CHAPTER 1. THE NEED FOR MULTILEVEL MODELING

I should venture to assert that the most pervasive fallacy of philosophic thinking goes back to neglect of context.

—John Dewey, 1931

Background and Rationale

When one considers almost any phenomenon of interest to social and health scientists, it is hard to overestimate the importance of context. For example, we know that the likelihood of developing depression is influenced by social and environmental stressors. The psychoactive effects of drugs can vary based on the social frame of the user. Early childhood development is strongly influenced by a whole host of environmental conditions: diet, amount of stimulation in the environment, presence of environmental pollutants, quality of relationship with mother, and so on. Physical activity is shaped by neighborhood environment; people who live in neighborhoods with sidewalks are much more likely to walk. The probability of teenagers engaging in risky behavior is related to being involved in structured activities with adult involvement. A child's educational achievement is strongly affected by classroom, school, and school system characteristics.

These examples can be extended to situations beyond where individuals are being influenced by their contexts. The likelihood of couples avoiding divorce is strongly related to certain types of religious and cultural backgrounds. Group decision-making processes can be influenced by organizational climate. Hospital profitability is strongly affected by reimbursement policies set by government and insurance companies.

What all these examples have in common is that characteristics or processes occurring at a higher level of analysis are influencing characteristics or processes at a lower level. Constructs are defined at different levels, and the hypothesized relations between these constructs operate across different levels. Table 1.1 presents an example of the interdependence among levels of analysis, here with an example from the area of tobacco control. Research programs on tobacco control exist at all levels of analysis, from the genetic up to the sociocultural and political (i.e., "from cells to society"). Moreover, although research can occur strictly

Table 1.1 Levels of analysis in health research with examples from tobacco control

Level of Analysis	*Examples From Tobacco Control*
Cultural/political	Measuring elasticity of the effect of cigarette taxation on population smoking rates
Social/environmental	Measuring the relative importance of family and peer environment on teen smoking incidence
Behavioral/psychological	Designing effective smoking prevention and cessation programs
Organ systems	Designing ways to block tumor formation in smokers
Cellular	Tracing metabolic pathways of nicotine uptake
Molecular/genetic	Examining the genetic basis of nicotine dependence

within any of these levels, much of the most important research will look at the links between the levels. For example, as we learn more about the genetic basis of nicotine dependence, we may be able to tailor specific preventive interventions to particular genotypes.

These types of *multilevel* theoretical constructs require specialized analytic tools to properly evaluate. These multilevel tools are the subject of this book.

Despite the importance of context, throughout much of the history of the health and social sciences, investigators have tended to use analytic tools that could not handle these types of multilevel data and theories. In earlier years, this was due to the lack of such tools. However, even after the advent of more sophisticated multilevel modeling approaches, practitioners have continued to use more simplistic single-level techniques (Luke, 2005).

Theoretical Reasons for Multilevel Models

The simplest argument, then, for multilevel modeling techniques is this: Because so much of what we study is multilevel in nature, we should use theories and analytic techniques that are also multilevel. If we do not do this, we can run into serious problems, including making incorrect causal claims.

For example, it is very common to collect and analyze health and behavioral data at the aggregate level. Epidemiologic studies, for example, have shown that in countries where fat is a larger component of the diet, the death rate from breast cancer is also higher (Carroll, 1975). It might seem reasonable to then assume that women who eat a lot of fat would be more likely to get breast cancer. However, this interpretation is an example of the *ecological fallacy*, where relationships observed in groups are assumed to hold for individuals (Freedman, 1999). Recent health studies, in fact, have suggested that the link between fat intake and breast cancer is not very strong at the individual level (Holmes et al., 1999).

This type of problem can also work the other way. It is very common in the behavioral sciences to collect data from individuals and then aggregate the data to gain insight into the groups to which those individuals belong. This can lead to the *atomistic fallacy,* where inferences about groups are incorrectly drawn from individual-level information (Diez-Roux, 1998). It is possible to be successful assessing ecological characteristics from individual-level data; for example, see Moos's (1996) work on social climates. However, as Shinn and Rapkin (2000) have argued, this approach is fraught with danger and a much more valid approach is to assess group and ecological characteristics using group-level measures and analytic tools.

It is useful here to consider the sociological distinction between properties of *collectives* and *members* (Lazarsfeld & Menzel, 1961). Members belong to collectives, but various properties (variables) of both collectives and their members may be measured and analyzed at the same time. Lazarsfeld and Menzel identify analytical, structural, and global properties of collectives. *Analytical* properties are obtained by aggregating information from the individual members of the collective (e.g., proportion of Hispanics in a city). *Structural* properties are based on the relational characteristics of collective members (e.g., friendship density in a classroom). Finally, *global* properties are characteristics of the collective itself that are not based on the properties of the individual members. Presence of an antismoking policy in a school would be a global property of the school, for example.

Using this framework, it becomes clear that fallacies are a problem of *inference,* not of *measurement.* That is, it is perfectly admissible to characterize a higher level collective using information obtained from lower level members. The types of fallacies described above come about when relationships discovered at one particular level are inappropriately assumed to occur in the same fashion at some other (higher or lower) level.

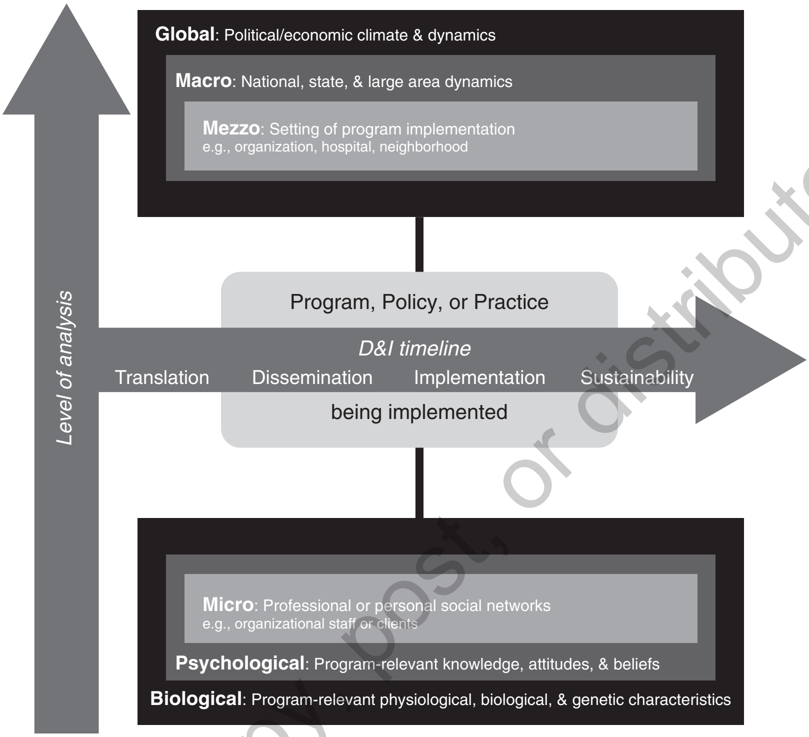
There is broad interest in social and physical context across the social sciences, and this can be seen most clearly in the ecological richness of various social science theories and conceptual frameworks. In sociology and criminology, the theory of *neighborhood disorder* proposes that various physical and social indicators of environmental disorder are related to a variety of individual and relational outcomes, including crime, violence, policing styles, and depression (Sampson, Morenoff, & Gannon-Rowley, 2002). Political scientists have consistently viewed political participation (e.g., voting, petitioning, contacting elected officials) as being driven by a number of contextual processes and factors, including organizational culture, media exposure, and peer influence (Uhlaner, 2015). Policy science generally views policy development and implementation as a process embedded in local, regional, and national political and geographic contexts. For example, the *political stream* of Kingdon's influential multiple streams framework is defined by referring to predominately contextual structures and processes including national mood, legislative body makeup, and interest group activities (Béland & Howlett, 2016). An alternative model of the policy process, the *advocacy coalition framework*, more explicitly positions policy activity as an output from the policy subsystem, which is in turn made up of three types of collectives: (1) coalitions, (2) governmental bodies, and (3) institutions (Sabatier & Weible, 2007).

Finally, implementation science (sometimes called dissemination and implementation science) is a relatively new social science that focuses on how evidence-based programs, practices, and policies can be better disseminated, implemented, and maintained to benefit population health. Early frameworks used in implementation science view implementation processes and outcomes as situated within social and organizational contexts. For example, Rogers's *diffusion of innovations* theory has been used extensively to study how new discoveries are passively diffused and actively disseminated across social and organizational networks and systems (Rogers, 2003; Valente, 1996). Figure 1.1 presents an expanded social–ecological framework for implementation science (Luke, Morshed, McKay, & Combs, 2017). This framework, adapted from Glass and McAtee (2006), suggests that relational, organizational, and social contexts are important for studying and understanding implementation outcomes.

Statistical Reasons for Multilevel Models

Despite the contextual richness of the theories and frameworks used by social scientists, they have often tended to utilize traditional

Figure 1.1 A social–ecological model for implementation science.



Source: Adapted from Glass and McAtee (2006).

individual-level statistical tools for their data, even if their data and hypotheses are multilevel in nature. One approach has been to disaggregate group-level information to the individual level so that all predictors in a multiple regression model are tied to the individual unit of analysis. This leads to at least two problems. First, all of the unmodeled contextual information ends up pooled into the single individual error term of the model (Duncan, Jones, & Moon, 1998). This is problematic because individuals belonging to the same context will presumably have correlated errors, which violates one of the basic assumptions of multiple regression (i.e., no autocorrelation of residuals). This is often called *clustering*, and it implies that observations within some larger unit are related to one another. For example, you might see clustering of student test scores within a classroom. If students are not randomly assigned to classes, if teacher ability varies across classrooms, or if classroom environments (e.g., class size) vary systematically, then you would expect to see clustering of scores.

The second statistical problem is that by ignoring context, the model assumes that the regression coefficients apply equally to all contexts, "thus propagating the notion that processes work out in the same way in different contexts" (Duncan et al., 1998, p. 98).

One partial solution to these statistical problems is to include an effect in the model that corresponds to the grouping of the individuals. This leads to an analysis of variance (ANOVA) or analysis of covariance (ANCOVA) approach to modeling. Unfortunately, there are still a number of issues with this approach. First, in the case where there are many groups, these models will have many more parameters, resulting in greatly reduced power and parsimony. Second, these group parameters are often treated as fixed effects, which ignores the random variability associated with group-level characteristics. Finally, ANOVA methods are not very flexible in handling missing data or greatly unbalanced designs.

Scope of This Book

Based on the previous discussion, the purpose of this monograph is to provide a relatively nontechnical introduction to multilevel modeling statistical techniques for social and health scientists. After this introduction, the book is split into two major sections. Chapters 2 through 4 discuss how to plan, build, and assess the basic two-level multilevel model, and describe the steps in fitting a multilevel model, including data preparation, model estimation, model interpretation, hypothesis testing, testing of model assumptions, centering, and power analysis. Chapters 5 through 7 constitute the second section of the book, covering extensions and more advanced topics. Chapter 5 covers useful extensions to the basic multilevel model, including modeling noncontinuous and nonnormal dependent variables, building three-level models, and building cross-classified models. Chapter 6 covers the use of mixed-effects models for longitudinal data. Finally, Chapter 7 provides some guidance on a couple of general multilevel modeling topics, including tips for presenting multilevel model results and general resources for analysts. The presentation of the topics covered in this book only assumes familiarity with multiple regression, and the text makes extensive use of example data and analyses.

The primary focus of this book is the application of mixed-effects modeling techniques for multilevel and longitudinal models. Mixed-effect models are powerful and flexible approaches that can handle a number of theoretical and statistical challenges. In particular, mixed-effects models can deal with the type of clustered data that more traditional multiple regression models cannot handle. Multilevel and longitudinal data sets are typified by clustered data (e.g., students nested in classrooms or observations nested within individuals), so that makes a mixed-effects modeling approach ideal in these situations.

A useful definition to serve as a basis for the rest of the presentation is as follows: A multilevel model is a statistical model applied to data collected at more than one level in order to elucidate relationships at more than one level. The statistical basis for multilevel modeling has been developed over the past several decades from a number of different disciplines and has been called various things, including hierarchical linear models (Raudenbush & Bryk, 2002), random coefficient models (Longford, 1995), mixed-effects models (Pinheiro & Bates, 2000), covariance structure models (Muthén, 1994), growth curve models (Rogosa, Brandt, & Zimowski, 1982), as well as multilevel models. All these specific types of multilevel models fall into one of two broad statistical categories: a mixed-effects multiple regression approach and a structural equation modeling (SEM) approach. This book will focus on the mixed-effects regression-based multilevel modeling. Generally, I will use "multilevel" or "longitudinal" models when I am talking about the purposes of a model; I will tend to used "mixed-effects" models when I am talking about the broad class of statistical models that can be used for multilevel or longitudinal analyses. For a good introduction to the SEM-based approach, see Chapter 6 in Heck and Thomas (2009).

Multilevel and longitudinal models are starting to appear more frequently in every area of the social and health sciences, as the techniques have become more widely known and integrated into the major statistical packages. There are as many specific types of multilevel models as there are scientific questions. However, there are certain types of overarching models that can be seen across the different research disciplines, as indicated in Table 1.2. This table is not meant to be exhaustive but to provide a catalyst for the reader and indicate the extremely wide applicability of multilevel methods. In particular, it is not immediately obvious to the new analyst that mixed-effects models can be extremely useful for modeling longitudinal data (where multiple observations are nested within an individual) and also for meta-analytic studies (where multiple effect statistics are nested within individual studies).

Table 1.2 Types of multilevel models and structures found in the health and social sciences

Type	*Multilevel Structure*	*Examples*
Physical	Entities nested within the immediate physical environment, including biological, ecological, and built environments	Diez-Roux, et al. (2001); Lottig, et al. (2014); O'Campo, et al. (2015)
Social	Entities nested within social structures, including families, peer groups, and social networks	Buka, et al. (2003); Koster, et al. (2015); Sampson, et al. (2005)
Organizational	Individuals and small groups nested within organizational and institutional contexts	Beidas, et al. (2015); Maes and Lievens (2003); Masuda, et al. (2018)
Political	Individuals, groups, and communities nested within specific sociopolitical, cultural, or historical contexts	Boehmer, et al. (2008); Luke and Krauss (2004); Skiple, et al. (2016)
Temporal	Multiple observations of a single entity taken over time	Boyle and Willms (2001); Curran, et al. (1997); Howe, et al. (2016)
Analytic	Multiple effect measures nested with individual studies (i.e., meta-analysis)	Goldstein, et al. (2000); Weisz, et al. (2017)

Online Book Resources

Finally, a variety of resources are available online to support the material covered in this book. These resources include all the data sets used for the examples, R and Stata code to reproduce most of the models, results, and figures, and an occasionally updated mixed-effects modeling resource list. These materials can be found at **https://www.douglasluke.com**.