



# INTRODUCTION TO MULTILEVEL MODELING

## OVERVIEW

---

Multilevel modeling goes back over half a century to when social scientists became attuned to the fact that what was true at the group level was not necessarily true at the individual level. The classic example was percentage literate and percentage African American. Using data aggregated to the state level for the 50 American states in the 1930s as units of observation, Robinson (1950) found that there was a strong correlation of race with illiteracy. When using individual-level data, however, the correlation was observed to be far lower. This difference arose from the fact that the Southern states, which had many African Americans, also had many illiterates of all races. Robinson described potential misinterpretations based on aggregated data, such as state-level data, as a type of “ecological fallacy” (see also Clancy, Berger, & Magliozzi, 2003). In the classic article by Davis, Spaeth, and Huson (1961), the same problem was put in terms of within-group versus between-group effects, corresponding to individual-level and group-level effects. A central function of multilevel modeling is to separate within-group individual effects from between-group aggregate effects.

Multilevel modeling (MLM) is appropriate whenever there is clustering of the outcome variable by a categorical variable such that error is no longer independent as required by ordinary least squares (OLS) regression but rather error is correlated. Clustering of level 1 outcome scores within levels formed by a level 2 clustering variable (e.g., employee ratings clustering by work unit) means that OLS estimates will be too high for some units and too low for others. As a corollary, errors of over- and underestimation will cluster by unit. It is this clustering of errors which is “correlated error.” Multilevel modeling is one of the leading approaches to dealing with correlated error.

While there are other methods of handling clustered data and correlated error, such as generalized estimating equations (GEE) or use of cluster-robust standard errors, multilevel modeling is the most ubiquitous approach (McNeish, Stapleton, & Silverman, 2017). If OLS regression is used for data with multilevel effects, the model is misspecified and parameter estimates (b coefficients, for example) will be in error, leading to possible significant errors of model interpretation. Put another way, the researcher contemplating a regression model of the usual monolevel type should first check to see if MLM modeling is needed instead.

It is common to refer to the outcome variable data (e.g., student scores) as being at level 1 and the clustering variable (e.g., classrooms) as being at level 2 (and there may be yet higher levels,

### CLUSTERING DEFINED

As used here, clustering of data occurs whenever measures of an outcome variable are conditional in part or whole on values of a categorical variable. For instance, student test scores may be conditional on the student's school. Students at better schools may tend to score higher. Thus higher scores may be concentrated in better schools. This concentration is "clustering." When clustering is ignored, estimates of student scores will be too low for better schools and too high for worse schools, on average. Therefore errors will tend to be positive values for better schools, reflecting underestimates (where error = observed minus estimated values) and will tend to be negative values for worse schools, reflecting overestimates. In this way, error clusters by school, meaning there is "correlated error."

Multilevel modeling handles the problem of correlated error due to the existence of a clustering variable. In multilevel analysis, typically the DV (test scores) is considered level 1 and the clustering variable (schools) is level 2. There may be higher levels of clustering (e.g., school district), calling for three-level or higher models. The clustering variable is also called the grouping variable, the level variable, or the link variable. It is also possible that the DV will cluster on other categorical variables not serving as grouping variables. This clustering over and above that associated with grouping variables is addressed in part through use of cluster-robust standard errors, discussed in Chapter 2.

like level 3 = schools). The notion of levels is contained in the term *hierarchical linear modeling* (HLM), often used interchangeably with multilevel modeling (MLM). In this book, however, we prefer MLM terminology for three reasons.

1. Not all multilevel models are strictly hierarchical, as we shall see when cross-classified models are discussed.
2. The clustering variable need not be one which is organizationally at a higher level as classrooms are with respect to students. Rather, potentially any categorical variable may be a clustering variable (e.g., nation of origin, which is not normally discussed in terms of "higher level" in the organization chart sense).
3. There is ambiguity about the term *level* in HLM terminology, where level may refer to the clustering variable (e.g., religious affiliation) but values of that variable (e.g., Catholic, Baptist, Orthodox Jewish) are also called the levels of that variable.

Linear mixed modeling (LMM) is another synonym for multilevel modeling. In LMM, *mixed* refers to a model having both fixed and random effects. In LMM, random effects are the effects of clustering of the dependent variable (DV) within categorical levels of a clustering variable. Fixed effects are those in the level 1 regression model, just as conventional OLS regression models are fixed effects models. SPSS and certain other statistical packages implement MLM using LMM terminology. In this book, though we certainly discuss fixed and random effects (see Chapter 4), we prefer MLM as the general term for our models since we have found it to be more intuitive for the introductory graduate-level students who are our target audience.

Numerous other synonyms for multilevel modeling exist. Depending on the discipline and research design, various types of multilevel model may be referred to as random intercept models, random coefficients models, random effects models, growth models, longitudinal models, and covariance components models, to name a few. These and other terms are treated in this book as they arise.

The reader may be given pause by the term *linear* being part of linear mixed modeling, which has just been described as equivalent to multilevel modeling. It is true that, by default, MLM models assume linear relationships, as do OLS regression models. Most of the examples in this book make this assumption. This is not an unreasonable assumption since a great many social science relationships are linear.

However, it is perfectly possible to extend MLM to nonlinear modeling. For instance, logistic regression may be substituted for OLS regression for a model in which the outcome variable is binary. Nonlinear MLM is called “generalized multilevel modeling” (GMLM). Synonyms include but are not limited to “generalized linear mixed modeling” (GLMM) and “generalized hierarchical linear modeling” (GHLM). Generalized modeling software allows OLS regression but also supports nonlinear forms of regression such as binary logistic regression for binary DVs, ordinal regression for ordinal DVs, multinomial regression for nominal DVs, gamma regression for skewed data, Poisson regression for count data, and other variants, to name a few. Generalized multilevel modeling is treated in Chapter 12 of this book.

## WHAT MULTILEVEL MODELING DOES

Multilevel models adjust estimates of the intercept (mean) of one or more dependent variables at level 1 based on grouping variables defining higher levels. Predictor variables at any level may also be incorporated in the model. Multilevel models may also adjust the slope (b coefficient) of one or more predictors (regressors) at lower levels. For instance, math score at the student level (level 1) may be predicted by student socioeconomic status (also level 1) as influenced by school (the level 2 clustering variable). Predicted mean math scores and/or the estimated slope (b coefficient) for socioeconomic status will be adjusted by MLM algorithms for the clustering of math scores at the school level.

As discussed above, the multiple levels of multilevel models are defined by categorical (factor) variables, such as classroomid for classrooms as the variable defining groups at level 2. Such a factor variable may be called the clustering variable, the grouping variable, the level variable, or the link variable. While the clustering variable often is a function of sampling (voters sampled within census tracts, census tracts sampled within voting districts, etc.), the clustering variable may be any categorical variable, the presence of which causes data and data residuals to be nonindependent, thereby violating a basic assumption of ordinary regression. For example, when math scores cluster by school because student scores within a given school are more similar than otherwise would be expected, both student scores and error terms associated with their predicted scores are not independent. Independence of error terms is an assumption of OLS regression and related techniques.

When a clustering variable is present and has an effect, lack of independence means that standard errors computed in ordinary regression will be in error. This in turn means that significance tests will be in error. Specifically, standard errors will be too low, resulting in the Type I error of spuriously “significant” effects (Maas & Hox, 2004, p. 428).

When a clustering variable is suspected, the researcher will want to compute significance for the variance component for the effect of the clustering variable in the null model. This

is mathematically equivalent to computing the significance of the intraclass correlation coefficient (ICC). Variance components and the ICC are discussed in Chapter 3 and later in this book. If that variance component or ICC is significant, then nonindependence is present to a significant degree and some form of multilevel modeling should be undertaken. Clustering is very common in social science data and therefore multilevel modeling has wide application.

It should be noted that multilevel modeling can change substantive conclusions in social science research. This point is illustrated by Raudenbush and Bryk (2002), pioneers in multilevel modeling, who studied elementary school students' math scores. They wrote of the difference made by multilevel modeling that

The results were startling—83% of the variance in growth rates was between schools. In contrast, only about 14% of the variance in initial status was between schools, which is consistent with results typically encountered in cross-sectional studies of school effects. This analysis identified substantial differences among schools that conventional models would not have detected because such analyses do not allow for the partitioning of learning-rate variance into within- and between-school components. (pp. 9–10)

## THE IMPORTANCE OF MULTILEVEL THEORY

As with other quantitative techniques, the researcher should not arrive at study conclusions on a data-driven basis. Rather, in the case of multilevel problems, the researcher should have a multilevel theory about the relationships among the variables in the study. There should be some plausible line of reasoning, supported by the literature, induction from examples, or deduction from principles, that links the upper level in a multilevel model to the lower level. For example, it is not enough to show that there is a significant school effect on students' math scores. There should also be some argument about how this effect comes about. For example, it may be reasoned that schools with higher budgets pay teachers more, motivating them to elicit better student performance. Or it may be reasoned that schools within which the average socioeconomic status (SES) of pupils is higher get better test results because the pupils have higher SES peer learning. There may be many multilevel theories, each of which will suggest additional variables which might be included in the model. In general, multilevel models require theories about the mediators between levels.

An “ecological fallacy” is thinking that because something is true at one level, it must be true at another level. A previously mentioned example is thinking that the correlation of race and illiteracy in aggregated state-level data means that race and illiteracy are correlated in the same way at the individual level. They are not. The higher correlation with state aggregate data is because states high on percentage of minority race are also high on illiteracy for all races. The implication for multilevel modeling is that level 1, which is the observation level, must be the level at which the actual causation is thought to occur. For the (false) theory that race causes illiteracy based on state aggregate data, we cannot argue that individual-level illiteracy is caused by state-level racial proportions.

In general, the researcher must have sound theory justifying whether a construct is to be measured at the individual level, the cluster level, or both. Stapleton, Yang, and Hancock (2016) have noted the pitfalls of construct validation in a multilevel context, such as the pitfall of spurious construct-irrelevant dependency and other spurious effects which may occur with hierarchical data, illustrating these through simple simulations. For instance, if the researcher

posits that a construct exists only at the individual level (e.g., student scores) but it is found that dependency exists at the cluster level (e.g., school level), then the researcher must show either that there are cluster-level causal mechanisms with a transfer effect to the individual level (e.g., higher paid, higher ability, and/or more motivating teachers affecting class performance) or that there are no real transfer effects, only spurious ones.

## TYPES OF MULTILEVEL DATA

In broad terms, multilevel modeling is applicable to five types of data:

1. *Hierarchical data:* As discussed above, multilevel modeling applies when data are organized hierarchically. This was the example of math test scores given over time and nested within student id; student id nested within classrooms; classrooms nested within schools, and so on. With hierarchical data, there may be variables at each level: test scores at the test level, student attributes at the student id level, classroom attributes such as teacher experience at the classroom level, school budget per pupil at the school level, and so on. Only categorical variables above level 1 (the test level in this example) may be the clustering variables (student id, classroom id, school id, etc.). Multilevel modeling will show how clustering variables and other variables at higher hierarchical levels affect the dependent variable at level 1 (e.g., math scores).

2. *Repeated measures data:* Repeated measures may be seen as a special case of hierarchical data. The repeated measures (e.g., math scores) become level 1 while the unit of analysis (student id) becomes level 2. Level 1 data need not be repeated measures as, for instance, would be the case of a cross-sectional study of budget variables at level 1 nested within the 50 American states at level 2. The difference is that with repeated measures, each unit of analysis (e.g., student id) has multiple rows of data—one row for each test administration.

3. *Random effects data:* Random effects data are another special case of hierarchical data. An example is a marketing study in which consumer attitudes at level 1 are nested within product brands at level 2, where the researcher is interested in whether the brand effect is significant. The researcher can generalize the findings from multilevel modeling under one of two conditions: (1) Level 2 includes all brands (an “enumeration” or “census”), or (2) level 2 includes a random sample of brands (sometimes weakened in practice to be a “representative” sample). In condition (2), the level 2 effect is called a “random effect.” Unfortunately, it is common to call all level 2 effects “random effects” and software often labels these effects as such. However, if the units in level 2 are a convenience sample of unknown representativeness of the population of all such units, the random effect of the level 2 clustering variable should not be interpreted the same way as for a random sample. In the example, it is no longer “the brand effect” but rather *the effect of a particular set of brands, where the effect may be different for a different set of brands.*

4. *Cross-classified data:* In some cases, data are not nested in a strict hierarchy. An example is where students are nested hierarchically within neighborhoods (any student is in just one neighborhood) but neighborhoods are cross-classified within schools (a given neighborhood may send students to multiple schools; a given school may draw from multiple neighborhoods). Another example is where subjects are interviewed by an interviewer who may use any of multiple forms. Subjects are nested within interviewers (hierarchical) but one interviewer will use many forms and a given form may be used by many interviewers (cross-classified).

Cross-classified examples are numerous, so much so that some assert that cross-classified data are more common in social science than strictly hierarchically nested data. Multilevel modeling can handle cross-classified data, but it must use a different algorithm.

The researcher must declare the data to be cross-classified<sup>1</sup> and must use software which supports multilevel modeling for cross-classified data. Failure to do this can result in serious bias. For further reading on cross-classification in IBM® SPSS® Statistics, see Heck, Thomas, and Tabata (2010, Ch. 8); in SAS<sup>®2</sup>, see Patterson (2013) and Beretvas (2008); in Stata<sup>3</sup>, see Leckie (2013); in HLM 7, see Raudenbush, Bryk, Cheong, Congdon, and Du Toit (2011); and in R, see Roberts and Bates (2010). Cross-classification in all five statistical packages is covered in West, Welch, and Galecki (2015, pp. 369–394).

5. *Multiple outcome data:* Multilevel modeling can also support models in which there is more than one dependent variable (DV) at level 1. Again, this must be declared by the researcher who must use software supporting multivariate multilevel modeling (MMLM), also known as multivariate linear mixed modeling (MLMM) or hierarchical multivariate linear modeling (HMLM). For nonlinear models, there is also generalized multivariate multilevel modeling (GMMLM), also called multivariate generalized linear mixed modeling (MGLMM). See the discussion of generalized models in Chapter 12. One use of such models is multilevel latent outcome modeling, as when the DVs are several math tests at level 1 seen as indicators of a latent construct called *math IQ* at level 2. Another use has been in the analysis of dyadic data (see Knafl et al., 2009). Multivariate multilevel modeling is discussed further in Snijders and Bosker (1999, pp. 282–288) and in Brant and Sheng (2013). For coverage of multivariate multilevel analysis in SPSS, see Heck et al. (2010, Ch. 7); in SAS, see Wright (1998); and in HLM 7, see Raudenbush et al. (2011). See Baldwin, Imel, Braithwaite, and Atkins (2014), especially the statistical programming supplement by Atkins (2014), for coverage of multivariate multilevel modeling in Stata, SAS, R, SPSS, and MPlus.

## COMMON TYPES OF MULTILEVEL MODEL

There are many models possible with multilevel modeling and, unfortunately, an even larger number of labels for these models. In this section, we briefly describe the most common model types. We use the five arrows in Figure 1.1 to illustrate possible combinations of effects and their relationship to model types. To simplify here, we assume only two levels and a maximum of one predictor variable at either level (not counting the level variable), but in real research there may well be more predictor variables at either level and there may be more hierarchical levels or cross-classification of levels.

There are six common types of multilevel model discussed below. These are the unconditional random intercept (null) model, the conditional random intercept model, the random coefficients model, the random intercept regression model, the random intercept ANCOVA model, and the random coefficients ANCOVA model. We base this common labeling scheme largely on Raudenbush and Bryk (2002), but it should be noted that a variety of typologies and labels are used by different authors in describing types of multilevel model. Readers are advised, therefore, not to overemphasize model labels but instead to focus on underlying differences in the models. As a further complexity not discussed in this introductory section, all six models may come in the usual nested (hierarchical) flavor or in the cross-classified flavor and all six may come in the usual monivariate (one outcome) or multivariate (multiple outcomes) flavors.

### The Null or Unconditional Random Intercept Model

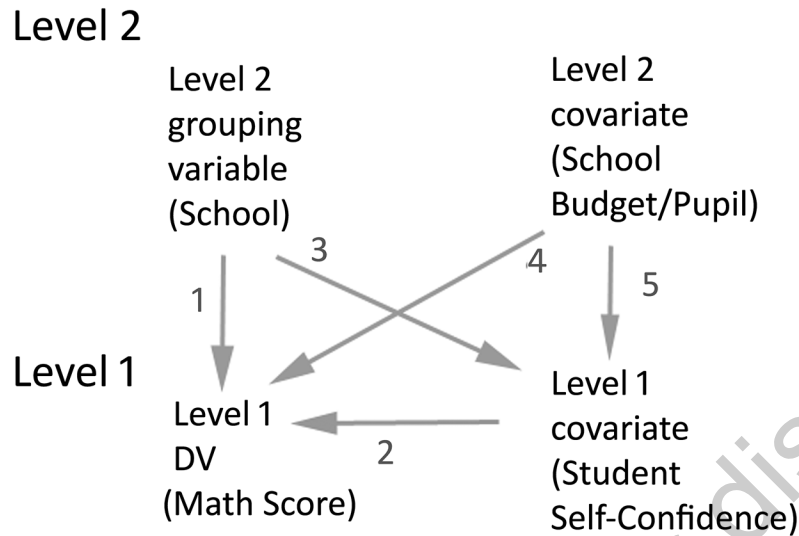
In multilevel modeling, the null model isn't one with just the intercept (constant) of the dependent variable, without any predictor variables, as is the case in OLS regression. Rather it is the model with only the grouping (clustering, level) variable as a determinant of the intercept of the dependent variable. That is, it is a model with just Arrow 1 in Figure 1.1. This is an

<sup>1</sup> SPSS is a registered trademark of International Business Machines Corporation.

<sup>2</sup> SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries.\*indicates USA registration.

<sup>3</sup> Stata is a registered trademarks of StataCorp LLC.

FIGURE 1.1 Common Types of Multilevel Effects



“unconditional” model in that there are no other predictor variables to condition the estimates. It is a “random intercept” model because it is predicting the level 1 intercept of the outcome variable and is not predicting any  $b$  coefficients at lower levels (there aren’t any). Note that all other models involve predictor variables and therefore are “conditional.”

The null model is often used to see if there is any need for multilevel modeling in place of some form of regression, which is monolevel. In Figure 1.1, if Arrow 1 is nonsignificant then there is no level 2 (here, school) effect and data do not cluster by school. Therefore OLS or other types of monolevel regression may be used without recourse to multilevel modeling. The null model is discussed with worked examples in Chapter 3.

There are several less used synonyms for the null model:

- the unconditional random intercept model (as discussed above)
- the unconditional model (the outcome is not conditioned by any predictor variables)
- the intercept-only model (because there are no level 1 predictors)
- the baseline model (we use the null model to compare model fit with later models which include predictor variables or other additional effects)
- the random intercept null model (the school effect is a random effect predicting average outcome values at level 1, where the intercept is the average, and null refers to not having other predictor variables)
- the one-way ANOVA with random effects model (one-way because there is only one independent variable, such as the clustering variable school; random effects because this variable is treated as a random effect; ANOVA, which centers on comparing means, because we are seeing if the predicted mean or intercept using the cluster variable differs from the mean not using it)

In contrast to the null model, which is unconditional, the researcher's more complex model is a conditional model. That is, estimates of the outcome variable are conditioned on the predictor variables which have been entered at any level of the model. The null model may be used as a baseline to compare the researcher's model using a measure of how much model error exists. Ideally, the researcher's model displays significantly less error. Since multilevel modeling uses maximum likelihood estimation (ML) rather than ordinary least squares (OLS), error is reflected in the likelihood statistic, where higher is more error (contrast OLS regression, where model error is  $1 - R^2$ ). The likelihood conforms to a chi-square distribution, making it useful for significance testing, when its log is multiplied by  $-2$ . The  $-2$  log likelihood statistic is labeled  $-2LL$  and is also called model chi-square or deviance.

The likelihood ratio test uses  $-2LL$  to test the significance of the amount of error in predicting the level 1 outcome (e.g., math score) compared to some nested model, such as the null model. This may be done by comparing  $-2LL$  in the researcher's full model with  $-2LL$  in the null model. The  $-2LL$  value will be lower in the researcher's model, assuming effects in the researcher's model have some explanatory effect and thereby reduce error as compared to knowing only the level 2 clustering variable as is the case in the null model. The likelihood ratio test is discussed in later chapters, starting in Chapter 3.

### The Conditional Random Intercept Model

A random intercept model, also called a conditional random intercept model, is random because it incorporates the random effect of the clustering variable (e.g., school in a model predicting math score at level 1). It is an intercept model because only the intercept of the outcome variable is adjusted for the random effect. It is conditional because predictor variables are present in addition to the clustering variable(s) which define level 2 or higher. The intercept, of course, represents the mean of the dependent variable (math score). The slopes (b coefficients) of any level 1 predictor variables are not modeled as random effects.

Note that here we follow the rather ambiguous labeling convention prevalent in multilevel modeling, where random effect may refer to either higher or lower level variables. Strictly speaking, of course, random effects refer to any variable at a lower level whose intercept or slope is adjusted to take into account correlated error at a higher level (this does not include fixed effects of higher level predictor variables used in the fixed effects regression at level 1). In widespread practice, however, the higher level variables which are the source of adjustment are also often called "random effects." The reader is advised not to worry too much about this ambiguity but rather to understand that random effects always involve at least two variables: a higher level variable which is the basis for adjustment and a lower level variable whose intercept and/or slope is adjusted by the multilevel algorithm.

In multilevel modeling, the regression model, or more precisely, the regression portion of the multilevel model, is the dependent variable at level 1 as predicted by any fixed effects. Fixed effects, in turn, include any level 1 predictor variables. However, in addition, a level 2 predictor variable may be used as a random effect at level 2 (it may be used along with the level variable for level 2 to condition the DV intercept at level 1) while still also serving as a fixed effect at level 1. In fact, it is usual in multilevel modeling to enter all higher level predictor variables as fixed effects. Fixed versus random effects are discussed in Chapter 4.

The conditional random intercept model is a regression model with one or more fixed effects of level 1 predictor variables on the intercept of the level 1 outcome (Arrow 2 in Figure 1.1) combined with the random effect on the DV intercept (e.g., on mean student math scores) of the level 2 clustering variable (e.g., school; this is Arrow 1 in Figure 1.1). Level 2 predictor variables



treated as random effects (e.g., school budget per pupil, in Figure 1.1) are not part of a basic random intercept model. However, below we discuss the random intercept regression model and the random intercept ANCOVA model. In all intercept models, the intercept of the outcome variable at level 1 is adjusted for higher level effects. If slopes are adjusted also, this makes the model a type of random coefficients model, also discussed below.

In Figure 1.1, student self-confidence is a level 1 predictor variable modeling mean math score, which also is at level 1. The random intercept model adjusts the intercept (mean) of math score due to the random effect of school as the clustering variable. A synonymous label for the random intercept model is the ANCOVA model with random effects. Random intercept models are discussed with worked examples in Chapter 5.

### The Conditional Random Coefficients Model

The random coefficients model, also called the conditional random coefficients model, is the model with Arrows 1, 2, and 3 in Figure 1.1.

- Arrow 1 adjusts the intercept (mean) of math score for the random effect of school at level 2.
- Arrow 2 is the fixed effects regression fixed effects regression portion of the model.
- Arrow 3 adjusts the  $b$  coefficient (slope) of student self-confidence, the level 1 predictor variable, for the random effect of school also.

The *coefficients* term in random coefficients model should not obscure the fact that a random coefficients model estimates the intercept (mean) as well as the slope (regression coefficient) at level 1. Synonyms for the random coefficients model are RC model and random coefficients regression model. Random coefficients models are treated in more detail in Chapter 5 and later chapters.

### The Random Intercept Regression Model

In terms of Figure 1.1, the random intercept regression model is the model with only Arrows 1 and 4. This and all remaining models mentioned in this chapter are also conditional models since one or more predictor variables in addition to the level variable are in the models. Only the mean (intercept) of math score is estimated, not the slope of any level 1 predictor variable. Unlike the basic random intercept model, adjustments to the level 1 intercept (representing mean math scores in our example) are made not only based on the clustering variable (school, reflected in Arrow 1) but also for a level 2 variable (school budget per pupil, reflected in Arrow 4). There are no level 1 regressors (adding a level 1 regressor would make it a random intercept ANCOVA model, discussed next). Synonyms for this model are the random intercept model with level 2 predictors or the means as outcomes regression model.

### The Random Intercept ANCOVA Model

A random intercept ANCOVA model is a random intercept regression model to which one or more level 1 regressors (predictor variables) have been added. In terms of Figure 1.1, this is the model adding Arrow 2 to Arrows 1 and 4. That is, the random intercept ANCOVA model combines the monolevel regression model (Arrow 2) with the random intercept regression model (Arrows 1 and 4). Synonyms are the random intercept model with level 1 and level 2 predictors or the means as outcomes ANCOVA model. The ANCOVA part refers to the fact that this model centers on seeing if the predicted intercept using the cluster variable and any level 1 predictor

variables differs from the mean not using them. In the random intercept ANCOVA model, only the intercept of math score is estimated. If level 1 slopes are also estimated, this becomes the random coefficients ANCOVA model, discussed next.

### The Random Coefficients ANCOVA Model

The random coefficients ANCOVA model may be either the model with Arrows 1 to 4 in Figure 1.1 or the model with Arrows 1 to 5. The central difference from the random intercept ANCOVA model discussed above is the addition of Arrow 3 and/or Arrow 5, both of which represent adjustments to the slope (coefficient) of student self-confidence, a level 1 predictor variable. At level 1, the intercept (mean) of math score is adjusted by taking into account the random effects of school as the grouping variable (Arrow 1) and of school budget per pupil as a level 2 predictor (Arrow 4). Also, the slope of student self-confidence is estimated taking into account the random effect of school as the grouping variable (Arrow 3).

The slope of student self-confidence may also be modeled by a level 2 predictor variable (Arrow 5). Note that modeling a lower level variable's slope by a higher level predictor variable (not the grouping variable) is equivalent to add the interaction term for the two variables to the model and most statistical packages other than HLM 7 represent Arrow 5 in this manner. Synonyms for this model are the full random coefficients model or the intercepts and slopes as outcomes model.

Of the six common types of multilevel model, none is best. Rather, the selection of model depends on the research purposes and causal assumptions of the researcher. Not uncommonly, researchers may find the need to investigate more than one type of model. And, of course, there are many additional types of model, particularly if there are more than two levels as is often the case, for example, for longitudinal data.

### Mediation and Moderation Models in Multilevel Analysis

Mediation refers to the causal chain from predictor to outcome passing through one or more intermediate variables. Moderation refers to the strength or direction of the causal path from predictor to outcome being contingent on values of a third variable (the moderator). It is also possible to have moderated mediation and mediated moderation. Mediation and moderation models in multilevel modeling is an advanced topic not covered in the present volume. Moreover, one will not find these terms in the indices of the great majority of other texts on multilevel modeling.

Nonetheless, considerable work has been done on the subject in the last two decades (e.g., Bauer, Preacher, & Gil, 2006; Kenny, Korchmaros, & Bolger, 2003; Krull & MacKinnon, 1999, 2001). The rationale for multilevel mediation analysis (analysis of indirect effects) is given in Tofghi, West, and MacKinnon (2013), who also present an online supplement of extended analytic proofs. These authors outline a multilevel procedure for analyzing mediation effects which involves running mixed models with and without an assumed correlation of a moderator and an outcome variable, and they illustrate through simulations employing the lme4 package in R (discussed in subsequent chapters of the current volume).

Perhaps the most cited work on mediation and moderation in regression models is that by Andrew Hayes (2013). Hayes has since published on multilevel mediation analysis, using the MPlus software package (Hayes, 2014). In 2017, Hayes's Mechanisms and Contingencies Lab at Ohio State University released an SPSS macro called MLMED, written by Hayes's doctoral student Nick Rockwood. MLMED implements multilevel mediation and moderated mediation analysis, and is available for download at <https://njrockwood.com/mlmed>.

For Stata, the *ml\_mediation* command is provided to implement the multilevel mediation method described by Krull and MacKinnon (2001). The Institute for Digital Research and Education (IDRE) at UCLA provides a description of this method.<sup>1</sup> An alternative method using the Stata command *xtmixed* (since superseded by *mixed*) based on the work of Bauer, Preacher, and Gil (2006) has also been described by IDRE.<sup>2</sup> See also Zhang, Zyphur, and Preacher (2009), who use Stata for their multilevel mediation simulations.

For SAS, IDRE has documentation of multilevel mediation analysis paralleling the two methods described for Stata above.<sup>3</sup>

For the R language, Elizabeth Page-Gould (see Sharples & Page-Gould, 2016) has made available R source code, example R syntax, example data, plus slides and a help file to estimate unbiased indirect effects in multilevel models.<sup>4</sup> Indirect effects, of course, is simply a synonym for mediation effects.

As mentioned at the outset of this subsection, multilevel mediation analysis is an advanced topic, which in turn means it is fraught with complexities and pitfalls. As Zhang et al. (2009) noted, for instance,

tests of multilevel mediation can be problematic when between-group variation in a Level-1 variable is not explicitly separated in a test of 2-1-1 mediation—the same would be true for 1-1-1 mediation tests. By adhering to traditional recommendations for testing mediation with multilevel data, researchers may be making one hypothesis (i.e., group-level mediation) while testing another (i.e., mediation that conflates between-group and within-group effects). (p. 717)

These and other methodological problems must be taken into account by researchers pursuing multilevel mediation models.

With regard to multilevel moderation analysis, heterogeneous multilevel models (discussed in Chapter 9) may be used to address the question of whether the multilevel model varies conditional on a categorical moderator variable. However, more complex moderation models may be better examined using multilevel structural equation modeling (MSEM), as advocated by Preacher, Zhang, and Zyphur (2016), who implement such models in MPlus. MSEM may also be implemented in Stata using the *gsem* command and in R using the *xxM* package.

As with multilevel mediation analysis, multilevel moderation analysis can be marked by complexities and pitfalls. Preacher et al. (2016) critique common tests of multilevel moderation as set forth, for instance, by Raudenbush and Bryk (1986, 2002), noting that “problems occur because most approaches to testing multilevel moderation do not separate lower- and higher-level effects into their orthogonal components, and instead conflate these effects by combining them into single coefficients” (p. 189; see also Preacher, Zyphur, & Zhang, 2010). Across multiple fields, such conflation is known to cause model misspecification (see Hausman, 1978), resulting in conceptual and statistical problems. Conceptually, researchers (a) create theories and hypotheses that are insensitive to the different yet theoretically meaningful ways that moderation can occur, and (b) specify models reflecting this insensitivity. Statistically, researchers test moderation by (a) unknowingly constraining effects to equality across levels and (b) introducing bias into estimates of moderation effects by not treating outcomes and predictors as latent variables at the levels of analysis stipulated in theory. As a result, researchers’ theories are often tested with conflated and potentially biased parameter estimates, while theoretically meaningful moderation effects go undetected. The latent variable approach to which these authors allude is integral to multilevel structural equation modeling, which is beyond the scope of the present volume.

## ALTERNATIVE STATISTICAL PACKAGES

In the examples in later chapters, an attempt has been made to show the parallel implementation of the same models for the same data in SPSS, Stata, SAS, HLM 7 software packages, and in the R statistical language. All these statistical programming environments support many more options and models than are described in this introductory volume. While any given researcher is apt to pick a particular software her or his favorite, nonetheless, the multipackage approach in this book is helpful for a number of reasons and it is recommended that the reader at least scan implementations in all packages.

1. Different packages use different labels and vocabulary for the same thing, and having some multipackage familiarity enhances the researcher's ability to read the literature.
2. Different packages have differing options and outputs, and the researcher may need at times to be aware of alternative options and to use an alternative package.
3. Different packages may use different algorithms, defaults, and assumptions. The default output for one package may differ from that for another package, something that may not be noticed by the researcher who only uses one package. Running the model in two packages may illuminate issues surrounding the choice of options and at a minimum will confirm findings.
4. When the researcher is attempting to replicate the work of others who have used a different statistical package, it may be necessary to understand the options of statistical packages other than one's favorite.
5. Taking the model one has implemented in one package and replicating it in another package is a good method for catching mistakes and for validating results.
6. Pedagogically, looking at a problem from the viewpoint of multiple approaches is often helpful to the learning process.
7. Knowledge of multiple packages helps the researcher become a statistics-literate reader of the professional literature, where diverse statistical packages are used.

No particular statistics package is endorsed in this book. Each has advantages and disadvantages, though this does not mean they are equivalent. For the novice reader, a few orienting remarks may be helpful.

- SPSS was long the favorite in the social sciences and is still widely used. It employs a very user-friendly menu system which makes accomplishment of most tasks easy, though a programming mode is also available. It is criticized for being less comprehensive than some other packages but it has sought to remedy this by making the process of adding third-party modules, even R modules, easier. It is also criticized for its expense. It has good support from its company and user communities, though not at the level of some other packages.

- SAS has been the dominant package for business and government. Often it has the most extensive options for any given procedure. SAS is not just a statistical package but is also an enterprise management solution. While a menu mode is available, most users employ the programming mode, writing SAS syntax. Syntax is widely shared on the web. Third-party modules can extend SAS's functionality. Extensive company and community support is available. In its Enterprise Miner solutions, it can handle what is called *big data*, including textual data.

- Stata has eclipsed SPSS as a teaching tool due to its lower expense, superb user support, versatile programming language, excellent graphics, and wealth of third-party extensions which are easily added, among other factors. It too provides a menu mode though most use the command mode. For the same reasons, Stata has become a leading tool for researchers. Very active user communities have developed in some lines of academic inquiry, including econometrics and survival analysis. Its free technical support is perhaps the best of the packages considered here.

- HLM 7 originates from founders of the multilevel modeling method and reflects classic literature on MLM. It is free software in its student version and its interface clearly reveals just what effects are being modeled at each level and what the corresponding equations are, making it a useful learning tool. It displays the equations underlying the model at each level whereas the other packages do not display separate equations for each level. Students wishing an equations-based approach to multilevel modelling should give particular attention to the HLM 7 sections in this volume. It offers some options which can be difficult to implement in other packages, though it is not as comprehensive as some and, indeed, is not intended to be a comprehensive statistical package, just an MLM package. Also, the student version has limited functionality.<sup>5</sup>

- R evolved as a favorite for advanced users who wished to program their own statistical procedures. With roots based on John Chambers's S language, R is an open-source statistical programming environment which has slowly become popular for teaching and general research, not just advanced data analytics, aided by the fact that it is free, with full functionality. While it has modules to accomplish almost any statistical task, including an array of data visualization possibilities, it is less user-friendly than most other packages. However, its thousands of third-party modules cover state-of-the-art procedures not available in other packages. It is emerging, for example, as the package of choice for collecting and analyzing big data, such as that from social media. Unlike the other packages, R is not hosted by a company which maintains quality control and provides centralized user help but rather relies on a very widespread and active user community to continually improve what is available (much like wikis do). R is a constantly changing and evolving environment. Getting started with R is described in Online Appendix 1.

It should be noted that there are, of course, many other statistical packages for multilevel modeling apart from the five covered in this volume. Among these are MPlus, MLWin, R2MLwiN (runs MLWin from within R), and STAT-JR, to name a few.

## MULTILEVEL MODELING VERSUS GEE

It is possible to examine hierarchical data using the alternative statistical procedure known as generalized estimating equations (GEE). For longitudinal data GEE is equivalent to population-averaged panel data regression (Garson, 2013b; free pdf at <http://www.statisticalassociates.com>). Both multilevel modeling and GEE address the problem of correlated error which arises when values of an outcome variable at level 1 cluster within groups defined by a categorical variable. However, the two procedures attack the problem in different ways.

Multilevel models incorporate both fixed and random effects while GEE allows only fixed effects. Fixed effects, which are also called marginal effects, generate coefficients assumed to be fixed for the entire population. For this reason, like regression weights, fixed effect coefficients are a type of population average. Population-averaged panel data regression models are the same

as GEE models. In contrast, multilevel models assume that the categorical grouping variable may involve context-specific (i.e., group-specific) effects on the outcome variable. Multilevel models also partition variance into between-group effects and within-group effects. The former reveal the impact of differences between groups (e.g., schools, in the classic case of predicting student achievement scores at level 1) while the latter reveal the impact of residual variation in values of the outcome variable due to variation among subjects within groups (e.g., variation of students within schools). Put another way, the effect of within-group variability is directly observable in output for the residual random effects variance component, whereas in GEE models variability within groups is treated as a nuisance term and is incorporated in the intercept, along with other sources of model error, preventing the researcher from analyzing the effect of variability within groups directly.

Thus GEE is a population-averaged (marginal) method rather than a subject-specific method. As a simple example, let days of unemployment be measured over a period of time and let the groups be a group receiving training and a group not receiving training. GEE does not focus on variation within groups of days of unemployment but rather focuses on the difference in average effects between the training and no-training groups. To make this comparison, GEE must assume that the within-group variation over time does not matter. What matters in GEE is the average response in each subpopulation (each group) based on the predictor variables. GEE maximizes the predictability of group means, not the predictability of subject-level outcome measures.

Because multilevel modeling and GEE handle subject-level variation differently, when the grouping variable does indeed provide contextual effects for the relationship of predictor variables to the outcome variable, both methods may provide similar estimates but the GEE method will give erroneous standard errors and the regression coefficients for predictor variables will tend to be lower than in subject-specific models (see Neuhaus, Kalbfleisch, & Hauck, 1991). This reflects the lowering of effect sizes due to averaging, aggregating, binning, and other approaches which do not take advantage of subject-specific information.

While more information is generally better than less information, making multilevel modeling with its incorporation of subject-specific effects a better choice than GEE, this is not the whole story. There are two oft-cited advantages of GEE, which may make it a better choice in certain contexts.

1. GEE is a nonparametric procedure (see Hardin & Hilbe, 2002, p. 55). The researcher need not assume that the outcome variable is normally distributed as in the case of ordinary multilevel modeling. Being nonparametric is achieved by using quasi-likelihood estimation rather than maximum likelihood estimation, and by using fit measures like QIC (the quasi-likelihood under independence criterion) instead of ML-based fit measures like AIC or BIC (Akaike information criterion and Bayesian information criterion). The estimation method and fit measures in GEE are attempts to approximate what maximum likelihood would estimate under conditions of fuller information about the data.
2. Likewise, homogeneity of variances of values of the outcome variable between groups is not assumed.

Advocates of multilevel modeling tend to view multilevel modeling as preferable to GEE because it incorporates more information. With regard to the parametric/nonparametric issue, generalized multilevel modeling is available (see Chapter 12), allowing the researcher to select an appropriate data distribution and link function. If the data distribution is unknown, alternative

distributions can be explored using likelihood ratio tests. With regard to the issue of homogeneity/heterogeneity of variances, multilevel modeling supports heterogeneous variance models (see Chapter 9) and again, likelihood ratio tests allow the researcher to explore the effects of different assumptions about variances.

One researcher put the difference between multilevel mixed modeling and GEE modeling this way:

GEEs appeal to people who don't like distributional assumptions, whereas MLMs appeal to people who like generative models. I prefer MLMs because I like to set up an explicit model for the data; others prefer GEEs because they like to have a procedure that estimates parameters in the absence of assumptions for how the coefficients vary. (Gelman, 2006; see also McNeish et al., 2017)

In summary, the multilevel mixed modeling vs. GEE choice remains a debatable topic with no “correct” answer. There is agreement that GEE is a population-averaged method unsuitable for analysis of subject-specific effects. The essential argument in favor of GEE for analysis of group-averaged effects is that its relaxed assumptions present the researcher fewer obstacles and make the issue of model misspecification less important (see Hubbard et al., 2010). The essential argument in favor of multilevel mixed modeling is that it allows the researcher to explore subject-specific as well as between-group effects and to use full information to analyze alternative model assumptions. Moreover, for larger samples, multilevel modeling is itself reliable and robust even when assumptions are violated to a non-extreme degree (cf. Hox, Moerbeek, & van de Schoot, 2018).

## Summary

Key concepts learned by the reader in this introductory chapter include the following points:

- Multilevel modeling is used for any data in which the outcome variable's values are not independent but rather cluster by the groups formed by categorical variables within which it is nested or by which it is cross-classified.
- Ignoring the clustering effects of multilevel data violates the data independence assumption of OLS regression and may lead to serious misinterpretation of regression effects. OLS regression will tend to yield too many spuriously significant findings (Type I errors).
- Whether a multilevel model is needed can be inferred from either of two mathematically equivalent tests: (1) The intraclass correlation (ICC) is significant, or (2) the random effect of the clustering variable component is significant.
- There are many synonyms for multilevel modeling, the primary ones being linear mixed modeling (LMM) and hierarchical linear modeling (HLM).
- Ordinary multilevel modeling uses linear regression but there is also generalized multilevel modeling for nonlinear relationships.
- The categorical variable by which values of the dependent variable(s) cluster is called the clustering variable. Synonyms are the grouping variable, link variable, or level variable.
- The clustering variable at level 2 may be used to adjust the intercept (estimated mean) of the DV at level 2. In addition it may be used to adjust the slope (b coefficient) of any level 1 predictor variables (regressors, predictor variables) of the DV. The same is true of any level 2 predictor variables. There may be more than two levels.

- If only the intercept of the DV is adjusted, the model is a type of multilevel intercept model. If one or more slopes are adjusted as well, the model is a type of multilevel coefficients model.
- Different statistical packages implement multilevel modeling in somewhat different ways, not infrequently using different labels for the same things. Moreover, different statistical packages offer different default settings, different estimation algorithms, different output options, and different assumptions. Familiarity with different packages helps with selecting the best tool for the job, replication of previous studies, and simply reading the diverse multilevel modeling literature. This book illustrates MLM models in SPSS, Stata, SAS, HLM 7, and R.
- It is not enough to show higher levels affect intercepts and slopes at lower levels. The researcher should be prepared to explain the mechanisms by which higher levels may affect lower levels. To avoid having data-driven results, the researcher should base the research design in multilevel theory.
- Multilevel models use five main types of multilevel data: hierarchical data, repeated measures data, random effects data, cross-classified data, and multiple outcome data.
- In multilevel modeling, the null model is the model with only the clustering variable as a modifier of the mean (intercept) of the DV, with no predictor variables at any level. A primary use of the null model is to test if there is a significant effect of the clustering variable, thereby justifying the need for multilevel modeling.
- The random intercept model is one in which the clustering variable models the intercept at level 1 (as does the null model) but there are also fixed effects of predictor variables at level 1. This is a type of ANCOVA model with random effects.
- The random coefficients model has the elements of a random intercept model but in addition the clustering variable models one or more slopes of lower level predictor variables.
- The random intercept regression model is one with no predictor variables at level 1 but in which the level 1 intercept is modeled both by the clustering variable at level 2 and by one or more level 2 predictor variables. This is sometimes called the means as outcomes regression model.
- The random intercept ANCOVA model is a random intercept regression model to which one or more predictor variables have been added at level 1.
- The random coefficients ANCOVA model is a random intercept ANCOVA model in which the slopes of one or more level 1 predictor variables is modeled by the clustering variable and possibly by one or more level 2 predictor variables.
- There are many other types of multilevel model.
- Hierarchical data may also be handled using generalized estimating equations (GEE), but that technique is unsuitable for analyzing subject-level effects.

## Glossary

### Multilevel modeling (MLM)

Common synonyms are linear mixed modeling (LMM) or hierarchical linear modeling (HLM). Multilevel modeling is used when there is clustering of the outcome variable by a categorical variable, such that error is correlated. When error is no longer independent, OLS regression would result in biased estimates.

Multilevel models have at least 2 levels. Level 1 includes the outcome variable while level 2 reflects groups defined by the categorical clustering (grouping, level) variable. Higher levels may also be in the model. Although commonly used for nested hierarchical data, MLM can also handle cross-classified data if the researcher declares the data as such.



Multilevel models include both fixed and random effects, with fixed effects referring to the level 1 regression model for the outcome variable. Predictor variables at any level may be fixed effects in this model. Random effects refer to the effects of the grouping variable on the intercept of the outcome variable or to its effects on slopes (b coefficients) of one or more level 1 fixed effects. Lower level slopes may also be conditioned by higher level predictor variables.

#### Fixed effects vs. random effects

Fixed effects refer to effects in the level 1 regression model. Fixed effects may be associated with predictors at any level. Random effects refer to the effects of clustering of the outcome variable within categorical levels of a clustering variable which defines a level in a multilevel model. Random effects are used to adjust estimates for the intercept of the level 1 outcome variable or slope (b) coefficients of one or more predictor variables. The same variable may be both a fixed effect (it is a predictor in the level 1 regression model) and a random effect (its slope estimates are conditioned by random effects of the grouping variable).

#### Within-group effects vs. between-group effects

Between-group effects refer to effects existing at the group level as defined by the grouping variable. Within-group effects refer to effects remaining at the individual level after controlling for the effects of the grouping variable. The residual variance component in MLM is the within-group effect and reflects unexplained variance. Multilevel modeling can compare within-group and between group effects. When random effects are uncorrelated, as in variance components (VC) models, MLM can partition the variance in the outcome variable into the percentage attributable to between-groups effects (e.g., effects of a grouping variable such as *schoolid*) and the percentage attributable to within-groups effects (e.g., residual individual-level effects after controlling for between-groups effects).

#### Null model

In OLS regression, the null model is an intercept-only model with no predictors of the outcome variable. In MLM, the null model is the similar but is not strictly an intercept-only model since there is one predictor: the categorical variable defining groups at level 2. Another synonym for the null model in MLM is the baseline model, referring to the fact that the null may

be used to assess improvement in model fit vis-à-vis later models which include additional effects.

#### Unconditional model vs. conditional model

The null model is unconditional since estimates of the outcome variable are not conditioned by any effects other than the effect of the grouping (level) variable. More complex models are considered conditional, as the estimates of the outcome variable are conditioned on the predictor variables added to the model.

#### Random intercept models vs. random coefficients models

A random intercept model includes the random effect of the clustering variable on the mean (intercept) of the outcome variable but does not include random effects on the slopes of any regressors in the fixed effects (regression) model for the outcome variable at level 1. In contrast, a random coefficients model estimates not only the level 1 intercept but also one or more regressor slopes.

#### Random intercept regression model

The random coefficients regression model estimates only the intercept of the outcome variable. No slopes of level 1 predictor variables are estimated in this model. In addition to the effect of the level 2 grouping variable on the intercept, there are one or more level 2 predictor variables which also influence the level 1 outcome variable.

In comparison, the random intercept ANCOVA model is similar to the random intercept regression model but adds one or more level 1 predictor variables as predictors of the level 1 intercept. This is distinguished from the random coefficients ANCOVA model, which is similar to the random intercept ANCOVA model but the slopes of one or more level 1 predictor variables are modeled by the grouping variable.

#### Generalized multilevel modeling (GMLM)

GMLM supports nonlinear multilevel modeling. Although by default ordinary multilevel modeling assumes linear relationships, generalized modeling software can be used to incorporate such nonlinear forms of regression as gamma regression for skewed data, ordinal regression for ordinal outcome variables, logistic regression for binary outcomes, and many others. Synonyms include generalized linear mixed modeling and generalized hierarchical linear modeling.

## Challenge Questions With Answers

### Questions

- 1-1. Apart from obtaining more accurate model coefficients, what is the primary function of multilevel modeling for two-level models?
  - a. to separate within-group individual effects from between-group aggregate effects
  - b. to separate between-group aggregate effects from the total effects
  - c. to combine the effects of variables at different levels
  - d. to combine the effects of variables at the group level
- 1-2. What OLS regression assumption often is violated by multilevel data?
- 1-3. What is one of the two mathematically equivalent tests which indicate whether a multilevel model is needed?
- 1-4. True or false: Multilevel modeling can only be done with linear relationships.
- 1-5. True or false: Multilevel modeling can handle more than two levels.
- 1-6. Which of the following statistical packages handles multilevel modeling?
  - a. HLM 7
  - b. SAS
  - c. SPSS
  - d. All of these
- 1-7. What is multilevel theory and why is it important in multilevel analysis?
- 1-8. Which of the following is NOT one of the five main types of multilevel data?
  - a. cross-classified data
  - b. cross-sectional data
  - c. hierarchical data
  - d. multiple outcome data
- 1-9. What are the two primary uses of the null model in multilevel modeling?
- 1-10. True or false: The random coefficients model is one in which the grouping variable models the intercept at level 1 (as does the null model) but there are also fixed effects of predictor variables at level 1.

### Answers

- 1-1. A. Apart from obtaining more accurate estimates, the primary function of multilevel modeling is to separate within-group individual effects at level 1 from between-group aggregate effects at level 2, for two-level models.
- 1-2. The OLS regression assumption of data independence often is violated by multilevel data, as dependent variable values at level 1 may well cluster within groups defining level 2.
- 1-3. For the null model, the researcher looks at the significance of the intraclass correlation (ICC) coefficient and/or the random effect of the variance component for the grouping (clustering, level) variable. If one is significant, the other will be also. Significance indicates that a multilevel model is necessary.
- 1-4. False. Although ordinary multilevel modeling uses linear regression, generalized multilevel modeling is also available to handle nonlinear relationships.

- 1-5. True. Multilevel modeling can handle more than two levels. More than two levels is often the case, for example, with longitudinal data. In practice, multilevel models rarely have more than four levels.
- 1-6. D. All of these. Of the packages discussed in this book, HLM 7, SAS, SPSS, Stata, and R all support multilevel modeling, although implementation may differ (e.g., by using different labels and offering different default settings).
- 1-7. Multilevel theory explains the mechanisms behind these effects in the multilevel model. That is, multilevel theories relate higher level effects to lower level effects. Without multilevel theory, findings may be data driven and therefore may be less valid and less stable.
- 1-8. B. Cross-sectional data is not a type of multilevel data. The five main types of multilevel data are hierarchical data, repeated measures data, random effects data, cross-classified data, and multiple outcome data.
- 1-9. A primary use of the null model is to test if there is a significant effect of the grouping variable which defines a higher level such as level 2. If this is significant, the need for multilevel modeling is justified. A second major use of the null model is as a baseline for comparison using likelihood ratio tests of model differences.
- 1-10. False. The model in which the clustering variable models the intercept at level 1 (as does the null model) but there may also be fixed effects of predictor variables at level 1 is the random intercept model. In contrast, the random coefficients model has elements of the random intercept model, but in addition the clustering variable models one or more slopes of lower level predictor variables.

## Notes

- <https://stats.idre.ucla.edu/stata/faq/how-can-i-perform-mediation-with-multilevel-data-method-1/>
- <https://stats.idre.ucla.edu/stata/faq/how-can-i-perform-mediation-with-multilevel-data-method-2/>
- <https://stats.idre.ucla.edu/sas/faq/how-can-i-perform-mediation-with-multilevel-data-method-1/>; <https://stats.idre.ucla.edu/sas/faq/how-can-i-perform-mediation-with-multilevel-data-method-2/>
- <http://www.page-gould.com/r/indirectmlm/>
- Information on the student version of HLM is found at <http://www.ssicentral.com/hlm/student.html>. Limitations: The STAT/Transfer utility used for the importation of data is not included. The student edition will only accept ASCII, SYSTAT, SPSS for Windows, or SAS transport data files. Note that SPSS data files created with SPSS 21 or earlier can be used with the student edition. For a Level-3 model, the maximum number of observations that may be used at Levels 1, 2, and 3 is approximately 8000, 1700, and 60, respectively. Note that the restriction applies to observations in the case of the level-2 file, for example, and not to the actual number of level-2 units to be included in the analysis. For a level-2 model, the maximum number of observations at the two levels is 8000 at level 1 and 350 at level 2 of the hierarchy. No more than five effects may be included in any HLM equation at any level of the model, and the grand total of effects cannot be 25 or higher.

Visit [study.sagepub.com/researchmethods/statistics/garson-multilevel-modeling](http://study.sagepub.com/researchmethods/statistics/garson-multilevel-modeling) for downloadable study resources to accompany this text!

- Online Appendix 1: Getting Started with R and R Studio
- Online Appendix 2: Additional Frequently Asked Questions
- Datasets and Codebooks from the book
- Figures & Tables from the book