

CHAPTER ONE

Cross-Sectional, Longitudinal, and Retrospective Designs: General Utility and Threats to Validity

Ekjyot K. Saini and Douglas M. Teti

*Department of Human Development and Family Studies, The Pennsylvania State University,
University Park, PA, USA*

The fundamental task of developmental science is to understand change over time. Baltes, Reese, and Nesselroade (1977) defined developmental science as the study of “intraindividual change in behavior and interindividual differences in such change.” Indeed, an understanding of the developmental function is often rooted in classic conceptions defined by Wohlwill (1970, 1973) where variables could be considered “developmental” when changes occur with age in a uniform and consistent manner across individuals and environments. This definition of developmental change is also related to the function and shape of age rather than simply age-related differences in specific behaviors or outcomes.

Bronfenbrenner and Evans (2000) expanded upon earlier conceptualizations of developmental science to emphasize the underlying processes and conditions that account for change and individual differences therein. The term “change” must also take into consideration whether a construct under study is continuous (i.e., conceptualized and measured in the same manner across two or more points in time) or discontinuous (i.e., when changes take place in the underlying structure of a

construct with development). When a construct itself develops, it may need to be measured differently at later points in time compared to earlier (Overton & Reese, 1981). Such fundamental discontinuity in construct composition is common in development, particularly during periods of rapid growth. A classic example of such a construct is intelligence, which traditionally had been conceptualized and operationally defined differently in infancy vs. at later points in development as infants move from sensorimotor modes of functioning to more overtly representational modes of functioning (Rose, Feldman, & Jankowski, 2009). When constructs themselves undergo reconstruction with age, the issue of assessing stability/instability over time, and predicting a developmental outcome from theoretically relevant predictors assessed at early points in time, becomes decidedly more complex. In such cases, assessments of stability/instability may require identifying an underlying component or components common to a construct at both developmental time points. In the case of infant intelligence, those components were speed of habituation, visual recognition, and other indices of information processing efficiency that, when assessed in early infancy, were found to predict children's scores on psychometric assessments of intelligence in the preschool years much more strongly than traditional assessments of infant sensorimotor functioning (Slater, 1995).

Developmental research has primarily been conducted with two methodological designs, with emergent iterations of these methods. *Longitudinal* methods make use of repeated measurements over a prescribed period of time. Often the primary purpose of longitudinal studies is to uncover associations between age-related changes in outcomes of interest and to identify factors that can theoretically account for individual differences in change. *Cross-sectional* methods, by contrast, involve measurement of differing ages simultaneously in order to assess age group differences in variables of interest, such as behavior and cognitive capacities. Both longitudinal and cross-sectional designs get incorporated in Schaie's (1965) presentation of the "general developmental model," which conceptualizes development along three individual dimensions: age, cohort, and time (see below). This chapter will provide an overview of research designs as informed by the general developmental model (Schaie, 1965) as well as discussion of variations or hybrids on these designs. A discussion of design and sampling strategies, and special considerations when working with marginalized populations, will follow.

The General Developmental Model

K. Warner Schaie (1965) laid out what has proven to be a useful organizational framework for addressing developmental change. Schaie's (1965) general developmental model identified three key parameters – *age*, *cohort*, and *time* of measurement, as inherent to any approach used to study development. *Age* is often indexed as chronological age; however, key issues concerning this variable are discussed

further below. *Cohort* is defined as a group of individuals or a population who experiences an event or a series of events that may define them. The most frequent cohort-defining event is the birth of an individual. *Time* is examined in a multitude of ways but is usually defined according to calendar time; however, other methods of assessment are available and often utilized.

Age

Age is the most commonly utilized variable in developmental research and is often thought to be “the central marker of development in biological and psychological research on developmental phenomena” (Bergman & Magnusson, 1990, p. 26). Though chronological age is heavily examined as an independent variable in developmental research, it does not always map onto processes indicative of maturation, social, and psychological aging (Hertzog, 1996). Developmental processes can indeed be linked to age-related maturation, but care should be taken not to use age as a proxy for processes and events/experiences that may be occurring independent of age (Miller, 1998). Age is also confounded with cohort and time of measurement, and thus, a longitudinal study, in which a single group is followed over time, cannot fully tease apart the effects of age from the effects of cohort and time of measurement on the developmental process of interest.

Cohort

Cohort is an integral variable in the study of development, and arguably, cohort effects have not been given the attention they deserve in understanding developmental trajectories. At best, cohort effects reflect truly underlying developmental change, although cohort differences are not a substitute for actual evidence of longitudinal change. In a traditional cross-sectional design, in which different birth cohorts are studied at the same time to examine age group differences in a developmental outcome, cohort effects are fully confounded with age and time of measurement effects. Cohorts that are simply defined by time of birth may not automatically indicate that members of the cohort have shared the same experiences, rendering the cohort effect less meaningful (Kosloski, 1986). Similarly, historical events may have timing effects and may impact individuals of different ages differently. A recent example of such effects includes the 9/11 terrorist attacks in the United States, where the impact of the event would be expected to be different for those who were school-aged and adolescents compared to those who may have been younger (e.g., attending preschool or kindergarten; Eisenberg & Silver, 2011; Greene et al., 2018). At a minimum, an examination of cohort effects requires explication of how cohort differences are developmental, as well as discussion of underlying theoretical mechanisms of change (Baltes, Cornelius, & Nesselroade, 1979).

Time

Schaie (1986) emphasizes that historical time is an important component of developmental research as it provides an important marker for contextualizing developmental change. The contribution of time to the developmental function is not well understood, however, and is fully confounded with age and cohort effects (Schaie, 1984, 1986). Time as an indicator of change may overlap with age in certain instances, but differentiating the effects of time of measurement from age may be difficult. It is useful to examine societal changes related to culture, technology, and customs that may “constrain behavior” (Schaie, 1986) to key points in historical time, although it may be difficult to determine precisely what those changes might be in real time without the advantages of hindsight. The importance of time of measurement on developmental change cannot be underestimated, however, because environmental affordances (e.g., the presence/absence of smart phones, artificial intelligence, and remote learning platforms) that characterize a specific time of measurement will frequently determine what one considers to be important to study and the rate of development of a phenomenon.

General conclusions: the general developmental model

The general developmental model suggests that most research addressing development can be done using a combination of these three variables and sequential designs (cross-sequential, cohort-sequential, or time-sequential; Schaie & Caskie, 2005) that integrate two of these three dimensions can be used to tease apart cohort effects from age effects, age effects from time of measurement effects, etc. However, it is critical to note that these indices of developmental change are not independent. Upon determining two of these parameters, the third is typically automatically defined (Baltes, 1968; Schaie, 1965); *age* is often confounded with either *cohort* or *time* or *period* of measurement. This is often due to the fact that most conceptualizations of these critical variables are often linked to calendar and chronological time, and it has been suggested that it is important to disentangle these concepts (Schaie, 1986).

However, Schaie’s general development model is not without limitations and has been subject to a number of critiques (e.g., Baltes, Reese, & Nesselroade, 1977). A notable limitation of the model is the assumption that change occurs incrementally over time in a linear manner (Kosloski, 1986). Others have noted that the general developmental model is only useful for describing change, rather than explaining it (Baltes, Reese, & Nesselroade, 1977). Despite these critiques, the general developmental model has provided much guidance for how development should be examined. As indices of developmental change via age, cohort, or time are proxies for other causal variables, the model has spurred complex and intricate methods to understand the developmental process. One suggestion to further understand the developmental function is to incorporate the contributions of age, cohort, and time

into research models and study design due to their inextricably linked nature within any given data set.

In response to these critiques, Baltes and Nesselroade (1979) proposed a model of developmental change that incorporates age, cohort, and time of measurement effects. Their model specifies three primary influences on behavioral development: normative age-graded, normative history-graded, and nonnormative influences. Normative age-graded influences are those that are highly correlated with chronological age, such as biological maturation (i.e., physical ability). Normative history-graded influences are those biological and social processes that are impacted by a social or culturally based phenomenon that impacts almost all members of cohort, such as entering school. Nonnormative influences are those biological and social processes that are less common and do not impact most individuals of a cohort. Examples of such nonnormative influences include illnesses, disability, and unemployment. Each of these influences, to whatever extent possible, must be taken into consideration by developmental scientists. These influences occur simultaneously over time and contribute to between-cohort differences in developmental change, as well as within-cohort differences (Baltes, Cornelius, & Nesselroade, 1979).

Developmental Research Designs

A number of different developmental designs can be derived from the combination of variables that describe developmental change – age, cohort, or time/period. Simple cross-sectional and longitudinal designs are often defined by the age of interest to researchers, the cohort where the sample is drawn from, and the time(s) of assessment. Age variables within cross-sectional research design tap into *interindividual differences*, whereas in longitudinal research designs, age taps into *intraindividual change* (Schaie, 1983, 1986). Within these research designs, cohort serves as variable of *individual differences*, while time of measurement or period is an *intraindividual change* variable.

Simple cross-sectional developmental designs

A cross-sectional “developmental” design incorporates at least two samples or two cohorts that are assessed at the same time but that differ in age. One example of this type of developmental study would be if a researcher examined the social behavior of children entering peer groups and observed children who were 6, 8, and 10 years old. The use of this type of design assumes that inferences can be made concerning behavior at younger and older age groups because the sample is drawn from the same general population without waiting for development or maturation to occur (Achenbach, 1978). Though many researchers will draw longitudinal conclusions

from this type of design, these are subject to questionable validity (Achenbach, 1978; Kraemer et al., 2000). It should be noted that studies that take place at single time point or snapshot of time with a single age group are often erroneously termed cross-sectional. These types of designs are best labeled concurrent or correlational designs.

Commonly utilized in developmental research, cross-sectional developmental designs are relatively inexpensive and easier and less time-consuming to conduct. Additionally, cross-sectional studies allow researchers to pilot new measures or technologies, generate and clarify potential hypotheses, and provide information to guide future and follow-up studies (Kraemer, 1994). These designs provide information about age group differences or interindividual differences; however, they cannot tell us about any age changes or interindividual differences in interindividual change (Miller, 1998; Wohlwill, 1973). Using the example given earlier, a cross-sectional study would provide information about the existence of any differences in socialization among 6-, 8-, and 10-year-olds, but it cannot provide information about how or when these changes emerged, nor how these behaviors may change across these ages.

Though cross-sectional studies can be of crucial utility, there are methodological concerns and limitations that must be kept in mind. A primary limitation is the inability to answer questions about the stability of a variable of interest or processes over time (Miller, 1998), as well as the loss of information through the use of averages to ascertain group differences (Wohlwill, 1973). Critical to cross-sectional design is the need to ensure that measurement instruments (e.g., survey and intellectual assessments) are appropriate and similarly valid for use at each age being investigated (Miller, 1998).

Cross-sectional studies are subject to issues with external validity (i.e., generalizability) as cohorts can be affected by historical or cultural differences (Achenbach, 1978). For example, if one were studying the development of mathematical problem-solving, the ability to compare second graders and fourth graders in the same school would be compromised if the mathematics curriculum were changed for the second graders. This situation would serve as a historical event that would make the cohorts nonequivalent and different from one another. This is often understood as the age-by-cohort confound, which presents a serious limitation of cross-sectional designs – the effect of age and belonging to a specific cohort cannot be distinguished in this design. According to Miller (1998), this cohort effect is less likely to be present when the dependent variable is a “biological” variable or a “basic” variable (e.g., heart rate and visual acuity), than if the dependent variable might be linked to social experiences. The age span of the sample can increase the risk of cohort effects – with samples that have larger age ranges being more vulnerable to cohort effects operating. These types of threats to the validity of cross-sectional designs create risk to researchers engaged in aging or lifespan research.

Another particular risk for cross-sectional methods is the possibility of bias in the sample through flawed selection procedures, especially since one cannot randomly

assign to age groups (Baltes & Nesselroade, 1979; Kosloski, 1986; Miller, 1998; Wohlwill, 1973). Methodologists (e.g., Cook & Campbell, 1979; Shadish, Cook, & Campbell, 2002) have long held the position that samples should be identical except for the variable that is of interest; in this case, that variable would be age. Matching on a variable other than age can result in a nonrepresentative sample. For example, comparing 25-year-olds and 75-year-olds matched on educational attainment (e.g., having a college degree) would result in a positively biased sample of 75-year-olds. Furthermore, if transitions into and out of the sampling population are not random, then there is a risk of making incorrect inferences about developmental processes being examined (Kraemer et al., 2000). An example of this would be if one is interested in the relationships between age and the move toward an assisted living facility, which is not a random occurrence, is most likely related to factors associated with age. One would need to account for this confound, and the use of a cross-sectional design would not allow for the identification of potential predictors. Longitudinal design may be more suited to answering such questions. Card and Little (2007) suggest that concurrent studies of developmental processes should consider evaluating direct and moderating effects of age; however, this will not eliminate confounding issues between age and cohort.

The utility of cross-sectional approaches may be limited, but they provide valuable information that allows researchers to generate hypotheses about intraindividual age-related changes from observing interindividual differences. Such inferences require assumptions that participants in all comparison groups are the same except for chronological age. Inferences from cross-sectional samples need to be drawn with the consideration of how factors such as the measurement of time, the type of developmental trajectory of the developmental process (i.e., fixed traits parallel or nonparallel trajectories), and the reliability and timing of the measurement (Kraemer et al., 2000) might threaten such conclusions. Drawing longitudinal inferences from cross-sectional research can result in misleading and perhaps invalid conclusions about characteristics and processes (Hertzog, 1996; Kraemer et al., 2000).

Example of a cross-sectional developmental study Belsky and Most (1981) utilized a cross-sectional design in a study where they observed the development of play behavior at multiple ages of infancy and toddlerhood. Twenty boys and 20 girls who were evenly spaced in 10 age groups between the ages of 7.5 months and 21 months were observed at home in free play interactions where observers narrated the children's play. The study revealed that younger children (between 7.5 and 15 months) used mouthing and simple manipulation play behaviors more than older children (16.5 months to 21 months). They also found that there were age group differences in the use of pretend and representation play that involved relating two objects together. The study fostered greater interest in understanding the developmental sequence of play and exploration behaviors.

Simple longitudinal developmental design

A known solution to the shortcomings of cross-sectional design is the utilization of repeated assessments of a participant sample over a period of time – a longitudinal research design. Longitudinal designs have been defined broadly. Miller (1998) described longitudinal designs as “repeated tests that span an appreciable length of time”; however, Miller did not specify the number or the frequency of the “repeated” tests. Additionally, what constitutes an “appreciable length of time” is also unclear. In some cases, an appreciable length of time could be linked to the developmental level of the sample being studied, or it may simply be the timing of observation or measurement of time as it relates to a specific event. Longitudinal research designs also can be examined on a multitude of time scales at both across measurement (e.g., monthly observation) and within measurement (e.g., multiple biomarker samples across a stressor experience).

Baltes and Nesselroade (1979) provided a comprehensive foundation to longitudinal design and the unique benefits of this research design. They identified and expounded upon key concepts that continue to be pertinent today and furthered the notion that longitudinal designs are critical and necessary for the examination of intraindividual change, developmental sequences, and co-occurring social and environmental change. Within-person change can be seen as incremental (e.g., change in magnitude of a construct over a specified interval of time) or transformational (reflected by changes or transitions in state; Ford & Lerner, 1992). Longitudinal designs also permit the investigation of interindividual differences in intraindividual consistency or change, while also allowing the examination of the degree to which earlier experiences predict later development (Farrington, 1991; Miller, 1998; Wohlwill, 1973).

Longitudinal analysis also allows for further insights into relationships between domains of development and also about possible causal explanations for changes in given domains over time. As development within given domains does not occur in isolation, the use of this type of multi-domain approach allows for us to understand potential co-occurring processes that may be happening simultaneously. Longitudinal approaches also allow researchers to consider whether within-person change in one domain is a precursor to change in another (Grimm, Davoudzadeh, & Ram, 2012).

As mentioned, a pertinent goal of longitudinal research is to provide means to identify possible causal processes or determinants of intraindividual change, as well as interindividual differences in intraindividual change. The use of such designs enables an examination of changes that may occur at different rates and different points in time. Precise identification of causal determinants of development from a longitudinal design alone cannot be accomplished. At the same time, however, a key methodological feature of longitudinal studies is the ordering of time, which is one of the three primary criteria for causality (Cook & Campbell, 1979; Shadish, Cook, & Campbell, 2002). This characteristic of longitudinal designs allows one to examine antecedents and consequences of a developmental phenomenon that enables

reasonable speculations about causality. Schaie (1988) cautioned, however, that whereas observations in a longitudinal study are time-ordered unidirectionally, that does not mean change is unidirectional. Whereas some developmental processes may be unidirectional over certain periods of time, other processes may be nonlinear (e.g., cyclical or recursive). Baltes and Nesselroade (1979) also indicated that interindividual differences in intraindividual change are not universal and may be explained by moderating effects such as demographic characteristics or additional constructs that may or may not have been measured.

Examining developmental trajectories provides opportunities to consider equifinality and multifinality, concepts that have been influential in developmental science for understanding how interindividual variability at specific points of development can be conceptualized (Cicchetti & Rogosch, 1996; Cicchetti & Toth, 2009). Equifinality refers to the convergence of trajectories from different starting points to a common endpoint or developmental outcome. Multifinality, by contrast, refers to divergence from common starting points toward multiple outcomes. Recent work from Schulenberg, Maslowsky, and Jager (2018) also introduced the concept of developmental fanning, suggesting heterogeneity in interindividual differences in trajectories where these differences may increase across development. These related concepts are influential and often serve as the crux of many developmental theories as they may provide information about pathways and change (Lerner, 2006).

Newer, time-intensive methodologies, which have been made possible by innovations in smartphone technology and other in-the-moment monitoring devices, have emerged that address some of these rationales for understanding growth, change, and development across time (Ferrer & McArdle, 2010; Little, 2013). Grimm, Davoudzadeh, & Ram (2017) provided a brief historical overview of developments in longitudinal methods that discuss complex modeling of these processes. Later chapters of this book will also provide greater detail about these innovative methods and their unique contributions to understanding development.

Variations in longitudinal developmental designs: retrospective and intensive longitudinal designs

Retrospective designs The aforementioned longitudinal designs presume an implicit forward-looking or prospective approach. Prospective studies are designed to examine effects based on the independent variable of interest. For example, a researcher interested in how early executive function predicts academic performance across elementary school could recruit children in preschool and follow them over time. However, longitudinal research can also utilize a retrospective or backward-looking approach to understand a phenomenon. Typically, retrospective approaches identify an outcome on which individuals differ at a specific point in time and then examine possible developmental antecedents of that difference that are theoretically defensible.

For example, a retrospective approach to study developmental antecedents of clinical depression in adulthood could initiate a search for early life predictors that have theoretical links to adult depression. Such a search could target familial or ecological antecedents earlier in life that might plausibly explain the difference, such as earlier episodes of depression, parent–child relationships, early experiences of loss or of trauma, and early life stress (Hipwell et al., 2023; Kushalnagar et al., 2017; Teicher et al., 2009).

Retrospective designs offer practical and efficient ways to examine long-term developmental trajectories. This type of design offers a unique opportunity to identify potential effects without the need to extensively design assessments as they are often based on the fact that participants have already experienced events that are of importance and interest to the researcher. Retrospective designs are particularly best utilized to examining indices that are difficult to capture and not feasible to study in a prospective manner. Retrospective designs offer utility as being cost-effective and less intensive than prospective designs. Additionally, retrospective studies are well-suited to questions where data may be readily available such as the use of documentation (e.g., electronic medical or school records) or secondary data analysis of publicly available data and may allow for larger sample sizes.

Despite these benefits, a number of drawbacks to retrospective designs include limitations on available data that are not systematically archived, missing data, or potential recall issues (e.g., inaccuracy and/or bias), as well as sample issues (e.g., lack of generalizability). Retrospective approaches that rely on participant report are subject to issues of inaccurate recall of events, experiences, and behavior (Raphael, 1987; Widom & Morris, 1997). Relatedly, the impact of current life conditions and context may also contribute to participant recall bias and memory distortions. These types of recall issues can contribute to issues of reliability, as well as incomplete and missing data. Furthermore, recall inaccuracy can reduce potential longitudinal associations and issues of recall bias can introduce unintended longitudinal associations. Careful considerations are also necessary due to the inability to examine potential confounding factors and limited control in the timing and frequency of when data were collected.

Intensive longitudinal designs Intensive longitudinal designs, sometimes referred to as micro-longitudinal or “burst” designs, have been used increasingly to capture longitudinal, dynamic, in-the-moment processes of change across short time frames. These designs enable researchers to examine the micro-level processes and mechanisms that may contribute to or index developmental changes within individuals. Data collection methods used will vary depending on the question(s) asked and include daily diaries, random or systematic sampling of specific behaviors over very brief time spans, passive data collection measurement (as obtained from wearable devices such as actigraphs), ecological momentary assessment (EMA), and other approaches (see Teti, this volume). Whatever method is used, the researcher must choose an assessment frequency adequate to capture the desired process without

overburdening the participant. Other considerations include whether assessment should be done randomly or systematically (perhaps in response to particular events or according to a timetable). It is not uncommon for intensive longitudinal designs to make use of several data collection methods that are used integratively to capture a specific dynamic process. EMA designs, for example, rely upon real-time data collection over an interval of time to assess particular experiences and events in participants' lives within their normal environments (Shiffman, Stone & Hufford, 2008). This type of data collection can use random sampling or fixed times for the collection of individual surveys/assessments and include the use of written (e.g., diary or log) or technological methods (e.g., texted survey, physiological sensor) to obtain data. The collection of multiple forms of assessment allows researchers to understand how experiences and behavior changes occur and vary over time.

Data collection bursts are functional in their own right as micro-longitudinal studies of change processes. It is not uncommon, however, to embed micro-longitudinal data collection bursts into larger, longitudinal time frames, and as such, these data bursts can be integrated with single-point-in-time measures. The data bursts in these "measurement-burst" designs, as they are called, can be used as predictors or outcomes, depending on the research questions posed (see below). Micro-longitudinal (or intensive longitudinal) study designs can also be embedded within cross-sectional designs or studies of single snapshot of time or period of development (e.g., the transition to formal schooling). This design allows for the researcher to examine specific dynamic phenomena in a cost-effective and timely manner while operating within a cross-sectional and cost-effective research design.

Measurement-burst designs are a type of longitudinal design where intensive bursts of data collection occur in a short time frame (e.g., days or week) and the data collection is repeated again over a longer or wider period of time (e.g., monthly or yearly; Stawski, MacDonald, & Sliwinski, 2015). This type of design was an effort to address limitations of single time point snapshots (Nesselroade, 1991); however, elaborations on this design have allowed for an examination of more dynamic processes that may be occurring (MacDonald & Stawski, 2016; Sliwinski, 2008). A particular issue of importance in measurement-burst designs revolves around the timing, or temporal intervals for the phenomenon of interest are being examined (Walls et al., 2012). Measurement-burst designs also allow for direct examinations of changes at a longer longitudinal period, as well as those that may transpire across shorter period of time. A number of challenges are pertinent to measurement-burst designs due to the complexities related to data collection.

The intensive nature of measurement-burst design can pose a logistical and labor-intensive challenge to researchers. Additionally, the intensive nature of the design may also contribute to participant burden and potential missing data issues. An important challenge relates to the temporal spacing of assessments (e.g., moments, daily, or weekly) because developmental phenomena of interest may not occur at the same time, especially when examining short-term and longer-term processes. At the same time, these intensive data collection strategies enable researchers to capture

changes in behavior, cognition, or other developmental domains as they occur, providing a more detailed understanding of developmental processes. A particular benefit of these types of design is that they allow for intraindividual change variability to be observed over short time spans, and if imbedded into a broader longitudinal frame (measurement-burst designs), over longer time spans as well. Alignment between design (including timing and assessments) and theoretical approaches for the developmental construct of interest allows for a more precise examination of within-person variability and how changes on a shorter scale may contribute to well-being (Cerino & Hooker, 2019).

Threats to validity in longitudinal developmental research Despite a multitude of benefits of longitudinal research designs and strategies, there are a number of drawbacks and limitations to take into account. Conducting longitudinal research is expensive, time-consuming, and often labor-intensive for researchers, as well as participants. In addition to these issues, longitudinal research is vulnerable to a number of threats to validity, particularly sample selection, attrition, instrumentation, and regression to the mean (Shadish, Cook, & Campbell, 2002).

Decisions about sampling are dependent upon what type of design is being employed – prospective or retrospective design. Prospective studies rely upon sampling that is reliant upon the independent variable of interest, whereas retrospective sample designs are assembled around the dependent variable. A number of issues involving the sample should be addressed. First of which, what type of sampling procedure should be used? For example, should the sample attempt to be representative (e.g., birth cohort, community sample cohort) or nonrepresentative (e.g., clinical populations, adoptees)? The use of a representative sample would allow for greater generalizability of the findings, understanding multiple types of phenomena (e.g., social, biophysiological variables), as well as document prevalence data (Baltes, Reese, & Nesselrode, 1977). With longitudinal designs, beyond the initial recruitment phase (i.e., sample attainment), it is critical that the research team develop and maintain cooperative and engaged relationships with members of the sample over a period of time. As it is relatively difficult to both sample a representative group of individuals and have these individuals continue to take part in the study across multiple waves of data collection, it is a challenge for developmental research projects to be representative (Achenbach, 1978; Miller, 1998). Jager, Putnick, & Bornstein (2017) have indicated that homogenous convenience samples are similar in sociodemographic distributions to the target population of interest and thus may provide accurate and valid estimations of phenomenon of interest in the target population. However, trying to retain the representative nature of such a sample across time, and not differentially attrit (based on sociodemographic attributes or general functioning), is difficult. It may be that a researcher will have to weigh whether obtaining data from a larger sample size where one may not be able to conduct as many assessments in order to reduce attrition is

more or less beneficial than obtaining data from a smaller but perhaps less representative or generalizable sample, which may contribute to external validity challenges.

Sample attrition is a common issue in longitudinal research and one that has been deemed a “virtually ubiquitous” problem that afflicts any attempt to follow a group of individuals across time (Graham, 2009, p. 567). A number of different strategies have been suggested to reduce sample attrition such as asking participants to provide contact information for a relative (Jordan, 1994). In addition, technological advances (e.g., email, text messages) have allowed for improvements in maintaining contact with and obtaining data from participants, as well as to be able to reestablish contacts (following ethical guidelines) that have been lost. Some contributions to attrition may be under the control of the researcher and can be addressed, such as designing assessments to reduce participant burden and provide appropriate incentives. However, there may be a number of impediments to participation that are not within researchers’ grasp.

Selective dropout and retention within a study can lead to biased findings, as individuals who remain in the study may differ systematically from those who drop out. Nonresponders in a study are likely to be different from those who actively maintain participation, and these differences may be related to variables being examined (Bergman & Magnusson, 1990; Nicholson, Doebeck, & Howard, 2017). Technological advancements in data collection methods have helped to reduce missing data through the use of online surveys, wearable technologies, and online collection of observational and task performance data. Participants are able to provide partial data that may be usable and would reduce noncompliance and potential bias. Threats to validity pertaining to missing data can also be mitigated through the use of planned missing data designs that alleviate participant burden, lower the cost of research, and help improve validity. Little and Rhemtulla (2013) provide an overview of research designs on planned missing data and their utility for developmental researchers, whereas Wu and Jia (2021) provide specific information about such methods for longitudinal panel studies. In addition to planned designs, extensive methodological work has emerged in the past decades to deal with the problem of missing data (Enders, 2013; Jeličić, Phelps, & Lerner, 2009; Little et al., 2014). Nicholson, Doebeck, and Howard (2017) provide a review of commonly used statistical methods in developmental science. This burgeoning area of developmental methodology provides statistical techniques that employ maximum-likelihood estimation and imputation as solutions to handle missing data, reduce sample bias, and reduce threats to validity.

Beyond issues of the sampling in longitudinal studies, testing and measurement or instrumentation also represent potential threats to validity. Longitudinal research relies upon repeated measurement of constructs, and repeated testing can introduce practice effects or sensitization, where participants become more familiar with the measures and alter their responses. It is also possible that the course or trajectory of development could change due to participation in a longitudinal study due to

greater participant awareness of the constructs being examined (Goldstein, 1979). These types of testing threats can be mitigated through the use of alternate forms or variations in the assessments (Wohlwill, 1973).

Issues of measurement can also extend to concerns with instrumentation (Baltes et al., 1977; Goldstein, 1979). One example of instrumentation is the use of different instruments to assess a construct that itself develops and changes with age, such as intelligence, peer relationships, child–parent attachment, or temperament. Threats to validity from instrumentation can continue to persist even if a variable is measured identically at two points in time because of the likely change in the distribution of scores, an indicator that the meaning of the score may have changed (Achenbach, 1978). Problems with measurement equivalence over time may also emerge due to developmental discontinuation of phenomena or behavior (Hartmann, 2005; Schaie, 1998). Other instrumentation issues can arise related to “aging” where assessment tools may no longer be valid or reliable, to changes in study personnel, and to changes/updates in our definitions and understanding of specific phenomena (Achenbach, 1978; van der Kamp, Bijlaved, & London, 1998). Indeed, it is possible that theories and hypotheses may become outdated as a study progresses over time and may require revising or reformulation (Jordan, 1994).

Another common threat to validity to longitudinal measurement is regression to the mean (Baltes & Nesselroade, 1979), or the tendency for extreme scores in a distribution to become less extreme over time, simply because whatever influences that led to an extreme performance at one time point may not be present in the same magnitudes with measurements taken at later time points. The researcher is left with separating out “true” developmental change (always approximated and never actually known) or changes due to an independent variable from changes in or errors in measurement, or changes in historical factors that impact individual performance. In intervention research, regression to the mean (independent of intervention effects) can be estimated when the intervention and control groups are assumed to be equivalent, as evidenced perhaps by similar error variances and regression trajectories in the dependent variables observed across groups.

Longitudinal research designs are susceptible to the issues of cohort, age, and time that were mentioned earlier in the chapter. A cohort effect may impact the internal and external validity of a study where unique experiences or characteristics of a cohort may make them dissimilar for another cohort at the same age (Baltes et al., 1979; Bergman & Magnusson, 1990). One way to mitigate such effects is to obtain cross-sectional data on relevant variables in order to limit cohort effects. Additionally, study validity can be impacted by the age-by-time or period-of-measurement confounds. It has been suggested that it is critical to consider disentangling these concepts and that longitudinal studies that have only one cohort are particularly susceptible to such effects because it is difficult to decouple effects of age and time (Schaie, 1983, 1986).

Planning and implementation of longitudinal research

The logistics concerning longitudinal research are an often overlooked and underestimated but necessary for conducting rigorous developmental research. Conceptualization, methodology, and data analyses are often integral to planning and implementing studies, and it is critical to keep a number of things in mind. Researchers are advised to plan longitudinal studies to have a broad scope, include eclectic measures, and incorporate constructs from different theories so that the resulting data can be analyzed from multidisciplinary perspectives and be used to answer an array of research questions (Bergman & Magnusson, 1990). Although the methods, constructs, and meaning of scores may change over the course of the study (Achenbach, 1978), designing longitudinal studies from such a broad and multidisciplinary approach will help insulate the data from potential aging effects and decreased relevancy overtime (Bergman & Magnusson, 1990). The timing of when to measure a behavior, construct, or phenomenon is essential to researchers to adequately capture development, and allowing for a large variation in time of measurement can pose problems (Goldstein, 1979). For example, if one is interested in assessing language acquisition skills during toddlerhood, it would be critical to measure these skills frequently, such as every few weeks rather than to measure them every few months, in order to observe changes in skills.

Compared to cross-sectional studies, longitudinal studies are time-consuming and expensive to execute and require long-term commitment from participants. It is critical for researchers to have flexible plans and implementation strategies to undertake work that may have a slow rate of return and take a considerable length of time to observe phenomena of interest (Wohlwill, 1973). It is critical that the sample population is considered carefully in order to accurately examine developmental phenomenon. A normal representative sample such as birth, school, or community cohorts allows for greater generalizability of findings, a wider array of behavior and phenomena to be examined, and investigation of potential group differences (e.g., gender, race within a general population) (Baltes, Reese, & Nesselroade, 1977; Bornstein, Jager, & Putnick, 2013; Mednick, Griffith, & Mednick, 1981). Nonrepresentative samples, such as those who experience specific conditions or specialized cohorts, can be important for investigating the unique factors that impact development within specific parts of our society, as well as be utilized for hard-to-reach populations (e.g., individuals experiencing social marginalization; Raifman et al., 2022). Bornstein et al. (2013) provide a brief overview of the pitfalls and solutions to common sampling issues in developmental science. Researchers seeking greater detail on issues with sampling strategies can also refer to Levy and Lemeshow (2013), as well as Cochran (1977).

A key deterrent to conducting longitudinal studies is that they tend to be expensive to conduct due to investment in necessary equipment, materials to implement research plans, and training of staff and personnel. Staff and personnel involved in longitudinal studies play a critical role as they often interact the most with participants,

compared to research investigators. The use of trained professional staff for data collection typically improves recruitment and retention of participants in longitudinal studies (Hill et al., 2016). In addition to investments in personnel and training, researchers should invest in multiple methods to ensure sample retention and data validity. Hill et al. (2016) recommend that researchers consider designing consent forms in a manner that allows for long-term follow-up, maintaining a relationship with the sample between waves of data collection, as well as being mindful of study branding and outreach (e.g., logos, social media presence, and newsletters). Emergent social media and technologies can be utilized to contact missing participants and maintain rapport; however, such contact should be ethical and within Institutional Review Board (IRB) approved protocols and guidelines.

Lastly, the timing of measurement and data management are key aspects of longitudinal design that need to be taken into consideration by researchers looking to conduct longitudinal research. Planning of research involves attention to theory, organization, and administration; however, considerable attention is needed to how to manage the data once generated (Bergman & Magnusson, 1990; Jordan, 1994). Researchers should be mindful of how questions and responses are structured, relevancy of measures to the developmental age, and how data can be collected (e.g., online survey or in-person) (Hill et al., 2016). It is also useful to collect data in a manner that would allow for it to be examined in different ways and from other theoretical perspectives, as well as maintaining collected data in a raw or basic format (rather than summary or composite scores) (Bergman & Magnusson, 1990).

Longitudinal Developmental Study Example

SIESTA – a longitudinal study

The Study of Infants' Emergent Sleep TrAjectories (SIESTA) is a longitudinal prospective study that was conducted using multiple informants and methods (Teti, 2017). The objectives of the study were to understand how families navigated the transition to parenthood and how parent- and family-level processes influence children's sleep and development across the first 2 years of life. The study was designed to assess families when the target infant was 1, 3, 6, 9, 12, 18, and 24 months. The sample is representative of the geographic area of Pennsylvania where data collection occurred. The sample consists of two-parent households that were socioeconomically diverse (median yearly family income was \$60,000; 60% of mothers completed high school) and predominantly White (86%), with the remaining sample split evenly between African American, Asian American, Latinx, or those who identified as "Other."

The study utilized a multi-method design where parents provided survey information about their well-being, relationship quality, parenting practices, and infant

behavior at each time point. In addition to survey measures, parents and infants wore actigraphs to assess multiple domains of sleep, filled out sleep diaries, and allowed for observation of infant sleep and bedtime practices. The use of observational assessments of parenting behavior, parent–child interactions, and other aspects of the family environment allowed for additional information to be obtained through reliable observer coding. The study has examined numerous longitudinal research questions, and an overview of longitudinal findings concerning parenting and family influences on infant sleep can be found in Teti (2017) and in Teti (this volume).

Considerations when conducting research with minoritized populations

Research in developmental science with populations that are historically or currently minoritized or disadvantaged has garnered significant attention in the past few decades. Specific focus has been devoted to racial, ethnic, and cultural identities (Boyce & Fuligni, 2007; Huston, McLoyd, & Coll, 1994; McLoyd, 1998; Schwartz et al., 2014; Syed et al., 2018), along with an increasing focus on sexual gender minority (SGM) individuals (Fish, 2020; Toomey, 2021). As greater emphasis has been placed on understanding development in diverse populations, emerging literature has started to provide additional guidance to assist researchers on specific issues pertaining to working with these groups (Syed et al., 2018; Toomey, 2021). It is important to incorporate these ideas into discussions of general developmental design in order to bring awareness to relevant issues.

Longitudinal research with populations that are minoritized or disadvantaged may face particular issues that should be attended to in order to reduce potential threats to validity and better understand pertinent developmental issues. Research that is ontogenic and sociogenic in nature is subject to a number of threats to validity, most notably aging effects. Such threats may be even more apparent as developmental science incorporates intersectional approaches that may not map onto previous research or methodologies. One example of such a threat concerns terminology and definitions pertaining to race, ethnicity, culture, and sexual gender identity, where evolving research aims to better understand individuals in these groups. Similarly, expansions and greater understanding in what is defined as “disability” or being “differently abled” may also not map onto previous research, and researchers should consider open-ended responses and alternative approaches to instrumentation and measurement (Davison et al., 2022; Odom et al., 2007).

Furthermore, research on populations of interest often focuses upon between-group distinctions; however, it is necessary that careful consideration needs to be given as to whether such a research design is necessary. Oftentimes, one group is assumed to be the “norm” and other group(s) are often viewed as “deficient or abnormal.” Within-group research designs are often underutilized and allow for a unique understanding of developmental phenomena without comparison to other

groups, as well as an examination of whether processes and theories operate in the same way. Developmental researchers should be cognizant of the research questions that are being examined and be aware of alternative approaches that may best encapsulate the development they wish to understand.

Racial and ethnic minority populations Across disciplines, issues of understanding how race and ethnicity are conceptualized and implemented in research have received a great deal of attention (Schwartz et al., 2014; Syed et al., 2018). As discussed above, researchers should be mindful of how individuals define themselves rather than to assign categories or force selection into predefined groups. Additionally, minoritized populations are often subject to social stressors such as socioeconomic adversity and poverty, discrimination, and mistrust of researchers and systems that may prevent their participation or lead to attrition and biased results in longitudinal studies. Measures and instrumentation tools that are reliable and valid for use in one population may not be salient for all racial and ethnic groups, and careful consideration should be given on how to best assess constructs (Harachi et al., 2006). These issues of instrumentation, recruitment, and measurement can serve as threats to validity to developmental research.

Social and historical events are likely to introduce cohort effects when examining individuals in a longitudinal and developmental framework (Okazaki & Sue, 2016). A recent example of potential cohort effects includes global social justice and political movements occurring since 2020 due to the increased violence and hate crimes against African American/Black individuals and those of Asian descent compared to other recent generations. In the United States specifically, these have served as pivotal events that may introduce cohort effects when examining questions around ethnic and racial identity, attitudes, and cognitions about issues pertaining to social justice, as well as a multitude of other developmental research areas.

Research focused on racial and ethnic minoritized individuals in developmental research needs to pay attention to the paradigm being utilized (Phinney & Landin, 1998), as well as consider how research constructs are conceptualized (Okazaki & Sue, 2016). Phinney and Landin (1998) report that research often uses general process models to understand development when examining racially and ethnically diverse populations; however, these general process models may overlook important cultural factors and not consider specific effects of race/ethnicity or culture on developmental processes. This oversight may inadvertently contribute to mischaracterized interpretations of findings. Researchers have advocated for a greater focus on cultural processes, such as cultural connectedness and experiences of discrimination, as relevant toward understanding development rather than merely using race/ethnicity, place of birth, and other sociodemographic markers (Boyce & Fuligni, 2007). In addition to these considerations, theoretical models that are specific to racial and ethnic minority individuals may need to be considered, such as the seminal work of the integrative model from Coll et al. (1996) that incorporated social and cultural considerations in the development in minority children.

Gender and sexual minority populations Research focused on SGM individuals has received greater attention in the past few decades, with a specific focus on understanding mental health disparities (Russell & Fish, 2016). Similar to other minoritized populations, longitudinal studies examining SGMs are limited, and Fish (2020) notes that studies that examine “developmental timing, sequencing, and constancy of specific experiences (e.g., ‘coming out’, victimization, hormone therapy)” are understudied areas despite being of significant importance to SGM individuals. Additionally, Russell and Fish (2016) note that limited work has focused on protective and promotive factors that underlie development, leading to potential bias in how we understand the lives and development of lesbian, gay, bisexual, trans, queer, and other identities (LGBTQ+).

Research with such populations must consider changes that occur to the socio-cultural environment and how these changes may impact developmental processes. Umberson et al. (2015) identify the shifting legal landscape as a particular challenge that may be important for researchers to consider as these types of issues may present unique challenges to researchers. Stigma, experiences of discrimination, and difficulties in accessing populations may also contribute to inadequate assessment of development in these populations. Furthermore, the experiences of those with intersectional identities (e.g., SGM individuals of color) may not be easily captured with existing methods and instrumentation (Russell & Fish, 2016; Syed et al., 2018). Each of these aspects may create threats to validity when working with SGM populations, and recent work provides some guidelines on how to approach such challenges (Fish, 2020). Additionally, Umberson et al. (2015) provide an overview of issues concerning sample size, recruitment, and relational and family research aims as they relate to sexual and gender minority individuals and families.

Socioeconomic adversity and poverty

Socioeconomic adversity and issues of poverty have been longstanding areas of inquiry among developmental researchers (Conger et al., 1994; Davis & Havighurst, 1946; Emerson & Hatton, 2009; Huston, McLoyd, & Coll, 1994). Sample recruitment and attrition issues are of particular concern when conducting longitudinal research with populations experiencing poverty and related adversity. Individuals may face barriers to participation such as lack of access to technology, unstable living and caregiving situations, and access to transportation. These barriers may also contribute to issues of attrition and transient participation that may contribute to inaccurate findings.

Similar to other minoritized groups, cohort considerations should be made around socioeconomic events, such as periods of economic recession, war/conflict, and other catastrophic events that have been experienced by communities and countries. These types of events may have significant influences on development, including alteration of trajectories of development. Additionally, researchers must keep in

mind that socioeconomic measures may not be valid across the period of study and may not be accurate representations of economic resources in that society and time period. One such well-known example of this is the use of the Hollingshead indices (1975) that were heavily utilized in research that was reliant upon occupation and levels of educational attainment. Critiques of these assessments have persisted (Gottfried, 1985), mainly for being outdated due to the changing social contexts concerning the prestige or value of occupations. Furthermore, Braveman et al. (2005) indicate that socioeconomic status (SES) measures are not interchangeable and assessment may not adequately capture the experiences and nuances of poverty, potentially leading to biased or inaccurate data.

To improve the validity of developmental and longitudinal research in populations experiencing poverty and socioeconomic adversity, it is crucial to employ appropriate methodologies, including using measures and assessment tools that consider the cultural and economic contexts. Conger and Donnellan (2006) advocate that researchers consider more dynamic and interactionist models, rather than static models, to best understand how socioeconomic adversity is associated with development.

Disabled and ability-challenged populations

A population that is often less discussed when discussing developmental research is individuals who are differently abled or have a disability (Odom et al., 2007). Individuals with disabilities may have impairments that could encompass sensory, cognitive, physical, or psychological domains. Oftentimes, individuals have varying levels of impairment and this issue should be considered when making sampling decisions. The validity of developmental research with these populations is dependent upon researchers being attuned to measurement issues as current developmental tools may not be sufficient for capturing experiences accurately. Davison and colleagues (2022) report techniques for the modification and adaptation of assessment tools and the use of participatory research with youth with intellectual disabilities. Disabilities might intersect with other factors like race, gender, and SES, which can impact individuals' development in unique ways (Syed et al., 2018). Research should employ inclusive methodologies that consider diverse abilities and modification of assessment tools to accommodate different communication styles for high-quality and valid research.

Conclusion

This chapter provides an overview on the fundamentals of developmental research design, discussing cohort, age, and time as variables in Schaie's (1965) general developmental model that are represented in cross-sectional, longitudinal, and retrospective

designs and the threats to validity baked into each. Although not covered extensively in this chapter, it is common for different developmental research approaches to be integrated into a larger developmental study, as when, for example, a study makes use of multiple birth cohorts of participants and follows them longitudinally or when a study makes use of a prospective longitudinal design (e.g., involving direct observation) and a retrospective design to obtain information on a developmental experience. It is also common for longitudinal designs to be incorporated into experimental designs to examine the effects of an independent variable on participant performance over time.

Also discussed in this chapter were micro-longitudinal designs that have come of age with the emergence of technologies that enable rapid, intensive collection of data bursts across brief periods of time. Such data bursts can provide researchers with insights about developmental processes of inherent theoretical interest in their own right, heretofore not subject to empirical study but that, when embedded into a larger longitudinal frame (measurement-burst designs), promote insights about how “in-the-moment” processes contribute to development over the longer term. It is interesting to speculate about the creative ways that measurement-burst designs can be used by developmental scientists going forward. Technologies (visual, auditory, tactile, and physiological) that capture in-the-moment information about the micro-contexts in which subjects participate may be useful for documenting historical experiences unique to an individual. Such a use is consistent with Baltes and Nesselroade (1979), who called for the systematic study of both normative and nonnormative historical experiences as predictors of developmental change in addition to developmental age-graded change. One can envision a myriad of ways in which longitudinal designs, interspersed with intensive micro-longitudinal data collection, can help more clearly characterize between-subject differences in historical experiences (e.g., in home, school, or work contexts) that can help account for changes over time and developmental outcomes more attributable to history than to age. Similarly, micro-longitudinal designs may enable a more fine-grained identification and examination of within-family individual differences and their influence that exist along interpersonal dimensions. Examples of differential within-family histories that burst designs might help elucidate are the different social, emotional, and learning environments that one sibling creates for another in the same family or in-the-moment individual differences in differential treatment of children growing up.

Lastly, this chapter discussed the application of developmental designs to understudied populations and the issues attendant to each. The dearth of developmental research in underserved, underrepresented populations is conspicuous, as is also the case in non-Western cultures across the world. Understanding developmental change and developmental processes within and across multiple cultural contexts, at the individual and family level (e.g., family formation), and the variation that exists within culture, is critically important for the viability of developmental science as a discipline.

The utility and application of developmental designs continue to grow and proliferate in areas of study both traditionally and not traditionally deemed as developmental (Biller et al., 2024; Følling et al., 2022; Murray et al., 2024; Pfund et al., 2022; Stolz et al., 2024). We look forward to seeing the new and creative ways in which developmental design can be used to inform different fields of inquiry as ways of measuring and analyzing change across multiple time scales continue to advance.

References

- Achenbach, T. M. (1978). *Research in developmental psychology: Concepts, strategies, methods*. Free Press.
- Baltes, P. B. (1968). Longitudinal and cross-sectional sequences in the study of age and generation effects. *Human Development*, 11(3), 145–171. <https://doi.org/10.1159/000270604>
- Baltes, P. B., Cornelius, S. W., & Nesselroade, J. R. (1979). Cohort effects in developmental psychology. In J. R. Nesselroade & P. B. Baltes (eds.), *Longitudinal research in the study of behavior and development* (pp. 61–87). Academic Press.
- 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.
- Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1977). *Life-span developmental psychology: Introduction to research methods*. Lawrence Erlbaum Associates, Inc.
- Belsky, J., & Most, R. K. (1981). From exploration to play: A cross-sectional study of infant free play behavior. *Developmental Psychology*, 17(5), 630–639. <https://doi.org/10.1037/0012-1649.17.5.630>
- Bergman, L. R., & Magnusson, D. (1990). General issues about data quality in longitudinal research. In D. Magnusson & L. R. Bergman (eds.), *Data quality in longitudinal research* (pp. 1–27). Cambridge University Press.
- Biller, A. M., Fatima, N., Hamberger, C., Hainke, L., Plankl, V., Nadeem, A., et al. (2024). The Ecology of Human Sleep (EcoSleep) Cohort Study: Protocol for a longitudinal repeated measurement burst design study to assess the relationship between sleep determinants and outcomes under real-world conditions across time of year. *Journal of Sleep Research*, e14225. <https://doi.org/10.1111/jsr.14225>. Epub ahead of print
- Bornstein, M. H., Jager, J., & Putnick, D. L. (2013). Sampling in developmental science: Situations, shortcomings, solutions, and standards. *Developmental Review*, 33(4), 357–370. <https://doi.org/10.1016/j.dr.2013.08.003>
- Boyce, C. A., & Fuligni, A. J. (2007). Issues for developmental research among racial/ethnic minority and immigrant families. *Research in Human Development*, 4(1–2), 1–17. <https://doi.org/10.1080/15427600701480972>
- Braveman, P. A., Cubbin, C., Egerter, S., Chideya, S., Marchi, K. S., Metzler, M., & Posner, S. (2005). Socioeconomic status in health research: One size does not fit all. *JAMA*, 294(22), 2879–2888. <https://doi.org/10.1001/jama.294.22.2879>
- Bronfenbrenner, U., & Evans, G. W. (2000). Developmental science in the 21st century: Emerging questions, theoretical models, research designs and empirical findings. *Social Development*, 9(1), 115–125. <https://doi.org/10.1111/1467-9507.00114>

- Card, N. A., & Little, T. D. (2007). Longitudinal modeling of developmental processes. *International Journal of Behavioral Development*, 31(4), 297–302. <https://doi.org/10.1177/0165025407077750>
- Cerino, E. S., & Hooker, K. (2019). *Intraindividual variability in lifespan developmental methodology*. Oxford University Press.
- Cicchetti, D., & Rogosch, F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology*, 8(4), 597–600. <https://doi.org/10.1017/S0954579400007318>
- Cicchetti, D., & Toth, S. L. (2009). The past achievements and future promises of developmental psychopathology: The coming of age of a discipline. *Journal of Child Psychology and Psychiatry*, 50(1–2), 16–25. <https://doi.org/10.1111/j.1469-7610.2008.01979.x>
- Cochran, W. G. (1977). *Sampling techniques*. John Wiley & Sons.
- Coll, C. G., Crnic, K., Lamberty, G., Wasik, B. H., Jenkins, R., García, H. V., & McAdoo, H. P. (1996). An integrative model for the study of developmental competencies in minority children. *Child Development*, 67(5), 1891–1914. <https://doi.org/10.1111/j.1467-8624.1996.tb01834.x>
- Conger, R. D., & Donnellan, M. B. (2006). An interactionist perspective on the socioeconomic context of human development. *Annual Review of Psychology*, 58(1), 175–199. <https://doi.org/10.1146/annurev.psych.58.110405.085551>
- Conger, R. D., Ge, X., Elder, G. H., Lorenz, F. O., & Simon, R. L. (1994). Economic stress, coercive family process, and developmental problems of adolescents. *Child Development*, 65, 541–561. <https://doi.org/10.1111/j.1467-8624.1994.tb00768.x>
- Cook, TD & Campbell, DT *Quasi-Experimentation: Design and Analysis for Field Settings*.Houghton Mifflin, Boston, 1979
- Davis, A., & Havighurst, R. J. (1946). Social class and color differences in child-rearing. *American Sociological Review*, 11(6), 698–710. <https://doi.org/10.2307/2087065>
- Davison, J., Maguire, S., McLaughlin, M., & Simms, V. (2022). Involving adolescents with intellectual disability in the adaptation of self-reported subjective well-being measures: Participatory research and methodological considerations. *Journal of Intellectual Disability Research*, 66(7), 628–641. <https://doi.org/10.1111/jir.12936>
- Eisenberg, N., & Silver, R. C. (2011). Growing up in the shadow of terrorism: Youth in America after 9/11. *American Psychologist*, 66(6), 468–481. <https://doi.org/10.1037/a0024619>
- Emerson, E., & Hatton, C. (2009). Socioeconomic position, poverty, and family research. In L. M. Glidden & M. M. Seltzer (eds.), *International Review of Research in Mental Retardation* (Vol. 37, pp. 97–129). Academic Press. [https://doi.org/10.1016/S0074-7750\(09\)37004-4](https://doi.org/10.1016/S0074-7750(09)37004-4)
- Enders, C. K. (2013). Dealing with missing data in developmental research. *Child Development Perspectives*, 7(1), 27–31. <https://doi.org/10.1111/cdep.12008>
- Farrington, D. P. (1991). Longitudinal research strategies: Advantages, problems, and prospects. *Journal of the American Academy of Child & Adolescent Psychiatry*, 30(3), 369–374. <https://doi.org/10.1097/00004583-199105000-00003>
- Ferrer, E., & McArdle, J. J. (2010). Longitudinal modeling of developmental changes in psychological research. *Current Directions in Psychological Science*, 19(3), 149–154. <https://doi.org/10.1177/0963721410370300>
- Fish, J. N. (2020). Future directions in understanding and addressing mental health among LGBTQ youth. *Journal of Clinical Child & Adolescent Psychology*, 49(6), 943–956. <https://doi.org/10.1080/15374416.2020.1815207>

- Følling, I. S., Klöckner, C., Devle, M. T., & Kulseng, B. (2022). Preventing type 2 diabetes, overweight and obesity in the Norwegian primary healthcare: A longitudinal design with 60 months follow-up results and a cross-sectional design with comparison of dropouts versus completers. *BMJ Open*, 12(3), e054841. <https://doi.org/10.1136/bmjopen-2021-054841>
- Ford, D. H., & Lerner, R. M. (1992). *Developmental systems theory: An integrative approach*. Sage Publications, Inc.
- Goldstein, H. (1979). *The design and analysis of longitudinal studies: Their role in the measurement of change*. Academic Press.
- Gottfried, A. W. (1985). Measures of socioeconomic status in child development research: Data and recommendations. *Merrill-Palmer Quarterly*, 31(1), 85–92. <http://www.jstor.org/stable/23086136>
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60, 549–576.
- Greene, T., Gross, R., Amsel, L., & Hoven, C. W. (2018). The mental health of children and adolescents exposed to 9/11: Lessons learned and still to be learned. In A. Szema (ed.), *World Trade Center pulmonary diseases and multi-organ system manifestations* (pp. 549–576). Cham: Springer. https://doi.org/10.1007/978-3-319-59372-2_9
- Grimm, K. J., An, Y., McArdle, J. J., Zonderman, A. B., & Resnick, S. M. (2012). Recent changes leading to subsequent changes: Extensions of multivariate latent difference score models. *Structural Equation Modeling: A Multidisciplinary Journal*, 19(2), 268–292. <https://doi.org/10.1080/10705511.2012.659627>
- Grimm, K. J., Davoudzadeh, P., & Ram, N. (2017). Developments in the analysis of longitudinal data. *Monographs of the Society for Research in Child Development*, 82(2), 46–66. <https://doi.org/10.1111/mono.12298>
- Harachi, T. W., Choi, Y., Abbott, R. D., Catalano, R. F., & Bliesner, S. L. (2006). Examining equivalence of concepts and measures in diverse samples. *Prevention Science*, 7(4), 359–368. <https://doi.org/10.1007/s11121-006-0039-0>
- Hartmann, D. P. (2005). Assessing growth in longitudinal investigations: Selected measurement and design issues. In D. M. Teti (ed.), *Handbook of research methods in developmental science* (pp. 320–339). <https://doi.org/10.1002/9780470756676.ch16>
- Hertzog, C. (1996). Research design in studies of aging and cognition. In E. Birren, K. W. Schaie, R. P. Abeles, M. Gatz, & T. A. Salthouse (eds.), *Handbook of the psychology of aging* (4th ed., pp. 24–37). Academic Press.
- Hill, K. G., Woodward, D., Woelfel, T., Hawkins, J. D., & Green, S. (2016). Planning for long-term follow-up: Strategies learned from longitudinal studies. *Prevention Science*, 17(7), 806–818. <https://doi.org/10.1007/s11121-015-0610-7>
- Hipwell, A. E., Tung, I., Krafty, R. T., Leong, A. W., Spada, M., Vaccaro, H., Homitsky, S. C., Moses-Kolko, E., & Keenan, K. (2023). A lifespan perspective on depression in the postpartum period in a racially and socioeconomically diverse sample of young mothers. *Psychological Medicine*, 53(10), 4415–4423. <https://doi.org/10.1017/S0033291722001210>
- Hollingshead, A. B. (1975). *Four factor index of social status*. Unpublished Working Paper, Department of Sociology Yale New Haven, CT
- Huston, A. C., McLoyd, V. C., & Coll, C. G. (1994). Children and poverty: Issues in contemporary research. *Child Development*, 65(2), 275–282. <https://doi.org/10.1111/j.1467-8624.1994.tb00750.x>

- Jager, J., Putnick, D. L., & Bornstein, M. H. (2017). II. More than just convenient: The scientific merits of homogeneous convenience samples. *Monographs of the Society for Research in Child Development*, 82(2), 13–30. <https://doi.org/10.1111/mono.12296>
- Jeličić, H., Phelps, E., & Lerner, R. M. (2009). Use of missing data methods in longitudinal studies: The persistence of bad practices in developmental psychology. *Developmental Psychology*, 45(4), 1195–1199. <https://doi.org/10.1037/a0015665>
- Jordan, T. E. (1994). The arrow of time: Longitudinal study and its applications. *Genetic, Social, and General Psychology Monographs*, 120(4), 469–531.
- Kosloski, K. (1986). Isolating age, period, and cohort effects in developmental research: A critical review. *Research on Aging*, 8(4), 460–479. <https://doi.org/10.1177/0164027586008004002>
- Kraemer, H. C. (1994). Special methodological problems of childhood developmental follow-up studies: Focus on planning. In S. L. Friedman & H. C. Haywood (eds.), *Developmental follow-up: Concepts, domains, and methods* (pp. 259–276). Academic Press.
- Kraemer, H. C., Yesavage, J. A., Taylor, J. L., & Kupfer, D. (2000). How can we learn about developmental processes from cross-sectional studies, or can we? *American Journal of Psychiatry*, 157(2), 163–171. <https://doi.org/10.1176/appi.ajp.157.2.163>
- Kushalnagar, P., Bruce, S., Sutton, T., & Leigh, I. W. (2017). Retrospective basic parent-child communication difficulties and risk of depression in deaf adults. *Journal of Developmental and Physical Disabilities*, 29(1), 25–34. <https://doi.org/10.1007/s10882-016-9501-5>
- Lerner, R. M. (2006). Developmental science, developmental systems, and contemporary theories of human development. In R. M. Lerner, W. Damon, & R. M. Lerner (eds.), *Handbook of Child Psychology* (pp. 1–17). John Wiley & Sons, Inc. <https://doi.org/10.1002/9780470147658.chpsy0101>
- Levy, P. S., & Lemeshow, S. (2013). *Sampling of populations: Methods and applications*. John Wiley & Sons.
- Little, T. D. (2013). *The Oxford handbook of quantitative methods: Statistical analysis* (Vol. 2). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199934898.001.0001>
- Little, T. D., Jorgensen, T. D., Lang, K. M., & Moore, E. W. G. (2014). On the joys of missing data. *Journal of Pediatric Psychology*, 39(2), 151–162. <https://doi.org/10.1093/jpepsy/jst048>
- Little, T. D., & Rhemtulla, M. (2013). Planned missing data designs for developmental researchers. *Child Development Perspectives*, 7(4), 199–204. <https://doi.org/10.1111/cdep.12043>
- MacDonald, S. W. S., & Stawski, R. S. (2016). Chapter 2 – Methodological considerations for the study of adult development and aging. In K. W. Schaie & S. L. Willis (eds.), *Handbook of the psychology of aging* (8th ed., pp. 15–40). Academic Press. <https://doi.org/10.1016/B978-0-12-411469-2.00002-9>
- McLoyd, V. C. (1998). Changing demographics in the American population: Implications for research on minority children and adolescents. In V. C. McLoyd & L. Steinberg (eds.), *Studying minority adolescents: Conceptual, methodological, and theoretical issues* (pp. 3–28). Lawrence Erlbaum Associates Publishers. <https://doi.org/10.4324/9781410601506>
- Mednick, S. A., Griffith, J. J., & Mednick, B. R. (1981). Problems with traditional strategies in mental health research. In F. Schulzinger, S. A. Mednick, & J. Knop (eds.), *Longitudinal research: Methods and uses in behavioral science* (pp. 3–15). Springer Netherlands. https://doi.org/10.1007/978-94-009-8147-8_1

- Miller, S. A. (1998). *Developmental research methods* (2nd ed.). Prentice-Hall, Inc.
- Murray, A. L., Xie, T., Power, L., & Condon, L. (2024). Recruitment and retention of adolescents for an ecological momentary assessment measurement burst mental health study: The MHIM engagement strategy. *Health Expectations*, 27(3), e14065. <https://doi.org/10.1111/hex.1406>
- Nesselroade, J. R. (1991). Interindividual differences in intraindividual change. In *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 92–105). American Psychological Association. <https://doi.org/10.1037/10099-006>
- Nicholson, J. S., Deboeck, P. R., & Howard, W. (2017). Attrition in developmental psychology: A review of modern missing data reporting and practices. *International Journal of Behavioral Development*, 41(1), 143–153. <https://doi.org/10.1177/0165025415618275>
- Odom, S. L., Horner, R. H., Snell, M. E., & Blacher, J. (2007). The construct of developmental disabilities. In S. L. Odom, R. H. Horner, M. E. Snell, & J. Blacher (eds.), *Handbook of developmental disabilities* (pp. 3–14). The Guilford Press.
- Okazaki, S., & Sue, S. (2016). *Methodological issues in assessment research with ethnic minorities*. American Psychological Association. <https://doi.org/10.1037/14805-015>
- Overton, W. F., & Reese, H. W. (1981). Conceptual prerequisites for an understanding of stability-change and continuity-discontinuity. *International Journal of Behavioral Development*, 4(1), 99–123. <https://doi.org/10.1177/016502548100400107>
- Pfund, G. N., Hofer, M., Allemann, M., & Hill, P. L. (2022). Being social may be purposeful in older adulthood: A measurement burst design. *The American Journal of Geriatric Psychiatry*, 30(7), 777–786. <https://doi.org/10.1016/j.jagp.2021.11.009>
- Phinney, J. S., & Landin, J. (1998). Research paradigms for studying ethnic minority families within and across groups. In V. C. McLoyd & L. Steinberg (eds.), *Studying minority adolescents: Conceptual, methodological, and theoretical issues* (pp. 89–109). Lawrence Erlbaum Associates Publishers. <https://doi.org/10.4324/9781410601506>
- Raifman, S., DeVost, M. A., Digitale, J. C., Chen, Y.-H., & Morris, M. D. (2022). Respondent-driven sampling: A sampling method for hard-to-reach populations and beyond. *Current Epidemiology Reports*, 9(1), 38–47. <https://doi.org/10.1007/s40471-022-00287-8>
- Raphael, K. (1987). Recall bias: A proposal for assessment and control. *International Journal of Epidemiology*, 16(2), 167–170. <https://doi.org/10.1093/ije/16.2.167>
- Rose, S. A., Feldman, J. E., & Jankowski, J. J. (2009). Information processing in toddlers: Continuity from infancy and persistence of preterm deficits. *Intelligence*, 37(3), 311–320. <https://doi.org/10.1016/j.intell.2009.02.002>
- Russell, S. T., & Fish, J. N. (2016). Mental health in Lesbian, Gay, Bisexual, and Transgender (LGBT) youth. *Annual Review of Clinical Psychology*, 12(1), 465–487. <https://doi.org/10.1146/annurev-clinpsy-021815-093153>
- Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin*, 64(2), 92–107. <https://doi.org/10.1037/h0022371>
- Schaie, K. W. (1983). What can we learn from the longitudinal study of adult psychological development? In K. W. Schaie (ed.), *Longitudinal studies of adult psychological development* (pp. 1–19). Guilford.
- Schaie, K. W. (1984). Historical time and cohort effects. In K. A. McCluskey & H. W. Reese (eds.), *Life-span developmental psychology: Historical and generational effects* (pp. 107–121). Academic Press.

- Schaie, K. W. (1986). Beyond calendar definitions of age, time, and cohort: The general developmental model revisited. *Developmental Review*, 6(3), 252–277. [https://doi.org/10.1016/0273-2297\(86\)90014-6](https://doi.org/10.1016/0273-2297(86)90014-6)
- Schaie, K. W. (1988). *Ageism in psychological research* <https://doi.org/10.1037/0003-066X.43.3.179>. American Psychological Association.
- Schaie, K. W. (1998). Advances in longitudinal research methodology. *Australasian Journal on Ageing*, 17(s1), 78–81. <https://doi.org/10.1111/j.1741-6612.1998.tb00869.x>
- Schaie, K. W., & Caskie, I. L. (2005). Methodological issues in aging research. In D. M. Teti (ed.), *Handbook of research methods in developmental science* (pp. 19–39). <https://doi.org/10.1002/9780470756676.ch2>
- Schulenberg, J., Maslowsky, J., & Jager, J. (2018). Substance use and abuse during adolescence and the transition to adulthood are developmental phenomena: Conceptual and empirical considerations. In H. E. Fitzgerald & L. I. Puttler (eds.), *Alcohol use disorders: A developmental science approach to etiology* (pp. 199–222). Oxford University Press. <https://doi.org/10.1093/oso/9780190676001.001.0001>
- Schwartz, S. J., Syed, M., Yip, T., Knight, G. P., Umaña-Taylor, A. J., Rivas-Drake, D., & Lee, R. M. (2014). Methodological issues in ethnic and racial identity research with ethnic minority populations: Theoretical precision, measurement issues, and research designs. *Child Development*, 85(1), 58–76. <https://doi.org/10.1111/cdev.12201>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Mifflin and Company: Houghton.
- Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4(1), 1–32. <https://doi.org/10.1146/annurev.clinpsy.3.022806.091415>
- Slater, A. (1995). Individual differences in infancy and later IQ. *Journal of Child Psychology and Psychiatry*, 36(1), 69–112. <https://doi.org/10.1111/j.1469-7610.1995.tb01656.x>
- Sliwinski, M. J. (2008). Measurement-burst designs for social health research. *Social and Personality Psychology Compass*, 2(1), 245–261. <https://doi.org/10.1111/j.1751-9004.2007.00043.x>
- Stawski, R. S., MacDonald, S. W. S., & Sliwinski, M. J. (2015). Measurement burst design. In S. K. Whitbourne (ed.), *The Encyclopedia of Adulthood and Aging* (pp. 1–5). <https://doi.org/10.1002/9781118521373.wbeaa313>
- Stolz, E., Schultz, A., Mayerl, H., Roller-Wirnsberger, R., & Andrew, C. (2024). Revisiting unstable disability and the fluctuations of frailty: A measurement burst approach. *Age and Ageing*, 53(8). <https://doi.org/10.1093/ageing/afae170>
- Syed, M., Santos, C., Yoo, H. C., & Juang, L. P. (2018). Invisibility of racial/ethnic minorities in developmental science: Implications for research and institutional practices. *American Psychologist*, 812–826. <https://doi.org/10.1037/amp0000294>
- Teicher, M. H., Samson, J. A., Polcari, A., & Andersen, S. L. (2009). Length of time between onset of childhood sexual abuse and emergence of depression in a young adult sample: A retrospective clinical report. *The Journal of Clinical Psychiatry*, 70(5), 684–691. <https://doi.org/10.4088/JCP.08m04235>
- Teti, D. M. (2017). Infant sleep in a family context. In S. M. McHale, V. King, & O. M. Buxton (eds.), *Family contexts of sleep and health across the life course* (pp. 3–23). Springer International Publishing. https://doi.org/10.1007/978-3-319-64780-7_1
- Toomey, R. B. (2021). Advancing research on minority stress and resilience in trans children and adolescents in the 21st century. *Child Development Perspectives*, 15(2), 96–102. <https://doi.org/10.1111/cdev.12405>

- Umberson, D., Thomeer, M. B., Kroeger, R. A., Lodge, A. C., & Xu, M. (2015). Challenges and opportunities for research on same-sex relationships. *Journal of Marriage and Family*, 77(1), 96–111. <https://doi.org/10.1111/jomf.12155>
- Van der Kamp, L. J. T., Bijleveld, C. C. J. H., & London, E. S. (1998). Methodological issues in longitudinal research. In C. C. J. H. Bijleveld, L. J. T. Van der Kamp, A. Mooijart, W. A. van der Kloot, R. van der Leeden, & E. Van Der Burg (eds.), *Longitudinal data analysis: Designs, models and methods* (pp. 1–44). Sage Publications Ltd.
- Walls, T. A., Barta, W. D., Stawski, R. S., Collyer, C. E., & Hofer, S. M. (2012). Time-scale-dependent longitudinal designs. In B. Laursen, T. D. Little, & N. A. Card (eds.), *Handbook of developmental research methods* (pp. 46–64). The Guilford Press.
- Widom, C. S., & Morris, S. (1997). Accuracy of adult recollections of childhood victimization, Part 2: Childhood sexual abuse. *Psychological Assessment*, 9(1), 34–46. <https://doi.org/10.1037/1040-3590.9.1.34>
- Wohlwill, J. F. (1970). Methodology and research strategy in the study of developmental change. In *Life-span developmental psychology* (pp. 149–191). Academic Press.
- Wohlwill, J. F. (1973). *The study of behavioral development*. Academic Press.
- Wu, W., & Jia, F. (2021). Applying planned missingness designs to longitudinal panel studies in developmental science: An overview. *New Directions for Child and Adolescent Development*, 2021(175), 35–63. <https://doi.org/10.1002/cad.20391>