Multilevel Thinking

Andrew Grogan-Kaylor

12/20/22

Table of contents

1 The Usefulness of Multilevel Modeling and Multilevel Thinking

For decades now, multilevel models have been an important quantitative tool for social research. While multilevel models have become ubiquitous in social research, there are dimensions of these models that are explored less frequently in published articles. This document arises from my experiences of teaching a course entitled *Multilevel and Longitudinal Modeling* that I have taught for over a decade in the *Joint Doctoral Program in Social Work and Social Science* at the University of Michigan.

My contention is that *multilevel modeling* offers powerful tools for understanding the *multilevel data* that social researchers often confront. For example, researchers are often interested in studying outcomes for diverse groups of children in different schools, residents of diverse and different neighborhoods, or individuals or families living in diverse and different countries. Such inherently multilevel data lead to analytic complexities, some of which appear to me to be well understood, while others seem to be much less often appreciated.

The point that I wish to make about multilevel data is that when presented with complex multilevel data, failure to use the appropriate multilevel model may lead to conclusions that are demonstrably incorrect. Fortunately, many of these difficulties can be avoided with applications of simple and straightforward multilevel models.

After presenting some initial ideas about multilevel modeling, I go on to explore some more complex ideas about multilevel models that I see less often in the published empirical literature. I focus especially on the idea of *multilevel models as the exploration of variation across countries and cultures.*

Certainly, none of the statistical ideas contained in this document are unique to me. There are thorough–and often much more mathematically rigorous–presentations of many of the idea contained in this document in some of the excellent foundational texts on multilevel modeling such as the early book by Raudenbush and Bryk (2002), the excellent book on longitudinal models by Singer and Willett (2003), and Rabe-Hesketh and Skrondal (2012)'s more recent and extremely comprehensive two volume text. Luke (2004), and Kreft and de Leeuw (1998), offer shorter, less mathematical, but still excellent introductions to the topic of multilevel modeling. Gelman et al. (2007) introduced me to the ideas that in this document I describe as "multilevel structure" using an example with voting patterns.

My intent in this document is to offer a kind of accessible tutorial for applied researchers, including especially those who see their research having some advocacy based component. My approach, while offering up some equations, is less mathematical than some of the above mentioned texts, and written with the intent of providing a clear and practically focused guide for the applied researcher who is attempting to carry out better research with diverse populations.

Acknowledgements

Some Preliminary Thoughts

Preliminaries

"Like you I
Love love, life, the sweet smell of things, the sky-blue landscape of January days.
…
I believe the world is beautiful.
And that poetry like bread, is for everyone.
And that my veins don't end in me.
But in the unanimous blood.
Of those who struggle for life,
Love, little things,
Landscape and bread, the poetry of everyone."
— Roque Dalton (tr. By Jack Hirschman)

"A lifetime is too narrow to understand it all, beginning with the huge rockshelves that underlie all that life.
No one ever told us we had to study our lives, make of our lives a study, as if learning natural history or music, that we should begin with the simple exercises first and slowly go on trying the hard ones, practicing till strength and accuracy became one with the daring …
But there come times—perhaps this is one of them—when we have to take ourselves more seriously or die, when we have to pull back from the incantations, rhythms we've moved to thoughtlessly, and disenthrall ourselves, bestow ourselves to silence, or a severer listening …"
— Adrienne Rich

"Research is formalized curiosity. It is poking and prying with a purpose."
— Zora Neale Hurston

2 Introduction

2.1 Quantitative Methods and Social Justice

There is clearly need for both qualitative and quantitative methods. Central to the argument of this document is the idea that advanced quantitative methods can be core contributors to the agenda of understanding issues of diversity and social justice more fully and thoroughly (Cokley and Awad 2013; Grogan-Kaylor et al. 2018). Quantitative methods, particularly in discussions comparing qualitative and quantitative methodologies, are sometimes labelled as inherently *positivist* methods. My argument regarding this point is twofold. There is nothing within the mathematics of quantitative methods that requires a positivist epistemology. Quantitative methodologies could as easily be conducted using a critical epistemology–that is aware of dynamics of power and privilege–as any other methodology (Stage and Wells 2014). Second, when we have samples of a hundred, several hundred, several thousand, or even hundreds of thousands of study participants, it is difficult to imagine a methodology other than a quantitative methodology that could accomplish the following:

1. Sift through thousands of responses, and determine the *overall, or average, pattern of relationships* between risk factors, protective factors, and outcomes.
2. Explore the *variation in these relationships* across social contexts.
3. Determine whether there is evidence that the relationships observed within the data are more than *statistical noise.*
4. Adjudicate the *complex multivariate relationships* of risk factors, protective factors and outcomes.

In **?@sec-pvalues** and **?@sec-multilevelstructure**, I explore the ways that *multilevel* data can contribute substantially to the complexity of this enterprise. I thus argue that quantitative methods can play an important role in contributing to liberatory ideas. I note that one of the pioneers of liberation psychology, Martin-Baró (Aron and Corne 1994), used both qualitative and quantitative methods (Martin-Baro 1994), including in the latter case, relatively sophisticated arguments about patterns of missing data across a survey data set (Aron and Corne 1994).

There is thus an ethical argument that is embedded in this document. Many of us do research with the hope of better understanding the relationship of risk and protective factors with outcomes in diverse, and often disadvantaged or marginalized, populations. Many of us further

hope that our work might be part of conversations about appropriate polices, programs, treatments or interventions. Given the frequent vulnerability and marginalization of the people with whom we work, when using quantitative methods, it is incumbent upon us to employ methods that adequately address the complexities of the data, that offer an appreciation of the variability and diversity within the data, that provide the most accurate estimates possible, and that increase the probability of obtaining correct answers to important substantive questions.

> "It is hard to imagine that anyone with a humanitarian worldview would argue against the need for a more quantitatively literate citizenry. Informed political decision-making, retirement planning, active parenting, and the vast majority of choices we make in our personal, occupational, and civic lives can be better served by improved quantitative understanding and reasoning, as well as accompanying action-oriented dispositions." (Wiest, Higgins, and Frost 2007)

The idea of this document is that a deeper study of multilevel modeling can result in an advanced "quantitative literacy" (Wiest, Higgins, and Frost 2007), or "principled argument" (Abelson 1995), that is appropriate for drawing accurate conclusions from multilevel data.

2.2 Are Answers from Social Science "Obvious"?

Closely related, I think to the the idea that quantitative research can advance issues of social justice, is the question of whether answers from social science are "obvious". If social science answers are obvious, then social science has limited abilities to make new discoveries, and to build scientific foundations for evidence.

I have been thinking a lot about the idea that *Everything Is Obvious, Once You Know The Answer*, as detailed in the book with this title by Duncan Watts (2011).

This seems to me especially true in social research. Arguably, some conclusions of social research may indeed be obvious. For example, it may be obvious that *Adverse Childhood Experiences* (ACEs) are associated with long term decreases in mental health. However, even obvious conclusions may need to be quantitatively documented, in order to legitimate programs and interventions, and to secure funding. I also observe that I think that there is often a *historical* dimension to what is considered "obvious": conclusions that are at first considered to be unlikely to be true, or even counter-intuitive, require the weight of accumulating evidence over time for these connections to become "obvious". It is likely that the "obviousness" of the relationship between ACEs and later physical and mental health problems did not become apparent until research began to document these relationships (e.g. Felitti et al. (1998)). As another example, Proctor (2012) documents the way which smoking was first considered to be an *unlikely* cause of lung cancer; only over the course of several decades of research and discussion to become an *obvious* cause of lung cancer. A similar *historical* dynamic seems to be playing out in some research on parenting and child development. Despite decades of evidence

indicating that corporal punishment has undesirable conequeqences for children (Gershoff and Grogan-Kaylor 2016b), corporal punishment remains a disciplinary strategy endorsed by the majority of the American population (Hines, Kalil, and Ryan 2022).

In contrast sometimes the conclusions of social research may not always be obvious. For example:

1. There has been an ongoing debate about whether corporal punishment is more or less harmful when used by parents in social contexts, or communities where it is more common, or normative. Eamon (2001) suggested that "when environmental risk is high, parenting practices that are firmer and higher in control result in lower levels of young adolescent antisocial behavior." This echoes similar research by (Deater-Deckard et al. 1996) suggesting that physical punishment was harmful for European-American children, but not for African-American children. Later, larger sample research has found that this appears not to be the case: physical punishment is harmful for children in all groups (Gershoff and Grogan-Kaylor 2016b, 2016a; Pace, Lee, and Grogan-Kaylor 2019).
2. Using MICS Data (UNICEF 2021), we conducted a study of the link between gender inequality and physical child abuse (Ma et al. 2022). We expected to find that higher levels of gender inequality led to higher levels of physical abuse for female children, but not for male children. Instead, we found that higher levels of gender inequality were associated with higher levels of physical abuse for both male and female children. Additionally, there was some slight evidence that male children were at higher risk of being abused than female children. Equally interesting was that we found that gender inequality was predictive of levels of child abuse, while country level GDP was not.
3. In a study of parenting during Covid-19 (Lee et al. 2022), we expected to find that households with children would experience *higher* levels of anxiety and depression than households without children. Instead, we found the opposite. Being in a household with children was generally *protective* against anxiety and depression.

In Section 2.6, **?@sec-pvalues** and **?@sec-multilevelstructure**, I provide specific examples of how multilevel data provides even more opportunity to present answers that are *not* obvious.

2.3 Presenting Advanced Statistical Ideas

In presenting advanced, statistical concepts, one is faced with a quandary. One can present statistical concepts in the most general terms, in terms of x and y. While perhaps the mathematically most general way to present ideas, a highly general (and abstract) presentation risks not being a good way of teaching the ideas, as it is sometimes difficult to apply abstract ideas to one's own specific area of research.

Alternatively, one can present statistical ideas in terms of specific substantive concepts. The risk of making use of a specific substantive concept is that while concrete examples are always

helpful, it may be difficult for the reader to generalize from a specific example to their own area of research.

I ground this presentation in research that we have conducted on parenting and child development in international context (Grogan-Kaylor et al. 2021; Ward, Grogan-Kaylor, Pace, et al. 2021; Ward et al. 2022; Ward, Grogan-Kaylor, Ma, et al. 2021; Pace, Lee, and Grogan-Kaylor 2019; Ma et al. 2022). For the presentation in this document, I use simulated data on these issues.

Using the simulated data, I refer to *predictors* and *outcomes*, and explore the ways that the multilevel model can contribute to understanding how relationships between predictors and outcomes might be similar, or might be different, across *social contexts*. In the examples presented below, I focus on two predictors, parental *warmth*, and parental use of *physical punishment* and focus on the *outcome* of *improved* mental health. I use the social context of different *countries* in our example.

It is my belief that while I use this specific set of examples, that the idea of studying *families in different countries* is generalizable enough to a multiplicity of diverse contexts, such that the reader can apply these ideas to their own area of interest, whether that be *children in schools*; *residents in neighborhoods*; or *people in different countries*.

2.4 Research on Parenting and Child Development in International Context

Research on parenting and child development has identified robust associations between parenting behaviors and child developmental outcomes. Broadly speaking, physical punishment is associated with increases in child aggression, child anxiety and child mental health problems (Gershoff and Grogan-Kaylor 2016b), while warm and supportive parenting is associated with decreases in these outcomes (Rothenberg et al. 2022; Khaleque and Rohner 2002). However, much of this research is conducted on North American samples (Henrich, Heine, and Norenzayan 2010; Draper et al. 2022).

Barth and Olsen (2020) have argued, that children constitute a class of oppressed persons. If children are oppressed, then it is imperative to empirically determine what factors are promotive of children's well-being, and what factors constitute risk factors that contribute to decreases in children's well-being. Equally imperative–given the North American focus of so much research on parenting and child development (Henrich, Heine, and Norenzayan 2010; Draper et al. 2022)–would be efforts to extend the study of parenting and child development to a broader, more global context. As part of such a research agenda, it is necessary to have quantitative tools that are able to determine the consistency of relationships in parenting and child development. That is, are the relationships between certain forms of parenting and child developmental outcomes, largely consistent across countries, largely different across countries, or somewhere in between?

This document will discuss the ways in which a multilevel statistical perspective not only allows one to appropriately analyze cross cultural or international data, but also the ways in which a multilevel perspective affords the opportunity for more precise quantitative thinking about cross cultural phenomena.

This document takes a very pragmatic and very advocacy oriented approach to improving research on families and children, with the aim of improving the well-being of families and children.

> "It shouldn't be theories that define the problems of our situation, but rather the problems that demand, and so to speak, select, their own theorisation." – Martin-Baro (1998) in Burton and Kagan (2005).

Following from this pragmatic and advocacy oriented emphasis, the document is largely oriented to the *doing* of quantitative social research with multilevel (or multi-country) data, and is therefore mostly statistical in nature.

The document moves quickly into detailed statistical arguments. Some of these statistical discussions may seem very technical, or even overly technical. However, an overarching theme of the document is that multilevel data contains hidden complexities. A lack of awareness of the complexities of multilevel data—e.g. complexities of multi-country data—might lead to statistical analyses that point in the wrong direction: yielding false positives; false negatives; or substantively wrong conclusions.

2.5 Universalism And Particularity

The specific domain of cross-cultural research on parenting and child development raises more general questions in cross-cultural research of *universalism* and *particularity*. With regard to child development it is universal that all children need some amount of emotional and material care to grow into healthy youth and healthy adults (Kottak 2021). Further it is broadly understood that children should be protected from violence (UNICEF 2014). This broad consensus is manifested in such documents as the Convention on the Rights of the Child (United Nations General Assembly 1989) and the United Nations Sustainable Development Goals (United Nations 2022), representing global efforts to ensure the children are cared for, and are protected against violence.

At the same time, broad international efforts to improve children's well-being must engage with important considerations of cultural uniqueness. Put simply, what is considered to be beneficial for children in one country or culture may not be considered to be beneficial in all countries or cultures. Similarly, what is considered to be detrimental in one country or culture may not equally be considered to be detrimental in all. Within the area of parenting and child development, most of the debate has focused around the question of whether physical punishment is equally detrimental in all settings, particularly whether physical punishment is

detrimental in countries where it is especially common, or normative (Gershoff et al. 2010). Much less attention has been focused on the study of positive parenting internationally, and the degree to which the outcomes of positive parenting are consistent across countries remains understudied (Ward, Grogan-Kaylor, Ma, et al. 2021).

However, as global initiatives to improve child well-being and family life move forward, it becomes increasingly important to continue to collect internationally relevant data about parenting and child outcomes. If recommendations are to be made for policies, interventions, or treatments, such recommendations must be base on accurate balancing of that which is universal against that which is unique to particular cultural contexts. Thus it is necessary to employ statistical methods that are able to adequately and accurately analyze data across countries.

As I will outline below–and is evident in the literature (Raudenbush and Bryk 2002; Singer and Willett 2003; Rabe-Hesketh and Skrondal 2012; Luke 2004; Kreft and de Leeuw 1998)–multilevel models are eminently suited for cross-cultural research in that they are not only able to *control for* the clustering of study participants within countries, but are also able to *explore the variation*–or *consistency*–of patterns of family life across countries.

Long ago, Cesaire, writing about liberatory movements wrote...

> "My conception of the universal is that of a universal enriched by all that is particular, a universal enriched by every particular: the deepening and coexistence of all particulars." (Cesaire 1956)

It is this sensibility that I hope to echo in my discussion of the multilevel model below.

2.6 Multilevel Models As The Study Of Variation

> "Every being cries out silently to be read differently."

— Simone Weil, *Gravity and Grace* as reported in Su (2017)

Multilevel models are sometimes seen as an analytic technique that *controls for* the clustering or nesting of individuals inside larger social units such as schools, neighborhoods, or countries. I will describe below how this ability to *control for* clustering is indeed an important and crucial aspect of multilevel models.

However, my argument here is that multilevel models are better seen as a method to *explore* the variation in inherent within nested or clustered data. Again, while these issues are well understood within the statistical literature (Raudenbush and Bryk 2002; Singer and Willett 2003; Rabe-Hesketh and Skrondal 2012; Luke 2004; Kreft and de Leeuw 1998), they are less often noted in applied research.

In the graph below, imagine that physical punishment, or some other risk factor, is associated with detrimental mental health outcomes. Each country in the data has its own *country specific regression line.*

In Panel A, there is some variation in the *intercept*, which is equivalent to saying that there is some variation in the average level of psychological well-being across countries. When we look at the slope of the country-specific regression lines in Panel A, we notice that there is little variation in these *slopes*. Put another way, there is a great amount of consistency in the slopes of the country-specific regression lines: parental use of physical punishment is consistently associated with decreases in child psychological wellbeing across countries.

In Panel B, the situation is different. There is more variation in the *intercept*, that is, more variation between countries in the initial or average amount of psychological well-being. There is also more variation in the *slopes* of the country-specific regression lines. While the average association between physical punishment and psychological well-being is very similar to that in Panel A, there is more variation across countries, in the relationship of physical punishment and child psychological wellbeing, which would likely merit exploration were one considering developing programs, policies or interventions for different countries.

Lastly, the pattern of variation in Panel C is considerably different from either Panel A or Panel B. The average association of physical punishment with psychological well-being in the hypothetical scenario represented by Panel C is approximately 0. There is some variation in the *intercepts* of the country-specific regression lines. Additionally, there is considerable variation in the *slopes* of the country-specific regression line, suggesting that the use of physical punishment might be beneficial in some countries, and detrimental in others.

Empirically, data generally suggest a scenario somewhere between Panel A and Panel B, but these different hypothetical scenarios afford us the opportunity to think about possible patterns of variation.

A second pedagogically helpful example might be obtained if we flip the slopes in the diagram, and consider a different set of independent variables, perhaps some kind of treatment or intervention designed to improve psychological well-being.

We see a similar pattern as before, but the use of a different substantive example may be illustrative.

In Panel A, there is relative consistency in the initial levels of psychological well-being across countries, as well as consistency in the degree to which the intervention is associated with improvements in psychological well-being across countries.

In Panel B, we see more variation in both initial levels of psychological well-being, but also more variation in the association of the intervention with improvements in psychological well-being.

Lastly, in Panel C, we note an overall association of the intervention with psychological well-being that is close to zero. However associations vary widely by countries. In some countries
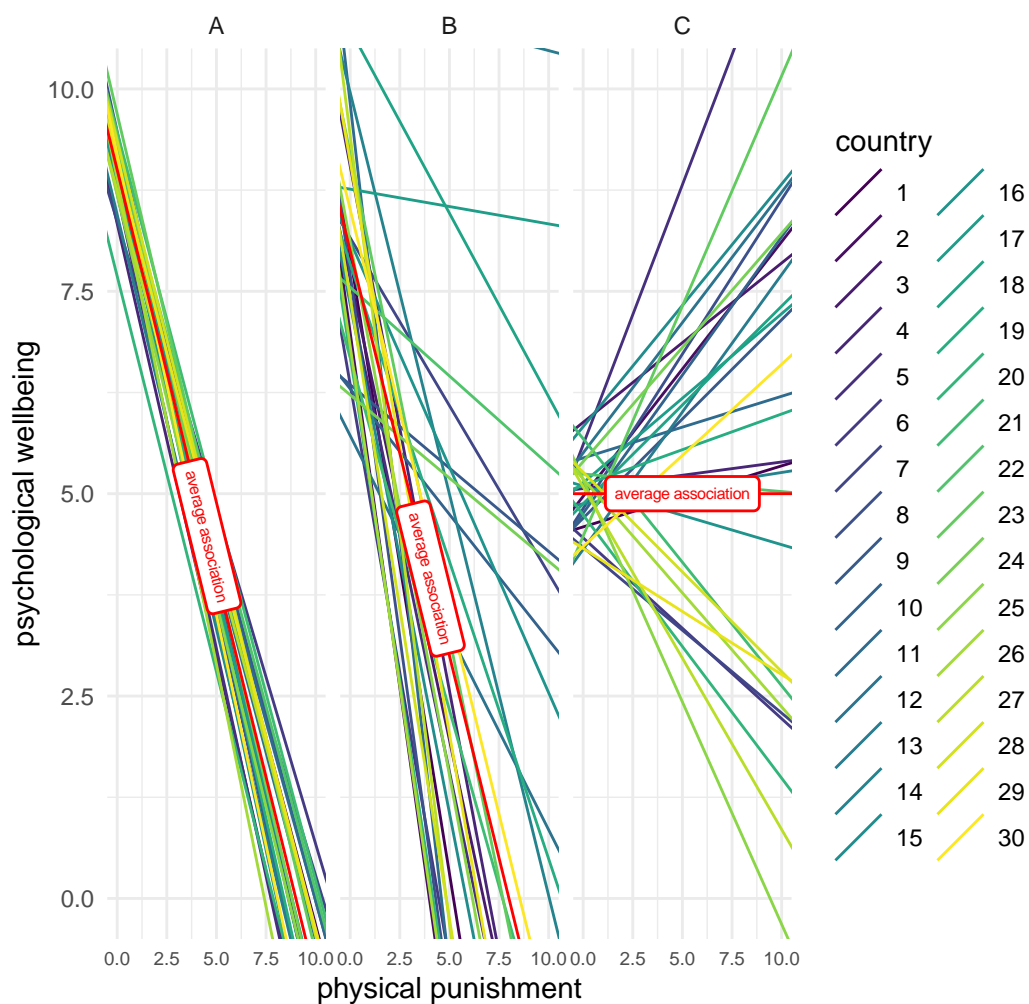
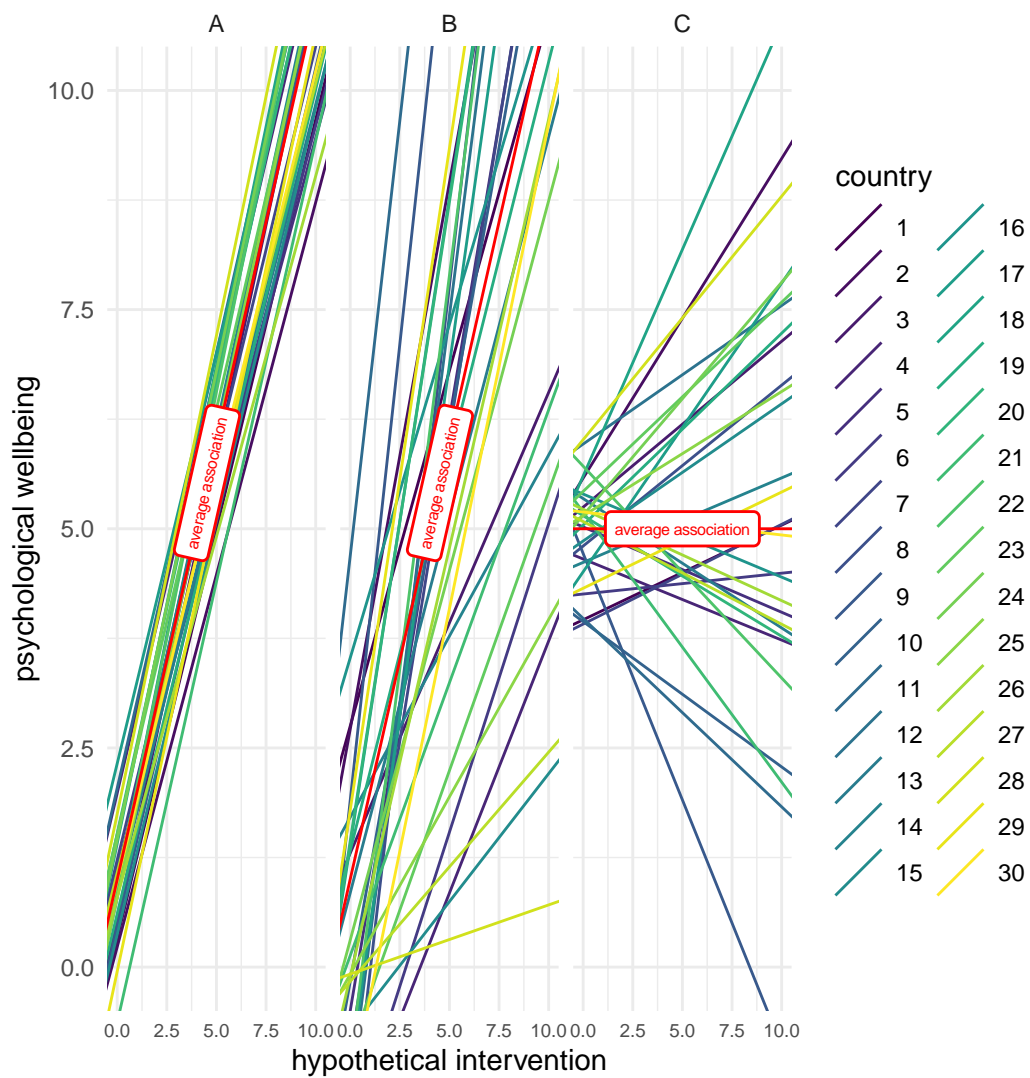Figure 2.1: Plausible Alternative Patterns of Between Country Variation

Figure 2.2: Considering an Intervention or Treatment Across Countries

there appears to be evidence that the intervention is beneficial, while in other countries there appears to be evidence that the intervention is not beneficial, or even possibly harmful.

Thus, I emphasize an approach to multilevel modeling that sees multilevel modeling as the *study of variation*, not simply *accounting for variation*, or *controlling for variation*.

> "…universal theorizing requires adequately sampled (i.e., diverse) data and better appreciation of issues of comparability and the most powerful theories ought to predict and explain variation, not sweep variation under the rug." (Blasi et al. 2022)

Again, sophisticated treatments of all of the ideas are available in one form or another across the excellent textbooks on multilevel modeling (Raudenbush and Bryk 2002; Singer and Willett 2003; Rabe-Hesketh and Skrondal 2012; Luke 2004; Kreft and de Leeuw 1998). However, some of these ideas appear less often in applied research, and my intention here is to make the application of these ideas to applied research more clear.

3 Simulated Multi-Country (Multilevel) Data

I use simulated data in this example. Data come from 30 hypothetical countries. Data contain measures of a few key aspects of parenting that have proven salient in the empirical literature on parenting to date: parental `warmth`, and `physical punishment`. Both parenting measures are normally distributed variables. Our `outcome` is conceptualized as a positive mental health outcome or behavioral outcome, and higher levels of `outcome` are considered to be better. Statistically, the data are clustered within countries.

In this simulation, I construct the data so that `warmth` is positively related to the `outcome`, while `physical punishment` is negatively related to the `outcome`.

Table 3.1: Simulated Multilevel Data

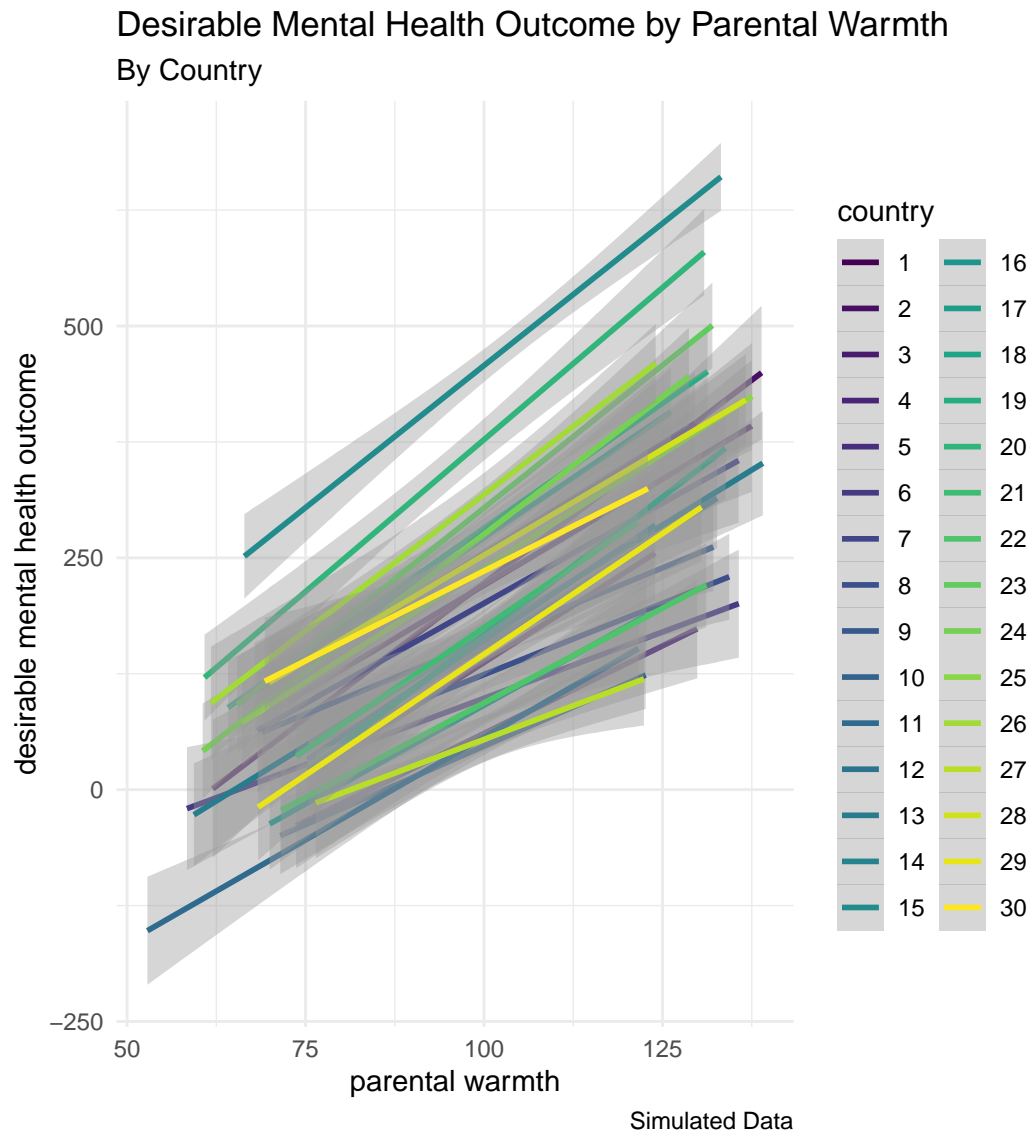| id	country	warmth	physical_punishment	outcome
1	15	129.6	84.86	685.3
2	24	106.2	90.91	330.5
3	11	83.41	95.94	-130.8
4	8	90.59	96	100.5
5	21	95.06	85.3	149.1
6	20	66.6	102.3	238.1

Figure 3.1: Graph of Simulated Data

References

Abelson, Robert P. 1995. *Statistics as Principled Argument. Statistics as Principled Argument.* Lawrence Erlbaum Associates, Inc.

Aron, Adrianne, and Shawn Corne. 1994. "Introduction." In *Writings for a Liberation Psychology*, edited by Adrianne Aron and Shawn Corne. Harvard University Press.

Barth, Richard P., and Amanda N. Olsen. 2020. "Are Children Oppressed? The Timely Importance of Answering This Question." *Children and Youth Services Review* 110: 104780. https://doi.org/https://doi.org/10.1016/j.childyouth.2020.104780.

Blasi, Dami'an E, Joseph Henrich, Evangelia Adamou, David Kemmerer, and Asifa Majid. 2022. "Over-Reliance on English Hinders Cognitive Science." *Trends in Cognitive Sciences*, October. https://doi.org/10.1016/j.tics.2022.09.015.

Burton, Mark, and Carolyn Kagan. 2005. "Liberation Social Psychology: Learning From Latin America Psychology of liberation: Learning from Latin America." *Journal of Community & Applied Social Psychology* 15 (January). https://doi.org/10.1002/casp.786.

Cesaire, Aime. 1956. "Letter to Maurice Thorez."

Cokley, Kevin, and Germine Awad. 2013. "In Defense of Quantitative Methods: Using the 'Master's Tools' to Promote Social Justice." *Journal for Social Action in Counseling & Psychology* 5 (May): 26. https://doi.org/10.33043/JSACP.5.2.26-41.

Deater-Deckard, Kirby, Kenneth A Dodge, John E Bates, and Gregory S Pettit. 1996. "Physical discipline among African American and European American mothers: Links to children's externalizing behaviors." *Developmental Psychology* 32 (6): 1065–72. https://doi.org/10.1037/0012-1649.32.6.1065.

Draper, Catherine E, Lisa M Barnett, Caylee J Cook, Jorge A Cuartas, Steven J Howard, Dana C McCoy, Rebecca Merkley, et al. 2022. "Publishing Child Development Research from Around the World: An Unfair Playing Field Resulting in Most of the World's Child Population Under-Represented in Research." *Infant and Child Development* n/a (October): e2375. https://doi.org/10.1002/icd.2375.

Eamon, Mary Keegan. 2001. "Poverty, Parenting, Peer, and Neighborhood Influences on Young Adolescent Antisocial Behavior." *Journal of Social Service Research* 28 (1): 1–23. https://doi.org/10.1300/J079v28n01_01.

Felitti, Vincent J, Robert F Anda, Dale Nordenberg, David F Williamson, Alison M Spitz, Valerie Edwards, Mary P Koss, and James S Marks. 1998. "Relationship of Childhood Abuse and Household Dysfunction to Many of the Leading Causes of Death in Adults: The Adverse Childhood Experiences (ACE) Study." *American Journal of Preventive Medicine* 14 (4): 245–58. https://doi.org/https://doi.org/10.1016/S0749-3797(98)00017-8.

Gelman, Andrew, Boris Shor, Joseph Bafumi, and David Park. 2007. "Rich State, Poor

State, Red State, Blue State: What's the Matter with Connecticut?" *Quarterly Journal of Political Science* 2 (November): 345–67. https://doi.org/10.2139/ssrn.1010426.

Gershoff, Elizabeth T., and Andrew Grogan-Kaylor. 2016a. "Race as a Moderator of Associations Between Spanking and Child Outcomes." *Family Relations* 65 (3): 490–501. https://doi.org/10.1111/fare.12205.

———. 2016b. "Spanking and Child Outcomes: Old Controversies and New Meta-Analyses." *Journal of Family Psychology* 30: 453–69. https://doi.org/10.1037/fam0000191.

Gershoff, Elizabeth T., Andrew Grogan-Kaylor, Jennifer E. Lansford, Lei Chang, Arnaldo Zelli, Kirby Deater-Deckard, and Kenneth A. Dodge. 2010. "Parent Discipline Practices in an International Sample: Associations with Child Behaviors and Moderation by Perceived Normativeness." *Child Development* 81: 487–502. https://doi.org/10.1111/j.1467-8624.2009.01409.x.

Grogan-Kaylor, Andrew, Berenice Castillo, Garrett T Pace, Kaitlin P Ward, Julie Ma, Shawna J Lee, and Heather Knauer. 2021. "Global perspectives on physical and nonphysical discipline: A Bayesian multilevel analysis." *International Journal of Behavioral Development*, January. https://doi.org/10.1177/0165025420981642.

Grogan-Kaylor, Andrew, Julie Ma, Shawna J. Lee, Berenice Castillo, Kaitlin P. Ward, and Sacha Klein. 2018. "Using Bayesian Analysis to Examine Associations Between Spanking and Child Externalizing Behavior Across Race and Ethnic Groups." *Child Abuse and Neglect* 86 (December): 257–66. https://doi.org/10.1016/j.chiabu.2018.10.009.

Henrich, Joseph, Steven J. Heine, and Ara Norenzayan. 2010. "The weirdest people in the world?" https://doi.org/10.1017/S0140525X0999152X.

Hines, Caitlin T, Ariel Kalil, and Rebecca M Ryan. 2022. "Differences in Parents' Attitudes Toward Spanking Across Socioeconomic Status and Region, 1986–2016." *Social Indicators Research* 160 (1): 133–58. https://doi.org/10.1007/s11205-021-02803-7.

Khaleque, Abdul, and Ronald P Rohner. 2002. "Perceived Parental Acceptance-Rejection and Psychological Adjustment: A Meta-Analysis of Cross-Cultural and Intracultural Studies." *Journal of Marriage and Family* 64 (February): 54–64. https://doi.org/https://doi.org/10.1111/j.1741-3737.2002.00054.x.

Kottak, Conrad. 2021. *Anthropology: Appreciating Human Diversity.* 19th ed. McGraw Hill.

Kreft, Ita, and Jan de Leeuw. 1998. *Introducing Multilevel Modeling.* SAGE Publications. https://doi.org/10.4135/9781849209366.

Lee, Shawna J, Kaitlin P Ward, Andrew Grogan-Kaylor, and Vijay Singh. 2022. "Anxiety and Depression During COVID-19: Are Adults in Households with Children Faring Worse?" *Journal of General Internal Medicine* 37: 1328–30. https://doi.org/10.1007/s11606-021-07256-9.

Luke, Douglas. 2004. *Multilevel Modeling.* SAGE Publications, Inc. https://doi.org/10.4135/9781412985147.

Ma, Julie, Andrew C. Grogan-Kaylor, Garrett T. Pace, Kaitlin P. Ward, and Shawna J. Lee. 2022. "The association between spanking and physical abuse of young children in 56 low- and middle-income countries." *Child Abuse & Neglect* 129 (July): 105662. https://doi.org/10.1016/j.chiabu.2022.105662.

Martin-Baro, Ignacio. 1994. "Public Opinion Research." In *Writings for a Liberation Psychology*, edited by Adrianne Aron and Shawn Corne. Harvard University Press.

———. 1998. "Retos y perspectivas de la psicología latinoamericana." In *Psicología de La Liberación*, edited by A. Blanco. Madrid, Spain: Trotta.

Pace, Garrett T., Shawna J. Lee, and Andrew Grogan-Kaylor. 2019. "Spanking and young children's socioemotional development in low- and middle-income countries." *Child Abuse and Neglect* 88: 84–95. https://doi.org/10.1016/j.chiabu.2018.11.003.

Proctor, Robert N. 2012. "The history of the discovery of the cigarette–lung cancer link: evidentiary traditions, corporate denial, global toll." *Tobacco Control* 21 (2): 87 LP–91. https://doi.org/10.1136/tobaccocontrol-2011-050338.

Rabe-Hesketh, Sophia, and Anders Skrondal. 2012. *Multilevel and Longitudinal Modeling Using Stata - Volume i: Continuous Responses. Stata Press.*

Raudenbush, Stephen W, and Anthony S Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Sage Publications.

Rothenberg, W Andrew, Sumbleen Ali, Ronald P Rohner, Jennifer E Lansford, Preston A Britner, Laura Di Giunta, Kenneth A Dodge, et al. 2022. "Effects of Parental Acceptance-Rejection on Children's Internalizing and Externalizing Behaviors: A Longitudinal, Multicultural Study." *Journal of Child and Family Studies* 31: 29–47. https://doi.org/10.1007/s10826-021-02072-5.

Singer, Judith D, and John B Willett. 2003. *Applied Longitudinal Data Analysis : Modeling Change and Event Occurrence. Applied Longitudinal Data Analysis : Modeling Change and Event Occurrence.* Oxford University Press.

Stage, Frances K, and Ryan S Wells. 2014. "Critical Quantitative Inquiry in Context." *New Directions for Institutional Research* 2013 (June): 1–7. https://doi.org/https://doi.org/10.1002/ir.20041.

Su, Francis Edward. 2017. "Mathematics for Human Flourishing." *The American Mathematical Monthly* 124 (June): 483–93. https://doi.org/10.4169/amer.math.monthly.124.6.483.

UNICEF. 2014. *Hidden in Plain Sign: A Statistical Analysis of Violence Against Children.* UNICEF.

———. 2021. "Multiple Indicator Cluster Surveys (MICS)." UNICEF. https://mics.unicef.org/.

United Nations. 2022. *Sustainable Development Goals.*

United Nations General Assembly. 1989. *Convention on the Rights of the Child.*

Ward, Kaitlin P, Andrew C Grogan-Kaylor, Garrett T Pace, Jorge Cuartas, and Shawna J. Lee. 2021. "A Multilevel Ecological Analysis of the Predictors of Spanking Across 65 Countries." *BMJ Open* 11 (e046075). https://doi.org/10.1136/bmjopen-2020-046075.

Ward, Kaitlin P, Andrew Grogan-Kaylor, Julie Ma, Garrett T Pace, and Shawna J Lee. 2021. "Associations Between 11 Parental Discipline Behaviors and Child Outcomes Across 60 Countries." PsyArXiv. https://doi.org/10.31234/osf.io/f5t8x.

Ward, Kaitlin P, Shawna J Lee, Andrew C Grogan-Kaylor, Julie Ma, and Garrett T Pace. 2022. "Patterns of Caregiver Aggressive and Nonaggressive Discipline Toward Young Children in Low- and Middle-Income Countries: A Latent Class Approach." *Child Abuse & Neglect* 128. https://doi.org/https://doi.org/10.1016/j.chiabu.2022.105606.

Watts, Duncan J. 2011. *Everything Is Obvious: *Once You Know the Answer.* Crown Business.

Wiest, Lynda R, Heidi J Higgins, and Janet Hart Frost. 2007. "Quantitative Literacy for Social Justice." *Equity & Excellence in Education* 40 (April): 47–55. https://doi.org/10.1080/10665680601079894.