CHAPTER 3.

# QUANTITATIVE PERSPECTIVES TO THE STUDY OF WRITING ACROSS THE LIFESPAN: A CONCEPTUAL OVERVIEW AND FOCUS ON STRUCTURAL EQUATION MODELING

**Matthew C. Zajic**
Teachers College, Columbia University

**Apryl L. Poch**
Duquesne University

As echoed throughout this edited collection, writing researchers are well aware of the complexities involved when adopting lifespan approaches to the study of written language. Writing researchers come from a wide array of fields (e.g., composition studies, rhetoric, psychology, education, and special education) that adopt different methodological approaches to answer a variety of research questions. A central issue to unpacking the complexities underlying the development of written language across the lifespan requires examining the available tools and methods offered by different research designs to pose and answer different types of research questions.

Typically, research approaches are categorized as quantitative or qualitative. Quantitative approaches generally focus on numbers (e.g., counting frequencies or measuring the associations between different skills) and reduce complex phenomena into measurable instances that can be analyzed using statistical analysis (Creswell & Creswell, 2018; Gelo et al., 2008). Qualitative approaches usually collect non-numerical data (e.g., texts, visuals, graphics, videos, or transcripts from interview and focus groups) that can be analyzed using inductive, interpretative analytical approaches (Creswell & Creswell, 2018; Gelo et al., 2008). These two approaches are often contrasted against each other as deductive vs. inductive, hypothesis-testing vs. hypothesis-generating, prediction vs. interpre-

DOI: https://doi.org/10.37514/PER-B.2020.1053.2.03

tation, generalization vs. contextualization, and explanation vs. comprehension (Creswell & Creswell, 2018; Gelo et al., 2008; Haig, 2013; Todd et al., 2004). Yet, juxtapositions aside, both approaches contribute to lifespan development research (Menard, 2008), suggesting that both approaches might inform ongoing lifespan writing research.

In this chapter, we offer a conceptual overview to quantitative research approaches with a focus on quantitative longitudinal research designs. While quantitative approaches will not answer all questions pertaining to lifespan writing development (see Bazerman, 2018), they are able to address many questions about how writing develops across the lifespan and, in some cases, consider research questions that qualitative approaches cannot. Furthermore, developmental methodologists have long applied quantitative approaches to issues concerning lifespan development (McArdle, 2010; Menard, 2008). We hope this chapter provides lifespan writing researchers with a starting point for mobilizing such methods to meet their research needs and a greater understanding of what such methods bring to lifespan writing research. Throughout the chapter, we emphasize conceptual understanding over technical jargon, especially as encouraging conceptual understanding fosters long-term statistical literacy (Harlow, 2013).

The first section broadly overviews quantitative longitudinal research approaches by drawing from recent longitudinal design frameworks (e.g., Bauer & Curran, 2019). In the next section, we introduce the concepts underlying structural equation modeling (SEM), a statistical, theory-driven framework commonly used to address both cross-sectional and longitudinal research questions. We discuss foundational SEM issues and provide examples based in writing research. The final section discusses longitudinal SEM through specific application of two different types of statistical models—autoregressive models and latent growth curve models (LGCMs).

## LONGITUDINAL QUANTITATIVE RESEARCH DESIGNS

Longitudinal quantitative research designs and their analytical choices are guided by collecting numerical data and fitting statistical models. These statistical models are informed by theory, the research questions asked, and the types of data collected to make inferences about populations based on representative sample data. Scientific fields also use statistical models for theory building and for exploring the relationships between different variables using predictive, explanatory, and descriptive approaches (Shmueli, 2010). Singer and Willet (2003) offer a non-technical description of what statistical models aim to represent:

> Statistical models are mathematical representations of population behavior; they describe salient features of the hypoth-

> esized process of interest among individuals in the target population. When you use a particular statistical model to analyze a particular set of data, you implicitly declare that this population model gave rise to these sample data. Statistical models are not statements about sample behavior; they are statements about the population process that generated the data. (Singer & Willet, 2003, p. 46)

Though statistical models underlie many quantitative approaches, the appropriate analytical approach differs based on the specific research questions asked and the sets of data collected. A robust set of available designs, methods, and tools fit under the umbrella of quantitative methods, and the following provides a categorized overview of different longitudinal research designs and some associated methodological approaches. Guided by the categorical approach taken by Bauer and Curran (2019), the remainder of this section introduces longitudinal quantitative approaches by focusing on two types of longitudinal data: time-to-event data and repeated measures data. These approaches can be useful to writing researchers to address research questions that may focus on whether or not an event occurred (e.g., a memorable writing experience during a particular time period of life); when it occurred (e.g., when do memorable writing experiences occur in postsecondary education?); and when changes occur in specific behaviors, attitudes, or feelings over time (e.g., does writing anxiety or apprehension change across time and context?).

## Time-to-Event Data

Research questions based on time-to-event data focus on evaluating whether a particular event happens or when that event might take place. One way of addressing these questions is survival analysis (also named event history analysis, failure time analysis, hazard analysis, transition analysis, and duration analysis), a collection of flexible statistical methods specifically for describing, explaining, and predicting the timing and occurrence of events (Allison, 2010, 2019). For example, researchers might be interested in understanding more about when an event occurs, such as when students begin formal cursive or typing instruction in schools or when individuals first start writing via social media or instant messaging.

The event of interest falls at the center and focal point of survival analysis and is generally a qualitative change that occurs at some specific, observed point in time (Allison, 2010). This event may be simply observed and require little additional formal operationalization (e.g., the purchase and subsequent use of a cellular phone or other device for text or instant messaging), it may re-

quire considering multiple criteria to determine the exact timing (e.g., development of emergent literacy skills based on multiple accounts of different reading and writing behaviors), or it may require considering underlying quantitative variables to better specify the event occurrence (e.g., high social media writing activity may be contextualized by looking at the amount of writing being done across different social media platforms). Additionally, if events can happen multiple times, further consideration can be made regarding which occurrence to focus on (i.e., the first occurrence or a later occurrence) and to what degree events can be considered similar (i.e., can two events be qualitatively similar but differ in degree of their impact on some additional factor?) (Allison, 2010). Furthermore, survival analysis requires specifying a given interval of time for an event to have occurred, and these intervals may be specified given the research questions but also may vary given the underlying interests of the questions (Allison, 2010, 2019). For example, if researchers were interested in modeling the event of the first meaningful writing experience undergraduate students have during their postsecondary education, they might specify the origin point as the start date of students' first quarter or semester at college. However, if researchers were interested in understanding to what degree a meaningful writing experience preceded college entry or if their postsecondary experience was related to an event prior to college entry, then an earlier origin point may need to be considered. Further consideration should be made if concerns about censoring—that is, when an event fails to occur or demonstrates an unknown event time—are warranted and if the presence of censoring might influence the data collection or analysis.

Methods for survival analysis differ depending upon how much a researcher knows about the timing of an event. If the exact timing of an event is known, then continuous-time methods are appropriate (i.e., time is treated as continuous when the occurrence of events is known with a very high rate of precision). These methods may be appropriate for examining questions pertaining to occurrences of events during specific writing activities or the tracking of daily writing habits. However, if the timing of an event is only coarsely known (i.e., in months or years rather than seconds or days), then discrete-time methods are more appropriate (i.e., time is not continuous, and the events are known with lower rates of precision). These methods may be more appropriate for answering questions about an event that takes place over longer periods of time, such as the likelihood of first enrolling in a writing in the disciplines course in postsecondary education. Though the difference between the two methods appears to be a matter of conceptual semantics (as days may sound coarse for one event but precise for another), the selection of continuous versus discrete time methods has methodological implications for treating, analyzing, and interpreting the data (Singer &

Willet, 2003). One approach to distinguishing between the two sets of analyses entails looking to the number of ties within the data (i.e., an occurrence of two individuals experiencing an event at the same recorded time; see Allison, 2010). Discrete-time methods are better designed for handling high rates of ties, as the presence of ties has an extremely low occurrence rate under continuous-time methods (Singer & Willet, 2003). Consequently, survival analysis is a useful approach for when your research questions pass the "whether and when test": If your research questions include either word—whether or when—you probably need to use survival methods (Singer & Willet, 2003).

## Repeated Measures Data

While research questions involving time-to-event data approaches focus on whether or when a particular event takes place, research questions involving repeated measures data focus instead on evaluating how abilities change over specific time periods. Like time-to-event data approaches, numerous methods and frameworks exist from which to study questions related to change in abilities over time. Bauer and Curran (2019) group approaches using repeated measures data into categories that depend upon the intensity of data collection (i.e., the number of times data are collected) and the number of units (i.e., abilities and items) being collected, resulting in three overarching research design categories: 1) time series analysis (intensive data collection involving few units); 2) intensive longitudinal data (intensive data collection involving many units); and 3) panel data (non-intensive data collection involving many units).

Both intensive data collection designs entail assessing one or more units on a high number of occurrences over a specified duration. The number of units included in these data collection points differentiates the focal point of these two approaches. Time series focuses on prediction or forecasting of a particular outcome and makes use of prior observations to predict expected change in the outcome at future time points. Intensive longitudinal data maintains a similar level of intensity but includes more units of data beyond a single outcome of interest to collect data on a higher number of individuals. Intensive longitudinal studies can include recording study units over many time points, often into the tens, hundreds, or thousands (see Walls & Schafer, 2006; Walls, 2013). Examples of such approaches include daily rating scales that might include self-reported ratings about different types of writing behaviors, like types of writing activities (e.g., text messaging or journal writing) or feelings about writing (e.g., instances of writing apprehension or motivation).

However, when only a handful of time points are included in a research design, a non-intensive panel data approach is most commonly used. This panel

data approach is often what researchers imagine when thinking more generally about longitudinal data collection (Bauer & Curran, 2019) and formed the basis for initial longitudinal research rationales (Baltes & Nesselroade, 1979). Panel data are often collected across a small number of time points on a relatively large number of units with the ultimate goal of describing change over time. As with previous models, the topic of change is another nuanced concept, as different frameworks exist for considering mean-level change versus individual-level change, both of which are briefly discussed next.

**Mean-level Change**

Examining mean-level change puts the focus on group-level average change for a specific outcome over multiple time points. For example, researchers might be interested in the extent to which handwriting abilities change from preschool through secondary school. As such, these approaches draw on marginal models (Heagerty & Zeger, 2000) that estimate linear mean-level change (i.e., repeated measures analysis of variance, repeated measures multivariate analysis of variance, and analysis of covariance) and non-linear mean-level change (i.e., generalized estimating equations).

**Individual-level Change**

Researchers are often interested in examining not only how change may happen across specific groups but also to what extent individuals demonstrate variability in their change over time. Individual-level change is often first understood as a simple model that includes three or more time points to estimate a unique starting point (the intercept) and the trend of change over the remaining time points (the slope). But examining differences in individual-level change over time requires another choice about what the underlying question is regarding the nature of change in individuals for a given outcome: Are the differences in change due to differences of degree (i.e., quantitative variation) or differences in kind (i.e., qualitative differences between different change trajectories) (Bauer & Curran, 2019)? Differences by degree to investigate quantitative individual-level differences include approaches like multilevel models, mixed effects models, and LGCMs. Differences in kind to investigate qualitative individual differences include approaches like growth mixture modeling. Furthermore, additional models like general growth mixture models allow for exploration of differences by degree simultaneously with differences in kind when researchers are interested in examining research questions related to both degree and kind of differences. See Bauer and Curran (2019); Hoyle (2012); Little et al. (2000); and Muthén and Shedden (1999) for more thorough overviews.

### Applications of Longitudinal Modeling Approaches for Lifespan Writing Researchers

Writing researchers interested in lifespan writing development may develop a variety of research questions requiring the use of time-to-event data or repeated-measures data approaches. Bazerman et al. (2018) offer a range of potential conceptual ideas to apply such methodological approaches, particularly as the complexity of the underlying factors involved with writing development are dynamic and not expected to develop in a rapid, linear fashion. Specifically, they emphasized that "writing needs time to mature, in fact decades, though at various moments motivated writers may make rapid progress on some dimensions. When and where those moments occur, however, may be hard to predict" (2018, p. 378). Survival analysis may be one useful approach to take to understand when these moments occur and how difficult they might be to predict among different writers. As survival analysis can take into consideration additional predictors of these moments (not discussed at length here; see Allison, 2010, 2019 and Singer & Willet, 2003), researchers can explore what factors may predict these experiences across distinct lifespan segments.

On the other hand, researchers might also use these moments to predict different types of writing outcomes. Graham's (2018) writer(s)-within-community model offers a range of potentially impactful factors that underlie writing development across different writing contexts. Though the theoretical basis for these factors is established, further empirical work is needed to examine how different underlying profiles based off these factors affect writing development differently over time and to what extent individual change in writing abilities may be measured using operationalized approaches to both the sociocultural and cognitive components of writing. Such an emphasis on connecting theory with data is at the heart of SEM, which we discuss next.

## SEM: A FLEXIBLE STATISTICAL FRAMEWORK

SEM is a flexible statistical modeling approach that allows for rigorous examination of specified hypotheses connected to research questions about both cross-sectional and longitudinal research designs, and it is applicable across a wide array of disciplines (Hoyle, 2012; McArdle & Nesselroade, 2014). SEM is not a single statistical technique or model. It is an analytical process that covers several related procedures that posit multiple structural equations (i.e., mathematical statements that represent the strength and nature of specified, hypothetical relationships among sets of variables) to depict relationships and effects between observed and unobserved (or latent) variables (Hoyle, 2012;

Kline, 2016; Mueller & Hancock, 2019; Newsom, 2015). In other words, SEM provides a framework for answering theory-based, researcher-specified questions about relationships between abilities measured at a single time point or across multiple time points. As a model-based approach, it confers unique advantages to researchers of various fields that are unavailable in more traditional statistical techniques. Furthermore, though its capabilities have become more advanced in recent years, SEM can still serve as a framework for conducting other well-known univariate analyses (e.g., t-test, analysis of variance, regression, and multiple regression) and several multivariate analyses (e.g., path analysis and confirmatory factor analysis) (Grimm et al., 2017).

## Some Necessary Terminology: Model Parameters, Covariances, and Latent Variables

Though our focus falls on the conceptual understanding of SEM, in order to clarify how SEM estimates relationships and effects between different variables, we briefly cover three statistical terms: model parameters, covariances, and latent variables. While SEM includes many other technical terms, these terms specifically cover some of the core terminology used across different SEM approaches.

First, a model parameter is a component of a statistical model that is generally not known to the researcher (i.e., a component that can be estimated) that can represent information about the relationships or effects between variables in that model (Raykov & Marcoulides, 2006). Parameters are not specific to SEM research, as parameters often generally reflect unknown aspects of statistical models that represent the phenomena under investigation. The goal of SEM is to estimate these parameters to answer underlying questions and hypotheses about the constructs under investigation (Raykov & Marcoulides, 2006). For example, in order to examine the relationship between two variables, a researcher would need to estimate their association given available data about those two variables.

Second, a covariance is a measure of the joint variance (the amount of shared variability present) between two (or more) variables that represents the strength of the linear association between variables and their variabilities (Kline, 2016). SEM primarily analyzes the variance-covariance matrix for a given dataset (i.e., a matrix that contains all of the variances and covariances of included variables). The importance of the variance-covariance matrix can be further seen by how SEM is often referred to as covariance structure analysis, covariance structure modeling, or analysis of covariance structures (Hoyle, 2012; Kline, 2016). However, non-covariance-based SEM approaches do exist, including latent class analysis (i.e., analysis of mixture models that contain exclusively observed categorical variables for a latent variable).

Third, a latent variable is an underlying characteristic that cannot be observed or measured directly, and instead requires at least one observed variable to estimate it (Bollen & Hoyle, 2012). Other terms for latent variables in the literature include unmeasured variables, latent factors, unobserved variables, or constructs, all generally meaning that they represent variables that are not immediately identifiable within a given dataset or that cannot be directly observed (or measured) from a sample of a specific population (Bollen, 2002; Raykov & Marcoulides, 2006). Latent variables can be both *a priori* and *a posteriori* and can be considered continuous, categorical, or hybrid depending on whether it is the presence of the latent trait that is the focus of the theory or if the latent trait has multiple gradations (Bollen, 2002). SEM has the capacity to measure relationships between multiple variables, regardless of whether they are observed or latent, while accounting for measurement error that is not accounted for when only investigating observed abilities (Raykov & Marcoulides, 2006). SEM accomplishes this by quantifying and removing the measurement error from the measurement of the latent variable, while simultaneously investigating relationships between distinct observed and unobserved abilities (Lei & Wu, 2007). Some examples of latent abilities include personality, attitudes, motives, emotions, and reading, as each are often measured using multiple observable measures to represent an underlying hypothetical construct (Bollen & Hoyle, 2012). Writing ability can similarly be thought of as a latent variable, as writing assessments are often tools that researchers use to make inferences about unobservable writing abilities. Multiple data points or assessments of specific writing abilities may provide a better estimate of an individual's unobservable writing ability, as assessments may capture different, smaller components of the larger unobservable ability.

## Steps to Implementing SEM

Different research designs require researchers to collect and analyze data using often very different approaches (see Gelo et al., 2008). SEM follows the research traditions of quantitative methodology, but different SEM approaches are used to answer different types of research questions, which requires different types of statistical models (see McArdle & Kadlec, 2013; Mueller & Hancock, 2019). However, most SEM approaches follow a similar overarching implementation framework, as described by Hoyle (2012). This brief-but-thorough overview of the implementation framework follows data acquisition and data preparation to include four required steps (specification, estimation, fit evaluation, and interpretation and reporting of findings) and an often required fifth step (respecification). We review these next in order to provide context to the types of questions researchers face when implementing SEM.

First, *specification* begins with a model (the formal statement positing the relationships to be explored within the given data) designed using theory-driven hypotheses. Model selection requires considering the different types of models available that would best fit the data collected (and how these data are related to the intended research hypotheses). Within the model, the researcher designates what variables will be included (both observed and latent variables) along with the relationships between the variables (i.e., unidirectional, allowed to covary, or unrelated) and their parameters (either fixed to a specific value or free to be estimated by information provided within the model). Specifying parameters requires attention to the need for model identification (i.e., all parameters must be identified by either being fixed or free, which is dependent on the number of observed variables included in the model).

Second, *estimates* are provided for parameters that are specified to be freely estimated as opposed to parameters that are fixed to specific values. The goal of estimation is to establish a model that minimizes the residual differences between the observed and the estimated covariance matrices given by the data and the model. Multiple estimation methods can be used depending on the characteristics of the data (e.g., the scales of the variables, distributional assumptions, and missing data), though most are iterative (i.e., they begin with one set of starting values for all free parameters and search for values that reduce the discrepancy between the model and the data).

Third, *fit evaluation* assesses how well the generated model represents the data by taking into consideration the discrepancy between the observed and implied covariance matrices. If fit appears poor (i.e., there is a large discrepancy in the covariance matrix) or is misspecified, then the model may be discarded or respecified (meaning that a new model may be generated to test a different underlying hypothesis; see fifth step). Different fit tests provide various fit statistics to make decisions regarding both absolute fit and comparative fit (i.e., how well one model fits in relation to other tested models) as well as corrections for parsimony.

Fourth and fifth are *interpretation and reporting of findings* and *respecification*, though the order in which researchers engage these steps depends on the results from fit evaluation. If a model does not demonstrate good fit, then respecification may be necessary to shift the focus to an exploratory approach to assess if alternative models may be better suited to the data. Choosing when to pursue respecification and what fit evaluation statistics to consider when deciding to move into either respecification or interpretation is a highly debated topic that cannot be thoroughly addressed in this brief overview. Nonetheless, readers need to be aware that the choice to consider respecification of a poorly fitting model requires a specific, theory-driven rationale.

If evaluation results in support for the originally specified model (or a respecified model), then interpreting and reporting the findings are done based on the stated hypotheses. Core components requiring interpretation include the basis for the model, the inclusion of and findings for specific parameters in the model, and how well the model accounts for uniqueness (i.e., variance) in the observed data. The way in which findings are interpreted depends upon the approaches taken during the implementation framework, as interpretation of a theory-driven model is more straightforward and meaningful than potentially uncertain rationales underlying exploratory models. Theoretical perspectives or previous empirical work should drive these interpretations, and researchers may need to consult further equivalent models (i.e., models that appear identical to the given model in terms of fit but include estimated parameters that contradict the chosen model).

This implementation framework outlines the overarching steps that researchers follow when using SEM to address specific research questions. Across these steps, researchers must make determinations (grounded in theory and empirical research) about their analyses beyond solely inputting numbers into a statistical program. Doing so allows researchers to understand better the relationships (or associations) between and among variables. However, while this framework briefly touched on some more technical aspects around planning and navigating the use of SEM, it did not adequately cover many of the technical decisions that researchers must make during the process (see Hoyle, 2012, and McArdle & Kadlec, 2013, for further technical discussions).

## Examples of Writing Research using SEM Approaches

SEM is not a novel technique to writing research, yet only a few studies have addressed research questions using cross-sectional SEM research designs with different groups of individuals. Parkin et al. (2020) modeled the effects of an oral language latent factor on different level of language factors (including writing) and evaluated the effects of lower language levels on higher levels of language. In doing so, they found that a psychoeducational assessment demonstrated expected theoretically driven relationships that showed some variability in the relationships between language levels when comparing students in general and special education. De Smedt et al. (2018) investigated gender and achievement effects within the context of how cognitive and motivational challenges mediate and correlate with students' writing performance across different groups of students (boys and girls, and low, average, and high achievers). Their results highlight group-level differences in the relationships between these skills and suggest that research take into consideration different learner characteristics when considering how these skills relate to and predict writing skills.

Numerous other studies have adopted SEM approaches to examine writing skill development predominantly across the school-age years. Kim and colleagues examined if data from kindergarten and first grade students supported the theoretical relationships between writing, oral language, reading, and cognitive abilities (Kim et al., 2011; Kim, Al Otaiba, Puranik, Folsom, & Greulich, 2014; Kim & Schatschneider, 2017). Limpo and colleagues examined relationships among transcription, higher-order writing processes, and writing performance in middle school students (Limpo et al., 2017) and compared relationships between transcription and self-regulation in late elementary and secondary students (Limpo & Alves, 2013). Berninger and colleagues investigated the relationships between writing with other language skills in typically developing writers (Abbott & Berninger, 1993; Berninger, Abbott et al., 2002; Graham et al., 1997) and writers with specific learning disabilities (Berninger, Nielsen et al., 2008; Nagy et al., 2003). Each of these studies generally sought to examine if theoretically driven questions about writing skills held for other skills among different groups when examined using highly specified modeling approaches. In all, they sought to examine if data supported the theoretically held beliefs about the relationships between writing skills and related linguistic, cognitive, and social cognitive skills. Though some studies included multiple samples from different age groups, these examples all discussed data collected from cross-sectional research designs.

## SEM APPLICATIONS FOR LONGITUDINAL DATA ON LIFESPAN WRITING

Writing, like many skills, does not simply develop at one point in time. Writing skills are shaped across time and context. Understanding the ways in which individuals develop and apply these skills over time is a focal point of interest to lifespan writing researchers. In addition to its flexibility for analyzing cross-sectional data, SEM can be equally useful and appropriate for analyzing longitudinal data. As with cross-sectional SEM, one of the goals of longitudinal SEM is to identify models composed of a minimal number of estimated parameters that fit the data well, ideally with the intention of making predictions about future actions of individuals and groups of individuals or that identify sample characteristics associated with the development of a construct. SEM is a powerful tool for researchers interested in modeling the relationships between observed and latent skills over time (see Wu et al., 2013), and many different analytical tools are available to researchers interested in modeling longitudinal data (e.g., Grimm et al., 2017; Little, 2013; Little et al., 2000; McArdle & Nesselroade, 2014; Newsom, 2015). From the available modeling approaches, we selected two approaches we consider to be foundational to SEM that serve as illustrative

introductions to the concepts underlying longitudinal SEM: autoregressive longitudinal models and LGCMs. Though presented separately, it is important to note that many SEM approaches for longitudinal data can incorporate features from both autoregressive models and LGCMs (Bollen & Curran, 2004; Curran & Bollen, 2001). However, for simplicity, we introduce and discuss them separately.

To assist with understanding how relationships between variables are modeled in the autoregressive model and LGCM examples, path diagrams are provided for each example (Figures 3.1–3.3). Though SEM is often represented using mathematical equations, path diagrams can also be used to visually depict these relationships (Ho et al., 2012; Little, 2013). Different path diagram components (i.e., parameter estimates and variables) are often labeled and named using Greek letters to convey their functions, though naming conventions can often differ. Little (2013) provides a cheat sheet for some of the commonly used Greek letters, and the conventions used for diagrams in this chapter draw from Little (2013) and Ho et al. (2012). The use of path diagrams was a deliberate choice for this chapter in order to visually depict the modeled relationships rather than rely on matrix algebra and mathematical equations, but path diagrams do not always provide as much detail as these mathematical representations. As cautioned by many methodologists (e.g., Kline, 2016; Little, 2013; McArdle, 2012; Mueller & Hancock, 2019), path diagrams are not a substitute for the equations they seek to represent, and researchers should be prepared to learn more about the mathematics underlying SEM after understanding the concepts (e.g., Harlow, 2013).

## Autoregressive Longitudinal Models

Are writing skills at one point in time predictive of writing skills at later points in time? Are specific writing skills predictive of other writing skills at different points in time? Questions specific to examining the degree that skills are predictive of themselves or other skills across time are well suited for autoregressive longitudinal models, a modeling approach used across disciplines for decades to investigate the relationships among specific variables over multiple time points (Biesanz, 2012; Little, 2013).

Autoregressive models conceptualize that performance at a specified time point is a function of earlier assessments of that variable plus new unique error that occurs with each time point (McArdle & Bell, 2000). Put differently, autoregressive models investigate the extent to which a future value for some variable is predicted from previous estimates of that variable. (Furthermore, *regressive* refers to the direct linear pathways between variables across time points,

and *auto* refers to the pathways between the same variables across timepoints.) Even with only one observed variable across multiple time points, this variable can be modeled as either only an observed variable or as a latent variable based on its observed variable, with the benefit of treating the variable of interest as latent to account for measurement error in the overall model (Biesanz, 2012). Autoregressive models can be useful to examine not only the predictive relationships within a single variable but also the cross-lagged relationships between multiple variables (i.e., the degree to which different variables can covary with or predict each other across multiple time points; Biesanz, 2012). Such cross-lagged approaches allow for temporal precedence in data collection to help assess for causal relations rather than correlational relations, as the cross-lagged specification sets up the framework for identifying causal relationships between abilities measured across multiple time points (Biesanz, 2012).

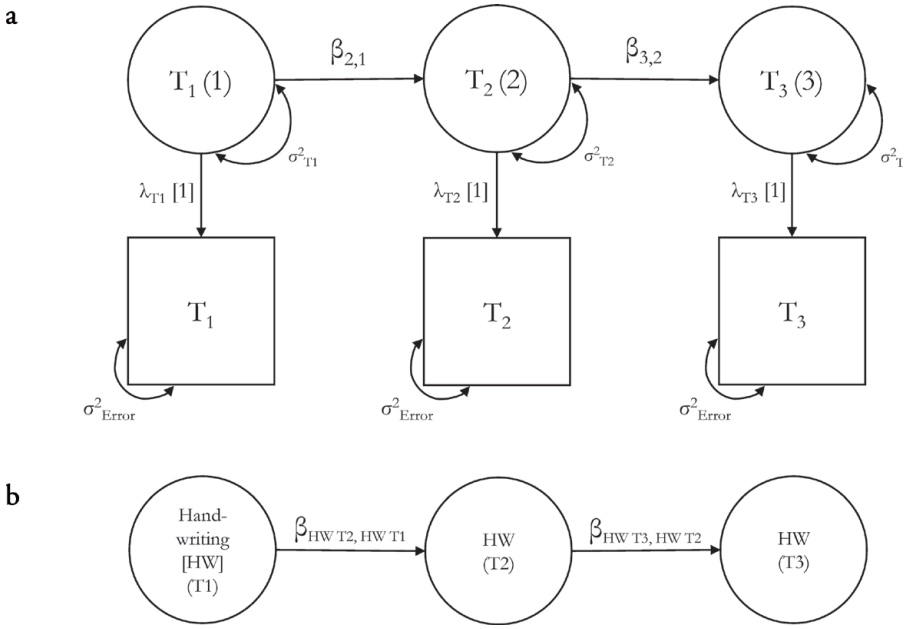


***Figure 3.1.*** *a) Path diagram example of an autoregressive model of one ability measured over three time points ($T_1$-$T_3$). Squares (□) represent the observed variables while circles (○) represent the unobserved latent variables. Curved, double-headed lines (↔) represent variances. Straight, single-headed lines (→) are directed, regressive relationships between observed or unobserved variables. See text for further information about specific parameter labels. b) Path diagram of handwriting skills assessed over three time points (showing only latent variable and autoregressive parameters).*

Path diagrams shown in Figures 3.1 and 3.2 represent a first-order, single measure autoregressive longitudinal model and a cross-lagged, dual measure autoregressive longitudinal model, respectively. In Figure 3.1.a, an ability is depicted as having been assessed across three separate time points ($T_1$-$T_3$). Each latent variable represents the time-point-specific ability of interest, which is accounted for by the observed variable and the unaccounted-for error (or variance). In this example, the time-point-specific latent variables are based off a single observed variable (and the relationship between these is set to 1, as the observed variable is functioning as an indicator for the latent variable that is not freely estimated). The focal point of Figure 3.1.a falls on the regressive parameters between time points (i.e., $\beta_{2,1}$ and $\beta_{3,2}$), as these represent the stability of individual differences across the two adjacent time points. For any given time point, the performance of a variable of interest is the product of this regressive parameter, its value at the earlier time point ($T_{1\text{-}3}$), and its unexplained variance ($\sigma^2_{1\text{-}3}$) and unaccounted-for error ($\sigma^2_{Error}$). Figure 3.1.b shows what this path model would look like if applied to the measurement of handwriting skills measured over three time points. For simplicity in this example, we have included visual representations of the latent variables and regressive parameters only.

In Figure 3.2.a, there are now two different abilities assessed at each time point ($T_{1,A}$-$T_{3,A}$ and $T_{1,B}$-$T_{3,B}$), and the focus falls on both the predictive association within variables (autoregressive parameters) and the predictive associations across variables between timepoints (cross-lagged parameters). The regressive parameters depicting the relationships between the same variable at different time points (i.e., $\beta_{B2,B1}$, $\beta_{B3,B2}$, $\beta_{A2,A1}$, and $\beta_{A3,A2}$) can be interpreted as was done with the regressive parameters shown in Figure 3.1 (i.e., they represent the stability of individual differences across the two adjacent time points for that variable). However, the cross-lagged regressive parameters focus on the relationships between the two different abilities across time points, as these parameters (i.e., $\beta_{B2,A1}$, $\beta_{B3,A2}$, $\beta_{A2,B1}$, and $\beta_{A3,B2}$) represent the predictive relationship of one variable assessed at an earlier time point on the second variable assessed at a later time point (while controlling for the first variable). Additionally, the covariances between the unexplained variance (i.e., $\sigma_{\sigma^2_{T1B,T1A}}$, $\sigma_{\sigma^2_{T2B,T2A}}$, $\sigma_{\sigma^2_{T3B,T3A}}$) capture the extent to which changes in one variable are associated with changes in the other variable for that given time point. Figure 3.2.b shows what this path model would look like if applied to the measurement of handwriting and spelling skills. Again, for simplicity, we have included only visual representations of the latent variables, regressive parameters (both autoregressive and cross-lagged), and the covariance parameter between $T_1$ skills.

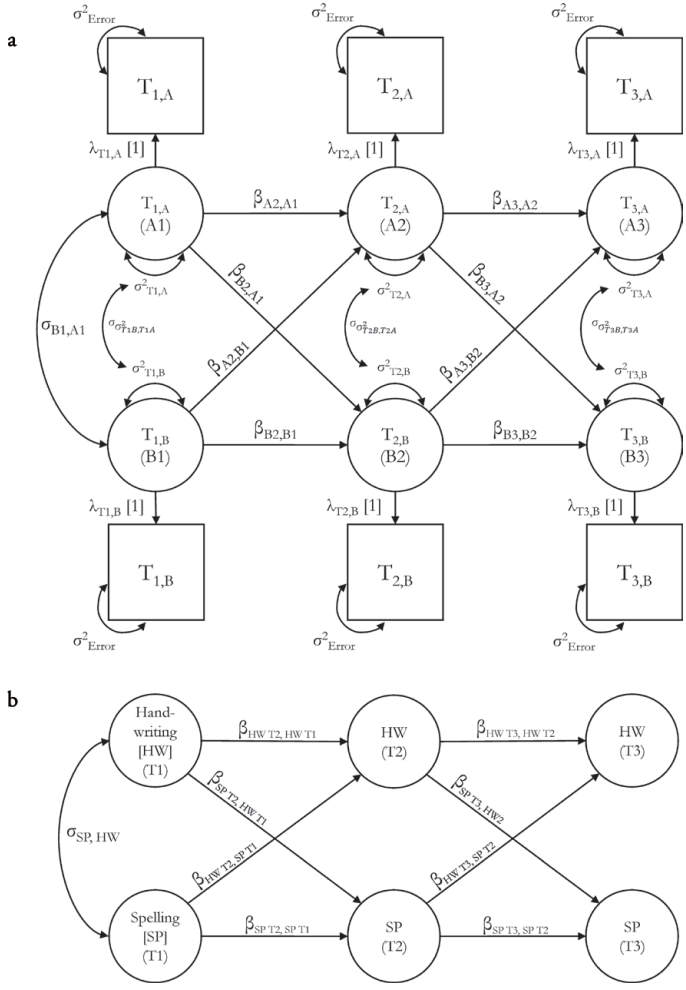


***Figure 3.2.*** *a) Path diagram example of a cross-lagged, autoregressive model of two abilities (A and B) measured over three time points ($T_1$-$T_3$). Squares (□) represent the observed variables while circles (○) represent the unobserved latent variables. Curved, double-headed lines (↔) can represent either variances (if they start and end within the same square or circle) or covariances (if they start and end on different squares or circles). Straight, single-headed lines (→) are directed, regressive relationships between observed or unobserved variables. See text for further information about specific parameter labels. b) Path diagram of a cross-lagged, autoregressive model of handwriting and spelling skills assessed simultaneously across three time points. Only the latent variables, the autoregressive and cross-lagged parameters, and the covariance parameter between $T_1$ skills are shown.*

**Using Autoregressive Models for Writing Research**

Not many studies have used autoregressive models to analyze longitudinal writing development. One example by Abbott et al. (2010) adopted the multiple levels of language theory to examine the relationships within writing (autoregressive) and between writing and reading (cross-lagged) using an overlapping cohort design that included students in first grade through seventh grade.

The authors examined the longitudinal development of five measures (handwriting, spelling, word reading, text composition, and reading comprehension) in two cohorts of students using three different models. Model 1 analyzed three measures—handwriting, spelling, and text composition—with both specified autoregressive and cross-lagged parameters. Model 2 analyzed three measures—handwriting, spelling, and word reading—with both specified autoregressive and cross-lagged parameters between time points. Model 3 analyzed four measures—word reading, spelling, text composition, and reading comprehension—with both specified autoregressive and cross-lagged parameters between time points. As such, each model included three distinct types of paths: a) between-measure correlations for each grade level, b) within-measure autoregressive longitudinal paths between adjacent grades, and c) longitudinal cross-lagged paths between each measure for each set of measures at adjacent time points. Additionally, the authors used observed rather than latent variables due to minimal measurement error and a high degree of measure reliability (Abbott et al., 2010, p. 286).

The authors reported results for both standalone autoregressive models and autoregressive models with additional cross-lagged components. For the autoregressive models, the authors reported that individual differences across measures appeared consistently associated longitudinally between adjacent years from grades 1 to 7. Additionally, they found that the magnitude (i.e., the strength) of the associations differed upon the level of language, in descending order from word-level (spelling and word reading), text-level (reading comprehension and text composition), and subword-level measures (handwriting).

The models that included specifications for both the cross-lagged and autoregressive parameters demonstrated better model fit (i.e., better represented the data) than models with only the autoregressive parameters specified. Model 1 estimates highlighted some stability in measure-specific individual differences across grade levels with some unreliable longitudinal relationships between certain skills (e.g., handwriting with spelling and text composition) and unexpected reliable relationships between other skills (e.g., spelling and composition). Model 2 estimates highlighted consistent measure-specific individual differences across grade levels for handwriting, spelling, and word

reading; a significant association (though small) between spelling and word reading; and no relationship between word reading and handwriting. Model 3 estimates highlighted consistent measure-specific individual differences and associations between different measures (e.g., spelling and word reading) across grade levels similar to those observed in Model 2 as well as new findings for consistent measure-specific individual differences (e.g., text composition and reading comprehension), associations between different measures (e.g., spelling and word reading; word reading and text comprehension), and no relationships between other measures (e.g., reading comprehension and spelling; text composition and word reading).

Abbott et al. (2010) provides one example as to how autoregressive models can be beneficial to longitudinal writing research. Their study focused explicitly on modeling the relationships within and between specific writing and reading skills across time to consider if data supported the multiple levels of language theory. Their findings offered a comprehensive examination of the relationships between multiple skills associated within writing and across reading and writing. These relationships highlighted not only the importance of multiple subskills within writing but also the extent to which different levels of language appear related across writing and reading domains at adjacent time points across the elementary and secondary school years (Abbott et al., 2010). However, this application of longitudinal SEM is but one of numerous approaches available to researchers.

## Latent Growth Curve Models (LGCMs)

Autoregressive models highlight relations between multiple variables over time but do not emphasize information about individual- or group-level performance. What if, instead, our research questions focused on the trajectories of change in writing skills over time? What if we wanted to model overall change between scores and ask questions about whether this change is related to an individual's initial skill level or to their growth in writing skills over time? LGCMs represent a different class of models that focus on the extent to which individuals demonstrate change in specific abilities over time rather than solely performance-related bidirectional effects. The LGCM framework allows for evaluating hypotheses specific to between-person differences in within-person change and goes by many different names (e.g., multilevel models of change, latent trajectory analysis, latent curve modeling, and mixed effects or random effects models of change) (Shiyko et al., 2012).

LGCMs treat multiple observed time points of the same variable as indicators of (usually) two latent constructs that represent how individuals change

over time. These latent factors include an intercept (i.e., the ability level at a single time point of interest and often the first time data were collected) and a slope (i.e., the change in an individual's ability over time). LGCMs use multiple time-point trajectories produced by individuals across different time points to provide a parsimonious representation of these trajectories via description of the average change trajectories and the degree to which inter-individual differences (i.e., between-person differences) in change occur. LGCMs allow for exploring a variety of different types of research questions related to the growth individuals demonstrate in a given ability measured over multiple time points and can range from simple to more complex models. LGCMs offer the flexibility of SEM with the advantage of modeling a variety of different random effects (e.g., means, variances, and covariances of individual differences for both the intercept and the slope) (Preacher, 2019).

Figure 3.3 depicts a path diagram of a simple LGCM representing the assessment of one ability across four time points. While the path diagram may share some visual similarities to the autoregressive models shown in Figures 3.1 and 3.2, the LGCM path diagram contains important distinctions. Working from the bottom of the diagram, the observed variables and associated errors are no different from those depicted in Figures 3.1 and 3.2 (i.e., a skill is measured four times). However, the two latent variables (labeled Intercept and Slope) are conceptually different from the latent variables depicted in Figures 3.1 and 3.2. Both latent variables contain unidirectional paths to each of the observed variables. Each path is labeled with a lower-case Lambda (λ) with subscripts to differentiate between different paths (e.g., $\lambda_{4,1}$ represents the path for the fourth observed variable for the intercept, while $\lambda_{4,2}$ represents the path for the fourth observed variable for the slope). The fixed values in brackets for this illustration represent what these path parameters would be set to when estimating a simple LGCM. While the fixing of each intercept path parameter to 1 follows the same rationale as used with the autoregressive models (in that each observed variable is acting as an indicator for the latent variable that is not freely estimated), the rationale behind fixing the slope paths is slightly different. In this example, each slope path parameter is fixed to a value between zero and three based on the time parameter (i.e., the fixed value represents the order of the time points beginning with zero as the first time point). These values may be fixed in this manner or left free to be estimated from the data (e.g., to assess for nonlinear growth, then these values would be freely estimated or partially fixed to allow for nonlinear estimation of time point slopes). The triangle represents that the initial intercept and slope values are assumed to be latent variables with fixed means ($\mu_{Intercept}$ and $\mu_{Slope}$) but random variances ($\sigma^2_{Intercept}$ and $\sigma^2_{Slope}$) and covariances ($\sigma_{Slope,\ Intercept}$).

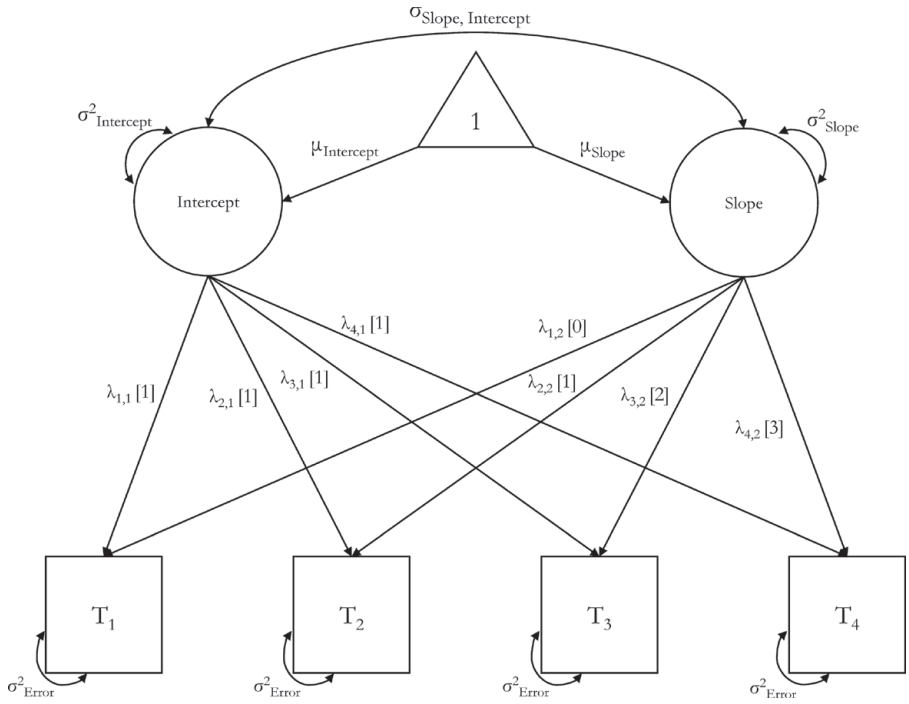


***Figure 3.3.*** *Path diagram example of a simple LCGM of one ability measured over four time points ($T_1$-$T_4$). Squares (□) represent the observed variables while circles (○) represent the unobserved latent variables. A triangle (△) represents a constant value. Curved, double-headed lines (↔) can represent either variances (if they start and end within the same square or circle) or covariances (if they start and end on different squares or circles). Straight, single-headed lines (→) are directed, regressive relationships between observed or unobserved variables. See text for further information about specific parameter labels.*

Bollen and Curran (2006) offer three guiding questions to assist with thinking about research questions involving trajectories of change for a given sample of individuals for a given skill. First, *what is the trajectory of the entire group?* This initial question seeks to characterize the entirety of the dataset and does not consider potential subgroups or other distinctions within the data. This approach is needed to help understand what potential underlying trajectories exist for the entire dataset before more specific questions are asked. Second, *are distinct trajectories needed for each case?* This question requires considering how subgroups may demonstrate trajectories different from the overall average trajectory. By accounting for subgrouping factors, potential distinctions between different in-

dividuals are considered alongside the differences observed within individuals across time. Third, *are there additional variables that can be used to predict individual trajectories?* After establishing both an average trajectory and the presence of meaningful variation in this overall average trajectory (particularly in terms of both the intercept and slope), what other information could be useful for predicting these observed distinctions? This approach takes into consideration additional information that may be meaningful to predicting and understanding why distinct trajectory patterns exist within a given dataset for a construct measured multiple times.

LGCMs further offer the capability of examining research questions related to a wide range of different types of longitudinally oriented research questions (Grimm et al., 2017). LGCMs deal with investigating individuals in the same abilities over a number of distinct time points, allowing for investigations into intra-individual differences (i.e., how do individuals change across time points with respect to this ability?). LGCMs allow for testing different modeling approaches that assume different patterns of change for observed abilities and provide a structured approach to investigating such inter-individual differences within the context of earlier mentioned intra-individual changes (Grimm et al., 2017). As changes in multiple constructs can occur both simultaneously and sequentially, approaching questions about these inter-relationships requires simultaneous analysis of multiple variables alongside evaluations of how variables may precede, covary, and/or follow changes observed in another variable (Grimm et al., 2017). Furthermore, the flexibility of the SEM framework allows for considering different predictors for both intra-individual change and inter-individual differences in intra-individual change, such as allowing for the inclusion of multiple groups or the specification of time-invariant covariates (i.e., variables that occur at specific points in time that are included at only specific time points rather than reassessed at multiple time points).

### Using LGCMs in Writing Research

Similar to autoregressive models, LGCMs have not yet been widely adopted for analyzing longitudinal writing development. However, Costa et al. (this volume) provide an investigation into the growth trajectories of written language and executive functions in 205 elementary-age children across first through fourth grade using LGCMs. Interested readers are directed to Chapter 10 for the full study, but what is of interest to this chapter is their consideration regarding key issues of model fit and estimation. Though their findings suggested interesting results relevant to the relationships between individual variability in written language (i.e., spelling, alphabet knowledge, and writing fluency) and executive functions (i.e., attentional control and planning) over time, they highlight par-

ticular concerns about model convergence and model fit, both of which showcase the complexity of issues that can occur when interpreting results from SEM. Even with their stated limitations, Costa et al. (this volume) present a worthwhile approach to using LGCMs for analyzing longitudinal writing data.

## CONCLUSION

As highlighted by the Lifespan Writing Development Group (Bazerman et al., 2018), writing develops in a complicated, multifaceted process that should not be expected to happen rapidly or linearly. To address research questions related to the development of writing and writing-related abilities across the lifespan, researchers need to be diverse in their questions and their methodologies. This chapter offered a review of longitudinal quantitative approaches that writing researchers may draw on to investigate writing development across the lifespan. In doing so, we highlighted the application of SEM as a comprehensive framework whose structure provides researchers with the means to answer quantitatively oriented research questions via deductive, theory-driven, hypothesis-testing, and predictive approaches.

SEM and longitudinal quantitative research approaches more generally provide lifespan writing researchers with valuable tools to test and answer research questions about how writing develops and changes from early development through late adulthood across cognitive and social contexts. Quantitative methods have a longstanding history and continue to guide much of the methodological foundations across a wide array of fields within the social sciences (see Haig, 2013). The consideration for the broader role of quantitative methods—as well as the specific role of advanced approaches like SEM—provides frameworks to researchers interested in analyzing data collected over many different types of research designs. Broader longitudinal quantitative approaches aligned with time-to-event data and repeated measures data allow researchers to postulate and analyze an array of questions ranging from the importance of carefully defined events on later development to ways in which skills predict performance in similar or associated skills over any set period of time. SEM further provides researchers with a robust framework to specify carefully articulated research questions about data based on theoretical beliefs (and provides researchers with the capabilities of considering nuances in the data that are only covered briefly here) (see Kline, 2016 and Wu et al., 2013).

We hope this chapter has sparked an interest in readers from different disciplines with different methodological backgrounds in the multitude of roles that longitudinal quantitative approaches may have to help answer questions about the development of writing across the lifespan. However, quantitative methods

do have their limitations and may not always be the best approach to take. Though SEM can be used for causal hypotheses, not all SEM approaches are causal in design, which can lead to overconfidence in data interpretation (see Hoyle, 2012; Jöreskog, 1993; or Kline, 2016). Additionally, the use of SEM does not magically transform correlational data into causal conclusions. Findings must be replicated across multiple datasets to avoid the capitalization of chance factors that might have been due to specific features of a dataset rather than the constructs under investigation (Raykov & Marcoulides, 2006). Furthermore, numerous researchers have called for additional considerations into the role of mixed methods in ongoing interdisciplinary research. These researchers argue that fewer distinctions may exist between quantitative and qualitative methods than many believe (see Haig, 2013) or that research designs may be strengthened by taking novel approaches that consider a wide array of methodologies (see Gelo et al., 2008 and Todd et al., 2004). Lifespan writing researchers should consider novel techniques across different approaches that may best answer their research questions and should build from findings across different lines of inquiry in the general pursuit of better understanding the ongoing development of writing abilities across the lifespan.

## ACKNOWLEDGMENTS

Matthew C. Zajic received support during the drafting of this chapter from a Postdoctoral Research Fellow Training Program Grant in Education and Autism Spectrum Disorders from the National Center for Special Education Research at the Institute of Education Sciences (R324B180034) awarded to the University of Virginia.

## REFERENCES

Abbott, R. D., & Berninger, V. W. (1993). Structural equation modeling of relationships among developmental skills and writing skills in primary-and intermediate-grade writers. *Journal of Educational Psychology*, *85*(3), 478-508. https://doi.org/10.1037/0022-0663.85.3.478

Abbott, R. D., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology*, *102*(2), 281-298. https://doi.org/10.1037/a0019318

Allison, P. D. (2010). Survival analysis. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 413-424). Taylor & Francis.

Allison, P. D. (2019). Event history and survival analysis. In G. R. Hancock, L, M. Stapleton, & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (2nd ed.) (pp. 86-97). Taylor & Francis.

Baltes, P. B., & Nesselroade, J. R. (1979). History and rationale of longitudinal research. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 1-39). Academic Press.

Bauer, D. J., & Curran, P. J. (2019). *Conducting longitudinal data analysis: Knowing what to do and learning how to do it* [PowerPoint slides]. https://curranbauer.org/how-do-you-choose-the-best-longitudinal-data-analytic-method-for-testing-your-research-questions/

Bazerman, C. (2018). Lifespan longitudinal studies of writing development: A heuristic for an impossible dream. In C. Bazerman, A. N. Applebee, V. W. Berninger, D. Brandt, S. Graham, J. V. Jeffery, P. K. Matsuda, S. Murphy, D. W. Rowe, M. Schleppegrell, & K. C. Wilcox (Eds.), *The lifespan development of writing* (pp. 326-368). National Council of Teachers of English.

Bazerman, C., Applebee, A. N., Berninger, V. W., Brandt, D., Graham, S., Jeffery, J. V., Matsuda, P. K., Murphy, S., Rowe, D. W., Schleppegrell, M., & Wilcox, K. C. (2018). *The lifespan development of writing*. National Council of Teachers of English.

Berninger, V. W., Abbott, R. D., Abbott, S. P., Graham, S., & Richards, T. (2002). Writing and reading: Connections between language by hand and language by eye. *Journal of Learning Disabilities*, *35*(1), 39-56. https://doi.org/10.1177/002221940203500104

Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Writing problems in developmental dyslexia: Under-recognized and under-treated. *Journal of School Psychology*, *46*(1), 1-21. https://doi.org/10.1016/j.jsp.2006.11.008

Biesanz, J. C. (2012). Autoregressive longitudinal models. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 459-471). The Guilford Press.

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*, *53*(1), 605-634. https://doi.org/10.1146/annurev.psych.53.100901.135239

Bollen, K. A., & Curran, P. J. (2004). Autoregressive latent trajectory (ALT) models a synthesis of two traditions. *Sociological Methods & Research, 32*(2), 336-383. https://doi.org/10.1177/0049124103260222

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. John Wiley & Sons.

Bollen, K. A., & Hoyle, R. (2012). Latent variables in structural equation modeling. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 56-67). Guilford.

Creswell, J. W., & Creswell, D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage.

Curran, P. J., & Bollen, K. A. (2001). The best of both worlds: Combining autoregressive and latent curve models. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 107-135). American Psychological Association.

De Smedt, F., Merchie, E., Barendse, M., Rosseel, Y., De Naeghel, J., & Van Keer, H. (2018). Cognitive and motivational challenges in writing: Studying the relation with writing performance across students' gender and achievement level. *Reading Research Quarterly*, *53*(2), 249-272. https://doi.org/10.1002/rrq.193

Gelo, O., Braakmann, D., & Benetka, G. (2008). Quantitative and qualitative research: Beyond the debate. *Integrative Psychological & Behavioral Science, 42*(3), 266-290. https://doi.org/10.1007/s12124-008-9078-3

Graham S. (2018). A writer(s)-within-community model of writing. In C. Bazerman, A. N. Applebee, V. W. Berninger, D. Brandt, S. Graham, J. V. Jeffery, P. K. Matsuda, S. Murphy, D. W. Rowe, M. Schleppegrell, & K. C. Wilcox (Eds.), *The lifespan development of writing* (pp. 272-325). National Council of Teachers of English.

Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology*, *89*(1), 170-182. https://doi.org/10.1037/0022-0663.89.1.170

Grimm, K., Ram, N., & Estabrook, R. (2017). *Growth modeling: Structural equation and multilevel modeling approaches*. The Guilford Press.

Haig, B. D. (2013). The philosophy of quantitative methods. In T. D. Little (Ed.), *The Oxford handbook of quantitative methods* (Vol. 1, pp. 7-31). Oxford University Press.

Harlow, L. L. (2013). Teaching quantitative psychology. In T. D. Little (Ed.), *The Oxford handbook of quantitative methods* (Vol. 1, pp. 105-117). Oxford University Press.

Heagerty, P. J., & Zeger, S. L. (2000). Marginalized multilevel models and likelihood inference. *Statistical Science, 15*(1), 1-26. https://doi.org/10.1214/ss/1009212671

Ho, M. R., Stark, S., & Chernyshenko, O. (2012). Graphical representation of structural equation models using path diagrams. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling (*pp. 43-55). The Guilford Press.

Hoyle, R. (2012). *Handbook of structural equation modeling*. The Guilford Press.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Lang (Eds.), *Testing structural equation models* (pp. 294-316). Sage.

Kim, Y. S., Al Otaiba, S., Puranik, C., Folsom, J. S., & Greulich, L. (2014). The contributions of vocabulary and letter writing automaticity to word reading and spelling for kindergartners. *Reading and Writing, 27*(2), 237-253. https:/doi.org/10.1007/s11145-013-9440-9

Kim, Y. S., Al Otaiba, S., Puranik, C., Folsom, J. S., Greulich, L., & Wagner, R. K. (2011). Componential skills of beginning writing: An exploratory study. *Learning and Individual Differences, 21*(5), 517-525. https://doi.org/10.1016/j.lindif.2011.06.004

Kim, Y. S., & Schatschneider, C. (2017). Expanding the developmental models of writing: A direct and indirect effects model of developmental writing (DIEW). *Journal of Educational Psychology, 109*(1), 35-50. https://doi.org/10.1037/edu0000129

Kline, R. B. (2016). *Principles and practice of structural equation modeling (4th ed.).* The Guilford Press.

Lei, P. W., & Wu, Q. (2007). Introduction to structural equation modeling: Issues and practical considerations. *Educational Measurement: Issues and Practice*, *26*(3), 33-43. https://doi.org/10.1111/j.1745-3992.2007.00099.x

Limpo, T., & Alves, R. A. (2013). Modeling writing development: Contribution of transcription and self-regulation to Portuguese students' text generation quali-

ty. *Journal of Educational Psychology*, *105*(2), 401-413. https://doi.org/10.1037/a0031391

Limpo, T., Alves, R. A., & Connelly, V. (2017). Examining the transcription-writing link: Effects of handwriting fluency and spelling accuracy on writing performance via planning and translating in middle grades. *Learning and Individual Differences*, *53*, 26-36. https://doi.org/10.1016/j.lindif.2016.11.004

Little, T. D. (2013). *Longitudinal structural equation modeling*. The Guilford Press.

Little, T. D., Schnabel, K., & Baumert, J. (2000). Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples. Lawrence Erlbaum Associates.

McArdle, J. J. (2010). What life-span data do we really need?. In R. L. Lerner & W. F. Overton (Eds.), *The handbook of life-span development* (Vol. 1, pp. 55-88). John Wiley & Sons.

McArdle, J. J. (2012). Latent curve modeling of longitudinal growth data. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 547-570). The Guilford Press.

McArdle, J. J., & Bell, R. Q. (2000). An introduction to latent growth models for developmental data analysis. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 69-107, 269-281). Lawrence Erlbaum Associates.

McArdle, J. J., & Kadlec, K. M. (2013). Structural equation models. In T. D. Little (Ed.), *The Oxford handbook of quantitative methods* (Vol. 2, pp. 295-337). Oxford University Press.

McArdle, J. J., & Nesselroade, J. (2014). *Longitudinal data analysis using structural equation models*. American Psychological Association.

Menard, S. (2008). *Handbook of longitudinal research design and analysis*. Elsevier.

Mueller, R. O., & Hancock, G. R. (2019). Structural equation modeling. In G. R. Hancock, L, M. Stapleton, & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (2nd ed., pp. 455-456). Taylor & Francis.

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, *55*(2), 463-469. https://doi.org/10.1111/j.0006-341X.1999.00463.x

Nagy, W., Berninger, V., Abbott, R., Vaughan, K., & Vermeulen, K. (2003). Relationship of morphology and other language skills to literacy skills in at-risk second-grade readers and at-risk fourth-grade writers. *Journal of Educational Psychology*, *95*(4), 730-742. https://doi.org/10.1037/0022-0663.95.4.730

Newsom, J. (2015). *Longitudinal structural equation modeling: A comprehensive introduction*. Taylor & Francis.

Parkin, J. R., Frisby, C. L., & Wang, Z. (2020). Operationalizing the simple view of writing with the Wechsler Individual Achievement Test (3rd ed.). *Contemporary School Psychology, 24*, 68-79. https://doi.org/10.1007/s40688-019-00246-z

Preacher, K. J. (2019). Latent growth curve models. In G. R. Hancock, L. M. Stapleton, & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (2nd ed., pp. 178-192). Taylor & Francis.

Raykov, T., & Marcoulides, G. A. (2006). *A first course in structural equation modeling* (2nd ed.). Lawrence Erlbaum Associates.

Shiyko, M. P., Ram, N., & Grimm, K. J. (2012). An overview of growth mixture modeling: A simple nonlinear application in OpenMx. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 532-546). The Guilford Press.

Shmueli, G. (2010). To explain or to predict? *Statistical Science, 25*(3), 289-310. https://doi.org/10.1214/10-STS330

Singer, J. & Willett, J. (2003). Applied longitudinal data analysis: Modeling change and event occurrence. Oxford University Press.

Todd, Z., Nerlich, B., & McKeown, S. (2004). Introduction. In Z. Todd, B. Nerlich, S. McKeown, & D. Clarke (Eds.), *Mixing methods in psychology: The integration of qualitative and quantitative methods in theory and practice*. Psychology Press.

Walls, T. A. (2013). Intensive longitudinal data. In T. D. Little (Ed.), *The Oxford handbook of quantitative methods* (Vol. 2, pp. 432-440). Oxford University Press.

Walls, T. A., & Schafer, J. L. (2006). *Models for intensive longitudinal data.* Oxford University Press.

Wu, W., Selig, J. P., & Little, T. D. (2013). Longitudinal data analysis. In T. D. Little (Ed.), *The Oxford handbook of quantitative methods* (Vol. 2, pp. 387-410). Oxford University Press.