Yes	0.89	0.39	0.65	0.06
how many times has someone stolen or deliberately damaged your property such as your car, clothing or books on school property?					
	0 times	0.21	0.59	0.59	0.86
	1 time	0.23	0.25	0.27	0.11
	2 or 3 times	0.31	0.14	0.12	0.03
	4 or 5 times	0.09	0.02	0.01	0.00
	6 or more times	0.15	0.01	0.01	0.00
has anyone offered, sold, or given you an illegal drug on school property?					
	No	0.28	0.73	0.67	0.91
	Yes	0.72	0.27	0.35	0.09

Table 3

Fit Statistics and Criteria by Number of Latent Classes for Level-2 Model Specification

	Number of Level-2 Classes		
Model	1 Class	2 Classes	3 Classes
Random effects model, nonparametric			
No. of free parameters	47	51	55
Log-likelihood	-101083	-100848	-100848
AIC	202260	201797	201806
AIC3	202307	201849	201861
BIC	202656	202227	202269
Sample adjusted BIC	202506	202065	202094
Entropy	0.77	0.78	0.78
Random effects model, nonparametric with Level-2 common factor for Level-1 indicators for chosen model			
No. of free parameters		60	
Log-likelihood		-100559	
AIC		201239	
AIC3		201299	
BIC		201744	
Sample adjusted BIC		201553	
Entropy		0.78	

Table 4
MLCA Odd Ratio Results for Student and School Predictors

	Comparison of Bullying Classes								
	High versus Moderate-Physical			Moderate-Sexual versus Moderate-Physical			Low versus Moderate-Physical		
		95% CI			95% CI			95% CI	
	Odds	-	+	Odds	-	+	Odds	-	+
Level 1 Predictors									
SES	0.71	0.45	0.92	0.96	0.78	1.12	1.18	1.04	1.30
Gender	0.54	0.39	0.75	6.74	3.46	13.12	6.92	6.15	7.78
American Indian	0.62	0.33	1.16	0.88	0.47	1.65	1.30	0.71	2.38
African American	0.96	0.73	1.26	1.12	0.78	1.62	1.06	0.75	1.49
Hispanic	0.98	0.73	1.30	1.10	0.84	1.43	1.07	0.82	1.39
Asian American	0.69	0.49	0.99	1.05	0.78	1.41	1.90	1.47	2.44
Other/Multiple	1.44	1.14	1.81	1.03	0.82	1.29	0.84	0.67	1.06
Age	0.99	0.88	1.10	1.00	0.92	1.09	1.15	1.06	1.25
Mental Distress	1.19	1.13	1.26	1.10	1.05	1.16	0.70	0.66	0.74
Family Support	0.96	0.91	1.01	0.98	0.93	1.02	1.02	0.98	1.06
Family Source of Information	1.03	0.99	1.07	1.01	0.98	1.04	0.99	0.96	1.01
Problems With Family	1.35	1.28	1.42	0.99	0.94	1.04	0.74	0.71	0.78
School Safety/Climate	0.72	0.69	0.76	0.95	0.91	0.99	1.23	1.18	1.28
Teacher/Community Support	0.92	0.85	0.98	1.10	1.02	1.18	1.23	1.17	1.30
GPA	0.98	0.91	1.06	1.26	1.15	1.38	1.24	1.13	1.37sc
Level 2 Predictors									
Attendance	0.99	0.98	1.01	1.00	0.97	1.03	1.01	0.99	1.04
City	0.62	0.51	0.75	1.01	0.70	1.44	2.20	1.68	2.88
Town	0.73	0.65	0.82	1.01	0.80	1.26	1.68	1.42	1.99
Rural	0.84	0.71	0.99	1.00	0.74	1.37	1.34	1.06	1.69
Low SES	--	--	--	--	--	--	--	--	--
Medium SES	1.21	1.08	1.37	1.00	0.80	1.24	0.73	0.62	0.86

Table 5
MLCA Odd Ratio Results for Level 1 Latent Class Indicator Intercepts

	School Predictor								
	Rural versus Suburban			Town versus Suburban			Medium SES versus High SES		
		95% CI			95% CI			95% CI	
	Odds	-	+	Odds	-	+	Odds	-	+
… has a student … kicked, bitten, or hit you?	1.94	1.34	2.80	1.37	1.07	1.75	1.17	1.02	1.35
… has a student … touched, grabbed or pinched you in a sexual way?	1.36	1.07	1.73	1.16	1.00	1.33	1.08	1.00	1.16
… has anyone offered, sold, or given you an illegal drug on school property?	0.55	0.37	0.82	0.76	0.60	0.95	--	--	--

References

Adcock, E. P. & Phillips, G. W. (1997). Measuring school effects with hierarchical linear modeling: Data Handling and Modeling Issues. *Multiple Linear Regression Viewpoints, 24*(1), 1-10. http://mlrv.ua.edu/1997/V24_N1_A7.pdf.

Arceneaux, K., & Nickerson, D. W. (2009). Modeling certainty with clustered data: A comparison of methods. *Political Analysis, 17*(2), 177-190. doi: 10.1093/pan/mpp004

Brown, J. R., Aalsma, M. C., & Ott, M. A. (2013). The experiences of parents who report youth bullying victimization to school officials. *Journal of Interpersonal Violence, 28*(3), 494-518. doi: 10.1177/0886260512455513

Cronbach, L. J. (1976). Research on classrooms and schools: Formulation of questions, design and analysis. Occasional paper of the Stanford Evaluation Consortium, Stanford University.

de Leeuw, J., & Kreft, I. G. G. (1995). Hierarchical linear models: Problems and prospects. *Journal of Educational and Behavioral Statistics, 20*(2), 171 – 189.

Dornfeld-Januzzi, J. (2006). Adult perceptions of bullying by boys and girls in middle school (PhD dissertation). Fordham University, NY.

Dunn, K. E., Airola, D. T., Lo, W. J., & Garrison, M. (2013). What teachers think about what they can do with data: Development and validation of the data driven decision-making efficacy and anxiety inventory. *Contemporary Educational Psychology, 38*(1), 87-98. doi:10.1016/j.cedpsych.2012.11.002

Hagenaars, J. A., & McCutcheon, A. L. (Eds.). (2002). *Applied latent class analysis*. Cambridge University Press.

Henry, K. L., & Muthén, B. O. (2010). Multilevel latent class analysis: An application of adolescent smoking typologies with individual and contextual predictors. *Structural Equation Modeling: A Multidisciplinary Journal, 17*(2), 193 - 215. doi: 10.1080/10705511003659342

Lukociene, O., Varriale, R. and Vermunt, J. K. 2010. The simultaneous decision(s) about the number of lower- and higher-level classes in multilevel latent class analysis. *Sociological Methodology, 40*, 247–283.

Lukociene, O. and Vermunt, J. K. 2010. "Determining the number of components in mixture models for hierarchical data". In *Advances in Data Analysis, Data Handling and Business Intelligence*, Eds.: Fink, A., Lausen, B., Seidel, W. and Ultsch, A. 241–250. Berlin: Springer.

MacLeod, I. R. (2007). A study of administrative policy responses to bullying in Illinois secondary schools (PhD dissertation). Loyola University, Chicago

McCoach, D. B. (2010). Dealing with dependence (Part II): A gentle introduction to hierarchical linear modeling. Gifted Child Quarterly, 54(3), 252-256.

McCoach, D. B., & Adelson, J. L. (2010). Dealing With Dependence (Part I): Understanding the Effects of Clustered Data. Gifted Child Quarterly, 54(2), 152-155.

Marsh, H. W., Ludtke, O., Nagengast, B., Trautwein, U., Morin, A. J. S., Abduljabbar, A. S., & Koller, O. (2012). Classroom climate and contextual effects: Conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist*, *47*, 106-124.

Means, B., Padilla, C., & Gallagher, L. (2010). Use of education data at the local level from accountability to instructional improvement. *U.S. Department of Education Office of Planning, Evaluation, and Policy Development.*

Messelt, J. (2004). Data-driven decision making: A powerful tool for school improvement [White paper]. Retrieved from: https://www.erdc.k12.mn.us/promo/sage/images/Analytics_WhitePaper.pdf

Morgan, H. (2012). What teachers and schools can do to control the growing problem of school bullying. *The Clearing House: A Journal of Educational Strategies, Issues, and Ideas* , *85*(5), 174-178. doi: 10.1080/00098655.2012.677075

Muthén, L. K., & Muthén, B. O. (2012). *Mplus statistical analysis with latent variables. User's guide*. (7th ed.) Los Angeles, CA: Muthén & Muthén.

Mutz, R., Bornmann, L., & Daniel, H. D. (2013). Types of research output profiles: A multilevel latent class analysis of the Austrian Science Fund's final project report data. *Research Evaluation*, *22*(2), 118-133.

National Society for the Prevention of Cruelty to Children. (2014). *Sexual bullying in schools: An NSPCC factsheet.* Retrieved from: http://www.nspcc.org.uk/Inform/research/questions/sexual_bullying_wda70106.html

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods.* (2nd ed.) Thousand Oaks, CA.: SAGE Publications, Inc.

Raykov, T. (2009). Estimation of maximal reliability for multiple-component instruments in multilevel designs. *British Journal of Mathematical and Statistical Psychology*, *62*(1), 129-142.

Raykov, T., & Penev, S. (2011). Evaluation of reliability coefficients for two-level models via latent variable analysis. *Structural Equation Modeling*, *17*(4), 629-641

Rodriguez, M. C., Matuska, K., Cabrera, J. C., & Karl, S. (2010). Improving the education of Minnesota's students from pre-K through college: Measuring student progress and using data to drive decision making. CURA Reporter, 40, 3-10.

U.S. Department of Health & Human Services (n.d.). *Bullying Definition.* Retrieved from http://www.stopbullying.gov/what-is-bullying/definition/index.html

Zhang, J., Zhang, M., Zhang, W., & Jiao, C. (2014). Model Selection for Complex Multilevel Latent Class Model. *Communications in Statistics - Simulation and Computation*, *43*(4), 838 – 850, DOI: 10.1080/03610918.2012.718836

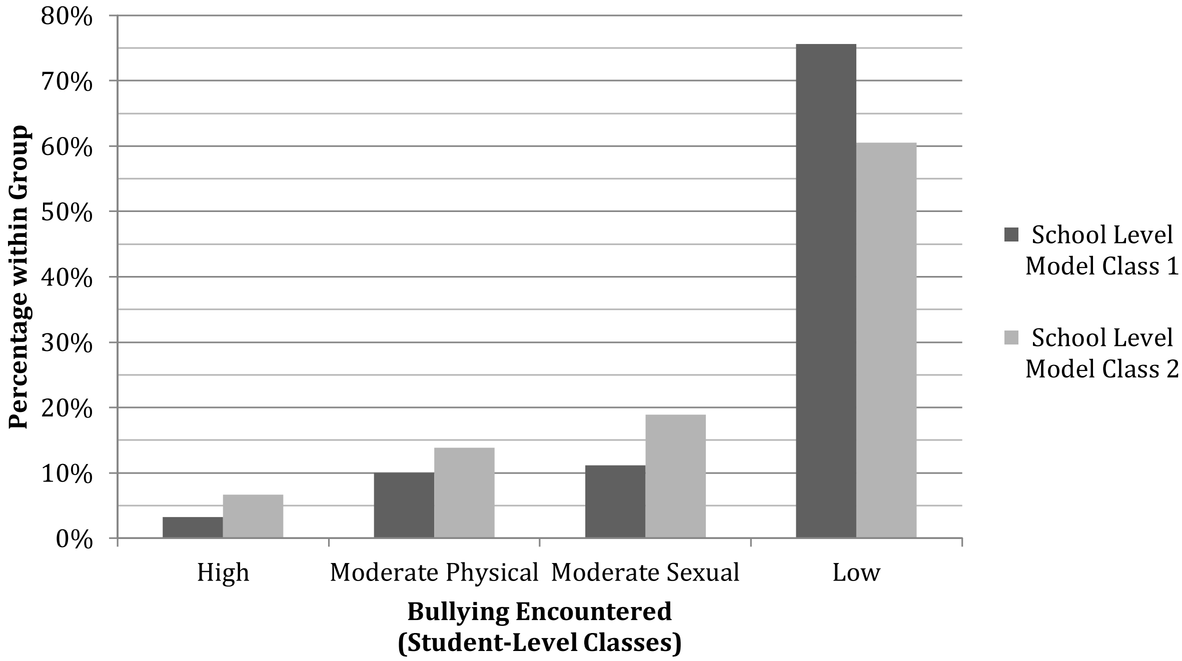

Figure 1. *Percentage of Students in Student-Level Latent Classes by School-Level Latent Classes*

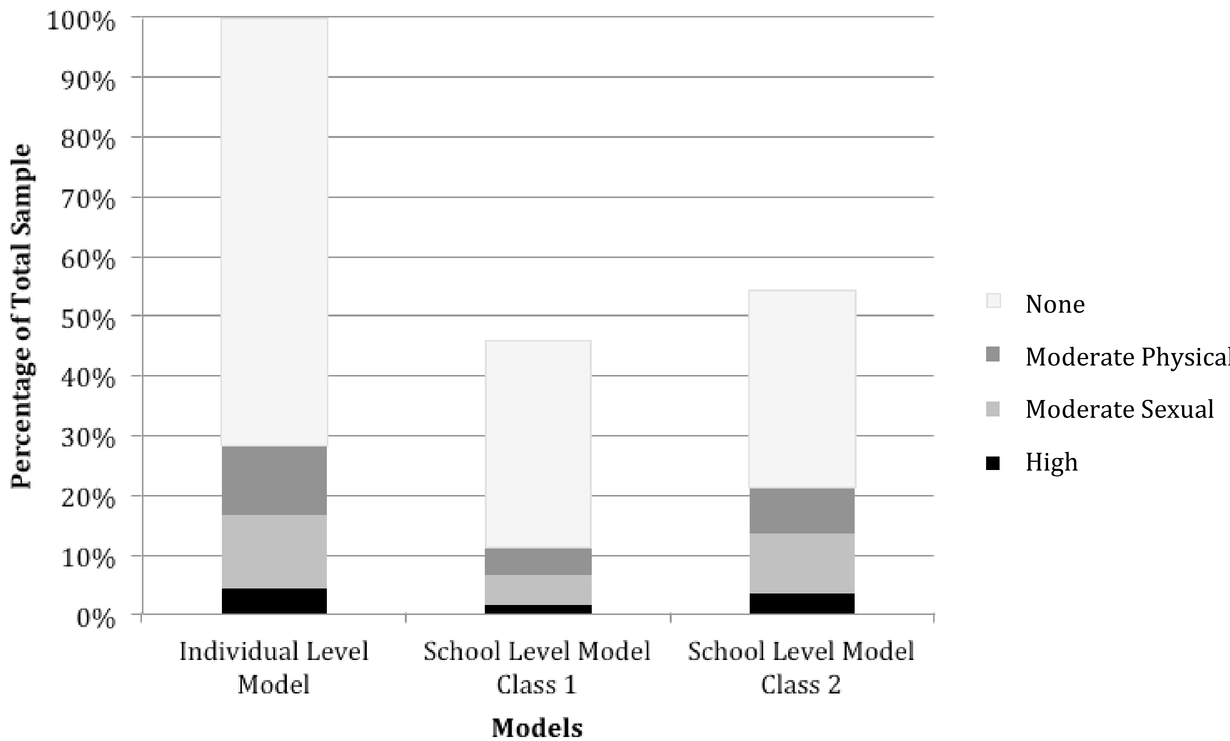

Figure 2. *Comparisons of the Structure of Level-1 Latent Classes between a Student-level model (Individual Level Model) ignoring data dependency versus a Multilevel Model (School Level Model) with a Common Factor on Student-Level Latent Class Indicators.*

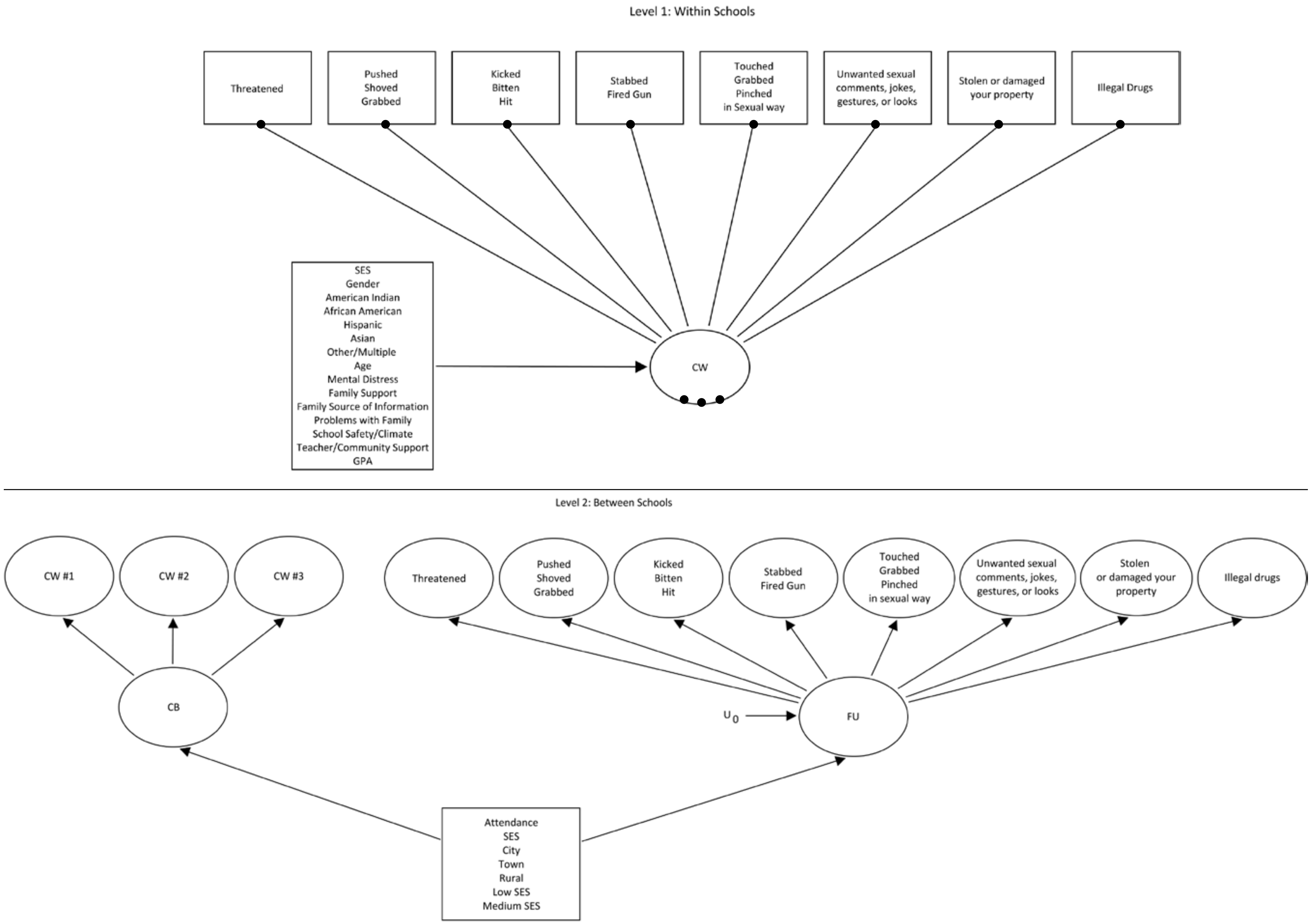

Figure 3. *Nonparametric Random Effects Multilevel Latent Class Analysis Model with Four Student-level Latent Class and Two School-level Latent Class model with a School-level Common Factor for Student-level Indicators*