Gifted Child Quarterly
54(3) 252–256
© 2010 National Association for Gifted Children
Reprints and permission: sagepub.com/journalsPermissions.nav
DOI: 10.1177/0016986210373475
http://gcq.sagepub.com

Dealing With Dependence (Part II): A Gentle Introduction to Hierarchical Linear Modeling

D. Betsy McCoach[1]

Abstract

In education, most naturally occurring data are clustered within contexts. Students are clustered within classrooms, classrooms are clustered within schools, and schools are clustered within districts. When people are clustered within naturally occurring organizational units such as schools, classrooms, or districts, the responses of people from the same cluster are likely to exhibit some degree of relatedness with each other. The use of hierarchical linear modeling allows researchers to adjust for and model this non-independence. Furthermore, it may be of great substantive interest to try to understand the degree to which people from the same cluster are similar to each other and then to try to identify variables that help us to understand differences both within and across clusters. In HLM, we endeavor to understand and explain between- and within-cluster variability of an outcome variable of interest. We can also use predictors at both the individual level (level 1), and the contextual level (level 2) to explain the variance in the dependent variable. This article presents a simple example using a real data set and walk through the interpretation of a simple hierarchical linear model to illustrate the utility of the technique.

Keywords

clustered data, hierarchical linear modeling, gifted education research

In this Methodological Brief, Dr. Betsy McCoach (University of Connecticut) continues an exploration of hierarchical linear modeling (HLM; see *Gifted Child Quarterly, 54*(2), 152-155) by providing a simple example of the use of HLM with nested data. This is the last of a two-part series that focused on HLM.

Tonya R. Moon, PhD
GCQ Methodological Briefs, Editor
Associate Editor, GCQ

In education, most naturally occurring data are clustered within contexts. Students are clustered within classrooms, classrooms are clustered within schools, and schools are clustered within districts. As mentioned in the previous Methodological Brief (McCoach & Adelson, 2010), most traditional statistical analyses assume that observations are independent of each other. When people are clustered within naturally occurring organizational units such as schools, classrooms, or districts, the responses of people from the same cluster are likely to exhibit some degree of relatedness with each other. In other words, they are not independent of each other. The degree of relationship among units from the same cluster is captured by the intraclass correlation coefficient (ICC; McCoach & Adelson, 2010). The use of hierarchical linear modeling (HLM) allows researchers to adjust for and model this nonindependence.

The advantages of HLM are not purely statistical. It may be of great substantive interest to try to understand the degree to which people from the same cluster are similar to each other and then to try to identify variables that help us understand differences both within and across clusters. In HLM, information is used from cluster samples to understand and explain between- and within-cluster variability of an outcome variable of interest. Predictors at both the individual level (Level 1) and the contextual level (Level 2) can be used to explain the variance in the dependent variable.

In traditional regression-based approaches, the relationship between two variables is estimated. Generally, it is assumed that this relationship is constant across the entire sample. However, when data are gathered from different contexts or clusters, it is plausible that the relationships among key variables of interest may vary by cluster or context. By using HLM, the relationship between an independent variable and the dependent variable can randomly vary across clusters. If the impact of the independent variable on the dependent variable varies across clusters, the variability in this

[1]University of Connecticut, Storrs, CT, USA

Corresponding Author:
D. Betsy McCoach, Educational Psychology Department, University of Connecticut, 249 Glenbrook Road, Unit 2064, Storrs, CT 06269-2064, USA
Email: betsy.mccoach@uconn.edu

Downloaded from gcq.sagepub.com at Serials Records, University of Minnesota Libraries on July 14, 2010

relationship using cluster-level variables can potentially be explained. For example, allowing the relationship between students' IQ and achievement to vary by class it might be found that the impact of IQ on achievement does vary by class. In this case, the variability might be explained by using class-level predictors, such as teacher experience or the teacher's use of differentiation strategies. If a Level-2 variable, such as teacher's use of differentiation strategies, does moderate the relationship between a Level-1 variable (IQ) and the dependent variable (achievement), this is called a *cross-level interaction.* Using HLM allows one to simultaneously model the impact of both individual (or Level-1) and contextual (or Level-2) variables on the dependent variable of interest, as well as model the cross-level interactions between higher level and lower level variables on the outcome of interest.

A simple HLM example using a real data set follows with the interpretation to illustrate the utility of the technique. For a more complete treatment of the analysis and interpretation of hierarchical data, the interested reader should consult Raudenbush and Bryk (2002), Snijders and Bosker (1999), or O'Connell and McCoach (2008).

Example of HLM Analysis

To illustrate the use of multilevel modeling, reading data collected as part of a federally funded Javits grant are used (Reis et al., 2005). The purpose of this analysis is to demonstrate a simple hierarchical linear model. Thus, the number of covariates at each level is small. Therefore, the reader is cautioned that the analysis presented here is not necessarily the most complete analysis of these data and is only used for illustration of the HLM technique.

For this analysis, there are 1,192 elementary school students nested within 70 classrooms. The dependent variable is the post–Iowa Test of Basic Skills (ITBS) reading comprehension score. The effects of two student-level variables, gifted status and pretest ITBS reading comprehension score, on students' posttest reading comprehension scores are considered. Gifted status is a dichotomous variable and is coded "1" for identified gifted students and "0" for all other students. The pretest score is the student's score on the ITBS reading comprehension subtest and is a continuous variable. In addition, the impact of two classroom-level variables, the percentage of gifted students in the classroom and the classroom mean pretest score, on predicted reading comprehension scores is examined. The effects of the two classroom-level variables on the relationship between pretest reading comprehension and posttest reading comprehension are also investigated. These represent cross-level interaction effects: They capture the effects of classroom-level variables on the effects of individual variables on the dependent variable. Although it is common to refer to "effects" in the multilevel literature, the "effects" that are referred to are correlational, not causal in nature, as the independent variables examined were not experimentally manipulated.

Model 1: Unconditional (Random Effects Analysis of Variance) Model

The process begins by estimating an unconditional model to determine the ICC. The unconditional model is

$$Y_{ij} = \beta_{0j} + r_{ij},$$
$$\beta_{0j} = \gamma_{00} + u_{0j}.$$

In this simplest, unconditional model, each person's score on the dependent variable consists of three elements: the overall mean (γ_{00}), the deviation of the cluster mean from the overall mean (u_{0j}), and the deviation of the person's score from his or her cluster mean (r_{ij}). The u_0 term allows the dependence of observations from the same cluster to be modeled because u_{0j} is the same for every observation within cluster j (Raudenbush & Bryk, 2002). In other words, every student in the same school will have the same value for u_{0j}. The u_0 term is referred to as a *random effect* for the intercept because of the assumption that the value of u_0 randomly varies across the Level-2 units (clusters) with a mean of 0 and a variance of τ_{00}.

Using this simple random-effects analysis of variance (ANOVA) model, the ICC is the ratio of the between class variance to the total variance. In this example, the between class variance in reading comprehension (τ_{00}) is 558.46. The within class variance (σ^2) in reading comprehension is 578.62. Therefore, the total variance in reading comprehension is $\tau_{00} + \sigma^2$, or 1137.08 (corresponding to a standard deviation of 33.72). Therefore, the intraclass correlation is the between-cluster variance divided by the total variance, $\tau_{00}/(\tau_{00} + \sigma^2)$, which is 0.491, meaning 49% of the variability in reading comprehension scores is accounted for by the cluster (classroom). Given this large ICC, it would be incorrect to treat these data as if they were independent. The parameter estimate for the intercept of this simple model, γ_{00}, is 209.62 (Table 1). This represents the mean posttest reading comprehension score.

Model 2: Level-1 Model

Next, a model that includes pretest reading comprehension scores as a control variable and gifted status as an additional Level-1 variable is estimated. To make the intercept of the multilevel equation interpretable, the pretest score is grand-mean centered.[1] However, gifted status is left in its raw metric. The set of equations for this Level-1 model is

$$Y_{ij} = \beta_{0j} + \beta_{1j}(GIFTED)_{ij} + \beta_{2j}(PRE\text{-}ITBS)_{ij} + r_{ij},$$
$$\beta_{0j} = \gamma_{00} + u_{0j},$$
$$\beta_{1j} = \gamma_{10},$$
$$\beta_{2j} = \gamma_{20} + u_{2j}.$$

Downloaded from gcq.sagepub.com at Serials Records, University of Minnesota Libraries on July 14, 2010

Table 1. Summary of REML Parameter Estimates for Two-Level Model of Reading Comprehension

Parameter	Unconditional Model		Level-1 Model		Full Level-2 Model	
	Parameter Estimate	*SE*	Parameter Estimate	*SE*	Parameter Estimate	*SE*
Fixed effects						
Intercept (γ_{00})	209.62*	2.92	207.60*	2.03	207.04*	1.60
Percentage gifted (γ_{01})	—	—	—	—	–0.10*	0.04
Class mean ITBS (γ_{02})	—	—	—	—	0.29*	0.06
Gifted status (γ_{10})	—	—	12.58*	1.79	13.88*	2.90
Pre-ITBS (γ_{20})	—	—	0.70*	0.03	0.65*	0.03
Percentage gifted (γ_{01})	—	—	—	—	0.002	0.001
Class mean ITBS (γ_{02})	—	—	—	—	–0.003	0.001
Variance estimates						
Within-class variance (σ^2)	578.62	—	308.24	—	309.95	—
Intercept variance (τ_{00})	558.46	—	41.13	—	24.95	—
Pre-ITBS slope variance (τ_{11})	—	—	0.021	—	0.014	—

Note. REML = restricted maximum likelihood; *SE* = standard error; ITBS = Iowa Test of Basic Skills.
*$p < .05$

In this model, γ_{00} (207.60) represents the predicted posttest reading comprehension score for a nongifted student when pretest is held constant at the mean. γ_{10}, the gifted identification slope (12.58; see Table 1), represents the differential between gifted and nongifted students on the posttest, after controlling for pretest scores. This indicates that after controlling for pretest scores, gifted students are expected to score 12.58 points higher than nongifted students on the reading comprehension posttest. γ_{20}, the pretest slope, is 0.70. This represents the expected change in reading comprehension achievement per unit change in pretest scores. In other words, for every point increase on the reading comprehension pretest, it is anticipated that students' scores would be 0.70 points higher on the posttest.

Percentage of variance explained. As might be expected, adding pretest and gifted status as student-level variables helps to explain a tremendous amount of both the between classroom variance and the within classroom variance in posttest reading achievement. Because the total variance in the dependent variable is partitioned into multiple pieces across the multiple levels of analysis, there is no real multilevel equivalent to R^2, or the proportion of variance explained in a single-level regression model. However, Raudenbush and Bryk (2002) have proposed a *proportional reduction in variance* statistic that can be computed for any variance component in the multilevel model. To compute this proportional reduction in variance statistic, the residual variance from the model with a greater number of variables is compared with the residual variance from a simpler baseline model. First, the variance of the larger model is subtracted from the variance of the simpler baseline model. The difference is then divided by the variance of the baseline model. This proportional reduction in variance provides a rough estimate of the proportion of variance that is explained using a set of independent variables (Raudenbush & Bryk, 2002).

To illustrate the proportional reduction in variance statistics, a comparison of the within-class variance in reading comprehension achievement and the between-class variance in reading comprehension achievement is done using the formulas provided by Raudenbush and Bryk (2002). The between-class variance for the unconditional model is 558.46. The between-class variance for the Level-1 model is 41.13. Thus the proportional reduction in Level-2 variance is (558.46 – 41.13)/558.46, or 0.926. In other words, adding pretest score and gifted status as student level predictors actually helps explain more than 90% of the between-class variance in students' reading posttest scores. The within-class variance for the unconditional model is 578.62. The within-class variance for the Level-1 model is 308.24. Thus, the proportional reduction in Level-1 variance is (578.62 – 308.24)/578.62, or 0.467. Therefore, adding the pretest and gifted identification variables also helps explain close to 50% of the within-class variance in reading comprehension achievement.

Model 3: Full Level-2 Model

Finally, a full two-level model is estimated, in which the percentage of identified gifted students and the classroom's mean pretest scores are included as predictors of both the intercept and the pretest slope. The set of equations for this full model is

$$Y_{ij} = \beta_{0j} + \beta_{1j}(GIFTED)_{ij} + \beta_{2j}(PRE\text{-}ITBS)_{ij} + r_{ij},$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\%GIFTED) + \gamma_{02}(MEAN\ PRE\text{-}ITBS) + u_{0j},$$

$$\beta_{1j} = \gamma_{10},$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}(\%GIFTED) + \gamma_{22}(MEAN\ PRE\text{-}ITBS) + u_{2j}.$$

Downloaded from gcq.sagepub.com at Serials Records, University of Minnesota Libraries on July 14, 2010

Both classroom-level variables are centered on their respective means to aid interpretation. Therefore, in this model, γ_{00}, the overall intercept, now represents the predicted posttest reading comprehension score for a nongifted student who scores at the mean on the pretest and who is in a class that scores at the class mean at the pretest and has an average number of identified gifted/talented students. Next, we consider the effects of the Level-2 variables on the intercept. γ_{01} represents the effect of increasing the percentage of gifted students in the classroom on that predicted reading comprehension score, after controlling for the class mean pretest scores. This coefficient is negative, indicating that holding classroom pretest scores constant at the mean, for every point increase in the percentage of gifted students in the class, the predicted posttest reading comprehension achievement level in the class would decrease by 0.10. This may seem odd; however, γ_{02} the effect of the class mean pretest score on the intercept is positive, indicating that for every point increase in the class's mean at pretest, the expected value of the reading posttest score increases by 0.29. It is important to remember that the percentage of gifted students in the class and the class mean pretest reading comprehension score are likely to be positively correlated. Thus, as the percentage of gifted students in the class increases, the average pretest reading comprehension score is also likely to increase. If the percentage of gifted students in the class increases, without a resulting increase in reading comprehension scores, then this model would actually predict a decrease in the predicted posttest reading comprehension achievement of the class, but this scenario is highly unlikely. The same-level interaction between class mean pretest reading comprehension scores and the percentage of gifted students in the class could also be considered. However, this interaction was tested and it was not statistically significant ($p > .05$).

Next, the Level-1 slopes for gifted status and pretest reading comprehension scores are examined. These represent the effects of the Level-1 variables on the predicted dependent variable, after controlling for all of the other variables in the model. The effect of gifted status on the intercept (γ_{10}) is 13.88, indicating that after controlling for pretest scores and the classroom level variables, gifted students are expected to outperform nongifted students by 13.88 points. γ_{20}, the intercept of the pretest reading comprehension slope is 0.65 (Table 1). This indicates that in a classroom with an average number of gifted students and an average class-mean pretest score, for every point increase in the pretest, the expected value for the posttest reading comprehension score increases by 0.65 points. Finally, consideration is given to the cross-level interaction terms. These are the moderational effects of the Level-2 variables on the relationship between the Level-1 variables and the dependent variable. In other words, cross-level interaction effects answer the question, "How do Level-2 variables moderate the impact of the Level-1 variables on the dependent variable? After controlling for class mean pretest reading comprehension scores, the percentage of gifted students in the class (γ_{21}) has a small positive influence on this slope. That is, for every percentage point above the mean a class is in terms of the number of students who are gifted, the effect of the pretest on the posttest increases by 0.002 points. In other words, after controlling for class mean pretest scores, the more gifted students that there are in a class, the stronger the relationship between pretest score and the posttest score. This cross-level interaction effect indicates that having more gifted students in a class appears to strengthen the relationship between students' pretest and posttest scores. In contrast, the coefficient for the effect of the class mean pretest score on the effect of the pretest on posttest is negative ($\gamma_{22} = -0.003$). This indicates that holding the percentage of gifted students in the class constant, as the average pretest score of the class increases, the relationship between the pretest score and the posttest score decreases. Pretest scores are less predictive of posttest scores in classrooms where the class average reading comprehension score is higher (again, after controlling for or holding constant the percentage of gifted students in the class). It is important to understand that this example is for illustrative purposes only. It may be that there are other omitted variables or unmodeled same-level interaction effects that could help better understand these data.

Percentage of variance explained. When compared with the model that includes only student level predictors, this final two-level model explains an additional 39% of the between class variance in the intercepts [(41.13 – 24.95)/41.13)] and it explains about 33% of the between class variance in the pretest slope [(0.021 – 0.014)/0.021)]. However, there is still additional between school variability remaining to be explained in both the intercept and pretest slope. There may be other classroom-level variables, such as the teacher's instructional style or classroom management skills that could help explain additional between class variance. Notice that the addition of the classroom-level variables does not help explain any additional within class variability in residual posttest reading scores. Conceptually, this makes sense as it is not expected that between-class variables would explain within-class variability.

Although there is much more that could be done with these analyses, it is hoped that this simple example provides the reader with the flavor of multilevel modeling and a sense of its ability to answer sophisticated research questions about the cross-level interaction effects of contextual variables on the dependent variable.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the authorship and/or publication of this article.

Funding

The data used for the example reported herein was supported under the Educational Research and Development Centers Program, PR/

Downloaded from gcq.sagepub.com at Serials Records, University of Minnesota Libraries on July 14, 2010

Award Number S206A040094, as administered by the Office of Elementary and Secondary Education, U.S. Department of Education. The findings and opinions expressed in this report do not reflect the position or policies of the Office of Elementary and Secondary Education or the U.S. Department of Education.

Note

1. A discussion of centering is beyond the scope of this article. However, the interested reader should consult Enders and Tofighi (2007) for an excellent discussion of the complex issues surrounding centering in multilevel models.

References

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*, 121-138.

McCoach, D. B., & Adelson, J. (2010). Dealing with dependence (Part I): Understanding the effects of clustered data. *Gifted Child Quarterly, 54*, 152-155.

O'Connell, A. A., & McCoach, D. B. (Eds.). (2008). *Multilevel modeling of educational data.* Charlotte, NC: Information Age Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage.

Reis, S. M., Eckert, R. D., Schreiber, F. J., Jacobs, J. K., Briggs, C., Gubbins, E. J., & Coyne, M. (2005). *The Schoolwide Enrichment Model–Reading study* (RM05214). Storrs: University of Connecticut, National Research Center on the Gifted and Talented.

Snijders, T., & Bosker, R. (1999). *Multilevel analysis.* Thousand Oaks, CA: Sage.

Bio

D. Betsy McCoach, PhD, is an associate professor in the Measurement, Evaluation and Assessment program at the University of Connecticut, where she teaches coursework in hierarchical linear modeling, structural equation modeling, instrument design, and research design. Her research interests include the underachievement of academically able students, growth curve modeling, and model fit issues. She has coedited the book *Multilevel Modeling of Educational Data,* and she is currently authoring a textbook on instrument design. She has published numerous peer review journal articles and book chapters in the areas of gifted education, research methodology, and educational research and currently serves as the research methodologist for several federally funded projects. She is the current coeditor of the *Journal of Advanced Academics*. She also serves on the editorial review boards for the *American Educational Research Journal*, the *Journal of Educational Psychology*, the *Journal of Educational Research*, and *Gifted Child Quarterly*. She is the 2007 recipient of the National Association for Gifted Children (NAGC) Early Scholar award and is currently the chair elect of NAGC's Research and Evaluation Network.

Downloaded from gcq.sagepub.com at Serials Records, University of Minnesota Libraries on July 14, 2010