Is Timing Everything? The Effects of Measurement Timing on the Performance of Nonlinear Longitudinal Models

by

Sebastian L.V. Sciarra

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of

Doctorate of Philosophy

in

Psychology

in the field of Industrial-Organizational Psychology

Guelph, Ontario, Canada

ABSTRACT

IS TIMING EVERYTHING? THE EFFECTS OF MEASUREMENT TIMING ON
THE PERFORMANCE OF NONLINEAR LONGITUDINAL MODELS

Sebastian L.V. Sciarra

University of Guelph, 2023

Advisor(s):

David Stanley

Despite the value that longitudinal research offers for understanding psychological processes, studies in organizational research rarely use longitudinal designs. One reason for the paucity of longitudinal designs may be the challenges they present for researchers. Three challenges of particular importance are that researchers have to determine 1) how many measurements to take, 2) how to space measurements, and 3) how to design studies when participants provide data with different response schedules (time unstructuredness). In systematically reviewing the simulation literature, I found that few studies comprehensively investigated the effects of measurement number, measurement spacing, and time structuredness (in addition to sample size) on model performance. As a consequence, researchers have little guidance when trying to conduct longitudinal research. To address these gaps in the literature, I conducted a series of simulation experiments. I found poor model performance across all measurement number/sample size pairings. That is, bias and precision were never concurrently optimized under any combination of manipulated variables. Bias was often low, however, with moderate measurement numbers and sample sizes. Although precision was frequently poor, the greatest improvements in precision resulted from using either seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$. With time-unstructured data, model performance systematically

decreased across all measurement number/sample size pairings when the model incorrectly assumed an identical response pattern across all participants (i.e., time-structured data). Fortunately, when models were equipped to handle heterogeneous response patterns using definition variables, the poor model performance observed across all measurement number/sample size pairings no longer appeared. Altogether, the results of the current simulation experiments provide guidelines for researchers interested in modelling nonlinear change.

DEDICATION

I dedicate this document to Don Cameron. Without you, none of this would have been possible.

ACKNOWLEDGEMENTS

First and foremost, I would like to extend my unending gratitude to my supervisor, David Stanley. Without your guidance, I would never have been able to put this document together. Whenever I had problems to solve, you always managed to help me find solutions, and whenever I set my sights on learning new topics, you always provided encouragement, regardless if my learning required a day or a month. You taught me how to write, how to code, and, most importantly, how to think.

Second, I would like to thank my committee members. To Drs. David Walters and Ian Newby-Clark, I appreciate all your feedback and contributions to my dissertation, and for unknowingly taking on the responsibility to read such a lengthy document. I also extend a similar gratitude to Dr. Jeffrey Spence for reading through a document that was certainly not a page-turner, and also for your excellent teaching of statistics and all our great discussions.

Third, and certainly not least, I would like to thank my family. I would like to thank my father, Massimo, for tolerating my unnecessarily demanding cooking expeditions, my mother, Gabriele, for her patience, my brother, Alessandro, for listening to my rants about cooking and statistics, and my grandmother, Annemarie, for telling my mother to calm down when dinner had not yet been served at exactly 7:30.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1 Introduction

> "Neither the behavior of human beings nor the activities of organizations can be defined without reference to time, and temporal aspects are critical for understanding them" (Navarro et al., 2015, p. 136).

The topic of time has received considerable attention in organizational psychology over the past 20 years. Examples of well-received articles published around the beginning of the 21st century have discussed how investigating time is important for understanding patterns of change and boundary conditions of theory (Zaheer et al., 1999), how longitudinal research is necessary for disentangling different types of causality (T. R. Mitchell & James, 2001), and explicated patterns of organizational change (or institutionalization; Lawrence et al., 2001). Since then, articles have emphasized the need to address time in specific areas such as performance (Dalal et al., 2014; C. D. Fisher, 2008), teams (Roe et al., 2012), and goal setting (Fried & Slowik, 2004) and, more generally, throughout organizational research (Aguinis & Bakker, 2021; George & Jones, 2000; Kunisch et al., 2017; Navarro et al., 2015; Ployhart & Vandenberg, 2010; Roe, 2008; Shipp & Cole, 2015; Sonnentag, 2012; Vantilborgh et al., 2018).

The importance of time has also been recognized in organizational theory. In defining a theoretical contribution, Whetten (1989) stated that time must be discussed in setting boundary conditions (i.e., under what circumstances does the theory apply) and in specifying relations between variables over time (George & Jones, 2000; T. R. Mitchell & James, 2001). Even if a considerable number of organizational theories do not adhere to the definition of Whetten (1989), theoretical models in organizational psychology consist of path diagrams that delineate the causal events of processes. Given that temporal

precedence is a necessary condition for establishing causality (Mill, 2011), time has a role, whether implicitly or explicitly, in organizational theory.

Despite the considerable attention given towards investigating processes over time and the ubiquity of time in organizational theory, the prevalence of longitudinal research has historically remained low. One study examined the prevalence of longitudinal research from 1970–2006 across five organizational psychology journals and found that 4% of articles used longitudinal designs (Roe, 2014, September 22–26). Another survey of two applied psychology journals in 2005 found that approximately 10% (10 of 105 studies) of studies used longitudinal designs (Roe, 2008). Similarly, two surveys of studies employing longitudinal designs with mediation analysis found that, across five journals, only about 10% (7 of 72 studies) did so in 2005 (Maxwell & Cole, 2007) and approximately 16% (15 of 92 studies) did so in 2006 (M. A. Mitchell & Maxwell, 2013).[1] Thus, the prevalence of longitudinal research has remained low.

In the seven sections that follow, I will explain why longitudinal research is necessary and the factors that must be considered when conducting such research. In the first section, I will explain why conducting longitudinal research is essential for understanding the dynamics of psychological processes. In the second section, I will overview patterns of change that are likely to emerge over time. In the third section, I will overview design and analytical issues involved in conducting longitudinal studies. In the fourth section, I will explain how design and analytical issues encountered in conducting longitudinal research can be investigated. In the fifth section, I will provide a systematic review of the research

_____

[1]Note that the definition of a longitudinal design in Maxwell and Cole (2007) and M. A. Mitchell and Maxwell (2013) required that measurements be taken over at least three time points so that measurements of the predictor, mediator, and outcome variables were separated over time.

that has investigated design and analytical issues involved in conducting longitudinal research. Finally, in the sixth and seventh sections, I will, respectively, discuss some methods for modelling nonlinear change and the frameworks in which they can be used. A summary of the three simulation experiments that I conducted in my dissertation will then be provided.

1.1 The Need to Conduct Longitudinal Research

Longitudinal designs provide several advantages over cross-sectional designs that allow them to more accurately investigate change (e.g., temporal precedence, testing reverse causality). Unfortunately, even though longitudinal studies often produce results that differ from those of cross-sectional studies, researchers commonly discuss the results of cross-sectional studies as if they have been obtained with a longitudinal design. One example of the assumption of equivalence between cross-sectional and longitudinal findings comes from the large number of studies employing mediation analysis. Given that mediation is used to understand chains of causality in psychological processes (Baron & Kenny, 1986), it would thus make sense to pair mediation analysis with a longitudinal design because understanding causality, after all, requires temporal precedence. Unfortunately, the majority of studies that have used mediation analysis have done so using cross-sectional designs—with estimates of approximately 90% (Maxwell & Cole, 2007) and 84% (M. A. Mitchell & Maxwell, 2013)—and often discuss the results as if they are longitudinal. Investigations into whether mediation results remain equivalent across cross-sectional and longitudinal designs have repeatedly concluded that using mediation analysis on cross-sectional data can return different, and sometimes completely opposite, results from using it on longitudinal data (Cole & Maxwell, 2003; Maxwell & Cole, 2007;

3

Maxwell et al., 2011; M. A. Mitchell & Maxwell, 2013; O'Laughlin et al., 2018). Therefore, mediation analyses based on cross-sectional analyses may be misleading.

The non-equivalence of cross-sectional and longitudinal results that occurs with mediation analysis is, unfortunately, not due to a specific set of circumstances that only arise with mediation analysis, but a consequence of a broader systematic cause that affects the results of many analyses. The concept of ergodicity explains why cross-sectional and longitudinal analyses seldom yield similar results. To understand ergodicity, it is first important to realize that variance is central to many statistical analyses—correlation, regression, factor analysis, and mediation are some examples. Thus, if variance remains unchanged across cross-sectional and longitudinal data sets, then analyses of either data set would return the same results. Importantly, variance only remains equal across cross-sectional and longitudinal data sets if two conditions put forth by ergodic theory are satisfied (homogeneity and stationarity; Molenaar, 2004; Molenaar & Campbell, 2009). If these two conditions are met, then a process is said to be ergodic. Unfortunately, the two conditions required for ergodicity are highly unlikely to be satisfied and so cross-sectional findings will frequently deviate from longitudinal findings (for a detailed discussion, see Appendix **??**).

Given that cross-sectional and longitudinal analyses are, in general, unlikely to return equivalent findings, it is unsurprising that several investigations in organizational research—and psychology as a whole—have found these analyses to return different results. Beginning with an example from Curran and Bauer (2011), heart attacks are less likely to occur in people who exercise regularly (longitudinal finding), but more likely to happen when exercising (cross-sectional finding). Correlational studies find differences in

correlation magnitudes between cross-sectional and longitudinal data sets (for a meta-analytic review, see A. J. Fisher et al., 2018; Nixon et al., 2011).[2] Moving on to perhaps the most commonly employed analysis in organizational research of mediation, several articles have highlighted that cross-sectional data can return different, and sometimes completely opposite, results than those obtained from longitudinal data (Cole & Maxwell, 2003; Maxwell & Cole, 2007; Maxwell et al., 2011; O'Laughlin et al., 2018). Factor analysis is perhaps the most interesting example: The well-documented five-factor model of personality seldom arises when analyzing person-level data consisting of personality measurements over 90 consecutive days (Hamaker et al., 2005). Therefore, cross-sectional analyses are rarely equivalent to longitudinal analyses.

With longitudinal analyses often producing results that differ from those of cross-sectional analyses, it is paramount that longitudinal designs be used to more accurately understand change. Fortunately, technological advancements have allowed researchers to more easily conduct longitudinal research in two ways. First, the use of the experience sampling method (Beal, 2015) in conjunction with modern information transmission technologies—whether through phone applications or short message services—allows data to often be sampled over time with relative ease. Second, the development of longitudinal analyses (along with their integration in commonly used software) that enable person-level data to be modelled such as multilevel models (Raudenbush & Bryk, 2002), growth mixture models (M. Wang & Bodner, 2007), and dynamic factor analysis (Ram et al.,

---

[2]Note that A. J. Fisher et al. (2018) also found the variability of longitudinal correlations to be considerably larger than the variability of cross-sectional correlations.

2013) provide researchers with avenues to explore the temporal dynamics of psychological processes. With one recent survey estimating that 43.3% of mediation studies (26 of 60 studies) used a longitudinal design (O'Laughlin et al., 2018), it appears that the prevalence of longitudinal research has increased from the 9.5% (Roe, 2008) and 16.3% (M. A. Mitchell & Maxwell, 2013) values estimated at the beginning of the 21st century. Although the frequency of longitudinal research appears to have increased over the past 20 years, several avenues exist where the quality of longitudinal research can be improved, and in my dissertation, I focus on investigating these avenues.

1.2 Understanding Patterns of Change That Emerge Over Time

Change can occur in many ways over time. One pattern of change commonly assumed to occur over time is that of linear change. When change follows a linear pattern, the rate of change over time remains constant. Unfortunately, a linear pattern places demanding restrictions on the possible trajectories of change. If change were to follow a linear pattern, then any pauses in change (or plateaus) or changes in direction could not occur: Change would simply grow over time. Unfortunately, effect sizes have been shown to diminish over time after peaking (for meta-analytic examples, see Cohen, 1993; Griffeth et al., 2000; Hom et al., 1992; Riketta, 2008; Steel et al., 1990; Steel & Ovalle, 1984). Moreover, many variables display cyclic patterns of change over time, with mood (Larsen & Kasimatis, 1990), daily stress (Bodenmann et al., 2010), and daily drinking behaviour (Huh et al., 2015) as some examples. Therefore, change over time is unlikely to follow a linear pattern.

A more realistic pattern of change to occur over time is a nonlinear pattern (for a review, see Cudeck & Harring, 2007). Nonlinear change allows the rate of change to

be nonconstant; that is, change may occur more rapidly during certain periods of time, stop altogether, or reverse direction. When looking at patterns of change observed across psychology, several examples of nonlinear change have been found in the declining rate of speech errors throughout child development (Burchinal & Appelbaum, 1991), rates of forgetting (Murre & Dros, 2015), development of habits (Fournier et al., 2017), and the formation of opinions (Xia et al., 2020). Given that nonlinear change appears more likely than linear change, my dissertation will assume change over time to be nonlinear.

1.3 Challenges Involved in Conducting Longitudinal Research

Conducting longitudinal research presents researchers with several challenges. Many challenges are those from cross-sectional research only amplified (for a review, see Bergman & Magnusson, 1990).[3] For example, greater efforts have to be made to prevent missing data which can increase over time (Dillman et al., 2014; Newman, 2008). Likewise, the adverse effects of well-documented biases such as demand characteristics (Orne, 1962) and social desirability (Nederhof, 1985) have to be countered at each time point. Outside of challenges shared with cross-sectional research, conducting longitudinal research also presents new challenges. Analyses of longitudinal data have to consider complications such as how to model error structures (Grimm & Widaman, 2010), check for measurement non-invariance over time (the extent to which a construct is measured with the same measurement model over time; Mellenbergh, 1989), and how to center/process data to appropriately answer research questions (Enders & Tofighi, 2007; L. Wang & Maxwell, 2015).

_____

[3]It should be noted that conducting a longitudinal study does alleviate some issues encountered in conducting cross-sectional research. For example, taking measurements over multiple time points likely reduces common method variance (Podsakoff et al., 2003; for an example, see Ostroff et al., 2002).

Although researchers must contend with several issues in conducting longitudinal research, three issues are of particular interest in my dissertation. The first issue concerns how many measurements to use in a longitudinal design. The second issue concerns how to space the measurements. The third issue focuses on how much error is incurred if the time structuredness of the data is overlooked. The sections that follow will review each of these issues.

1.3.1 Number of Measurements

Researchers have to decide on the number of measurements to include in a longitudinal study. Although using more measurements increases the accuracy of results—as noted in the results of several studies (e.g., Coulombe et al., 2016; Finch, 2017; Fine et al., 2019; Timmons & Preacher, 2015)—taking additional measurements often comes at a cost that a researcher may be unable to absorb given a limited budget. One important point to mention is that a researcher designing a longitudinal study must take at least three measurements to allow a reliable estimate of change and, perhaps more importantly, to allow a nonlinear pattern of change to be modelled (Ployhart & Vandenberg, 2010). In my dissertation, I hope to determine whether an optimal number of measurements exists when modelling a nonlinear pattern of change.

1.3.2 Spacing of Measurements

Additionally, a researcher must decide on the spacing of measurements in a longitudinal study. Although discussions of measurement spacing often recommend that researchers use theory and previous studies to determine measurement spacing (Cole & Maxwell, 2003; Collins, 2006; Dormann & Griffin, 2015; Dormann & van de Ven, 2014; T. R. Mitchell & James, 2001), organizational theories seldom delineate periods of time

over which processes unfold, and so the majority of longitudinal research uses intervals of convention and/or convenience to space measurements (Dormann & van de Ven, 2014; T. R. Mitchell & James, 2001). Unfortunately, using measurement spacings that do not account for the temporal pattern of change of a psychological process can lead to inaccurate results (e.g., Chen et al., 2014). As an example, Cole and Maxwell (2009) show how correlation magnitudes are affected by the choice of measurement spacing intervals. In my dissertation, I hope to determine whether an optimal measurement spacing schedule exists when modelling a nonlinear pattern of change.

1.3.3 Time Structuredness

Last, and perhaps most pernicious, latent variable analyses of longitudinal data are likely to incur error from an assumption they make about data collection conditions. Latent variable analyses assume that, across all collection points, participants provide their data at the same time. Unfortunately, such a high level of regularity in the response patterns of participants is unlikely: Participants are more likely to provide their data over some period of time after a data collection window has opened. As an example, consider a study that collects data from participants at the beginning of each month. If participants respond with perfect regularity, then they would all provide their data at the exact same time (e.g., noon on the second day of each month). If the participants respond with imperfect regularity, then they would provide their data at different times after the beginning of each month. The regularity of response patterns observed across participants in a longitudinal study determines the time structuredness of the data and the sections that follow will provide an overview of time structuredness.

1.3.3.1 Time-Structured Data

Many analyses assume that data are *time structured*: Participants provide data at the same time at each collection point. By assuming time-structured data, an analysis can incur error because it will map time intervals of inappropriate lengths onto the time intervals that occurred between participant's responses.[4] As an example of the consequences of incorrectly assuming data to be time structured, consider a study that assessed the effects of an intervention on the development of leadership by collecting leadership ratings at four time points each separated by four weeks (Day & Sin, 2011). The employed analysis assumed time-structured data; that is, each each participant provided ratings on the same day—more specifically, the exact same moment—each time these ratings were collected. Unfortunately, it is unlikely that the data collected from participants were time structured: At any given collection point, some participants may have provided leadership ratings at the beginning of the week, while others may only provide ratings two weeks after the survey opened. Importantly, ratings provided two weeks after the survey opened were likely influenced by changes in leadership that occurred over the two weeks. If an analysis incorrectly assumes time-structured data, then it assumes each participant has the same response pattern and, therefore, will incorrectly attribute the amount of time that elapses between most participants' responses. For instance, if a participant only provides a leadership rating two weeks after having received a survey (and six weeks after providing their previous rating), then using an analysis that assumes time-structured data would incorrectly assume that each collection point of this participant is separated

---

[4]It should be noted that, although seldom implemented, analyses can be accessorized to handle time-unstructured data by using definition variables (Mehta & West, 2000; Mehta & Neale, 2005).

by four weeks (the interval used in the experiment) and would, consequently, model the observed change as if it had occurred over four weeks. Therefore, incorrectly assuming data to be time structured leads an analysis to overlook the unique response rates of participants across the collection points and, as a consequence, incur error (Coulombe et al., 2016; Mehta & Neale, 2005; Mehta & West, 2000).

1.3.3.2 Time-Unstructured Data

Conversely, other analyses assume that data are *time unstructured*: Participants provide data at different times at each collection point. Given the unlikelihood of one response pattern describing the response rates of all participants in a given study, the data obtained in a study are unlikely to be time structured. Instead, and because participants are likely to exhibit unique response patterns in their response rates, data are likely to be time unstructured. One way to conceptualize the distinction between time-structured and time-unstructured data is on a continuum. On one end of the continuum, participants all provide data with identical response patterns, thus giving time-structured data. When participants exhibit unique response patterns, the resulting data are time unstructured, with the extent of time-unstructuredness depending on the average uniqueness of all response patterns. For example, if data are collected at the beginning of each month and participants only have one day to provide data at each time point, then the resulting data will have a low amount of time structuredness because response patterns can only differ from each other over the course of one day. Alternatively, if data are collected at the beginning of each month and participants have 30 days to provide data at each time point, then the resulting data will have a high amount of time structuredness because response patterns can differ from each other over the course of 30 days. Therefore, the continuum

11

of time struturedness has time-structured data on one end and time-unstructured data with long response windows on another end. In my dissertation, I hope to determine how much error is incurred when time-unstructured data of varying degrees are assumed to be time structured.

1.3.4 Summary

In summary, researchers must contend with several issues when conducting longitudinal research. In addition to contending with issues encountered in conducting cross-sectional research, researchers must contend with new issues that arise from conducting longitudinal research. Three issues of particular importance in my dissertation are the number of measurements, the spacing of measurements, and incorrectly assuming time-unstructured data to be time structured. These issues will be serve as a basis for a systematic review of the simulation literature.

1.4 Using Simulations To Assess Model Performance

In the next section, I will present the results of a systematic review of the literature that has investigated the issues of measurement number, measurement spacing, and time structuredness. Before presenting the results of the systematic review, I will provide an overview of the Monte Carlo method used to investigate the issues involved in conducting longitudinal research.

To understand how the effects of longitudinal issues on model performance can be investigated, the inferential method commonly employed in psychological research will first be reviewed with an emphasis on its shortcomings (see Figure 1.1). Consider an example where a researcher wants to understand how sampling error affects the accuracy with which a sample mean ($\bar{x}$) estimates a population mean (μ). Using the inferential

12

method, the researcher samples data and then estimates the population mean (μ) by computing the mean of the sampled data ($\bar{x}_1$). Because collected samples are almost always contaminated by a variety of methodological and/or statistical deficiencies (such as sampling error, measurement error, assumption violations, etc.), the estimation of the population parameter is likely to be imperfect. Unfortunately, to estimate the effect of sampling error on the accuracy of the population mean estimate ($\bar{x}_1$), the researcher would need to know the value of the population mean; without knowing the value of the population mean, it is impossible to know how much error was incurred in estimating the population mean and, as as a result, impossible to know the extent to which sampling error contributed to this error. Therefore, a study following the inferential approach can only provide estimates of population parameters.

The Monte Carlo method has a different goal. Whereas the inferential method focuses on estimating parameters from sample data, the Monte Carlo method is used to understand the factors that influence the accuracy of the inferential approach. Figure 1.1 shows that the Monte Carlo method works in the opposite direction of the inferential approach: Instead of collecting a sample, the Monte Carlo method begins by assigning a value to at least one parameter to define a population. Many sample data sets are then generated from the defined population ($s_1, s_2, ..., s_n$) and the data from each sample are then modelled by computing a sample mean ($\bar{x}_1, \bar{x}_2, ..., \bar{x}_n$). Importantly, manipulations can be applied to the sampling and/or modelling of the data. In the current example,the population estimates of each statistical model are averaged ($\bar{\bar{x}}$) and compared to the pre-determined parameter value (μ). The difference between the average of the estimates and the known population value constitutes bias in parameter estimation (i.e., parameter

13

bias). In the current example, the manipulation causes a systematic underestimation, on average, of the population parameter. By randomly generating data, the Monte Carlo method can estimate how a variety of methodological and statistical factors affect model performance (for a review, see Robert & Casella, 2010).

Monte Carlo simulations have been used to evaluate the effects of a variety of methodological and statistical deficiencies for several decades. Beginning with an early use of the Monte Carlo method, Boneau (1960) used it to evaluate the effects of assumption violations on the fidelity of t-value distributions. In more recent years, implementations of the the Monte Carlo method have shown that realistic values of sample size and measurement accuracy produce considerable variability in estimated correlation values (Stanley & Spence, 2014). Monte Carlo simulations have also provided valuable insights into more complicated statistical analyses. In investigating more complex statistical analyses, simulations have shown that mediation analyses are biased to produce results of complete mediation because the statistical power to detect direct effects falls well below the statistical power to detect indirect effects (Kenny & Judd, 2014). Given the ability of the Monte Carlo method to evaluate statistical methods, the experiments in my dissertation used it to evaluate the effects of measurement number, measurement spacing, and time structuredness on model performance.[5]

1.5 Systematic Review of Simulation Literature

To understand the extent to which issues involved in conducting longitudinal research had been investigated, I conducted a systematic review of the simulation literature.

---

[5]My simulation experiments also investigated the effects of sample size and nature of change on model performance.

Figure 1.1
Depiction of Monte Carlo Method
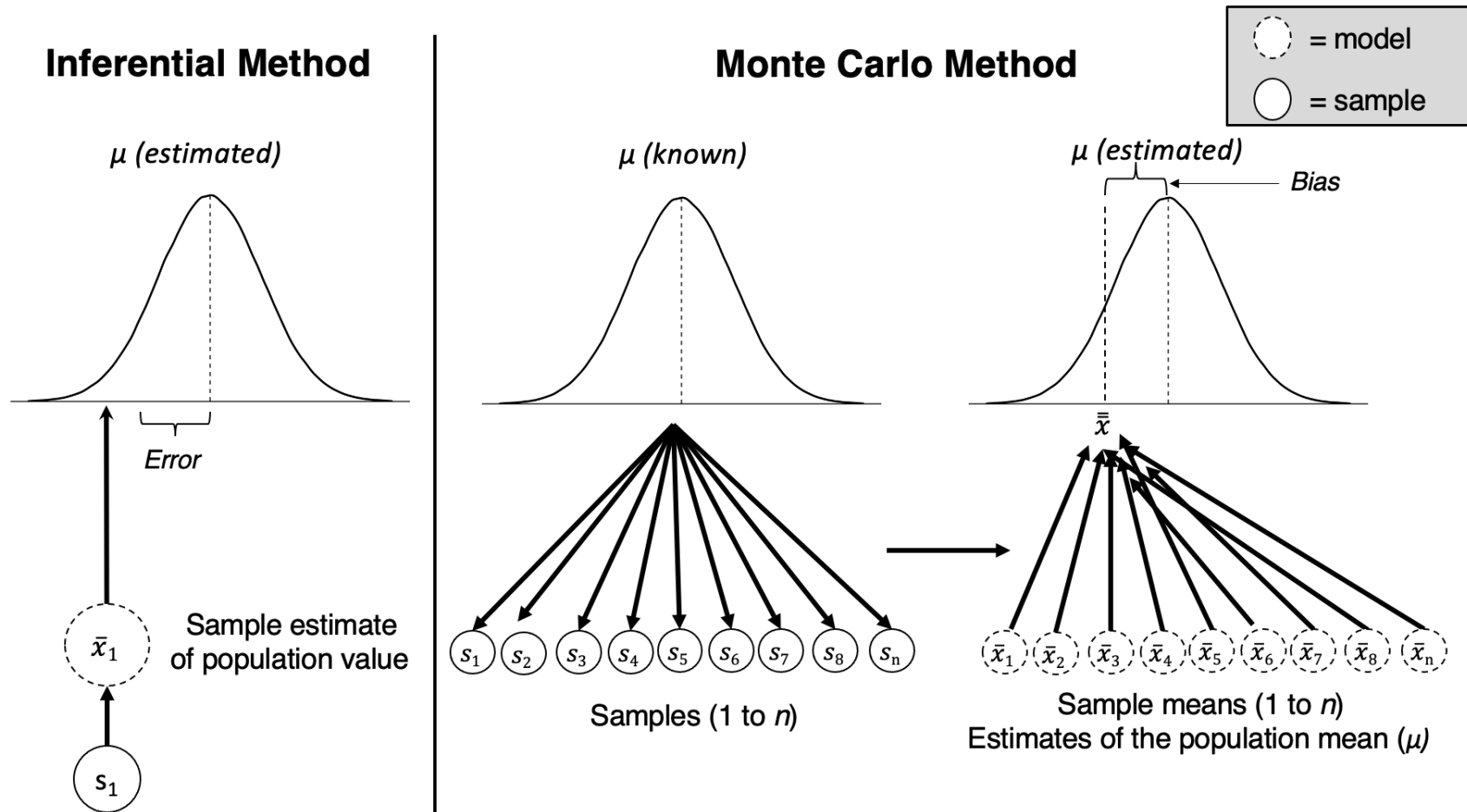
Note. Comparison of inferential approach with the Monte Carlo approach. The inferential approach begins with a collected sample and then estimates the population parameter using an appropriate statistical model. The difference between the estimated and population value can be conceptualized as error.

Because the population value is generally unknown in the inferential approach, it cannot estimate how much error is introduced by any given methodological or statistical deficiency. To estimate how much error is introduced by any given methodological or statistical deficiency, the Monte Carlo method needs to be used, which constitutes four steps. The Monte Carlo method first defines a population by setting parameter values. Second, many samples are generated from the pre-defined population, with some methodological deficiency built in to each data set (in this case, each sample has a specific amount of missing data). Third, each generated sample is then analyzed and the population estimates of each statistical model are averaged and compared to the pre-determined parameter value. Fourth, the difference between the estimate average and the known population value defines the extent to which the missing data manipulation affected parameter estimation (the difference between the population and average estimated population value is the parameter bias).

The sections that follow will first present the method I followed in systematically reviewing the literature and then summarize the findings of the review.

1.5.1 Systematic Review Methodology

I identified the following keywords through citation searching and independent reading: "growth curve", "time-structured analysis", "time structure", "temporal design", "individual measurement occasions", "measurement intervals", "methods of timing", "longitudinal data analysis", "individually-varying time points", "measurement timing", "latent difference score models", "parameter bias", and "measurement spacing". I entered these keywords entered into the PsycINFO database (on July 23, 2021) along with the word "simulation" in any field and considered any returned paper a viable ppaper (see Figure 1.2 for a PRISMA diagram illustrating the filtering of the reports). The search returned 165 reports, which I screened by reading the abstracts. Initial screening led to the removal of 60 reports because they did not contain any simulation experiments. Of the remaining 105 papers, I removed 2 more papers because they could not be accessed (Stockdale, 2007; Tiberio, 2008). Of the remaining 103 identified simulation studies, I deemed a paper as relevant if it investigated the effects of any design and/or analysis factor related to conducting longitudinal research (i.e., number of measurements, spacing of measurements, and/or time structuredness) and did so using the Monte Carlo simulation method. Of the remaining 103 studies, I removed 89 studies because they did not meet the inclusion criteria, leaving fourteen studies to be included in the review. I also found an additional 3 studies through citation searching, giving a total of 17 studies.

The findings of my systematic review are summarized in Tables 1.1–1.2. Tables 1.1–1.2 differ in one way: Table 1.1 indicates how many studies investigated each effect,
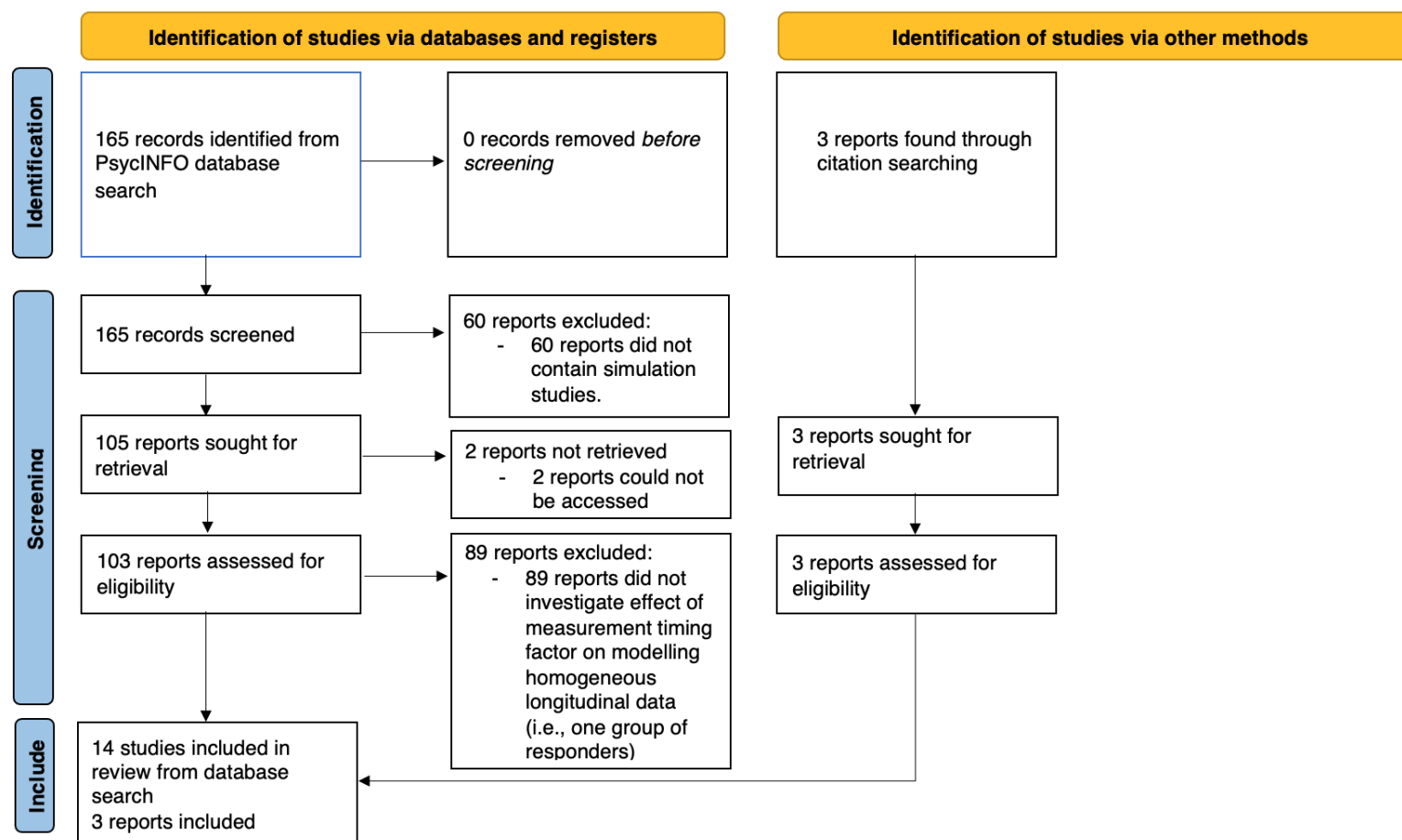
whereas Table 1.2 provides the reference of each study and detailed information about each study's method. Otherwise, all other details of Tables 1.1–1.2 are identical. The first column lists the longitudinal design factor (alongside with sample size) and the corresponding two- and three-way interactions. The second and third columns list whether each effect has been investigated with linear and nonlinear patterns of change, respectively. Shaded cells indicate effects that have not been investigated, with cells shaded in light grey indicating effects that have not been investigated with linear patterns of change and cells shaded in dark grey indicating effects that have not been investigated with nonlinear patterns of change.[6]

1.5.2 Systematic Review Results

Although previous research appeared to sufficiently fill some cells of Table 1.1, two patterns suggest that arguably the most important cells (or effects) have not been investigated. First, it appears that simulation research has invested more effort in investigating the effects of longitudinal design factors with linear patterns than with nonlinear patterns of change. In counting the number of effects that remain unaddressed with linear and nonlinear patterns of change, a total of five cells (or effects) have not been investigated, but a total of seven cells have not been investigated with nonlinear patterns of

---

[6]Table 1.2 lists the effects that each study (identified by my systematic review) investigated and notes the following methodological details (using superscript letters and symbols): the type of model used in each paper, assumption and/or manipulation of complex error structures (heterogeneous variances and/or correlated residuals), manipulation of missing data, and/or pseudo-time structuredness manipulation. Across all 17 simulation studies, 5 studies (29%) assumed complex error structures (Gasimova et al., 2014; Liu & Perera, 2022; Y. Liu et al., 2015; Miller & Ferrer, 2017; Murphy et al., 2011), 1 study (6%) manipulated missing data (Fine et al., 2019), and 2 studies (12%) contained a pseudo-time structuredness manipulation (Fine et al., 2019; Fine & Grimm, 2020). Importantly, the pseudo-time structuredness manipulation used in Fine et al. (2019) and Fine and Grimm (2020) differed from the manipulation of time structuredness used in the current experiments (and from previous simulation experiments of Coulombe et al., 2016; Miller & Ferrer, 2017) in that it randomly generated longitudinal data such that a given person could provide all their data before another person provided any data.

Figure 1.2

PRISMA Diagram Showing Study Filtering Strategy

	Identification of studies via databases and registers		Identification of studies via other methods
Identification	165 records identified from PsycINFO database search	0 records removed *before screening*	3 reports found through citation searching
Screening	165 records screened	60 reports excluded: - 60 reports did not contain simulation studies.	
	105 reports sought for retrieval	2 reports not retrieved - 2 reports could not be accessed	3 reports sought for retrieval
	103 reports assessed for eligibility	89 reports excluded: - 89 reports did not investigate effect of measurement timing factor on modelling homogeneous longitudinal data (i.e., one group of responders)	3 reports assessed for eligibility
Include	14 studies included in review from database search 3 reports included		

Note. PRISMA diagram for systematic review of simulation research that investigates longitudinal design and analysis factors.

19

Table 1.1

Number of Simulation Studies That Have Investigated Longitudinal Issues with Linear and Nonlinear Change Patterns (n = 17)

Effect	Linear pattern	Nonlinear pattern
Main effects		
Number of measurements (NM)	11 studies	6 studies
Spacing of measurements (SM)	1 study	1 study
Time structuredness (TS)	2 studies	1 study
Sample size (S)	11 studies	7 studies
Two-way interactions		
NM x SM	1 study	1 study
NM x TS	1 study	**Cell 1 (Exp. 3)**
NM x S	9 studies	5 studies
SM x TS	**Cell 2**	**Cell 3**
SM x S	**Cell 4**	**Cell 5 (Exp. 2)**
TS x S	1 study	2 studies
Three-way interactions		
NM x SM x TS	**Cell 6**	**Cell 7**
NM x SM x S	**Cell 8**	**Cell 9 (Exp. 2)**
NM x TS x S	1 study	**Cell 10 (Exp. 3)**
SM x TS x S	**Cell 11**	**Cell 12**

Note. Cells are only numbered for effects that have not been investigated. Cells shaded in light and dark grey, respectively indicate effects that have not been investigated with linear and nonlinear patterns of change.

Table 1.2

Summary of Simulation Studies That Have Investigated Longitudinal Issues with Linear and Nonlinear Change Patterns (n = 17)

Effect	Linear pattern	Nonlinear pattern
Main effects		
Number of measurements (NM)	(Timmons & Preacher, 2015, [a]; Murphy et al., 2011, [b↺]; Gasimova et al., 2014, [c↺]; Wu et al., 2014, [a]; Coulombe, 2016, [a]; Ye, 2016, [a]; Finch, 2017, [a]; O'Rourke et al., 2022, [d]; Newsom & Smith, 2020, [a]; Coulombe et al., 2016, [a])	(Timmons & Preacher, 2015, [a]; Finch, 2017, [a]; Fine et al., 2019, [e∘▽]; Fine & Grimm, 2020, [e,f▽]; J. Liu et al., 2022, [g]; Liu & Perera, 2022, [h↺]; Y. Liu et al., 2015, [g↺])
Spacing of measurements (SM)	(Timmons & Preacher, 2015, [a])	(Timmons & Preacher, 2015, [a])
Time structuredness (TS)	(Aydin et al., 2014, [a]; Coulombe et al., 2016, [a])	(Miller & Ferrer, 2017, [a↺]; Y. Liu et al., 2015, [g↺])
Sample size (S)	(Murphy et al., 2011, [b↺]; Gasimova et al., 2014, [c↺]; Wu et al., 2014, [a]; Coulombe, 2016, [a]; Ye, 2016, [a]; Finch, 2017, [a]; O'Rourke et al., 2022, [d]; Newsom & Smith, 2020, [a]; Coulombe et al., 2016, [a]; Aydin et al., 2014, [a])	(Finch, 2017, [a]; Fine et al., 2019, [e∘▽]; Fine & Grimm, 2020, [e,f▽]; J. Liu et al., 2022, [g]; Liu & Perera, 2022, [h↺]; Y. Liu et al., 2015, [g↺]; Miller & Ferrer, 2017, [a↺])
Two-way interactions		
NM x SM	(Timmons & Preacher, 2015, [a])	(Timmons & Preacher, 2015, [a])
NM x TS	(Coulombe et al., 2016, [a])	**Cell 1 (Exp. 3)**

Table 1.2

Summary of Simulation Studies That Have Investigated Longitudinal Issues with Linear and Nonlinear Change Patterns (n = 17) (continued)

Effect	Linear pattern	Nonlinear pattern
NM x S	(Murphy et al., 2011, [b][ʊ]; Gasimova et al., 2014, [c][ʊ]; Wu et al., 2014, [a]; Coulombe, 2016, [a]; Ye, 2016, [a]; Finch, 2017, [a]; O'Rourke et al., 2022, [d]; Newsom & Smith, 2020, [a]; Coulombe et al., 2016, [a])	(Finch, 2017, [a]; Fine et al., 2019, [e][○][▽]; Fine & Grimm, 2020, [e,f][▽]; J. Liu et al., 2022, [g]; Liu & Perera, 2022, [h][ʊ])
SM x TS	**Cell 2**	**Cell 3**
SM x S	**Cell 4**	**Cell 5 (Exp. 2)**
TS x S	(Aydin et al., 2014, [a])	(Y. Liu et al., 2015, [g][ʊ]; Miller & Ferrer, 2017, [a][ʊ])
Three-way interactions		
NM x SM x TS	**Cell 6**	**Cell 7**
NM x SM x S	**Cell 8**	**Cell 9 (Exp. 2)**
NM x TS x S	(Coulombe et al., 2016, [a])	**Cell 10 (Exp. 3)**
SM x TS x S	**Cell 11**	**Cell 12**

Note. Cells are only numbered for effects that have not been investigated. Cells shaded in light and dark grey indicate effects that have not, respectively, been investigated with linear and nonlinear patterns of change.

[a] Latent growth curve model. [b] Second-order latent growth curve model. [c] Hierarchical Bayesian model. [d] Bivariate latent change score model. [e] Functional mixed-effects model. [f] Nonlinear mixed-effects model. [g] Bilinear spline model. [g] Parallel bilinear spline model.

[○] Manipulated missing data. [ʊ] Assumed complex error structure (heterogeneous variances and/or correlated residuals). [▽] Contained pseudo-time structuredness manipulation.

change. Given that change over time is more likely to follow a nonlinear than a linear pattern (for a review, see Cudeck & Harring, 2007), it could be argued that most simulation research has investigated the effect of longitudinal design factors under unrealistic conditions.

Second, all the cells corresponding to the three-way interactions with nonlinear patterns of change have not been investigated (Cells 7, 9, 10, and 12 in Table 1.1), meaning that almost no study has conducted a comprehensive investigation into measurement timing. Given that longitudinal research is needed to understand the temporal dynamics of psychological processes—as suggested by ergodic theory (Molenaar, 2004)—it is necessary to understand how longitudinal design and analysis factors interact with each other (and with sample size) in affecting the performance of longitudinal models. Given that no simulation study identified in my systematic review conducted a comprehensive investigation into the effects of longitudinal design and analysis factors on modelling nonlinear change, I designed simulation studies to address these gaps.

1.6 Methods of Modelling Nonlinear Patterns of Change Over Time

Because my simulation experiments assumed change over time to be nonlinear, it is important to provide an overview of how nonlinear change can be modelled. On this note, I will provide an overview of two commonly employed methods for modelling nonlinear change: 1) the polynomial approach and 2) the nonlinear function approach.[7,8]

---

[7]It should be noted that nonlinear change can be modelled in a variety of ways, with latent change score models (e.g., O'Rourke et al., 2022) and spline models (e.g., Fine & Grimm, 2020) offering some examples.

[8]The definition of a nonlinear function is mathematical in nature. Specifically, a nonlinear function contains at least one parameter that exists in its corresponding partial derivative (at any order). For example, in the logistic function $\theta + \frac{\alpha - \theta}{1 + exp(\frac{\beta - t}{\gamma})}$ is nonlinear because β exists in $\frac{\partial y}{\partial \beta}$ (in addition to γ existing

Importantly, the simulation experiments in my dissertation will use the nonlinear function approach to model nonlinear change.

Consider an example where an organization introduces a new incentive system with the goal of increasing the motivation of its employees. To assess the effectiveness of the incentive system, employees provide motivation ratings every month over a period of 360 days. Over the 360-day period, the motivation levels of the employees increase following an s-shaped pattern of change over time. One analyst decides to model the observed change using a *polynomial function* shown below in Equation 1.1:

$$y = a + bx + cx^2 + dx^3. \tag{1.1}$$

A second analyst decides to model the observed change using a *logistic function* shown below in Equation 1.2:

$$y = \theta + \frac{\alpha - \theta}{1 + e^{\frac{\beta - time}{\gamma}}} \tag{1.2}$$

Figure 1.3A shows the response pattern predicted by the polynomial function of Equation 1.1 with the estimated values of each parameter (a, b, c, and d) and Figure 1.3B shows the response pattern predicted by the logistic function (Equation 1.2) along with the values estimated for each parameter (θ, α, β, and γ). Although the logistic and polynomial

---

in its corresponding partial derivative). The nth order polynomial function of $y = a + bx + cx^2 + ... + nx^n$ is linear because the partial derivatives with respect to any of the parameters (i.e., $1, x^2, ..., x^n$) never contain the associated parameter.

Figure 1.3

Response Patterns Predicted by Polynomial (Equation 1.1) and Logistic (Equation 1.2) Functions

A: Response Pattern Predicted by Polynomial (Linear) Function

$$y = a + bx + cx^2 + dx^3$$

$a = 3.04$
$b = -0.00244$
$c = 2.58\text{e}{-}05$
$d = -4.77\text{e}{-}08$

B: Response Pattern Predicted by Logistic (Nonlinear) Function

$$y = \theta + \frac{\alpha - \theta}{1 + e^{\left(\frac{\beta - time}{\gamma}\right)}}$$

$\theta = 3.00$
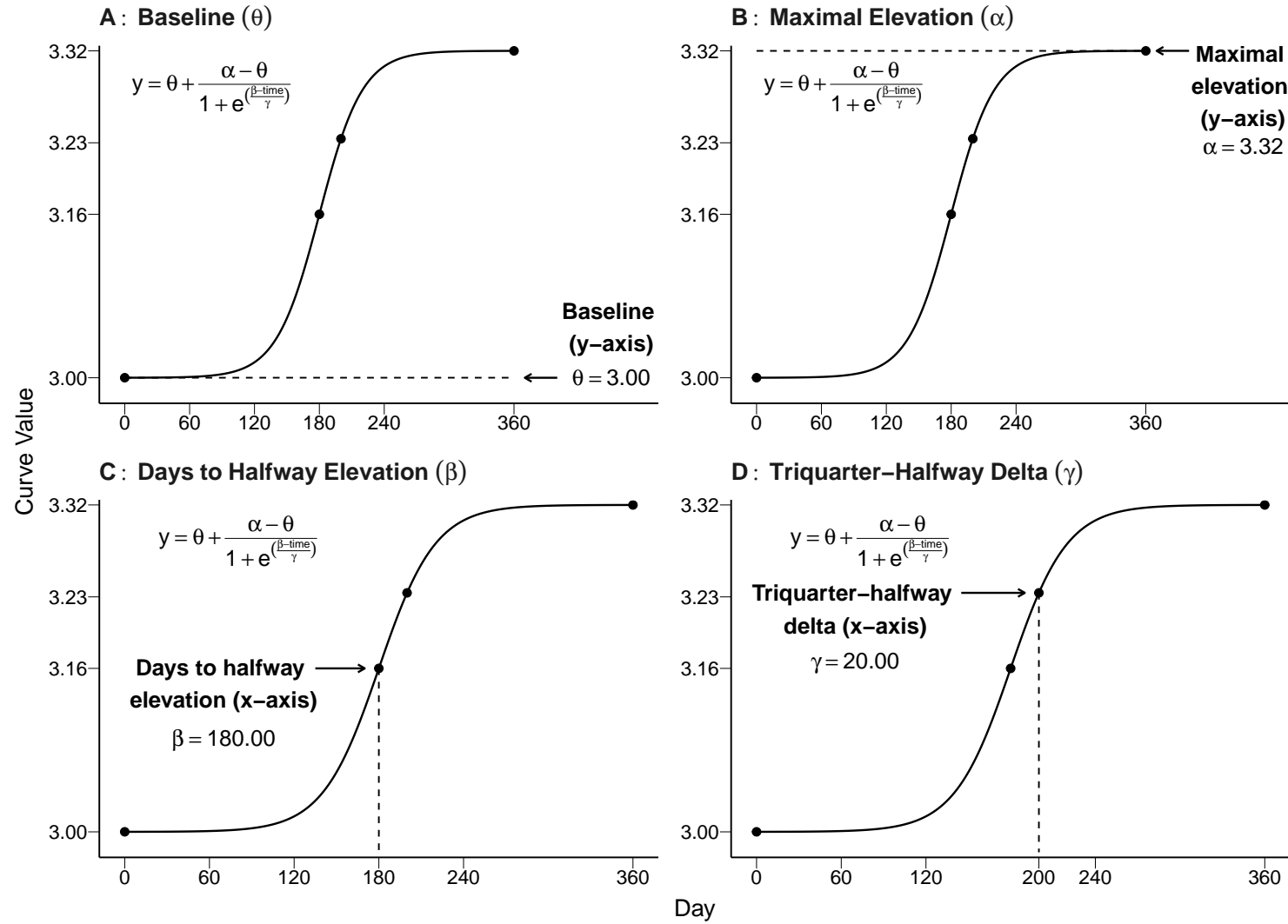$\alpha = 3.32$
$\beta = 180.00$
$\gamma = 20.00$
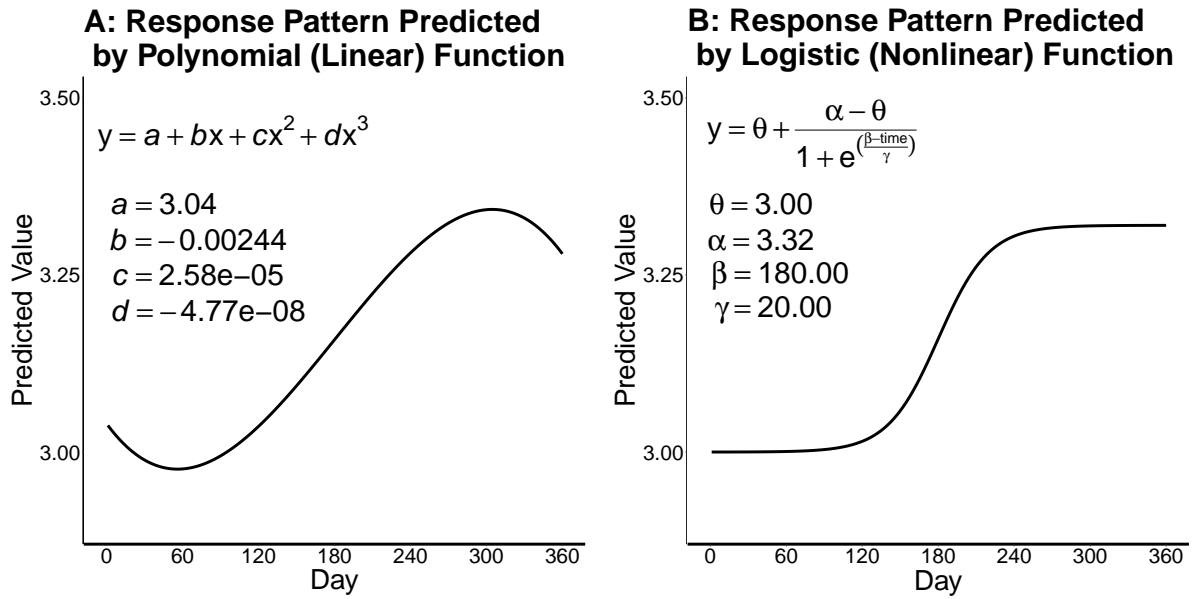
Note. Panel A: Response pattern predicted by the polynomial function of Equation (1.1). Panel B: Response pattern predicted by the logistic function of Equation (1.2).

functions predict nearly identical response patterns, the parameters of the logistic function have the following meaningful interpretations (see Figure 1.4):

- θ specifies the value at the first plateau (i.e., the starting value), and so is called the *baseline* parameter (see Figure 1.4A).

- α specifies the value at the second plateau (i.e., the ending value), and so is called the the *maximal elevation* parameter (see Figure 1.4B).

- β specifies the number of days required to reach half the difference between the first and second plateau (i.e., the midway point), and so is called the *days-to-halfway-elevation* parameter (see Figure 1.4C).

- γ specifies the number of days needed to move from the midway point to approximately 73% of the difference between the starting and ending values (i.e., satiation point), and so is called the *triquarter-halfway delta* parameter (see Figure 1.4D).

Figure 1.4

Description Each Parameters Logistic Function (Equation 1.2) Functions

A : **Baseline** (θ)

$$y = \theta + \frac{\alpha - \theta}{1 + e^{\left(\frac{\beta - time}{\gamma}\right)}}$$

Baseline (y–axis)

$\leftarrow$ $\theta = 3.00$

B : **Maximal Elevation** (α)

$$y = \theta + \frac{\alpha - \theta}{1 + e^{\left(\frac{\beta - time}{\gamma}\right)}}$$

$\leftarrow$ **Maximal elevation (y–axis)**

$\alpha = 3.32$

C : **Days to Halfway Elevation** (β)

$$y = \theta + \frac{\alpha - \theta}{1 + e^{\left(\frac{\beta - time}{\gamma}\right)}}$$

Days to halfway $\longrightarrow$
elevation (x–axis)

$\beta = 180.00$

D : **Triquarter–Halfway Delta** (γ)

$$y = \theta + \frac{\alpha - \theta}{1 + e^{\left(\frac{\beta - time}{\gamma}\right)}}$$

Triquarter–halfway $\longrightarrow$
delta (x–axis)

$\gamma = 20.00$

Curve Value

Day

Note. Panel A: The baseline parameter (θ) sets the starting value of the of curve, which in the current example has a value of 3.00 $(\theta = 3.00)$. Panel B: The

26

maximal elevation parameter (α) sets the ending value of the curve, which in the current example has a value of 3.32 (α = 3.32). Panel C: The days-to-halfway elevation parameter (β) sets the number of days needed to reach 50% of the difference between the baseline and maximal elevation values. In the current example, the baseline-maximal elevation difference is 0.32 ($\alpha - \theta$ = 3.32 - 3.00 = 0.32), and so the days-to-halfway elevation parameter defines the number of days needed to reach a value of 3.16. Given that the days-to-halfway elevation parameter is set to 180 in the current example ($\beta = 180.00$), then 180 days are needed to go from a value of 3.00 to a value of 3.16. Panel D: The triquarter-halfway delta parameter (γ) sets the number of days needed to go from halfway elevation to approximately 73% of the baseline-maximal elevation difference of 0.32 ($\alpha - \theta$ = 3.32 - 3.00 = 0.32). Given that 73% of the baseline-maximal elevation difference is 0.23 and the triquarter-halfway delta is set to 20 days ($\gamma = 20.00$), then 20 days are needed to go from the halfway point of 3.16 to the triquarter point of approximately 3.23).

Applying the parameter meanings of the logistic function to the parameter values estimated by using the logistic function (Equation 1.2), the predicted response pattern begins at a value of 3.00 (baseline) and reaches a value of 3.32 (maximal elevation) by the end of the 360-day period. The midway point of the curve is reached after 180.00 days (days-to-halfway elevation) and the satiation point is reached 20.00 days later (triquarter-halfway delta; or 200.00 days after the beginning of the incentive system is introduced). When looking at the polynomial function, it is almost impossible to meaningfully interpret the values of any of the other parameter values (aside from the 'a' parameter, which indicates the starting value). Therefore, using a nonlinear function such as the logistic function provides a meaningful way to interpret nonlinear change.

1.7 Multilevel and Latent Variable Approach

In addition to using the logistic function to model nonlinear change, another modelling decision concerns whether to do so using the multilevel or latent growth curve framework. In my dissertation, I opted for the latent growth curve framework for two reasons. First, the latent growth curve framework allows data to be more realistically modelled than the multilevel framework. As some examples, the latent growth curve framework allows the modelling of measurement error, complex error structures, and time-varying covariates (for a review, see McNeish & Matta, 2017). Second, and perhaps more important, the likelihood of convergence with multilevel models decreases as the number of random-effect parameters increases due to nonpositive definitive covariance matrices (for a review, see McNeish & Bauer, 2020). With the model I used in my simulation experiments having four random-effect parameters, it is likely that my simulation experiments

would have considerable convergence issues if they use the multilevel framework. Therefore, given the convergence issues of multilevel models and the shortcoming realistically modelling data, I decided, on balance, that the strengths of the multilevel framework (e.g., more options for modelling small samples) were outweighed by its shortcomings, and decided to use a latent growth curve framework in my simulation experiments.

1.7.1 Next Steps

Given that longitudinal research is needed to understand the temporal dynamics of psychological processes, it is necessary to understand how longitudinal design and analysis factors interact with each other (and with sample size) in affecting the performance of nonlinear longitudinal models. With no study to my knowledge having conducted a comprehensive investigation into how longitudinal design and analysis factors affect the modelling of nonlinear change patterns, my simulation experiments are designed to address these gaps in the literature. Specifically, my simulation experiments investigate how measurement number, measurement spacing, and time structuredness affect the performance of a longitudinal model of nonlinear change (see Cells 1, 5, 9, and 10 of Table 1.1/Table 1.2).

1.8 Overview of Simulation Experiments

To investigate the effects of longitudinal design and analysis factors on model performance, I conducted three Monte Carlo experiments. Before summarizing the simulation experiments, one point needs to be mentioned regarding the maximum number of independent variables used in each experiment. No simulation experiment manipulated more than three variables because of the difficulty associated with interpreting interactions

between four or more variables. Even among academics, the ability to correctly interpret interactions sharply declines when the number of independent variables increases from three to four (Halford et al., 2005). Therefore, none of my simulation experiments manipulated more than three variables so that results could be readily interpreted.

To summarize the three simulation experiments, the independent variables of each simulation experiment are listed below:

- Experiment 1: number of measurements, spacing of measurements, and nature of change.

- Experiment 2: number of measurements, spacing of measurements, and sample size.

- Experiment 3: number of measurements, sample size, and time structuredness.

The sections that follow will present each of the simulation experiments and their corresponding results.