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 low, 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

[To be completed after defence]

ACKNOWLEDGEMENTS

[To be completed after defence]

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 (Y. 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;

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 A).

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 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 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 a 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 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 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 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 Modelling Accuracy

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 modelling accuracy 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 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 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 the accuracy of a model (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 modelling accuracy.[5]

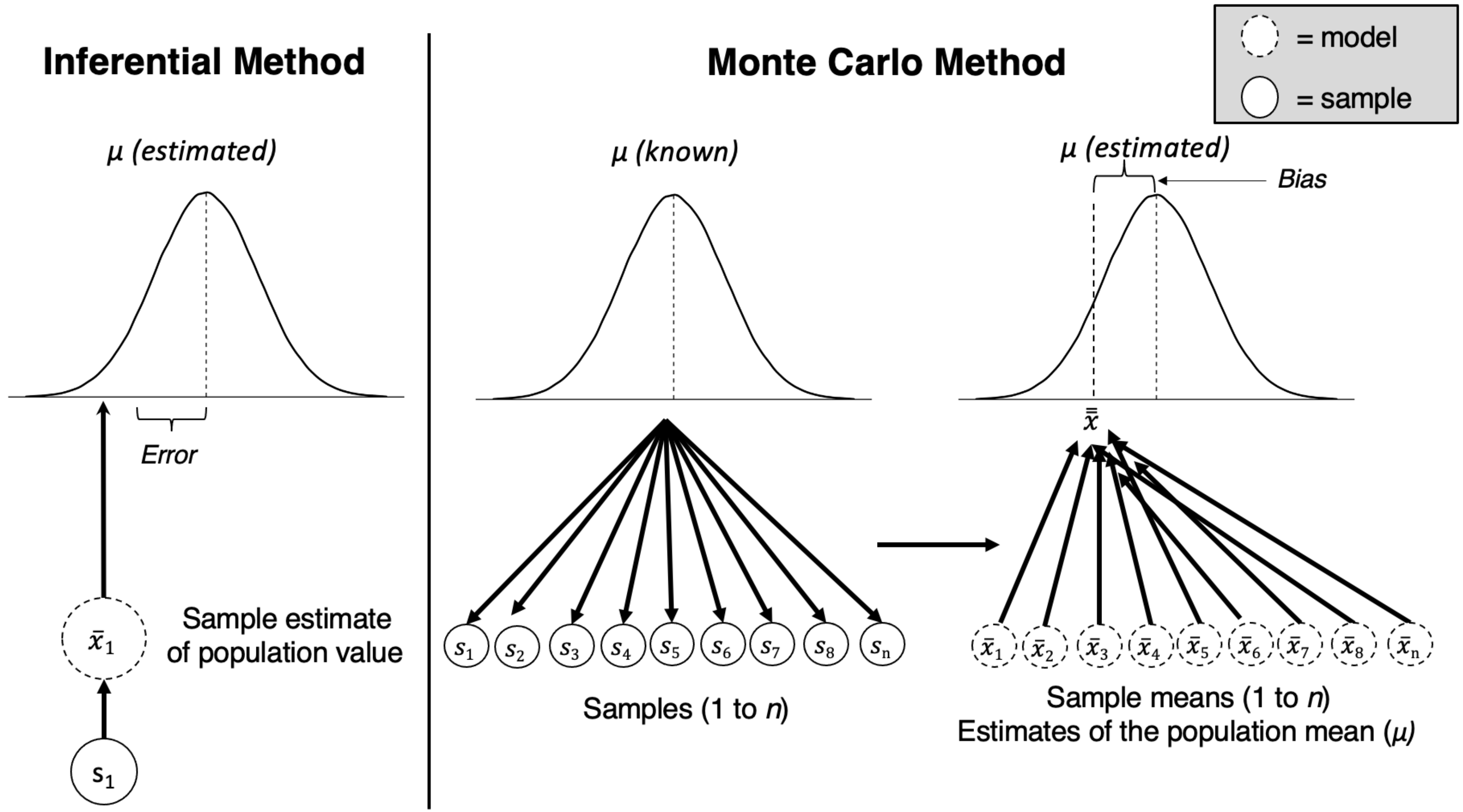
1.5 Systematic Review of Simulation Literature

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 modelling accuracy.

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]

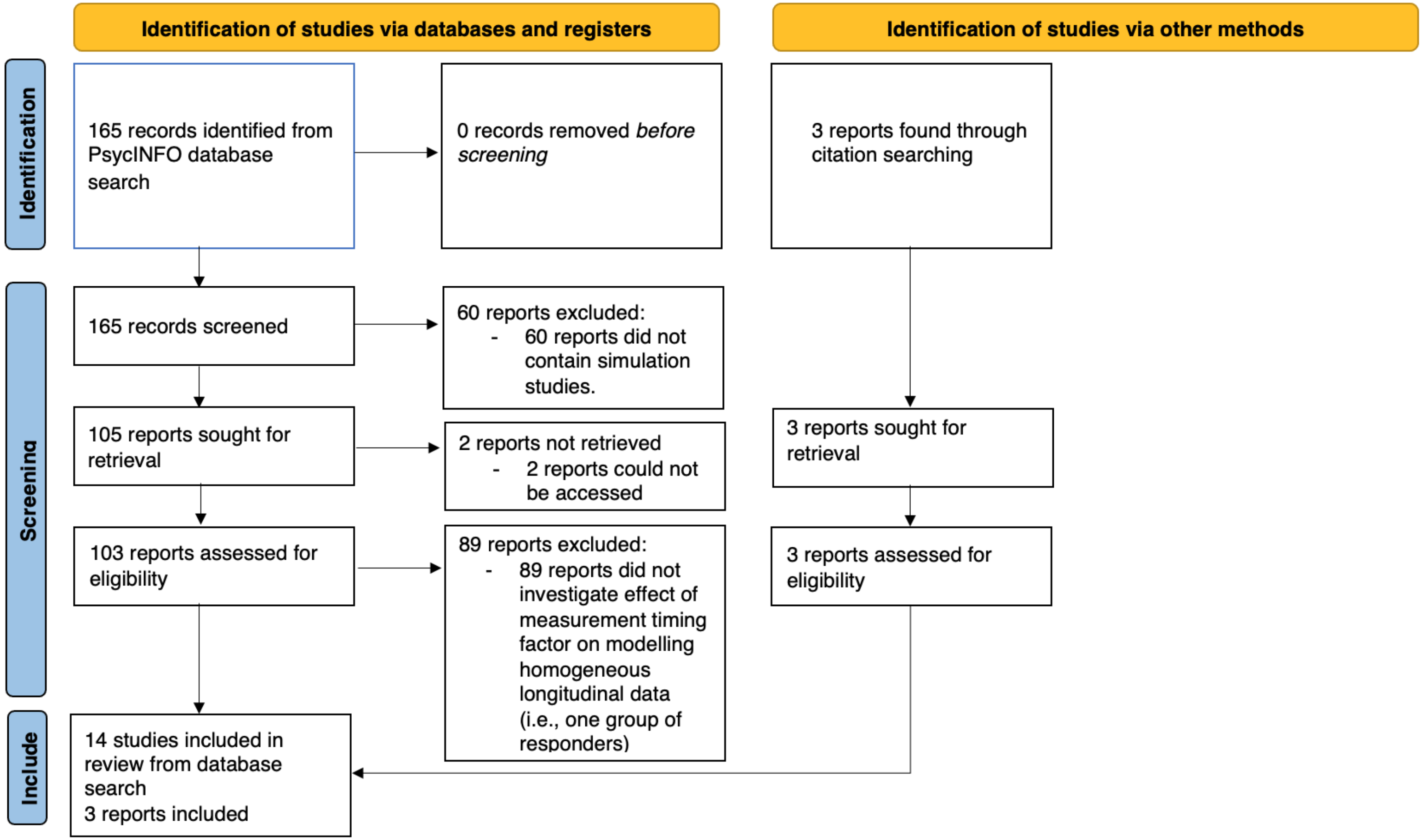
1.5.2 Systematic Review Results

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



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

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℧]; 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])	(Finch, 2017, [a]; Fine et al., 2019, [e∘▿]; Fine & Grimm, 2020, [e,f▿]; J. Liu et al., 2022, [g]; Liu & Perera, 2022, [h℧])
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℧]; Miller & Ferrer, 2017, [a℧])
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. [℧] 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 modelling accuracy of temporal dynamics. 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

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:

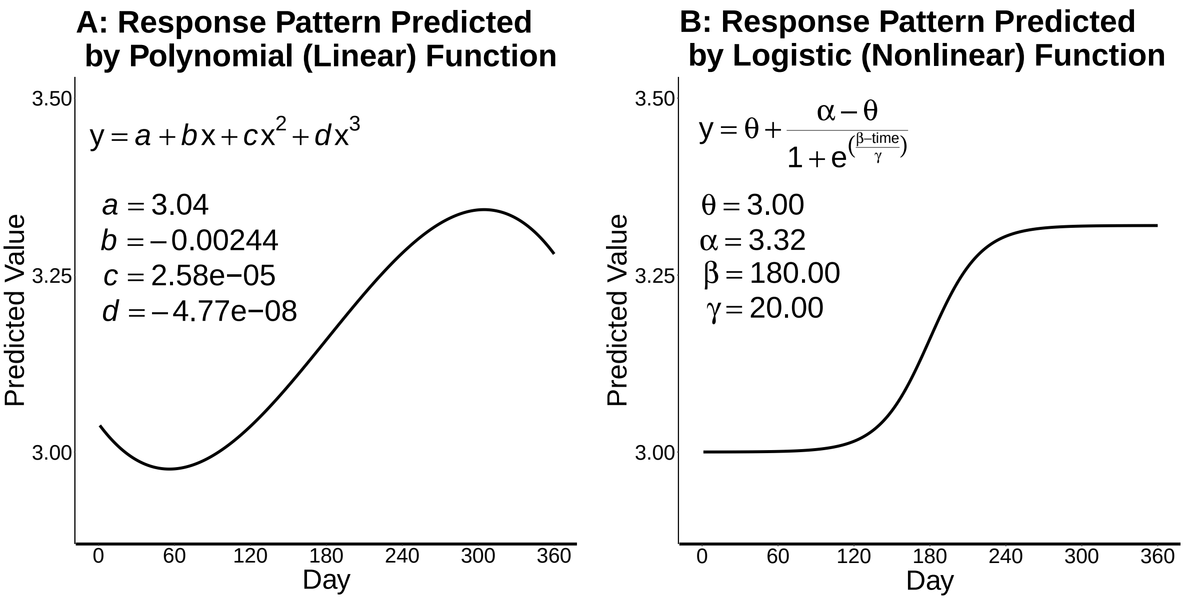
$$y = \theta + \frac{\alpha - \theta}{1 + e^{\frac{\beta - time}{\gamma}}} \tag{1.2}$$

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

existing in its corresponding partial derivative). The n^{th} 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



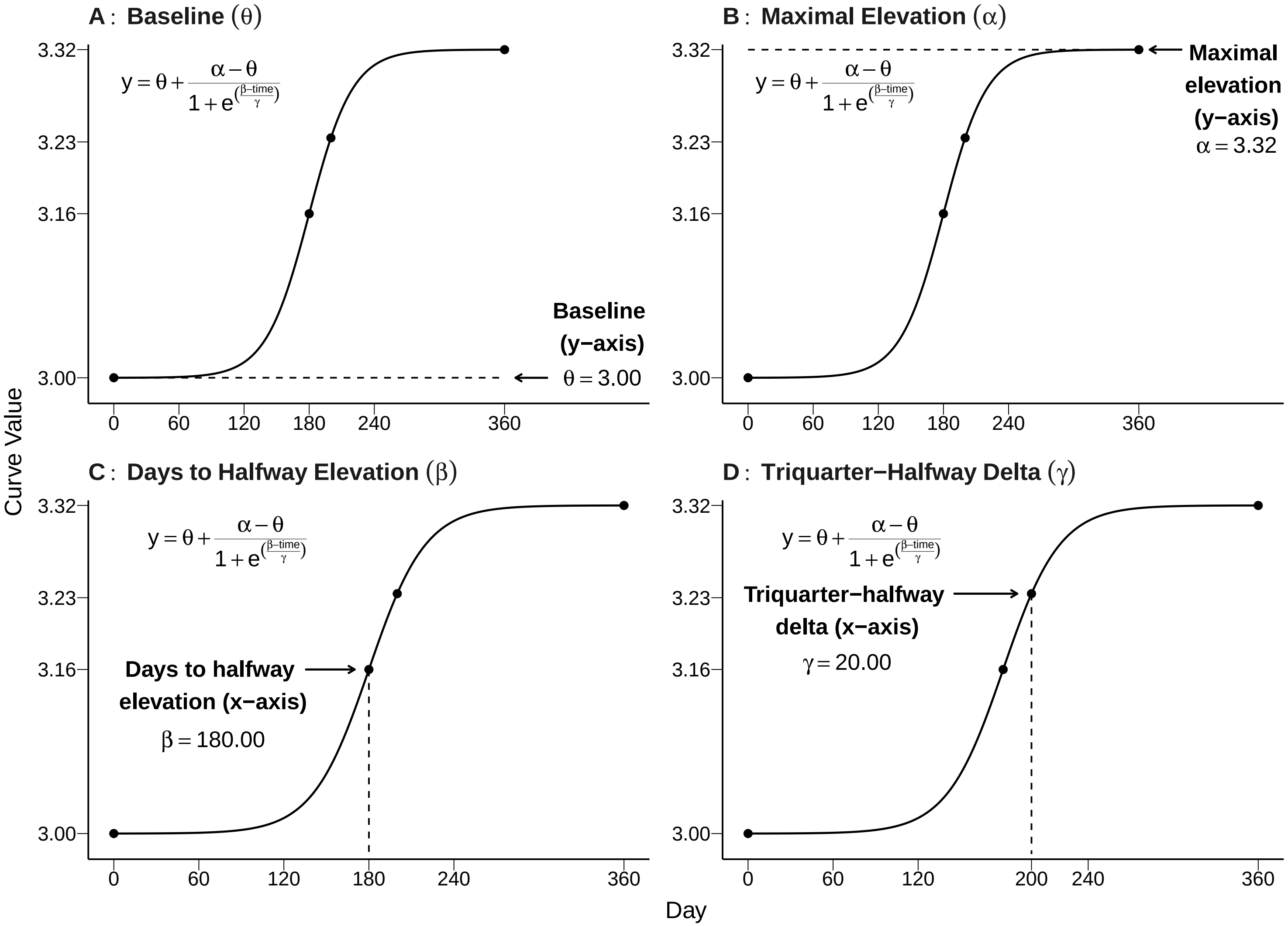
Note. Panel A: Response pattern predicted by the polynomial function of Equation (1.1). Panel B: Response pattern predicted by the logistic function of Equation (1.2).

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



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 (θ = 3.00). Panel B: The 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 accuracy with which nonlinear patterns of change are modelled. 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 accuracy with which a nonlinear change pattern is modelled (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 modelling accuracy, 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.

2 Experiment 1

In Experiment 1, I investigated the number of measurements needed to obtain high model performance for the estimation of each logistic function parameter (i.e., unbiased and precise estimation) under different spacing schedules and natures of change. Before presenting the results of Experiment 1, I present my design and analysis goals. For my design goals, I conducted a 4 (measurement spacing:equal, time-interval increasing, time-interval decreasing, middle-and-extreme) x 4 (number of measurements: 5, 7, 9, 11) x 3 (nature of change: population value for the fixed-effect days-to-halfway elevation parameter [β_{fixed}] of 80, 180, or 280) study. For my analysis goals, I was interested in answering two questions. First, I was interested in whether placing measurements near periods of change increases model performance. To answer my first question, I determined whether model performance under each spacing schedule increased when measurements were taken closer to periods of change.

Second, I was interested in how to space measurements when the nature of change is unknown. When the nature of change is unknown, this translates to a situation where a researcher has little to no knowledge of how change unfolds over time, and so any nature of change is a viable candidate for the true change. Therefore, to determine how to space measurements when the nature of change is unknown, I averaged the model performance of each spacing schedule across all possible nature-of-change curves and considered the spacing schedule with the highest model performance to be the best one.

2.1 Methods

2.1.1 Overview of Data Generation

2.1.1.1 Function Used to Generate Each Data Set

Data for each simulation experiment were generated using R (RStudio Team, 2020). To run the simulations, I created the `nonlinSims` package, which is available at the following GitHub repository: https://github.com/sciarraseb/nonlinSims. The code used to run the simulations and create the data set can be found in Appendix B and the data file (`exp_1_data.csv`) can be found at the following GitHub repository: https://github.com/sciarraseb/dissertation. To generate the data, the *multilevel logistic function* shown below in Equation (2.1) was used:

$$y_{ij} = \theta_j + \frac{\alpha_j - \theta_j}{1 + e^{\frac{\beta_j - time_i}{\gamma_j}}} + \epsilon_{ij}, \tag{2.1}$$

where θ represents the baseline parameter, α represents the maximal elevation parameter, β represents the days-to-halfway elevation parameter, and γ represents triquarter-halfway delta parameter. Note that, values for θ, α, β, and γ were generated for each j person across all i time points, with an error value being randomly generated at each i time point(ϵ_{ij}; see Figure 1.4 for a review of each parameter). In other words, unique response patterns were generated for each person in each generated data set. Importantly, 1000 data sets were generated per cell.

The logistic growth function (Equation 2.1) was used because it is a common pattern of organizational change (or institutionalization; Lawrence et al., 2001). Institutionalization curves follow an s-shaped pattern (i.e., logistic growth), and so their rates of change can be represented by the days-to-halfway elevation and triquarter-halfway delta parameters (β, γ, respectively), and the success of the change can be defined by the magnitude of the difference between baseline and maximal elevation parameters (α - θ, respectively).

2.1.1.2 Population Values Used for Function Parameters

Table 2.1 lists the parameter values that were used for the population parameters. Given that the decisions for setting the values for the baseline, maximal elevation, and residual variance parameters were informed by past research, the discussion that follows highlights how these decisions were made. The difference between the baseline and maximal elevation parameters (θ and α, respectively) corresponded to the effect size most commonly observed in organizational research (i.e., the 50$^{\text{th}}$ percentile effect size value; Bosco et al., 2015). Because the meta-analysis of Bosco et al. (2015) computed effect sizes as correlations, the 50$^{\text{th}}$ percentile effect size value of r = .16 was computed to a standardized effect size using the following conversion function shown in Equation 2.2

(Borenstein et al., 2009, Chapter 7):

$$d = \frac{2r}{\sqrt{1 - r^2}}, \tag{2.2}$$

where r is the correlation effect size. Using Equation 2.2, a correlation value of $r = .16$ becomes a standardized effect size value of $d = 0.32$. For the value of the residual variance parameter (ϵ), Coulombe et al. (2016) set it to the value used for the intercept variance parameter. In the current context, the intercept of the logistic function (Equation 2.1) is the baseline parameter (θ).[9] Given that the value for the variability of the baseline parameter was 0.05 (albeit in standard deviation units), the value used for the residual variance parameter was 0.05 ($\epsilon = 0.05$). Importantly, because Coulombe et al. (2016) set covariances between parameters to zero, all the simulation experiments used zero-value covariances. Because justification for the other parameters could not be found in any of the simulation studies identified in my systematic review, values set for the other parameters were largely arbitrary.

Two last brief points need to be mentioned about how data were generated to facilitate the interpretation of the results. First, data were generated to take on units similar to that of a Likert scale (range of 1–5) by assuming a standard deviation of 1.00. Thus, previously established effect size of $d = 0.32$ standard deviations implies an effect size of 0.32 units. Second, change was assumed to occur over a period of 360 days because many organizational processes are often governed by annual events (e.g.,

[9] The definition of an intercept parameter is the value of a curve when no time has elapsed, and this is precisely the definition of the baseline parameter (θ). Therefore, the variance of the intercept parameter carries the same meaning as the variance of the baseline parameter (θ_{random}).

performance reviews, annual returns, regulations, etc.).

2.1.2 Modelling of Each Generated Data Set

Previously, I described how data were generated. Here, I describe how the generated data were modelled.

Each data set generated by the multilevel logistic function (Equation 2.1) was analysed using a modified latent growth curve model known as a structure latent growth curve model (K. J. Preacher & Hancock, 2015).

Table 2.1
Values Used for Multilevel Logistic Function Parameters

Parameter Means	Value
Baseline, θ	3.00
Maximal elevation, α	3.32
Days-to-halfway elevation, β	180.00
Triquarter-halfway delta, γ	20.00
Variability and Covariability Parameters (in Standard Deviations)	
Baseline standard deviation, ψ_{θ}	0.05
Maximal elevation standard deviation, ψ_{α}	0.05
Days-to-halfway elevation standard deviation, ψ_{β}	10.00
Triquarter-halfway delta standard deviation, ψ_{γ}	4.00
Baseline-maximal elevation covariability, $\psi_{\theta\alpha}$	0.00
Baseline-days-to-halfway elevation covariability, $\psi_{\theta\beta}$	0.00
Baseline-triquarter-halfway delta covariability, $\psi_{\theta\gamma}$	0.00
Maximal elevation-days-to-halfway elevation covariability, $\psi_{\alpha\beta}$	0.00
Maximal elevation-triquarter-halfway delta covariability, $\psi_{\alpha\gamma}$	0.00
Days-to-halfway elevation-triquarter-halfway delta covariability, $\psi_{\beta\gamma}$	0.00
Residual standard deviation, ψ_{ϵ}	0.05

Note. The difference between α and θ corresponds to the 50$^{\text{th}}$ percentile Cohen's d value of 0.32 in organizational psychology (Bosco et al., 2015).

Importantly, the model fit to each generated data set estimated nine parameters: A fixed-effect parameter for each of the four logistic function parameters, a random-effect parameter for each of the four logistic function parameters, and an error parameter. As with a multilevel model, a fixed-effect parameter has a constant value across all individuals, whereas a random-effect parameter represents the variability of values across all modelled people.[10] To fit the logistic function to a given data set (Equation 2.1), a linear approximation of the logistic function was needed so that it could fit within the linear nature of structural equation modelling framework.[11] To construct a linear approximation of the logistic function, a first-order Taylor series was constructed for the logistic function. For a detailed explanation of how the logistic function was fit into the structural equation modelling framework, see Appendix D for an explanation of the model and Appendix E for the code used to create the model.

2.1.3 Variables Used in Simulation Experiment

2.1.3.1 Independent Variables

To build on current research, Experiment 1 used independent variable manipulations from a select number of previous studies. In looking at the summary of the simulation literature in Table 1.2, the study by Coulombe et al. (2016) was the only one to investigate three longitudinal issues of interest to my dissertation, and so represented the most comprehensive investigation. Because I was also interested in investigating measurement

[10] Estimating a random-effect for a parameter allows person- or data-point-specific values to be computed for the parameter.

[11] The logistic function (Equation 2.1) is a nonlinear function and so cannot be directly inserted into the structural equation modelling framework because this framework only allows linear computations of matrix-matrix, matrix-vector, and vector-vector operations. Unfortunately, the algebraic operations permitted in a linear framework cannot directly reproduce the operations in the logistic function (Equation 2.1) and so a linear approximation of the logistic function must be constructed so that the logistic function can be inserted into the structural equation modelling framework.

spacing, manipulations were inspired from the only other simulation study identified by my systematic review to manipulate measurement spacing (the study by Timmons & Preacher, 2015). The sections that follow will discuss each of the variables manipulated in Experiment 1.

2.1.3.1.1 Spacing of Measurements

The only simulation study identified by my systematic review that manipulated measurement spacing was Timmons and Preacher (2015). Measurement spacing in Timmons and Preacher (2015) was manipulated in the following four ways:

1) *Equal spacing*: measurements are divided by intervals of equivalent lengths.
2) *Time-interval increasing spacing*: intervals that divide measurements increase in length over time.
3) *Time-interval decreasing spacing*: intervals that divide measurements decrease in length over time.
4) *Middle-and-extreme spacing*: measurements are clustered near the beginning, middle, and end of the data collection period.

To maintain consistency with the established literature, I manipulated measurement spacing in the same way as Timmons and Preacher (2015) presented above. Importantly, because Timmons and Preacher (2015) did not create their measurement spacing schedules with any systematicity, I developed a novel and replicable procedure for generating measurement schedules for each of the four measurement spacing conditions, which is described in Appendix C. I also automated the generation of measurement schedules by creating a set of functions in R (RStudio Team, 2020).

Table 2.2 lists the measurement days that were used for all measurement spacing-measurement number cells. The first column lists the type of measurement spacing (i.e., equal, time-interval increasing, time-interval decreasing, or middle-and-extreme); the second column lists the number of measurements (5, 7, 9, or 11); the third column lists the measurement days that correspond to each measurement number-measurement spacing condition; and the fourth column lists the interval lengths between the measurements. Note that the interval lengths are equal for equal spacing, increase over time for time-interval increasing spacing, and decrease over time for time-interval decreasing spacing. For cells with middle-and-extreme spacing, the measurement days and interval lengths in the middle of the measurement window have been emboldened.

2.1.3.1.2 Number of Measurements

The smallest measurement number value in Coulombe et al. (2016) of three measurements could not be used in Experiment 1 (or any other simulation experiment that manipulated measurement number in my dissertation) because doing so would have created non-identified models The model used in my simulations estimated 9 parameters (p = 9; 4 fixed-effects + 4 random-effects + 1 error)[12] and so the minimum number of measurements (or observed variables) required for model identification (and to allow model comparison) was 4. Although a measurement number of three could not be used in my manipulation of measurement number, the next highest measurement number values in Coulombe et al. (2016) of 5, 7, and 9 were used. Importantly, a larger value of 11 was

[12]Degrees of freedom is calculated by multiplying the number of observed variables (p) by p + 1 and dividing it by 2 ($\frac{p[p+1]}{2}$; Loehlin & Beaujean, 2017).

Table 2.2

Measurement Days Used for All Measurement Number-Measurement Spacing Conditions

Spacing Schedule	Number of Measurements	Measurement Days	Interval Lengths
Equal	5	0, 90, 180, 270, 360	90, 90, 90, 90
	7	0, 60, 120, 180, 240, 300, 360	60, 60, 60, 60, 60, 60
	9	0, 45, 90, 135, 180, 225, 270, 315, 360	45, 45, 45, 45, 45, 45, 45, 45
	11	0, 36, 72, 108, 144, 180, 216, 252, 288, 324, 360	36, 36, 36, 36, 36, 36, 36, 36, 36, 36
Time-interval increasing	5	0, 30, 100, 210, 360	30, 70, 110, 150
	7	0, 30, 72, 126, 192, 270, 360	30, 42, 54, 66, 78, 90
	9	0, 30, 64.29, 102.86, 145.71, 192.86, 244.29, 300, 360	30, 34.29, 38.57, 42.86, 47.14, 51.43, 55.71, 60
	11	0, 30, 61.33, 94, 128, 163.33, 200, 238, 277.33, 318, 360	30, 31.33, 32.67, 34, 35.33, 36.67, 38, 39.33, 40.67, 42
Time-interval decreasing	5	0, 150, 260, 330, 360	150, 110, 70, 30
	7	0, 90, 168, 234, 288, 330, 360	90, 78, 66, 54, 42, 30
	9	0, 60, 115.71, 167.14, 214.29, 257.14, 295.71, 330, 360	60, 55.71, 51.43, 47.14, 42.86, 38.57, 34.29, 30
	11	0, 42, 82.67, 122, 160, 196.67, 232, 266, 298.67, 330, 360	42, 40.67, 39.33, 38, 36.67, 35.33, 34, 32.67, 31.33, 30
Middle-and-extreme	5	1, **150, 180, 210**, 360	150, **30, 30**, 150

Table 2.2

Measurement Days Used for All Measurement Number-Measurement Spacing Conditions (continued)

Spacing Schedule	Number of Measurements	Measurement Days	Interval Lengths
	7	1, 30, **150, 180, 210**, 330, 360	30, 120, **30, 30**, 120, 30
	9	1, 30, 60, **150, 180, 210**, 300, 330, 360	30, 30, 90, **30, 30**, 90, 30, 30
	11	1, 30, 60, **120, 150, 180, 210, 240,** 300, 330, 360	30, 30, 60, **30, 30, 30, 30**, 60, 30, 30

Note. For middle-and-extreme spacing levels, the measurement days and and interval lengths corresponding to the middle of measurement windows have been emboldened.

added to test for a possible effect of a high measurement number. Therefore, my simulation experiments used the following values in manipulating the number of measurements: 5, 7, 9, and 11.

2.1.3.1.3 Population Values Set for The Fixed-Effect Days-to-Halfway Elevation Parameter β_{fixed} (Nature of Change)

The nature of change was manipulated by setting the days-to-halfway elevation parameter (β_{fixed}) to a value of either 80, 180, or 280 days (see Figure 1.4A). Note that no other study in my systematic review manipulated nature of change using logistic curves and so its manipulation in Experiment 1 is, to the best of my knowledge, unique. Importantly, nature of change was manipulated to simulate situations where uncertainty exists in how change unfolds over time.

2.1.3.2 Constants

Given that each simulation experiment manipulated no more than three independent variables so that results could be readily interpreted (Halford et al., 2005), other variables had to be set to constant values. In Experiment 1, two important variables were set to constant values: sample size and time structuredness. For sample size, I set the value across all cells to the average sample size used in organizational research ($n = 225$; Bosco et al., 2015). For time structuredness, data across all cells were generated to be time structured (i.e., all participants provide data according to one response pattern; that is, at each time point, participants provide their data at the exact same moment).

2.1.3.3 Dependent Variables

2.1.3.3.1 Convergence Success Rate

The proportion of iterations in a cell where models converged defined the *convergence success rate*.[13] Equation (4.5) below shows the calculation used to compute the convergence success rate:

$$\text{Convergence success rate} = \frac{\text{Number of models that successfully converged in a cell}}{n}, \tag{2.3}$$

where n represents the total number of models run in a cell.

2.1.3.3.2 Model Performance

Model performance was the combination of two metrics: bias and precision. More specifically, two questions were of importance in the estimation of a given logistic function parameter: 1) How well was the parameter estimated on average (bias) and 2) what was a range of values that could be expected for an estimate from the output of a single model (precision). In the two sections that follow, I will discuss each metric of model performance and the cutoffs used to determine whether estimation was unbiased and precise.

2.1.3.3.2.1 Bias

Bias was calculated to evaluate the accuracy with which each logistic function parameter was estimated in each experimental cell. As shown below in Equation (2.4), *bias*

[13] Specifically, convergence was obtained if the convergence code returned by OpenMx was 0.

was obtained by summing the differences between the population value set for a parameter and the value estimated for the parameter by each i converged model and then dividing the sum by the number of N converged models.

$$\text{Bias} = \frac{\sum_i^N (\text{Population value for parameter} - \text{Average estimated value}_i)}{N} \tag{2.4}$$

Bias was calculated for the fixed- and random-effect parameters of the baseline (θ_{fixed}, θ_{random}), maximal elevation (α_{fixed}, α_{random}), days-to-halfway elevation (β_{fixed}, β_{random}), and the triquarter-halfway delta parameters (γ_{fixed}, γ_{random}) and the error parameter (ϵ).

2.1.3.3.2.2 Precision

In addition to computing bias, precision was calculated to evaluate the variability with which each parameter was estimated. Importantly, metrics used to evaluate precision in previous studies aoften ssume estimates are normally distributed (e.g., mean-squared error and empirical standard error). Because some parameters in my simulations had skewed distributions, using a metric that assumed a normal distribution would likely yield inaccurate results. Correspondingly, I used a distribution-independent definition of precision. In my simulations, *precision* was defined as the range of values covered by the middle 95% of values estimated for a logistic parameter.

2.1.4 Analysis of Data Modelling Output and Accompanying Visualizations

To analyse and visualize modelling performance, I calculated values for convergence success rate, bias, and precision in each experimental cell (see dependent variables). The sections that follow provide details on how I analysed each dependent variable and constructed plots to visualize bias and precision.

2.1.4.1 Analysis of Convergence Success Rate

For the analysis of convergence success rate, the mean convergence success rate was computed for each cell in each experiment (see section on convergence success rate). Because convergence rates exhibited little variability across cells due to the nearly unanimous high rates (almost all cells across all experiments had convergence success rates above 90%), examining the effects of any independent variable on these values would have provided little information. Therefore, I only reported the average convergence success rate for each cell (see Appendix G).

2.1.4.2 Analysis and Visualization of Bias

In accordance with several simulation studies, an estimate with a bias value within a ±10% margin of error of the parameter's population value was deemed unbiased (Muthén et al., 1997). To visualize bias, I constructed bias/precision plots. Figure 2.1 shows a bias/precision plot for the fixed-effect triquarter-halfway parameter (γ_{fixed}) for each measurement number and nature of change. The dots (squares, circles, triangles, diamonds) indicate the average estimated value (see bias). The horizontal blue line indicates the population value ($\gamma_{fixed} = 4.00$) and the gray band indicates the acceptable margin of error of ±10% of the parameter's population value. Dots that lie within the gray margin of error are filled and dots that lie outside of the margin remain unfilled. In the current example, the average value estimated for the fixed-effect triquarter-halfway delta parameter (γ_{fixed}) is only biased (i.e., lies outside the margin of error) with five measurements with a nature-of-change value of 80 ($\beta_{fixed} = 80$). Therefore, estimates are unbiased in almost all cells.

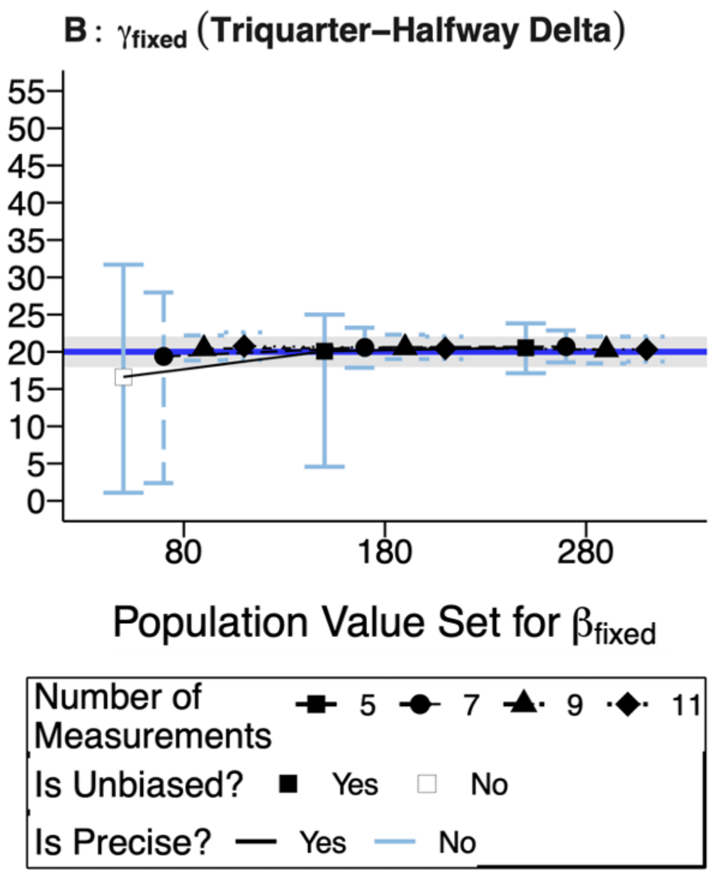
2.1.4.3 Analysis and Visualization of Precision

As discussed previously, precision was defined as the range of values covered by the middle 95% of estimated values for a given parameter (see precision). The cutoff value used to estimate precision directly followed from the cutoff value used for bias. Given that bias values within a ±10% of a parameter's population value were deemed acceptable, an acceptable value for precision should not allow any bias value above the ±10% cutoff. That is, the range covered by the middle 95% of estimated values should not contain a bias value outside the ±10% cutoff. If the range of values covered by the middle 95% of estimate values is conceptualized as an error bar centered on the population value, then an acceptable value for precision implies that neither the lower nor upper whiskers have a length greater than 10% of the parameter's population value. In summary, I deemed precision acceptable if no estimate within the range of values covered by the middle 95% of estimated values had a bias value greater than 10% of the population value, which also means that neither the lower nor upper whiskers of the error bar have a length greater than 10% of the population value.

Like bias, I also depicted precision in bias/precision plots using error bars. Each error bar in the bias/precision plot of Figure 2.1 indicates the range of values covered by the middle 95% of estimated values in the given cell for the fixed-effect triquarter-halfway delta parameter (γ_{fixed}). Importantly, if estimation is not precise, then at least one of the lower and/or upper whisker lengths exceeds 10% of the parameter's population value. When estimation is not precise, the error bar is light blue. When estimation is precise (i.e., neither of the lower or upper whisker lengths exceed 10% of the parameter's population value), the corresponding error bar is black. In the current example, all error bars are light blue and so precision is low in all cells.

Figure 2.1

Bias/Precision Plot for the Fixed-Effect Days-to-Halfway Elevation Parameter (γ_{fixed})



Note. Dots (squares, circles, triangles, diamonds) indicate the average estimated value and error bars show the range of values covered by the middle 95% of the estimated values (see Precision). The horizontal blue line indicates the population value (γ_{fixed} = 4.00) and the gray band indicates the acceptable margin of error (i.e., $\pm$10% of the population value) for bias. Dots that lie outside of the margin of error are unfilled and are considered biased estimates. Dots that lie inside the margin of error are filled and considered unbiased estimates. Error bars whose upper and/or lower whisker lengths exceed 10% of the parameter's population value are light blue and indicate parameter estimation that is not precise. Error bars whose upper and/or lower whisker lengths do not excced 10% of the parameter's population value are black and indicate parameter estimation that is precise.

2.1.4.3.1 Effect Size Computation for Precision

One last statistic I calculated was an effect size value to estimate the variance in parameter estimates accounted for by each effect. Among the several effect size metrics—at a broad level, effect size metrics can represent standardized differences or variance-accounted-for measures that are corrected or uncorrected for sampling error—the corrected variance-accounted-for effect size metric of partial ω^2 was chosen because of three desirable properties. First, partial ω^2 provides a less biased estimate of effect size than other variance-accounted-for measures (Okada, 2013). Second, partial ω^2 is more robust to assumption violations of normality and homogeneity of variance (Yigit & Mendes, 2018). Given that parameter estimates were often non-normally distributed across cells, effect size values computed with partial ω^2 should be relatively less biased than other variance-accounted-for effect size metrics (e.g., η^2). Third, partial ω^2 provides an effect size estimate that is not diluted by the inclusion of unaccountable variance in the denominator. To compute partial ω^2 value for each experimental effect, Equation 2.5 shown below was used:

$$\text{partial}\omega^2 = \frac{\sigma^2_{effect}}{\sigma^2_{effect} + MSE} \tag{2.5}$$

where σ^2_{effect} represents the variance accounted by an effect and MSE is the mean squared error. Importantly, σ^2_{effect} values were corrected values obtained by using the following formula in Equation 2.6 for a two-way factorial design with fixed variables (Howell, 2009):

$$\sigma^2_{effect} = \frac{(a-1)(MS_{effect} - MS_{error})}{nab}, \tag{2.6}$$

where a is the number of levels in the effect, b is the number of levels in the second effect, and n is the cell size. The variance accounted by the interaction was computed using the following formula in Equation 2.7:

$$\sigma^2_{AxB} = \frac{(a-1)(b-1)(MS_{AxB} - MS_{error})}{nab}. \tag{2.7}$$

To compute partial ω^2 values for effects, a Brown-Forsythe test was computed and the appropriate sum-of-squares terms were used to compute partial ω^2 values. A Brown-Forsythe test was used to protect against the biasing effects of skewed distributions (Brown & Forsythe, 1974), which were observed in the parameter estimate distributions in the current simulation experiments. To compute the Brown-Forsythe test, median absolute deviations in each cell were computed by calculating the absolute difference between each i estimate and the median estimated value in the given experimental cell as shown in Equation 2.8 below:

$$\text{Median absolute deviation}_i = |\text{Parameter estimate}_i - \text{Median parameter estimate}_{cell}|. \tag{2.8}$$

An ANOVA was then computed on the median absolute deviation values (using the independent variables of the experiment and the associated interactions as predictors), with the terms in Equation 2.5 extracted from the ANOVA output to compute partial ω^2 values.

2.2 Results and Discussion

In the sections that follow, I organize the results by presenting them for each spacing schedule (equal, time-interval increasing, time-interval decreasing, middle-and-extreme). The results are presented for each spacing schedule because answering my research questions first requires knowledge of these results. To answer my first question of whether model performance increases from placing measurements during periods of change, I need to determine whether model performance under each spacing schedule increases when measurements are placed near periods of change. To answer my second question of how to space measurements when the nature of change is unknown, model performance across all manipulated nature-of-change values must first be calculated for each spacing schedule. The spacing schedule that obtains the highest model performance across all nature-of-change values can then be determined as the best schedule to use.

For each spacing schedule, I will first present a concise summary table of the results and then provide a detailed report for each column of the summary table. Because the detailed reports are of considerable length, I provide concise summaries before the detailed reports to establish a framework to help interpret the detailed reports. The detailed report of each spacing schedule presents the results of each day-unit's bias/precision plot, model performance under each nature-of-change value, and then provides a qualitative summary. After providing the results for each spacing schedule, I then use the results to answer my research questions.

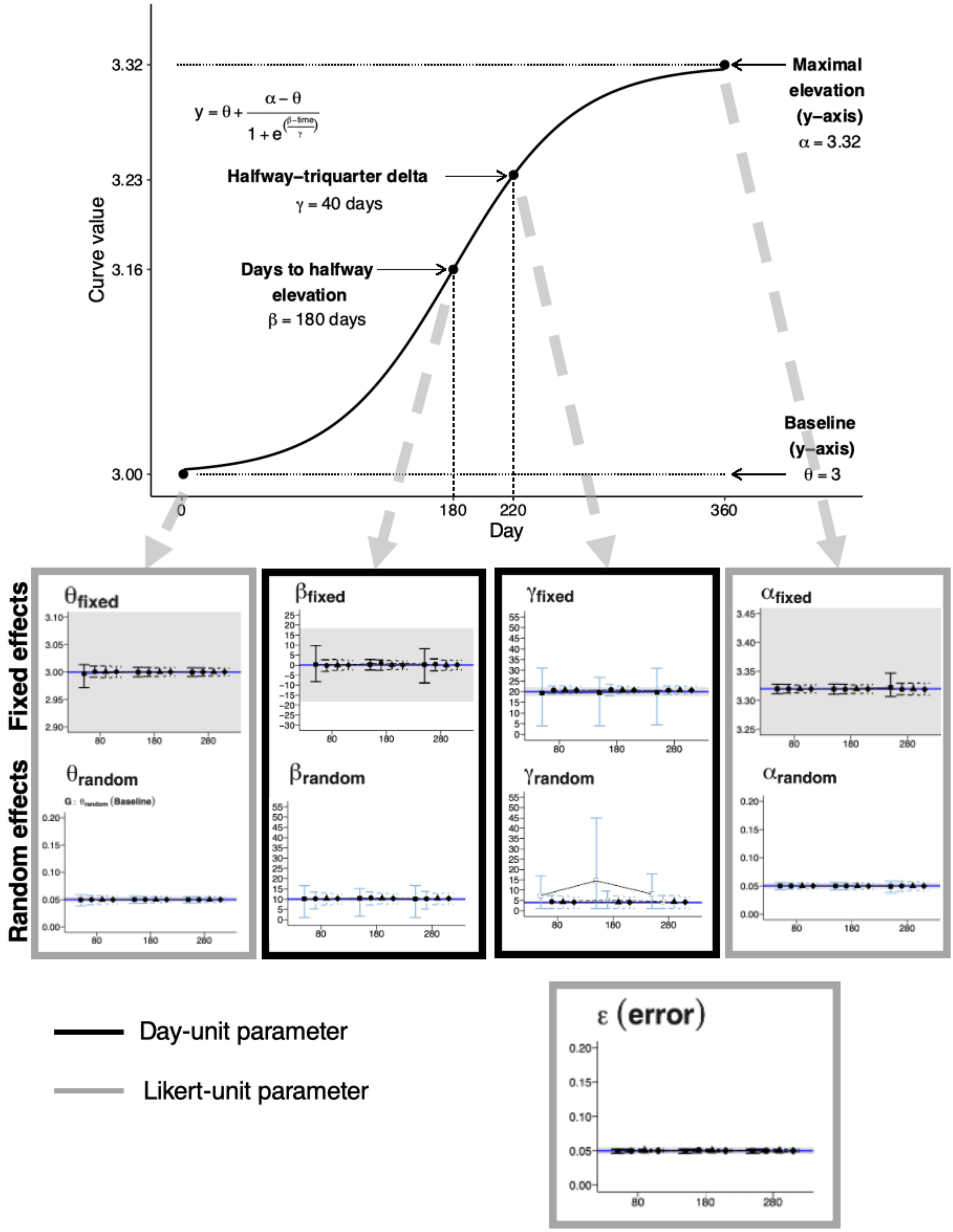
2.2.1 Framework for Interpreting Results

To conduct Experiment 1, the three variables of number of measurements (4 levels), measurement spacing (4 levels), and nature of change (3 levels) were manipulated, which yielded a total of 48 cells. Importantly, within each cell, bias and precision values were also computed for each of the nine parameters estimated by the structured latent growth curve model (for a review, see modelling of each generated data set). Thus, because the analysis of Experiment 1 computed values for many dependent variables, interpreting the results can become overwhelming. Therefore, I will provide a framework to help the reader efficiently navigate the results section.

Because I will present the results of Experiment 1 by each level of measurement spacing, the framework I will describe in Figure 2.2 shows a template for the bias/precision plots that I will present for each spacing schedule. The results of each spacing schedule contain a bias/precision plot for each of the nine estimated parameters. Each bias/precision plot shows the bias and precision for the estimation of one parameter across all measurement number and nature-of change levels. Within each bias/precision plot, dots indicate the average estimated value (which indicates bias bias) and error bars represent the middle 95% range of estimated values (which indicates precision). Bias/precision plots with black outlines show the results for day-unit parameters and plots with gray outlines show the results for Likert-unit parameters. Importantly, only the results for the day-unit parameters will be presented (i.e., fixed- and random-effect days-to-halfway elevation and triquarter-halfway delta parameters [β_{fixed}, β_{random}, γ_{fixed}, γ_{random}, respectively]). The results for the Likert-unit parameters (i.e., fixed- and random-effect baseline and maximal elevation parameters [θ_{fixed}, θ_{random}, α_{fixed}, α_{random}, respectively]) were largely trivial and so are presented in Appendix F. Therefore, the results of each spacing schedule will only present the bias/precision plots for four parameters (i.e., the day-unit parameters).

Figure 2.2

Set of Bias/Precision Plots Constructed for Each Spacing Schedule in Experiment 1



Note. A bias/precision plot is constructed for each parameter of the logistic function (see Equation 2.1). Bias/precision plots with black borders show the results for day-unit parameters and plots with gray border show the results for Likert-unit parameters. For each parameter, bias and precision are shown across each combination of measurement number and nature of change.

2.2.2 Pre-Processing of Data and Model Convergence

After collecting the output from the simulations, non-converged models (and their corresponding parameter estimates) were removed from subsequent analyses. Table G.1 in Appendix G provides the convergence success rates for each cell in Experiment 1. Model convergence was almost always above 90% and convergence rates, with rates only going below 90% in two cells (or instances) with five measurements.

2.2.3 Equal Spacing

For equal spacing, Table 2.3 provides a concise summary of the results for the day-unit parameters (see Figure 2.5 for the corresponding bias/precision plots). The sections that follow will present the results for each column of Table 2.3 and provide elaboration when necessary.

Before presenting the results for equal spacing, I provide a brief description of the concise summary table created for each spacing schedule and shown for equal spacing below in Table 2.3. Text within the 'Highest Model Performance' column indicates the nature-of-change value that resulted in the highest model performance for each day-unit parameter. Text within the 'Unbiased' and 'Precise' columns indicates the number of measurements that were needed to, respectively, obtain unbiased and precise parameter estimation across all manipulated nature-of-change values. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the measurement number that, respectively, resulted in unbiased estimation and the greatest improvements in bias and precision across all day-unit parameters and manipulated nature-of-change values. The 'Error Bar Length' column indicates the average error bar length across all manipulated nature-of-change values that resulted from using the measurement number listed in the

Table 2.3

Concise Summary of Results for Equal Spacing in Experiment 1

				Summary	
Parameter	Highest Model Performance	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 2.5A)	β_{fixed} = 180	All cells	All cells	Largest improvements in precision with **NM = 7**	5.64
γ_{fixed} (Figure 2.5B)	β_{fixed} = 180	All cells	No cells	Largest improvements in precision with **NM = 7**	4.37
β_{random} (Figure 2.5C)	β_{fixed} = 180	All cells	No cells	Largest improvements in precision with **NM = 7**	7.74
γ_{random} (Figure 2.5D)	β_{fixed} = 180	**NM $\geq$ 9**	No cells	Largest improvements in bias and precision with **NM = 7**	7.02

Note. 'Highest Model Performance' indicates the curve that resulted in the highest model performance (largely determined by precision; see nature of change). Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the number of measurements that, respectively, resulted in unbiased estimation and the greatest improvements in bias and precision across all day-unit parameters (note that acceptable precision was not obtained in the estimation of all day-unit parameters with equal spacing). 'Error Bar Length' indicates the average error bar length value across all nature-of-change values that resulted from using the measurement number in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect triquarter-halfway delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect triquarter-halfway delta parameter = 4. NM = number of measurements.

‘Qualitative Description’ column.

2.2.3.1 Nature of Change That Leads to Highest Model Performance

For equal spacing, Table 2.4 lists the precision values (i.e., error bar lengths) for each day-unit parameter across each nature-of-change value. The ‘Total’ column indicates the total error bar length, which is a sum of the lower (‘Lower’) and upper (‘Upper’) whisker lengths. Given that the lower and upper whisker lengths were largely equivalent for each

Table 2.4

Error Bar Lengths Across Nature-of-Change Values Under Equal Spacing in Experiment 1

	Population Value of β_{fixed}								
	80			180			280		
Parameter	Lower	Upper	Total	Lower	Upper	Total	Lower	Upper	Total
β_{fixed} (Figure 2.5A)	4.42	4.12	8.54	2.46	2.32	4.78	4.09	4.16	8.25
γ_{fixed} (Figure 2.5B)	4.84	4.69	9.53	4.95	3.7	8.65	4.79	4.65	9.44
β_{random} (Figure 2.5C)	4.74	3.88	8.62	3.96	3.55	7.51	4.77	4.05	8.82
γ_{random} (Figure 2.5D)	3.00	5.52	8.52	3.00	13.05[a]	16.05	3.00	5.78	8.78

Note. ‘Total’ indicates the total error bar length, which is a sum of the lower (‘Lower’) and upper (‘Upper’) whisker lengths. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect triquarter-halfway delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect triquarter-halfway delta parameter = 4. Note that error bar lengths were calculated by computing the average error bar length value across all number-of-measurement (NM) values (NM $\in \{5, 7, 9, 11\}$). Columns shaded in gray indicate the nature-of-change value that results in the shortest error bar and whisker lengths.

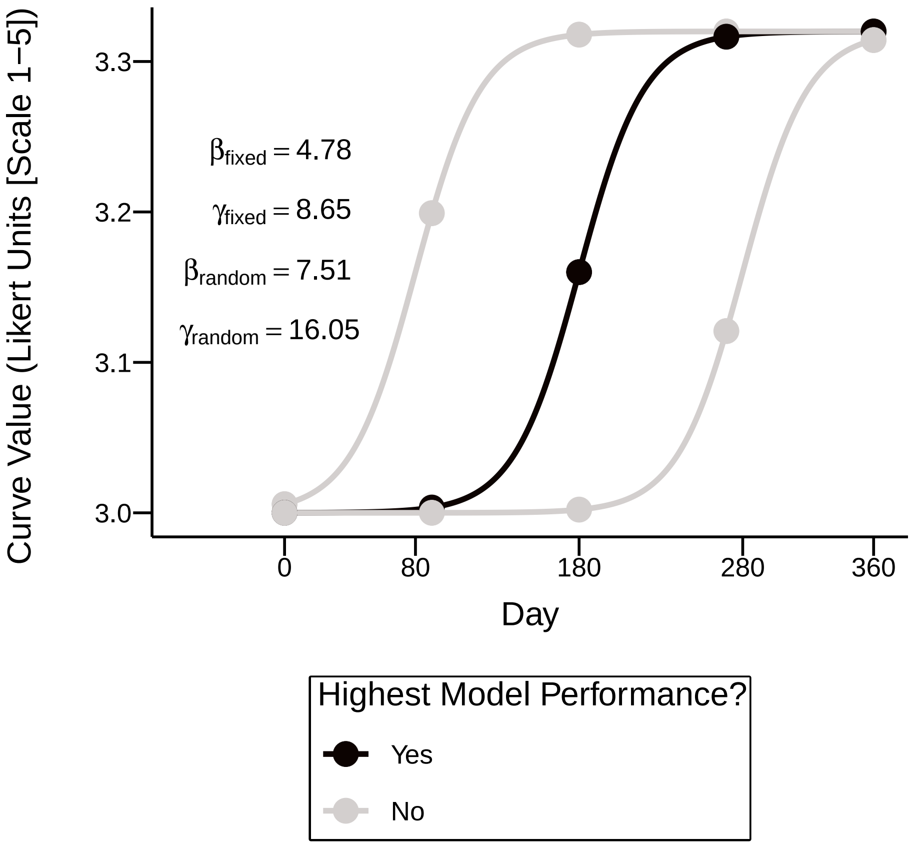
[a] Error bar length is longest in this case because of the existence of high-value outliers (see Figure 2.4).

parameter, they were largely redundant and so were not reported for equal spacing. Although model performance was determined by bias and precision, results for bias were not reported because the differences in bias across the nature-of-change values were negligible. Note that error bar lengths were obtained by computing the average length across all manipulated number of measurements. The column shaded in gray indicates the nature of change where precision was highest (i.e., shortest error bar lengths), which occurred with a nature-of-change value of 180 across all day-unit parameters under equal spacing with one exception (see the 'Highest Model Performance' in Table 2.3). Importantly, with a nature-of-change value of 180, measurements were taken closer to periods of change under equal spacing than with other nature-of-change values (see Figure 2.10). Therefore, it appears that placing measurements closer to periods of change increased model performance with equal spacing.

To understand why precision for the random-effect triquarter-halfway elevation parameter (γ_{random}) was lower with a nature-of-change value of 180, I looked at the distribution of estimated values. Figure 2.4 shows the distribution of values (i.e., density plots) estimated for the random-effect triquarter-halfway elevation parameter (γ_{random}) for each nature-of-change level with five measurements. Importantly, the error bars in the bias/precision plot of Figure 2.5D with five measurements are created from the density plots shown in Figure 2.4. Panel A shows the density plot with a nature-of-change value of 80 ($\beta_{fixed} = 80$). Panel B shows the density plot with a with a nature-of-change value of 180 ($\beta_{fixed} = 180$). Panel C shows the density plot with a with a nature-of-change value of 280 ($\beta_{fixed} = 280$). Regions shaded in gray represent the middle 95% of estimated

Figure 2.3

Model Performance Status Across Nature-of-Change Values With Equal Spacing



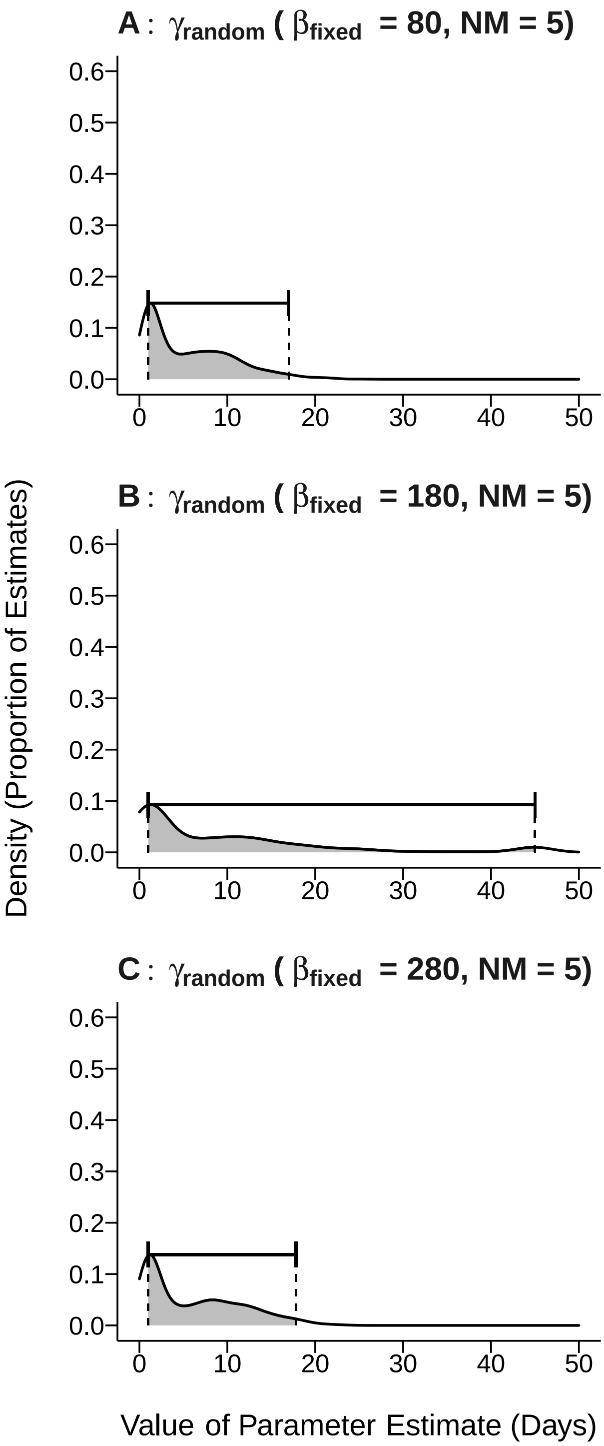
Note. Model performance was highest when measurements were taken closer to periods of greater change, which resulted with a nature-of-change value of 180 with equal spacing. Text prints error bar lengths that resulted when model performance was highest (see Table 2.4).

values and the width of the shaded regions is indicated by the length of the horizontal error bars. As originally confirmed by Table 2.4, Figure 2.4B shows that precision was indeed lowest (i.e., longer error bars) with a nature of change of 180. In looking across the density plots in Figure 2.4, precision was lowest (i.e., longest error bars) for the random-effect triquarter-halfway parameter (γ_{random}) with a nature-of-change value of 180 because of the existence of high-value outliers.

In summary, under equal spacing, model performance for all the day-unit parameters was greatest when the nature-of-change value set by the fixed-effect days-to-halfway

Figure 2.4

Density Plots of the Random-Effect Triquarter-Halfway Delta (γ_{random}; Figure 2.5D) With Equal Spacing in Experiment 1 (95% Error Bars)



Note. Regions shaded in in gray represent the middle 95% of estimated values and the width of the shaded regions is indicated by the length of the horizontal error bars. The error bar length if longest when the nature-of-change value is 180. γ_{random} = random-effect triquarter-halfway delta parameter, with population value of 4.00, NM = number of measurements.

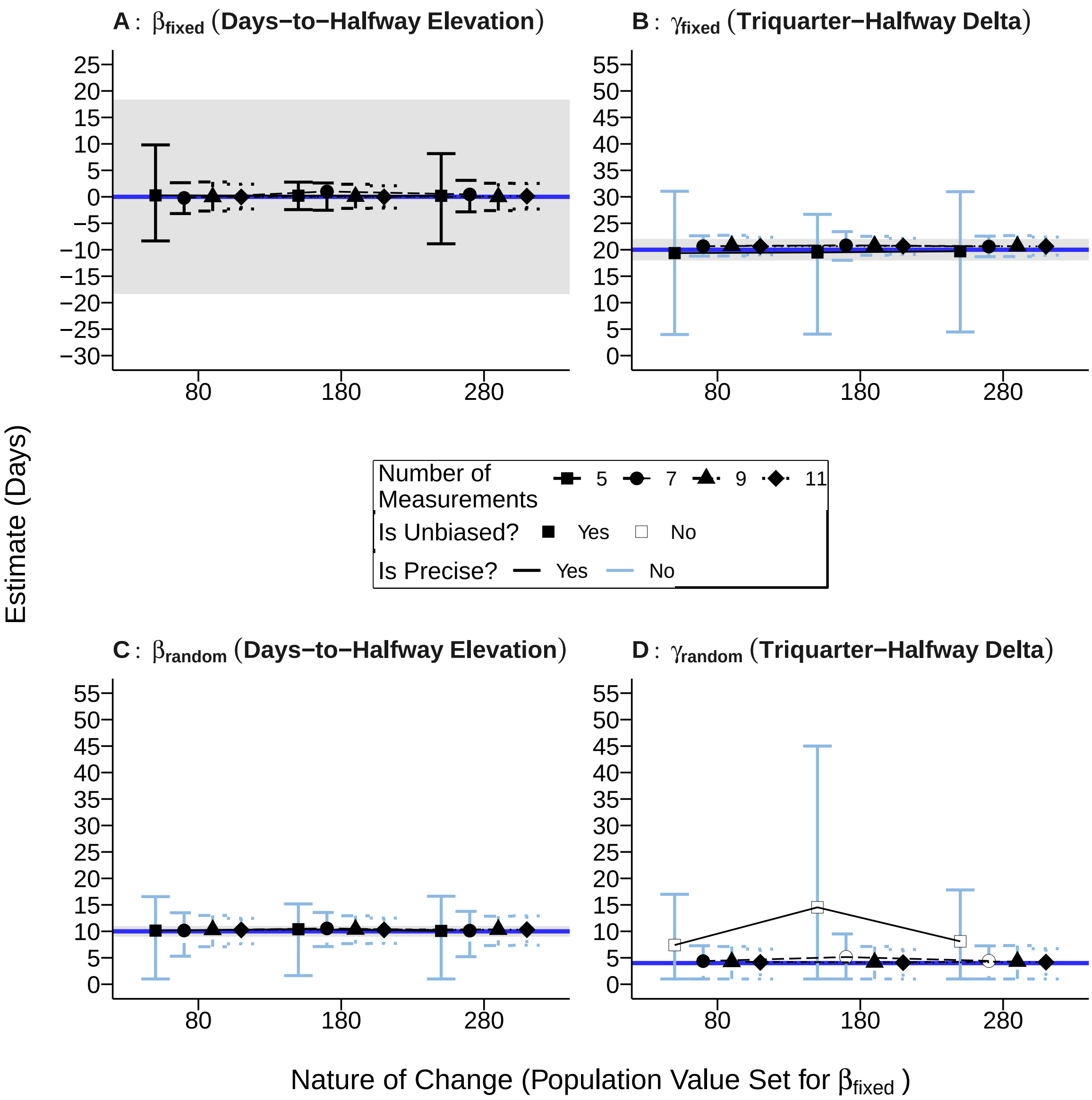
elevation parameter (β_{fixed}) had a value of 180. The one exception to this result was that model performance (as indicated by precision) was lower for the random-effect triquarter-halfway elevation parameter (γ_{random}) with a nature-of-change value of 180 because of high-value outliers.

2.2.3.2 Bias

Before presenting the results for bias, I provide a description of the set of bias/precision plots shown in Figure 2.5 and in the results sections for the other spacing schedules in Experiment 1. Figure 2.5 shows the bias/precision plots for each day-unit parameter and Table 2.5 provides the partial ω^2 values for each independent variable of each day-unit parameter. In Figure 2.5, blue horizontal lines indicate the population values for each parameter (with population values of $\beta_{fixed} \in \{80, 180, 280\}$, β_{random} = 10.00, γ_{fixed} = 20.00, and γ_{random} = 4.00). Gray bands indicate the ±10% margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Panels A–B show the bias/precision plots for the fixed- and random-effect days-to-halfway elevation parameters (β_{fixed} and β_{random}, respectively). Panels C–D show the bias/precision plots for the fixed- and random-effect triquarter-halfway delta parameters (γ_{fixed} and γ_{random}, respectively). Note that random-effect parameter units are in standard deviation units. Importantly, across all population values used for the fixed-effect days-to-halfway elevation parameter (β_{fixed}), the acceptable amount of bias and

Figure 2.5

Bias/Precision Plots for Day-Unit Parameters With Equal Spacing in Experiment 1



Note. Panel A: Bias/precision plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Bias/precision plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Bias/precision plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Bias/precision plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table H.1 for specific values estimated for each parameter and Table 2.5 for ω^2 effect size values.

Table 2.5

Partial ω^2 Values for Manipulated Variables With Equal Spacing in Experiment 1

	Effect		
Parameter	NM	NC	NM x NC
β_{fixed} (Figure 2.5A)	0.02	0.00	0.01
β_{random} (Figure 2.5B)	0.29	0.02	0.02
γ_{fixed} (Figure 2.5C)	0.36	0.01	0.03
γ_{random} (Figure 2.5D)	0.21	0.03	0.04

Note. NM = number of measurements $\in \{5, 7, 9, 11\}$, NC = nature of change (population value set for $\beta_{fixed} \in \{80, 180, 280\}$), NM x NC = interaction between number of measurements and population value set for β_{fixed}. β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect triquarter-halfway delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect triquarter-halfway delta parameter.

precision was based on a population value of 180.

With respect to bias for equal spacing, estimates were biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.5A): no cells.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 2.5B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.5C): no cells.
- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 2.5D): five measurements with all manipulated nature-of-change values and seven measurements with nature-of-change values of 180 and 280.

In summary, with equal spacing, estimation of all the day-unit parameters across all manipulated nature-of-change values was unbiased using nine or more measurements, which is indicated by the emboldened text in the 'Unbiased' column of Table 2.3.

2.2.3.3 Precision

With respect to precision for equal spacing, estimates were imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.5A): no cells.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 2.5B): all cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.5C): all cells.
- random-effect triquarter-halfway delta parameter [γ_{random}] in Figure 2.5D): all cells.

In summary, with equal spacing, estimation across all manipulated nature-of-change values was only precise for the fixed-effect days-to-halfway elevation parameter (β_{fixed}) with five or more measurements. No manipulated measurement number resulted in precise estimation of the fixed-effect triquarter-halfway delta parameter (γ_{fixed}) or the random-effect day-unit parameters (see the 'Precise' column of Table 2.3).

2.2.3.4 Qualitative Description

Although no manipulated measurement number resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) resulted from using moderate measurements numbers. With respect to bias under equal spacing, the largest improvements in bias across all manipulated nature-of-change values resulted from using the following measurement numbers for the following day-unit parameters (note that only the random-effect triquarter halfway delta parameter [γ_{random}] had instances of high bias):

- random-effect triquarter-halfway delta parameters (γ_{random}): seven measurements.

With respect to precision under equal spacing, the largest improvements precision in the estimation of all day-unit parameters (except the fixed-effect days-to-halfway elevation parameter [β_{fixed}]) were obtained with following measurement numbers:

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements, which resulted in a maximum error bar length of 4.37 days.
- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements, which resulted in a maximum error bar length of 7.74 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements, which resulted in a maximum error bar length of 7.02 days.

Therefore, for equal spacing, seven measurements led to the greatest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values (see the emboldened text in the 'Qualitative Description' column of Table 2.3).

2.2.3.5 Summary of Results With Equal Spacing

In summarizing the results for equal spacing, model performance was highest across all day-unit parameters (with the random-effect days-to-halfway elevation parameter (γ_{random}) being an exception) when measurements were placed closer to periods of change, which occurred with a nature-of-change value of 180 (see highest model performance). Unbiased estimation of all the day-unit parameters across all manipulated nature-of-change values resulted from using nine or more measurements (see bias). Precise estimation of all the day-unit parameters was never obtained with any manipulated measurement number (see precision). Although it may be discouraging that no manipulated measurement number under equal spacing resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters were obtained with moderate measurement numbers. With equal spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values were obtained using seven measurements (see Qualitative Description).

2.2.4 Time-Interval Increasing Spacing

For time-interval increasing spacing, Table 2.6 provides a concise summary of the results for the day-unit parameters (see Figure 2.7 for the corresponding bias/precision plots). The sections that follow will present the results for each column of Table 2.6 and provide elaboration when necessary (for a description of Table 2.6, see concise summary table).

2.2.4.1 Nature of Change That Leads to Highest Model Performance

For time-interval increasing spacing, Table 2.7 lists the precision values (i.e., error bar lengths) for each day-unit parameter across each nature-of-change value. The 'Total' column indicates the total error bar length, which is a sum of the the lower ('Lower') and upper ('Upper') whisker lengths. Given that the lower and upper whisker lengths were largely equivalent for each parameter, they were largely redundant and so were not reported for the remainder of the results for time-interval increasing spacing. Although model performance was determined by bias and precision, results for bias were not reported because the differences in bias across the nature-of-change values were negligible.

Table 2.6

Concise Summary of Results for Time-Interval Increasing Spacing in Experiment 1

				Description	
Parameter	Highest Model Performance	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 2.7A)	β_{fixed} = 80	All cells	NM $\geq$ 7	Largest improvement in precision with NM = 7	8.38
γ_{fixed} (Figure 2.7B)	β_{fixed} = 80	All cells	No cells	Largest improvement in precision with **NM = 9**	3.45
β_{random} (Figure 2.7C)	β_{fixed} = 80	NM $\geq$ 7	No cells	Largest improvement in bias and precision with NM = 7	9.47
γ_{random} (Figure 2.7D)	β_{fixed} = 80	**NM $\geq$ 9**	No cells	Largest improvements in bias and precision with **NM = 9**	5.97

Note. 'Highest Model Performance' indicates the curve that resulted in the highest model performance (largely determined by precision; see nature of change). Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the number of measurements that, respectively, resulted in unbiased estimation and the greatest improvements in bias and precision across all day-unit parameters (note that acceptable precision was not obtained in the estimation of all day-unit parameters with time-interval increasing spacing). 'Error Bar Length' indicates the average error bar length value across all nature-of-change values that resulted from using the measurement number in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect triquarter-halfway delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect triquarter-halfway delta parameter = 4. NM = number of measurements.

Note that error bar lengths were obtained by computing the average length across all manipulated number of measurements. The column shaded in gray indicates the nature of change where precision was highest (i.e., shortest error bar lengths), which occurred with a nature-of-change value of 80 across all day-unit parameters under time-interval increasing spacing (see the 'Highest Model Performance' in Table 2.6). Importantly, with a nature-of-change value of 80, measurements were taken closer to periods of change under time-interval increasing spacing than with other nature-of-change values (see Figure 2.6). Therefore, it appears that placing measurements closer to periods of change increased model performance with time-interval increasing spacing.

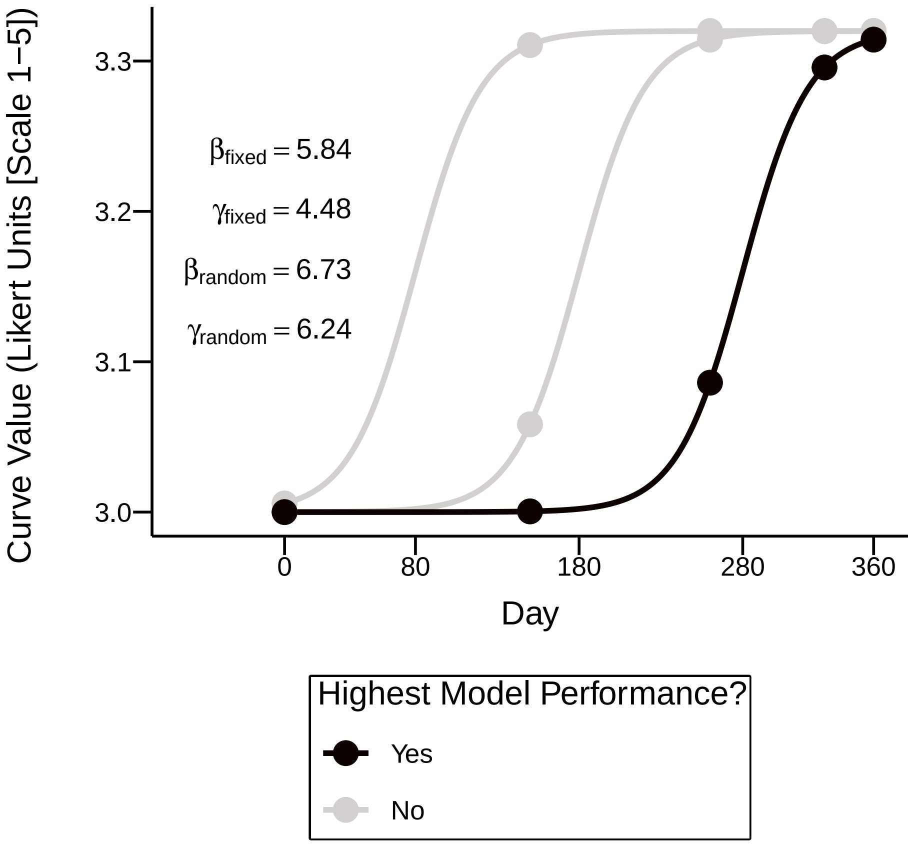
Table 2.7
Error Bar Lengths Across Nature-of-Change Values Under Time-Interval Increasing Spacing in Experiment 1

	Population Value of β_{fixed}								
	80			180			280		
Parameter	Lower	Upper	Total	Lower	Upper	Total	Lower	Upper	Total
β_{fixed} (Figure 2.7A)	3.04	2.76	5.80	3.90	6.72	10.62	17.87	14.84	32.71
γ_{fixed} (Figure 2.7B)	1.59	2.81	4.40	4.39	3.21	7.60	9.00	6.38	15.38
β_{random} (Figure 2.7C)	3.55	3.25	6.80	4.41	4.18	8.59	6.20	9.60	15.81
γ_{random} (Figure 2.7D)	3.00	3.34	6.34	3.00	4.10	7.10	3.00	7.09	10.09

Note. ‘Total’ indicates the total error bar length, which is a sum of the lower (‘Lower’) and upper (‘Upper’) whisker lengths. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect triquarter-halfway delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect triquarter-halfway delta parameter = 4. Note that error bar lengths were calculated by computing the average error bar length value across all number-of-measurement (NM) values (NM $\in \{5, 7, 9, 11\}$). Columns shaded in gray indicate the nature-of-change value that results in the shortest error bar and whisker lengths.

Figure 2.6

Model Performance Status Across Nature-of-Change Values With Time-Interval Increasing Spacing



Note. Model performance was highest when measurements were taken closer to periods of greater change, which resulted with a nature-of-change value of 80 with equal spacing. Text prints error bar lengths that resulted when model performance was highest (see Table 2.7).

2.2.4.2 Bias

With respect to bias for time-interval increasing spacing, estimates were biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.7A): no cells.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 2.7B): no cells
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.7C): five measurements with a nature-of-change value of 280.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.7C): five measurements with all nature-of-change values and seven measurements with nature-of-change values of 180 and 280.

In summary, with time-interval increasing spacing, estimation of all the day-unit parameters across all manipulated nature-of-change values was unbiased using nine or more measurements, which is indicated by the emboldened text in the 'Unbiased' column of Table 2.6.

2.2.4.3 Precision

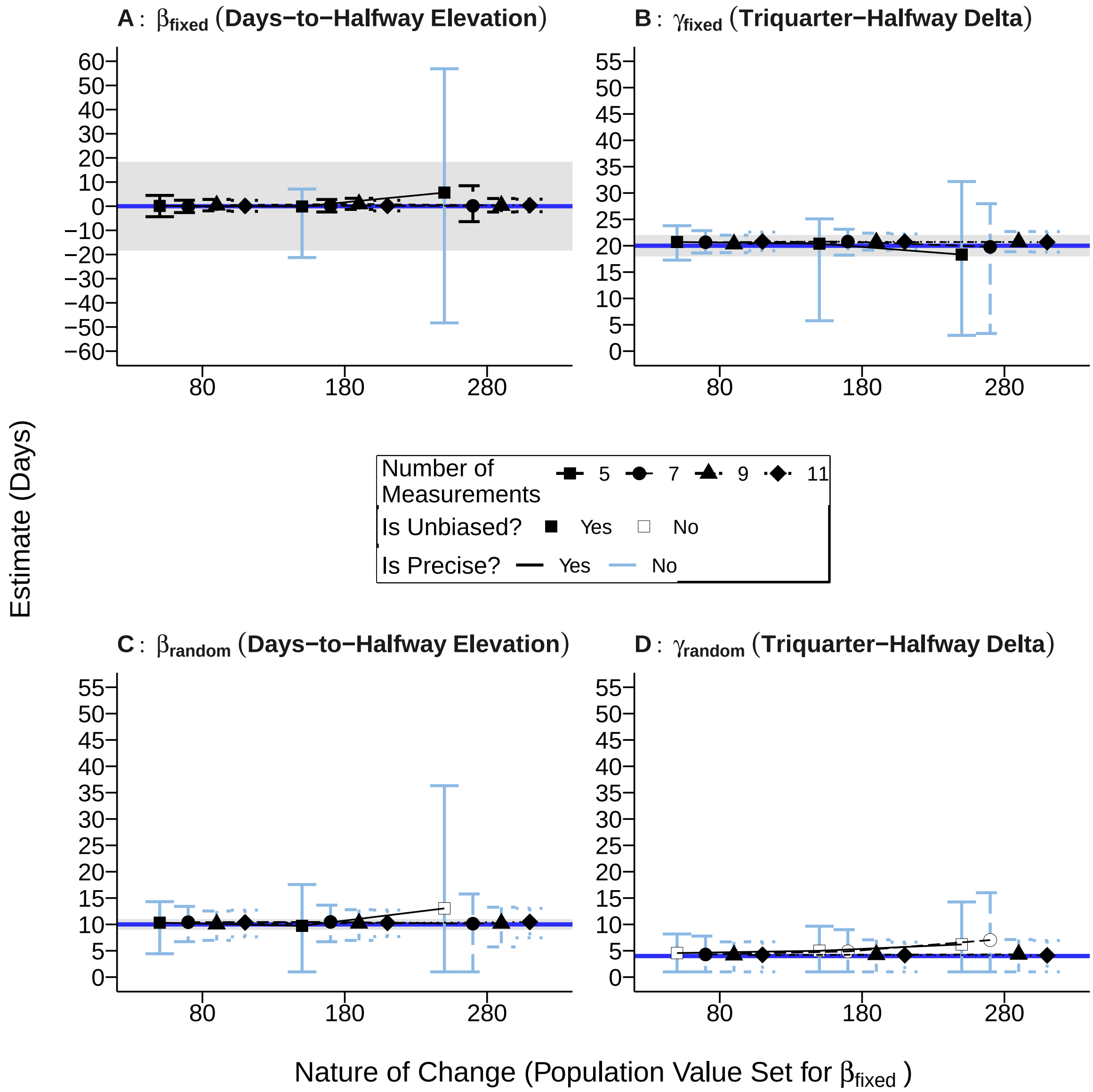
With respect to precision for time-interval increasing spacing, estimates were imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.7A): five measurements with nature-of-change values of 180 and 280.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 2.7B): all cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.7C): all cells.
- random-effect triquarter-halfway delta parameter (γ_{random}; Figure 2.7D): all cells.

In summary, with time-interval increasing spacing, estimation across all manipulated nature-of-change values was only precise for the fixed-effect days-to-halfway elevation parameter (β_{fixed}) with seven or more measurements. No manipulated measurement number resulted in precise estimation of the fixed-effect triquarter-halfway delta parameter (γ_{fixed}) or the random-effect day-unit parameters (see the 'Precise' column of Table 2.6).

Figure 2.7

Bias/Precision Plots for Day-Unit Parameters With Time-Interval Increasing Spacing in Experiment 1



Note. Panel A: Bias/precision plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Bias/precision plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Bias/precision plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Bias/precision plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80.00, 180.00, 280.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table H.1 for specific values estimated for each parameter and Table 2.8 for ω^2 effect size values.

Table 2.8

Partial ω^2 Values for Manipulated Variables With Time-Interval Decreasing Spacing in Experiment 1

	Effect		
Parameter	NM	NC	NM x NC
β_{fixed} (Figure 2.7A)	0.43	0.30	0.50
β_{random} (Figure 2.7B)	0.12	0.04	0.05
γ_{fixed} (Figure 2.7C)	0.26	0.21	0.22
γ_{random} (Figure 2.7D)	0.12	0.05	0.04

Note. NM = number of measurements $\in \{5, 7, 9, 11\}$, NC = nature of change (population value set for $\beta_{fixed} \in \{80, 180, 280\}$), NM x NC = interaction between number of measurements and population value set for β_{fixed}. β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect triquarter-halfway delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect triquarter-halfway delta parameter.

2.2.4.4 Qualitative Description

For time-interval increasing spacing in Figure 2.7, although no manipulated measurement number resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) resulted from using moderate measurements numbers. With respect to bias under time-interval increasing spacing, the largest improvements across all manipulated nature-of-change values in bias occurred with the following measurement numbers for the random-effect day-unit parameters:

- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements.
- random-effect triquarter-halfway delta parameters (γ_{random}): nine measurements.

With respect to precision under time-interval increasing spacing, the largest improvements precision in the estimation of all day-unit parameters across all manipulated nature-of-change values resulted with the following measurement numbers:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}): seven measurements, which results in an average error bar length of 8.38 days.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): nine measurements, which results in an average error bar length of 3.45 days.
- random-effect days-to-halfway elevation parameter (β_{random}): nine measurements, which results in an average error bar length of 9.47 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): nine measurements, which results in an average error bar length of 5.97 days.

Therefore, for time-interval increasing spacing, nine measurements resulted in the greatest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values (see the ’Qualitative Description‘ column in Table

2.6).

2.2.4.5 Summary of Results With Time-Interval Increasing Spacing

In summarizing the results for time-interval increasing spacing, model performance was highest across all day-unit parameters when measurements were placed closer to periods of change, which occurred with a nature-of-change value of 80 (β_{fixed} = 80; see highest model performance). Estimation of all day-unit parameters was unbiased across all manipulated nature-of-change values using nine or more measurements (see bias). Precise estimation was never obtained in the estimation of all day-unit parameters with any manipulated measurement (see precision). Although it may be discouraging that no manipulated measurement number under time-interval increasing spacing resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters were obtained with moderate measurement numbers. With time-interval increasing spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values resulted from using nine measurements (see qualitative description).

2.2.5 Time-Interval Decreasing Spacing

For time-interval decreasing spacing, Table 2.9 provides a concise summary of the results for the day-unit parameters (see Figure 2.9 for the corresponding bias/precision plots). The sections that follow will present the results for each column of Table 2.9 and provide elaboration when necessary (for a description of Table 2.9, see concise summary table).

Table 2.9

Concise Summary of Results for Time-Interval Decreasing Spacing in Experiment 1

				Description	
Parameter	Highest Model Performance	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 2.9A)	β_{fixed} = 280	All cells	NM $\geq$ 9	Largest improvements in precision with **NM = 9**	4.88
γ_{fixed} (Figure 2.9B)	β_{fixed} = 280	NM $\geq$ 7	No cells	Largest improvement in precision with **NM = 9**	3.40
β_{random} (Figure 2.9C)	β_{fixed} = 280	NM $\geq$ 7	No cells	Largest improvement in bias and precision with **NM = 9**	6.15
γ_{random} (Figure 2.9D)	β_{fixed} = 280	**NM $\geq$ 9**	No cells	Largest improvements in bias and precision with **NM = 9**	5.96

Note. ‘Highest Model Performance’ indicates the curve that resulted in the highest model performance (largely determined by precision; see nature of change). Emboldened text in the ‘Unbiased’ and ‘Qualitative Description’ columns indicates the number of measurements that, respectively, resulted in unbiased estimation and the greatest improvements in bias and precision across all day-unit parameters (note that acceptable precision was not obtained in the estimation of all day-unit parameters with time-interval decreasing spacing). ‘Error Bar Length’ indicates the average error bar length value across all nature-of-change values that resulted from using the measurement number in the ‘Qualitative Description’ column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect triquarter-halfway delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect triquarter-halfway delta parameter = 4. NM = number of measurements.

2.2.5.1 Nature of Change That Leads to Highest Model Performance

For time-interval decreasing spacing, Table 2.10 lists the error bar lengths for each day-unit parameter and nature-of-change value. The 'Total' column indicates the total error bar length, which is a sum of the the lower ('Lower') and upper ('Upper') whisker lengths. Given that the lower and upper whisker lengths were largely equivalent for each parameter, they were largely redundant and so were not reported for the remainder of the results for time-interval decreasing spacing. Although model performance was determined by bias and precision, results for bias were not computed because the differences in bias across the nature-of-change values were negligible. Note that error bar lengths were obtained by computing the average length across all manipulated measurement number values. The column shaded in gray indicates the nature of change where precision was highest (i.e., shortest error bar lengths), which occurred with a nature-of-change value of 280 across all day-unit parameters under time-interval decreasing spacing (see the 'Highest Model Performance' in Table 2.9). Importantly, with a nature-of-change value of 280, measurements were taken closer to periods of change under time-interval decreasing spacing than with other nature-of-change values (see Figure 2.8). Therefore, it appears that placing measurements closer to periods of change increased model performance with time-interval decreasing spacing.

2.2.5.2 Bias

With respect to bias for time-interval decreasing spacing, estimates were biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.9A): no cells.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 2.9B): five

Table 2.10

Error Bar Lengths Across Nature-of-Change Values Under Time-Interval Decreasing Spacing in Experiment 1

	Population Value of β_{fixed}								
	80			180			280		
Parameter	Lower	Upper	Total	Lower	Upper	Total	Lower	Upper	Total
β_{fixed} (Figure 2.9A)	30.51	15.73	46.24	7.64	3.67	11.31	3.28	2.56	5.84
γ_{fixed} (Figure 2.9B)	9.70	6.11	15.81	4.88	3.14	8.02	1.79	2.69	4.48
β_{random} (Figure 2.9C)	6.09	11.26	17.35	4.70	3.90	8.60	3.60	3.13	6.73
γ_{random} (Figure 2.9D)	3.00	6.57	9.57	3.00	4.20	7.20	3.00	3.24	6.24

Note. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect triquarter-halfway delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect triquarter-halfway delta parameter = 4. Note that error bar lengths were calculated by computing the average error bar length value across all number-of-measurement (NM) values (NM $\in$ {5, 7, 9, 11}). Columns shaded in gray indicate the nature-of-change value that results in the shortest error bar and whisker lengths.

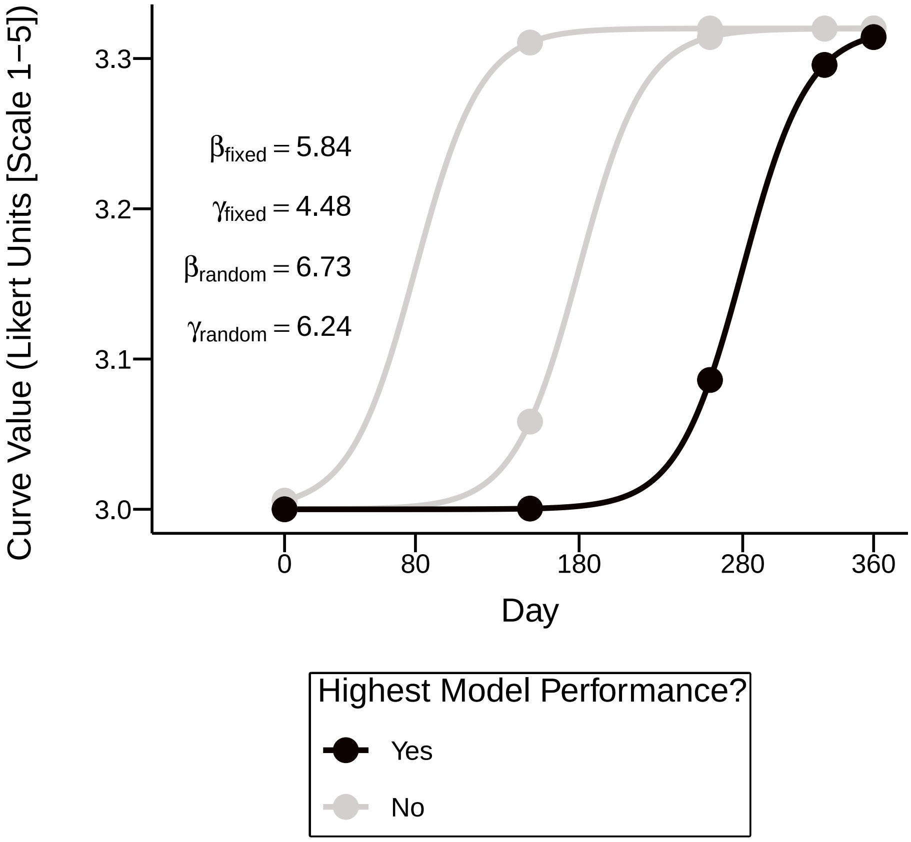
measurements with a nature-of-change value of 80.

- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.9C): five measurements with a nature-of-change value of 80.
- random-effect triquarter-halfway delta parameter (γ_{random}; Figure 2.9D): five measurements across all manipulated nature-of-change values and seven measurements with nature-of-change values of 80 and 180.

In summary, with time-interval decreasing spacing, unbiased estimation was obtained for all day-unit parameters across all manipulated nature-of-change values using nine or more measurements, which is indicated by the emboldened text in the 'Unbiased' column of Table 2.9.

Figure 2.8

Model Performance Status Across Nature-of-Change Values With Time-Interval Decreasing Spacing



Note. Model performance was highest when measurements were taken closer to periods of greater change, which resulted with a nature-of-change value of 280 with equal spacing. Text prints error bar lengths that resulted when model performance was highest (see Table 2.10).

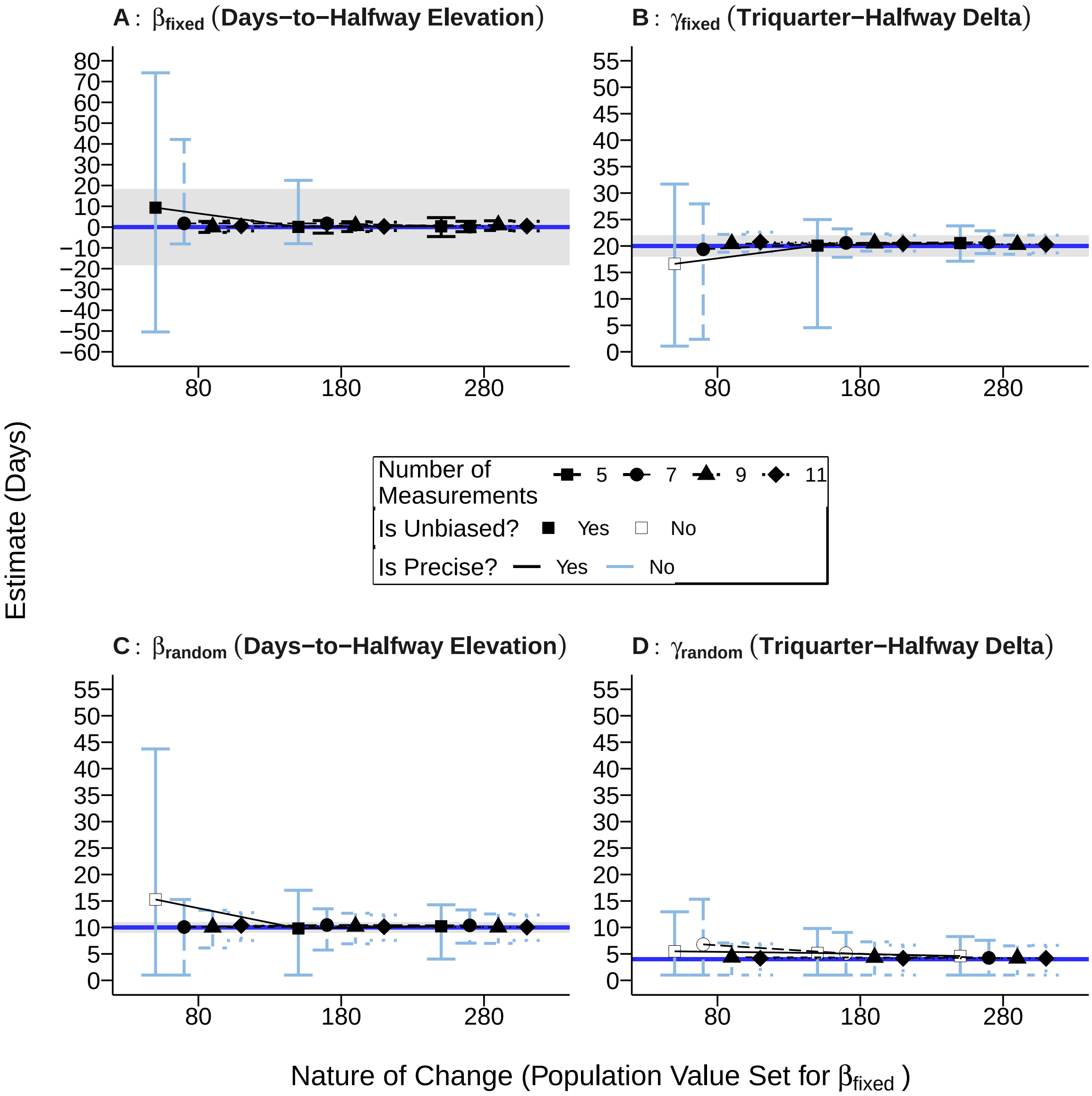
2.2.5.3 Precision

With respect to precision for time-interval decreasing spacing, estimates were imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.9A): five measurements with nature-of-change values of 80 and 180 an seven measurements with a nature-of-change value of 80.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 2.9B): all cells.

Figure 2.9

Bias/Precision Plots for Day-Unit Parameters With Time-Interval Decreasing Spacing in Experiment 1



Note. Panel A: Bias/precision plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Bias/precision plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Bias/precision plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Bias/precision plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80.00, 180.00, 280.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table H.1 for specific values estimated for each parameter and Table 2.11 for ω^2 effect size values.

Table 2.11
Partial ω^2 Values for Manipulated Variables With Time-Interval Decreasing Spacing in Experiment 1

	Effect		
Parameter	NM	NC	NM x NC
β_{fixed} (Figure 2.9A)	0.20	0.10	0.22
β_{random} (Figure 2.9B)	0.13	0.04	0.05
γ_{fixed} (Figure 2.9C)	0.27	0.19	0.21
γ_{random} (Figure 2.9D)	0.11	0.03	0.03

Note. NM = number of measurements $\in \{5, 7, 9, 11\}$, NC = nature of change (population value set for $\beta_{fixed} \in \{80, 180, 280\}$), NM x NC = interaction between number of measurements and population value set for β_{fixed}. β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect triquarter-halfway delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect triquarter-halfway delta parameter.

- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.9C): all cells.
- random-effect triquarter-halfway delta parameter (γ_{random}; Figure 2.9D): all cells.

In summary, with time-interval increasing spacing, estimation across all manipulated nature-of-change values was only precise for the estimation of the fixed-effect days-to-halfway elevation parameter (β_{fixed}) with nine or more measurements. No manipulated measurement number resulted in precise estimation of the fixed-effect triquarter-halfway delta parameter (γ_{fixed}) or the random-effect day-unit parameters (see the 'Precise' column of Table 2.9).

2.2.5.4 Qualitative Description

For time-interval decreasing spacing in Figure 2.9, although no manipulated measurement number resulted in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) were obtained using moderate measurements numbers. With respect to bias under time-interval decreasing spacing, the largest improvements across all manipulated nature-of-change values in bias occurred with the following measurement numbers for the random-effect day-unit parameters:

- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements
- random-effect triquarter-halfway delta parameters (γ_{random}): nine measurements

With respect to precision under time-interval decreasing spacing, the largest improvements precision in the estimation of all day-unit parameters across all manipulated nature-of-change values were obtained with the following measurement numbers:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}): seven measurements, which results in a maximum error bar length of 20.42 days.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): nine measurements, which results in a maximum error bar length of 3.4 days.
- random-effect days-to-halfway elevation parameter (β_{random}): nine measurements, which results in a maximum error bar length of 9.45 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): nine measurements, which results in a maximum bar length of 5.96 days.

Therefore, for time-interval decreasing spacing, nine measurements resulted in the greatest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values (see the emboldened text in the 'Qualitative Description' column in Table 2.9).

2.2.5.5 Summary of Results Time-Interval Decreasing Spacing

In summarizing the results for time-interval decreasing spacing, model performance was highest across all day-unit parameters when measurements were placed closer to periods of change, which occurred with a nature-of-change value of 280 (β_{fixed} = 280; see highest model performance). Unbiased estimation of the day-unit parameters across all manipulated nature-of-change values resulted from using nine or more measurements (see bias). Precise estimation of all the day-unit parameters was never obtained using any of the manipulated measurement numbers (see precision). Although it may be discouraging that no manipulated measurement number under time-interval decreasing spacing resulted in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters were obtained with moderate measurement numbers. With time-interval decreasing spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values were obtained using nine measurements (see qualitative description).

2.2.6 Middle-and-Extreme Spacing

For middle-and-extreme spacing, Table 2.12 provides a concise summary of the results for the day-unit parameters (see Figure 2.11 for the corresponding bias/precision plots). The sections that follow will present the results for each column of Table 2.12 and provide elaboration when necessary (for a description of Table 2.12, see concise summary table).

Table 2.12

Concise Summary of Results for Middle-and-Extreme Spacing in Experiment 1

				Description	
Parameter	Highest Model Performance	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 2.11A)	β_{fixed} = 180	All cells	NM $\geq$ 7	Largest improvements in precision with NM = 7	14.10
γ_{fixed} (Figure 2.11B)	β_{fixed} = 180	NM $\geq$ 7	No cells	Largest improvements in bias and precision with NM = 7	6.27
β_{random} (Figure 2.11C)	β_{fixed} = 180	NM $\geq$ 9	No cells	Largest improvements in bias and precision with **NM = 9**	9.02
γ_{random} (Figure 2.11D)	β_{fixed} = 180	**NM = 11**	No cells	Largest improvements in bias and precision with NM = 7	7.92

Note. 'Highest Model Performance' indicates the curve that resulted in the highest model performance (largely determined by precision; see nature of change). Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the number of measurements that, respectively, resulted in unbiased estimation and the greatest improvements in bias and precision across all day-unit parameters (note that acceptable precision was not obtained in the estimation of all day-unit parameters with middle-and-extreme spacing). 'Error Bar Length' indicates the average error bar length value across all nature-of-change values that resulted from using the measurement number in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect triquarter-halfway delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect triquarter-halfway delta parameter = 4. NM = number of measurements.

2.2.6.1 Nature of Change That Leads to Highest Model Performance

For middle-and-extreme spacing, Table 2.13 lists the error bar lengths for each day-unit parameter and nature-of-change value. The 'Total' column indicates the total error bar length, which is a sum of the the lower ('Lower') and upper ('Upper') whisker lengths. Given that the lower and upper whisker lengths were largely equivalent for each parameter, they were largely redundant and so were not reported for the remainder of the results for middle-and-extreme spacing. Although model performance was determined by bias and precision, results for bias were not reported because the differences in bias across the nature-of-change values were negligible. Note that error bar lengths were obtained by computing the average length across all manipulated number-of-measurement values. The column shaded in gray indicates the nature of change where precision was highest (i.e., shortest error bar lengths), which occurred with a nature-of-change value of 180 across all day-unit parameters under middle-and-extreme spacing (see the 'Highest Model Performance' in Table 2.12). Importantly, with a nature-of-change value of 180, measurements were taken closer to periods of change under middle-and-extreme spacing than with other nature-of-change values (see Figure 2.10). Therefore, it appears that placing measurements closer to periods of change increased model performance with middle-and-extreme spacing.

2.2.6.2 Bias

With respect to bias for middle-and-extreme spacing, estimates were biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.9A): no cells.

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 2.11B): five measurements with nature-of-change values of 80 and 280.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.11C): five and seven measurements with nature-of-change values of 80 and 280.
- random-effect triquarter-halfway delta parameter (γ_{random}; Figure 2.11D): five, seven, and nine measurements with nature-of-change values of 80 and 280.

In summary, with middle-and-extreme spacing, estimation of all the day-unit parameters across all manipulated nature-of-change values were unbiased using 11 measurements, which is indicated by the emboldened text in the 'Unbiased' column of Table 2.12.

Table 2.13

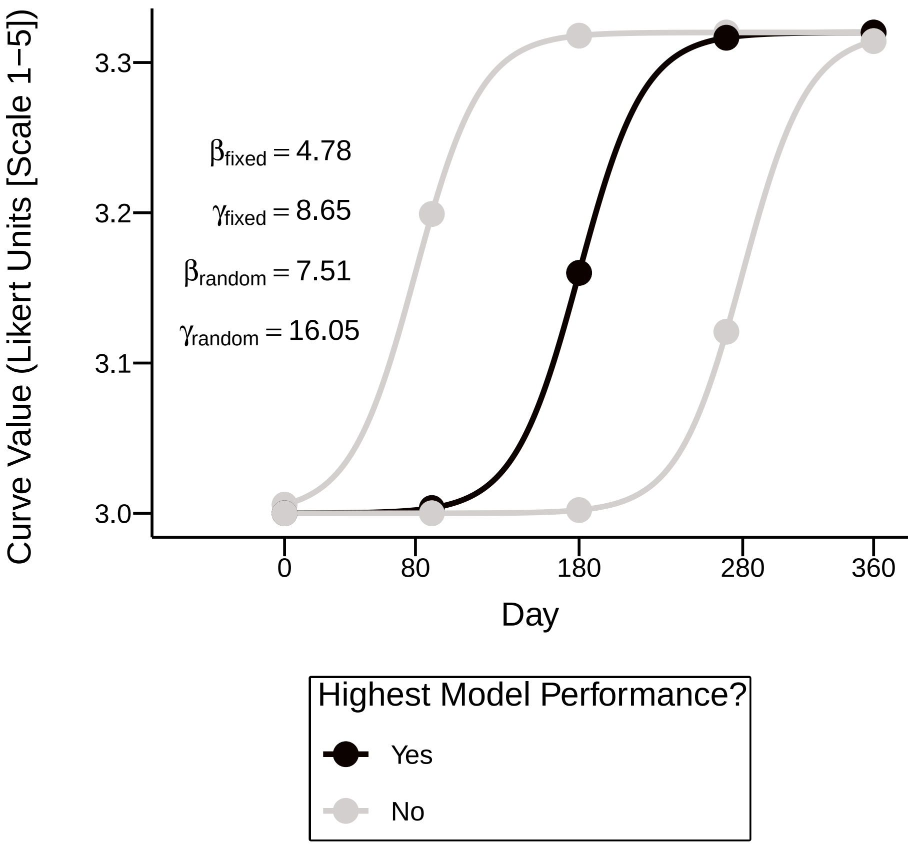
Error Bar Lengths Across Nature-of-Change Values Under Middle-and-Extreme Spacing in Experiment 1

	Population Value of β_{fixed}								
	80			180			280		
Parameter	Lower	Upper	Total	Lower	Upper	Total	Lower	Upper	Total
β_{fixed} (Figure 2.11A)	22.13	19.89	42.02	2.25	2.21	4.46	20.32	21.74	42.06
γ_{fixed} (Figure 2.11B)	6.50	5.77	12.27	0.87	2.22	3.09	6.73	6.11	12.84
β_{random} (Figure 2.11C)	7.14	16.84	23.97	2.28	2.48	4.76	7.27	15.69	22.96
γ_{random} (Figure 2.11D)	3.00	6.20	9.20	3.00	2.73	5.73	3.00	6.77	9.77

Note. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect triquarter-halfway delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect triquarter-halfway delta parameter = 4. Note that error bar lengths were calculated by computing the average error bar length value across all number-of-measurement (NM) values (NM $\in$ {5, 7, 9, 11}). Columns shaded in gray indicate the nature-of-change value that results in the shortest error bar and whisker lengths.

Figure 2.10

Model Performance Status Across Nature-of-Change Values With Middle-and-Extreme Spacing



Note. Model performance was highest when measurements were taken closer to periods of greater change, which resulted with a nature-of-change value of 180 with middle-and-extreme spacing. Text prints error bar lengths that resulted when model performance was highest (see Table 2.13).

2.2.6.3 Precision

With respect to precision for middle-and-extreme spacing, estimates were imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

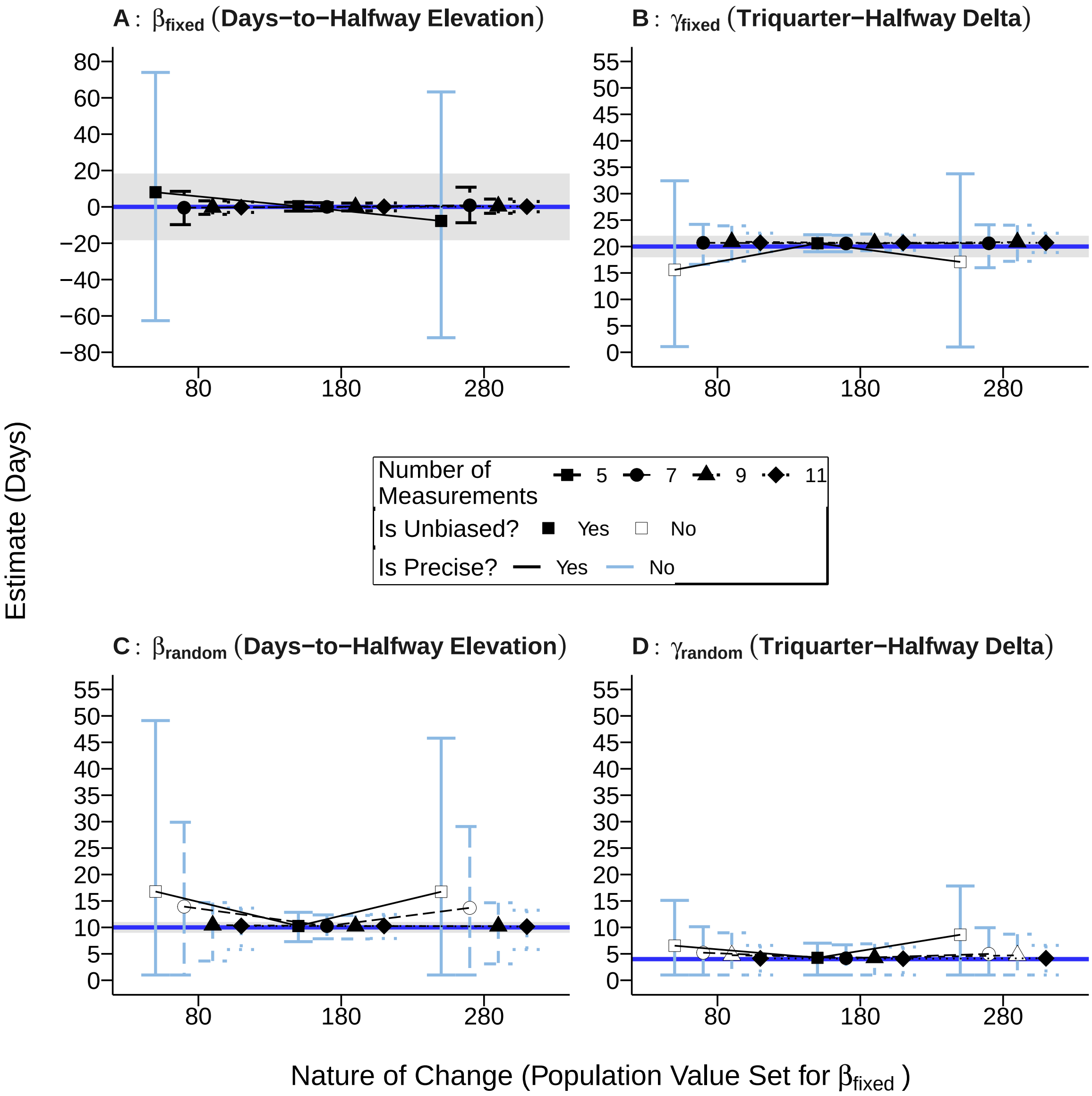
- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.11A): five measurements with nature-of-change values of 80 and 280.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 2.11B): five and seven, an nine measurements with nature-of-change values of 80 and 280 (shown on x-axis).

- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.11C): all cells.
- random-effect triquarter-halfway delta parameter (γ_{random}; Figure 2.11D): all cells.

In summary, with middle-and-extreme spacing, precise estimation of the fixed-effect day-unit parameters across all manipulated nature-of-change values was obtained with 11 measurements, but no manipulated measurement number resulted in precise estimation of the random-effect day-unit parameters (see the 'Precise' column of Table 2.12).

Figure 2.11

Bias/Precision Plots for Day-Unit Parameters With Middle-and-Extreme Spacing in Experiment 1



Note. Panel A: Bias/precision plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Bias/precision plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Bias/precision plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Bias/precision plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80.00, 180.00, 280.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table H.1 for specific values estimated for each parameter and Table 2.14 for ω^2 effect size values.

Table 2.14

Partial ω^2 Values for Manipulated Variables With Middle-and-Extreme Spacing in Experiment 1

	Effect		
Parameter	NM	NC	NM x NC
β_{fixed} (Figure 2.11A)	0.32	0.09	0.19
β_{random} (Figure 2.11B)	0.12	0.09	0.06
γ_{fixed} (Figure 2.11C)	0.49	0.20	0.32
γ_{random} (Figure 2.11D)	0.07	0.05	0.03

Note. NM = number of measurements $\in \{5, 7, 9, 11\}$, NC = nature of change (population value set for $\beta_{fixed} \in \{80, 180, 280\}$), NM x NC = interaction between number of measurements and population value set for β_{fixed}. β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect triquarter-halfway delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect triquarter-halfway delta parameter.

2.2.6.4 Qualitative Description

For middle-and-extreme spacing in Figure 2.11, although no manipulated measurement number resulted in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) were obtained using moderate measurements numbers. With respect to bias under middle-and-extreme spacing, the largest improvements across all manipulated nature-of-change values in bias occurred with the following measurement numbers for the following day-unit parameters:

- random-effect days-to-halfway elevation parameter (γ_{fixed}): seven measurements
- random-effect days-to-halfway elevation parameter (β_{random}): nine measurements
- random-effect triquarter-halfway delta parameters (γ_{random}): 11 measurements

With respect to precision under middle-and-extreme spacing, the largest improvements precision in the estimation of all day-unit parameters across all manipulated nature-of-change values result with the following measurement numbers:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}): seven measurements, which results in a maximum error bar length of 14.1 days.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements, which results in a maximum error bar length of 5.55 days.
- random-effect days-to-halfway elevation parameter (β_{random}): nine measurements, which results in a maximum error bar length of 20.49 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements, which results in a maximum error bar length of 7.2 days.

Therefore, for middle-and-extreme spacing, nine measurements resulted in the greatest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values (see the emboldened text in the 'Qualitative Description' column in Table 2.12).

2.2.6.5 Summary of Results With Middle-and-Extreme Spacing

In summarizing the results for middle-and-extreme spacing, model performance was highest across all day-unit parameters when measurements were placed closer to periods of change, which occurred with a nature-of-change value of 180 ($\beta_{fixed} = 180$); see highest model performance). Unbiased estimation of the day-unit parameters across all manipulated nature-of-change values resulted from using nine or more measurements (see bias). Precise estimation of all the day-unit parameters was never obtained using any of the manipulated measurement numbers (see precision). Although it may be discouraging that no manipulated measurement number under time-interval decreasing spacing resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters were obtained with moderate measurement numbers. With time-interval decreasing spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values resulted from using nine measurements (see qualitative description).

2.2.7 Addressing My Research Questions

2.2.7.1 Does Placing Measurements Near Periods of Change Increase Model Performance?

In Experiment 1, one question I had was whether placing measurements near periods of change increases model performance. To answer this question, I have recorded the nature of change values that result in the highest model performance for each spacing schedule in Table 2.15. Text in the 'Highest Model Performance' column indicates the

Table 2.15

Nature-of-Change Values That Lead to the Highest Model Performance for Each Spacing Schedule in Experiment 1

		Error Bar Summary			
Spacing Schedule	Highest Model Performance	β_{fixed}	γ_{fixed}	β_{random}	γ_{random}
Equal (see Figure 2.5 and Table 2.4)	β_{fixed} = 180	4.78	8.65	7.51	16.05
Time-interval increasing (see Figure 2.7 and Table 2.7)	β_{fixed} = 80	5.80	4.40	6.80	6.34
Time-interval decreasing (see Figure 2.9 and Table 2.10)	β_{fixed} = 280	5.84	4.48	6.73	6.24
Middle-and-extreme (see Figure 2.11 and Table 2.13)	β_{fixed} = 180	4.46	3.09	4.76	5.73

Note. 'Highest Model Performance' indicates the curve that results in the highest model performance. 'Error Bar Summary' columns lists error bar lengths for each day-unit parameter such that error bar lengths are computed by taking the average error bar length value across all the number-of-measurement (NM) values (NM $\in$ {5, 7, 9, 11}). Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter $\in$ {80, 180, 280}; γ_{fixed} = fixed-effect triquarter-halfway delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect triquarter-halfway delta parameter = 4.

nature-of-change with which each spacing schedule obtains its highest model performance. The 'Error Bar Summary' columns list the error bar lengths obtained for each day-unit parameter using the nature-of-change value listed in the 'Highest Model Performance' column.[14] Note that the error bar lengths are obtained by computing the average error bar length across all manipulated measurement numbers for the optimal nature-of-change value. Model performance for each spacing schedule is highest with the following nature-of-change values:

- equal spacing: $\beta_{fixed} = 180$
- time-interval increasing spacing: $\beta_{fixed} = 80$
- time-interval decreasing spacing: $\beta_{fixed} = 280$
- middle-and-extreme spacing: $\beta_{fixed} = 180$

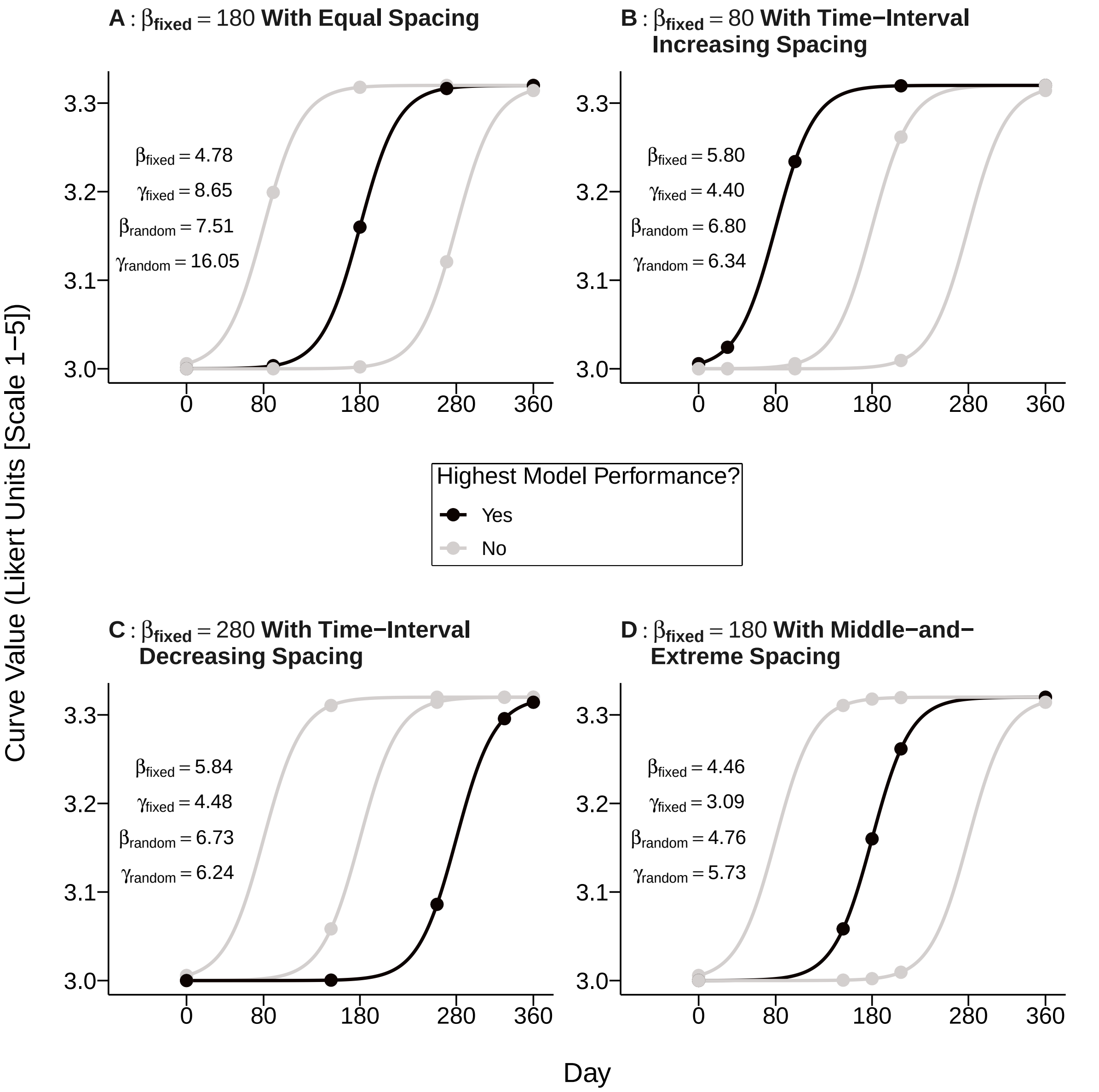
To understand why the model performance of each spacing schedule is highest with a specific nature of change, it is important to consider the locations on the curve where each schedule samples data. Figure 2.12 shows the measurement locations (indicated by dots) where each spacing schedule samples data for each manipulated nature of change ($\beta_{fixed} \in \{80, 180, 180\}$). In Figure 2.12A, data are sampled according to the equal spacing schedule. In Figure 2.12B, data are sampled according to the time-interval increasing spacing schedule. In Figure 2.12C, data are sampled according to the time-interval decreasing spacing schedule. In Figure 2.12D, data are sampled according to the middle-and-extreme spacing schedule. Black curves indicate curves for which model performance is highest and gray curves indicating curves where model performance is not at its highest. Error bar lengths (i.e., precision) for the estimation of each day-unit parameter are

[14]Bias values are not presented because the differences across the schedules are negligible.

copied

Figure 2.12

Nature-of-Change Curves for Each Spacing Schedule Have Highest Model Performance When Measurements are Taken Near Periods of Change



Note. Panel A: Measurement sampling locations on each manipulated nature-of-change curve under equal spacing. Panel B: Measurement sampling locations on each manipulated nature-of-change curve under time-interval increasing spacing. Panel C: Measurement sampling locations on each manipulated nature-of-change curve under time-interval decreasing spacing. Panel D: Measurement sampling locations on each manipulated nature-of-change curve under middle-and-extreme spacing. Black curves indicate the natures of change that lead to the highest model performance for each spacing schedule, and so are optimal. Gray curves indicate the natures of change that lead to suboptimal model performance for each spacing schedule, and so are not optimal. Text on each panel indicates the error bar lengths when model performance is highest (see Table 2.15).

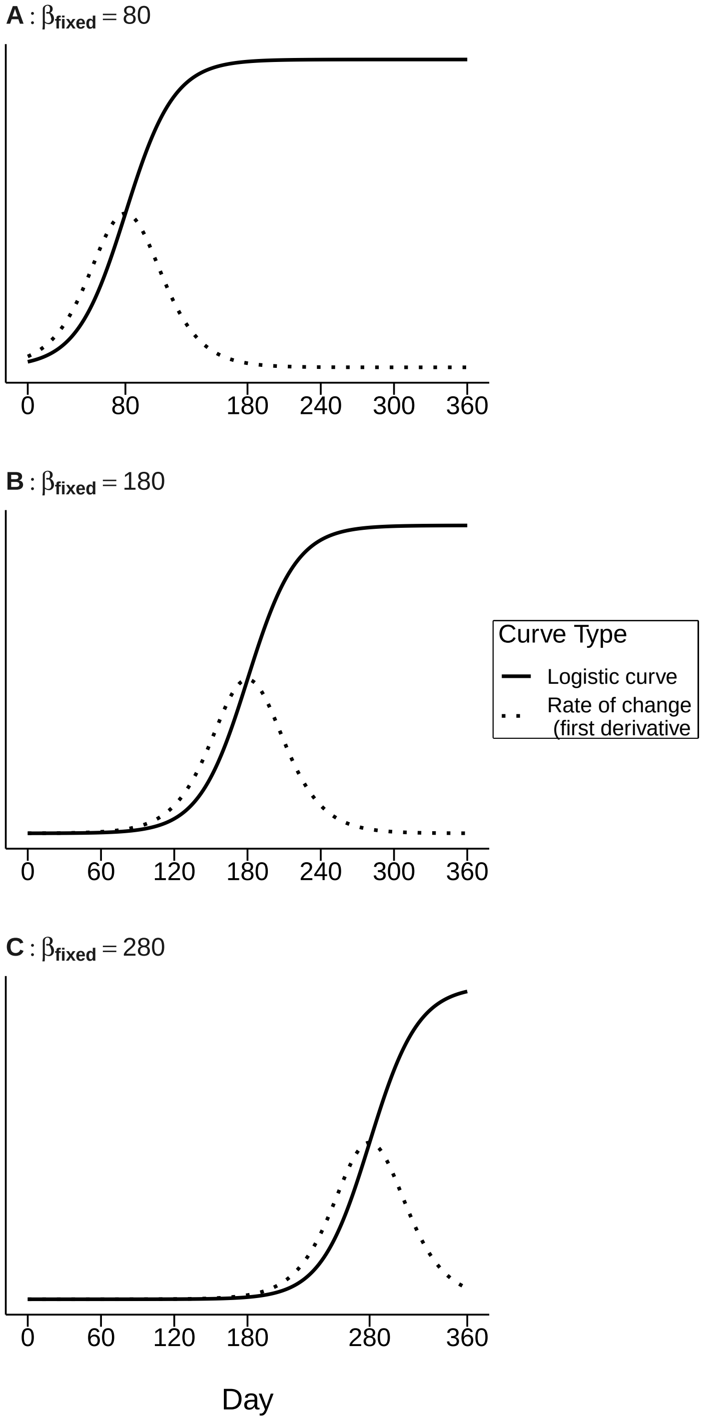
over from Table 2.15 to provide a reference with which to compare model performance between the spacing schedules with the optimal nature of change.

To investigate whether placing measurements near periods of change increases model performance, it is first important to define change. For the purpose of this discussion, change occurs when the first derivative of the logistic function has a nonzero value, with a larger (absolute) first derivative value implying greater change. Figure 2.13 shows each nature of change used in Experiment 1 (solid line) along with its corresponding first derivative curve (dotted line). For each nature of change, the first derivative value reaches its peak at the value set for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). In Figure 2.13A, the first derivative is greatest at day 80. In Figure 2.13B, the first derivative is greatest at day 180. In Figure 2.13C, the first derivative is greatest at day 280. Therefore, for each manipulated nature of change, change is greatest at the value set for the fixed-effect days-to-halfway elevation parameter (β_{fixed}).

Revisiting the question of whether placing measurements near periods of change increases model performance, I believe there are reasons to support this idea, and each reason is depicted in Figure 2.12. Figure 2.12 shows the measurement locations where each spacing schedules samples its measurements. Black curves indicate the curve that leads to the highest model performance for each spacing schedule and gray curves indicate the curves that lead to suboptimal model performance. In looking at the black curves (i.e., curves that lead to the highest model performance) for each spacing schedule, more measurements lie closer to the period of greatest change on the black curves than on

Figure 2.13

Rate of Change (First Derivative Curve) for Each Nature of Change Curve Manipulated in Experiment 1



1360 *Note.* Panel A: Logistic curve defined by β_{fixed} = 80, with first-derivative curve peaking at day 80. Panel B:
1361 Logistic curve defined by β_{fixed} = 180, with first-derivative curve peaking at day 180. Panel C: Logistic curve
1362 defined by β_{fixed} = 280, with first-derivative curve peaking at day 280.

the respective gray curves the gray curves that result in lower model performance. One clear example can be observed for the measurement locations under middle-and-extreme spacing (see Figure 2.12D). In looking across the nature-of-change curves, only the measurement locations of the middle three measurements on each curve are different. For the optimal black nature of change, the middle three measurements are centered on the period of greatest change. For the gray suboptimal nature-of-change curves, the middle three measurements are taken near regions of little change (near-zero first derivative value). Therefore, model performance is highest when spacing schedules place measurement near periods of greatest change.

Second, model performance under time-interval increasing and decreasing spacing is nearly identical because each spacing schedule samples data at the exact same regions of change. In looking at Table 2.15, it is important to realize that the precision (i.e., error bar lengths) obtained with time-interval increasing and decreasing spacing are nearly identical when model performance is highest. As an example of the precision obtained when model performance is highest, the average error bar length obtained for the fixed-effect days-to-halfway elevation parameter (β_{fixed}) is 5.80 days with time-interval increasing spacing and 5.84 days with time-interval decreasing spacing. The nearly equivalent precision obtained with time-interval increasing and decreasing spacing occurs because the rates of change (i.e., first derivative values) at the sampled locations are the exact same. As an example with five measurements, Table 2.16 lists the curve values and measurement days when the time-interval increasing and decreasing spacing schedules sample the each of five unique first-derivative values. Note that the time-interval increasing and decreasing spacing schedules sample the first-derivative values in opposite orders. In summary, although the time-interval increasing and decreasing spacing schedules sample data on different days on their respective optimal curves, they result in (nearly) identical model performance because they place measurements at the same periods of change.

Table 2.16
Identical First-Derivative Sampling of Time-Interval Increasing and Decreasing Spacing Schedules

	Time-Interval Increasing		Time-Interval Decreasing	
First Derivative Value	Curve Value	Measurement Day	Curve Value	Measurement Day
2.00e-06	3.00	0	3.32	360
8.80e-06	3.00	30	3.32	330
2.83e-04	3.01	100	3.31	260
2.39e-03	3.26	210	3.06	150
2.00e-06	3.32	360	3.00	0

Third, middle-and-extreme spacing obtains higher model performance than equal spacing by sampling data at periods of greater change. Importantly, both equal and middle-and-extreme spacing obtain their highest model performance with a with a nature-of-change value of 180 ($\beta_{fixed} = 180$), with middle-and-extreme spacing obtaining higher precision (i.e., shorter error bars) than equal spacing (see Figure 2.12 and Table 2.15). An inspection of Figures 2.12A and 2.12D reveals that middle-and-extreme spacing samples measurements at moments of greater change. As an example, consider the measurement locations of equal and middle-and-extreme spacing with five measurements, where only second and fourth measurement locations differ between the schedules. For equal spacing, the second and fourth measurements are respectively sampled on days 90 and 270. For middle-and-extreme spacing, the second and fourth measurements are respectively taken on days 150 and 210. By consulting the first-derivative curve in Figure 2.13, change is greater on days 150 and 210 than on days 90 and 270. Therefore, precision across all manipulated measurement numbers is greater (i.e., shorter error bars) with middle-and-extreme spacing than with equal spacing because middle-and-extreme spacing takes measurements closer to periods of change than equal spacing (see Figures 2.12A and 2.12D and Table 2.15).

The idea that model performance increases when data are sampled during periods of greater change has received considerable discussion and preliminary support. Over the past 20 years, researchers have recommended that measurements be sampled during periods of greater change (Ployhart & Vandenberg, 2010; Siegler, 2006), with one recent simulation study finding evidence to support this idea (Timmons & Preacher, 2015). Unfortunately, the evidence from Timmons and Preacher (2015) is preliminary for two reasons. First, the model used to estimate nonlinear change only ever included one random-effect parameter. Given that multilevel models often include several random-effect parameter in practice, the model employed in Timmons and Preacher (2015) may not necessary be realistic. Second, the estimates were obtained by using an impractical starting value procedure: Population values were used as starting values. Because practitioners never know the population value, it is not known whether the results of Timmons and Preacher (2015) replicate with a realistic starting value procedure.

My simulations in Experiment 1 replicated the finding that model performance increases from measuring change near periods of change under more realistic conditions. In contrast to the one-random-effect-parameter models used in Timmons and Preacher (2015), my simulations used a four-parameter model where each parameter was modelled as a fixed and random effect. For the starting value procedure, my simulations did not use the population values as starting values, but used the starting value procedure available in OpenMx, which uses an unweighted lease squares model to compute starting values.

Therefore, three results in Experiment 1 suggest that sampling data closer to periods of change leads to higher model performance. First, for each spacing schedule, model performance is highest when measurements are taken closer to periods of change. Second, the time-interval increasing and decreasing spacing schedules obtain nearly identical modelling accuracies for different curves because the sampled locations have the exact same rates of change. Third, middle-and-extreme spacing results in higher model performance than equal spacing by sampling measurements at periods of greater change. Although several researchers have posited model performance increases by sampling data closer to periods of change, with one simulation study (to my knowledege) having found support for this idea under unrealistic modelling conditions, my simulations in Experiment 1 support it under realistic modelling conditions.

2.2.7.2 When the Nature of Change is Unknown, How Should Measurements be Spaced?

A second question I had in Experiment 1 was how to space measurements when the nature of change is unknown. To answer this question, I first recorded the number of measurements needed to obtain the greatest improvements in model performance for each spacing schedule in Table 2.17. Text within the 'Qualitative Description' column indicates the number of measurements needed to obtain the largest improvements in bias and precision across all manipulated nature-of-change values for each spacing schedule. The 'Error Bar Summary' columns list the error bar lengths obtained for each day-unit parameter using the measurement number listed in the 'Qualitative Description' column. Note that the error bar lengths in the 'Error Bar Summary' column are obtained by computing the average length across all manipulated nature-of-change values for the measurement number listed Qualitative Description' column. For comprehensiveness, I also recorded the number of measurements needed to obtain unbiased and precise estimation of all the day-unit parameters across all manipulated nature-of-change values in the 'Unbiased' and 'Precise' columns.

The following number of measurements are needed to obtain unbiased estimation and the greatest improvements in bias and precision across all manipulated nature-of-change values for all day-unit parameters under each spacing schedule:

- equal spacing: nine or more measurements to obtain unbiased estimation and seven measurements to obtain the greatest improvements in bias and precision.
- time-interval increasing spacing: nine or more measurements to obtain unbiased estimation and nine measurements to obtain the greatest improvements in bias and precision.
- time-interval decreasing spacing: nine or more measurements to obtain unbiased estimation and nine measurements to obtain the greatest improvements in bias and precision.
- middle-and-extreme spacing: 11 measurements to obtain unbiased estimation and nine measurements to obtain the greatest improvements in bias and precision.

Table 2.17

Concise Summary of Results Across All Spacing Schedule Levels in Experiment 1

				Error Bar Summary			
Spacing Schedule	Unbiased	Precise	Qualitative Description	β_{fixed}	γ_{fixed}	β_{random}	γ_{random}
Equal (see Figure 2.5 and Table 2.3)	NM $\geq$ 9	No cells	Largest improvements in bias and precision with NM = 7	5.64	4.37	7.74	7.02
Time-interval increasing (see Figure 2.7 and Table 2.6)	NM $\geq$ 9	No cells	Largest improvements in bias and precision with NM = 9	4.97	3.45	6.31	5.97
Time-interval decreasing (see Figure 2.9 and Table 2.9)	NM $\geq$ 9	No cells	Largest improvements in bias and precision with NM = 9	4.88	3.40	6.15	5.96
Middle-and-extreme (see Figure 2.11 and Table 2.9)	NM = 11	No cells	Largest improvements in bias and precision with NM = 9	6.51	5.55	9.02	7.20

Note. Row shaded in gray indicates the spacing schedules that results in the highest modelling accurac across all manipulated nature-of-change curves. 'Qualitative Description' column indicates the number of measurements that obtains the greatest improvements in bias and precision across all day-unit parameters and manipulated nature-of-change values. 'Error Bar Summary' columns list the error bar lengths that result for each day-unit parameter using the measurement number listed in the 'Qualitative Description' column. Note that error bar lengths were calculated by computing the average length across all manipulated measurement numbers for the nature-of-change value listed in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter $\in \{80, 180, 280\}$; γ_{fixed} = fixed-effect triquarter-halfway delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect triquarter-halfway delta parameter = 4. NM = number of measurements.

An important point to mention is that the error bar lengths for each day-unit parameter across each spacing schedule are comparable. That is, each spacing schedule obtains similar model performance when using the number of measurements listed in the 'Qualitative Description' column. Because model performance is similar across the spacing schedules, then the schedule that requires the fewest number of measurements to obtain the greatest improvements in bias and precision can be said to model change most accurately when the nature of change is unknown. With equal spacing using fewer measurements than all the other manipulated spacing schedules to obtain similar model performance—using seven measurements instead of the nine measurements use by all other spacing schedules—equal spacing is the most effective schedule to use when the nature of change is unknown.

The finding that equal spacing results in the highest model performance when the nature of change is unknown is not unexpected. Given the previous finding that model performance increases by sampling data closer to periods of change, then, if the nature of change is unknown, change may occur at any point in time, and so it is prudent to space measurements equally over time so maximize the probability of sample measurements during a period of change.

2.3 Summary of Experiment 1

I designed Experiment 1 to investigate two questions. The first question was whether placing measurements near periods of change increases model performance. For each spacing schedule, model performance was highest when measurements were sampled at periods of greater change. Therefore, when a researcher has some knowledge of the nature of change, measurements should be placed near periods of change to increase model performance.

The second question was how to space measurements when the nature of change is unknown. Although no manipulated measurement number under any spacing schedule resulted in accurate estimation of all parameters, the improvements in model performance began to diminish under each spacing schedule at a specific measurement number and plateaued at similar level of model performance. With each spacing schedule plateauing at a similar level of model performance at a specific measurement number, I concluded that the spacing schedule that used the fewest number of measurements to arrive at this plateau was most the effective schedule to use when the nature of change was unknown. With equal spacing using the fewest number of measurements to obtain the greatest improvements in model performance and reach its plateau, I concluded that equal spacing was the most effective schedule to use when the nature of change was unknown.

3 Experiment 2

In Experiment 2, I investigated the measurement number and sample size combinations needed to obtain high model performance (i.e., unbiased and precise parameter estimation) under different spacing schedules. Before presenting the results of Experiment 2, I present my design and and analysis goals. For my design goals, I conducted a 4 (spacing schedule: equal, time-interval increasing, time-interval decreasing, middle-and-extreme) x 4(number of measurements: 5, 7, 9, 11) x 6(sample size: 30, 50, 100, 200, 500, 1000) study. For my analysis goals, I was interested in determining, for each spacing schedule, the measurement number and sample size combinations needed to obtain high model performance (i.e., unbiased and precise parameter estimation). For parsimony, I present model performance across all manipulated combinations of measurement number and sample size for each spacing schedule.

3.1 Methods

3.1.1 Overview of Data Generation

Data generation was computed the same way as in Experiment 1 (see data generation). Note that the code used to run the simulations and create the data set can be found in Appendix B and the data file (`exp_2_data.csv`) can be found in the following GitHub repository: https://github.com/sciarraseb/dissertation.

3.1.2 Modelling of Each Generated Data Set

Each generated data set was modelled using the structured latent growth curve model outlined in Experiment 1 (see data modelling and explicated in Appendix D.

3.1.3 Variables Used in Simulation Experiment

3.1.3.1 Independent Variables

3.1.3.1.1 Spacing of Measurements

For the spacing of measurements, I used the same measurement days as in Experiment 1 for equal, time-interval increasing, time-interval decreasing, and middle-and-extreme spacing (see spacing of measurements for more discussion).

3.1.3.1.2 Number of Measurements

For the number of measurements, I used the same values as in Experiment 1 of 5, 7, 9, and 11 measurements (see number of measurements} for more discussion).

3.1.3.1.3 Sample Size

Sample size values were adopted from Coulombe et al. (2016) with one difference. Because my experiments investigated the effects of measurement timing factors on the ability to model nonlinear patterns, which are inherently more complex than linear patterns of change, a sample size value of $N = 1000$ was added as the largest sample size. Therefore, the following values were used for my sample size manipulation: 30, 50, 100, 200, 500, and 1000.

3.1.3.2 Constants

Given that each simulation experiment manipulated no more than three independent variables so that results could be readily interpreted (Halford et al., 2005), other variables had to be set to constant values. In Experiment 2, two important variables were set to constant values: nature of change and time structuredness. For nature of change, I set the value for the fixed-effect days-to-halfway elevation parameter (β_{fixed}) across all cells to 180. For time structuredness, data across all cells were generated to be time structured (i.e., all participants provide data according to one response pattern; that is, at each time point, participants provide their data at the exact same moment).

3.1.3.3 Dependent Variables

3.1.3.3.1 Convergence Success Rate

The proportion of iterations in a cell where models converged defined the *convergence success rate*.[15] Equation (4.5) below shows the calculation used to compute the convergence success rate:

$$\text{Convergence success rate} = \frac{\text{Number of models that successfully converged in a cell}}{n}, \tag{3.1}$$

[15]Specifically, convergence was obtained if the convergence code returned by OpenMx was 0.

where n represents the total number of models run in a cell.

3.1.3.3.2 Model Performance

Model performance was the combination of two metrics: bias and precision. More specifically, two questions were of importance in the estimation of a given logistic function parameter: 1) How well was the parameter estimated on average (bias) and 2) what was a range of values that could be expected for an estimate from the output of a single model (precision). In the two sections that follow, I will discuss each metric of model performance and the cutoffs used to determine whether estimation was unbiased and precise.

3.1.3.3.2.1 Bias

Bias was calculated to evaluate the accuracy with which each logistic function parameter was estimated in each experimental cell. As shown below in Equation (3.2), *bias* was obtained by summing the differences between the population value set for a parameter and the value estimated for the parameter by each i converged model and then dividing the sum by the number of N converged models.

$$\text{Bias} = \frac{\sum_i^N (\text{Population value for parameter} - \text{Average estimated value}_i)}{N} \tag{3.2}$$

Bias was calculated for the fixed- and random-effect parameters of the baseline (θ_{fixed}, θ_{random}), maximal elevation (α_{fixed}, α_{random}), days-to-halfway elevation (β_{fixed}, β_{random}), and the triquarter-halfway delta parameters (γ_{fixed}, γ_{random}) and the error parameter (ϵ).

3.1.3.3.2.2 Precision

In addition to computing bias, precision was calculated to evaluate the variability with which each parameter was estimated. Importantly, metrics used to evaluate precision in previous studies aoften ssume estimates are normally distributed (e.g., mean-squared error and empirical standard error). Because some parameters in my simulations had skewed distributions, using a metric that assumed a normal distribution would likely yield inaccurate results. Correspondingly, I used a distribution-independent definition of precision. In my simulations, *precision* was defined as the range of values covered by the middle 95% of values estimated for a logistic parameter.

3.1.4 Analysis of Data Modelling Output and Accompanying Visualizations

Analysis and visualization was conducted as outlined in Experiment 1 (see analysis and visualization).

3.2 Results and Discussion

In the sections that follow, I organize the results by presenting them for each spacing schedule (equal, time-interval increasing, time-interval decreasing, middle-and-extreme). Importantly, only the results for the day-unit parameters will be presented (i.e., fixed- and random-effect days-to-halfway elevation and triquarter-halfway delta parameters [β_{fixed}, β_{random}, γ_{fixed}, γ_{random}, respectively]). The results for the likert-unit parameters (i.e., fixed- and random-effect baseline and maximal elevation parameters [θ_{fixed}, θ_{random}, α_{fixed}, α_{random}, respectively]) were largely trivial and so are presented in Appendix F.

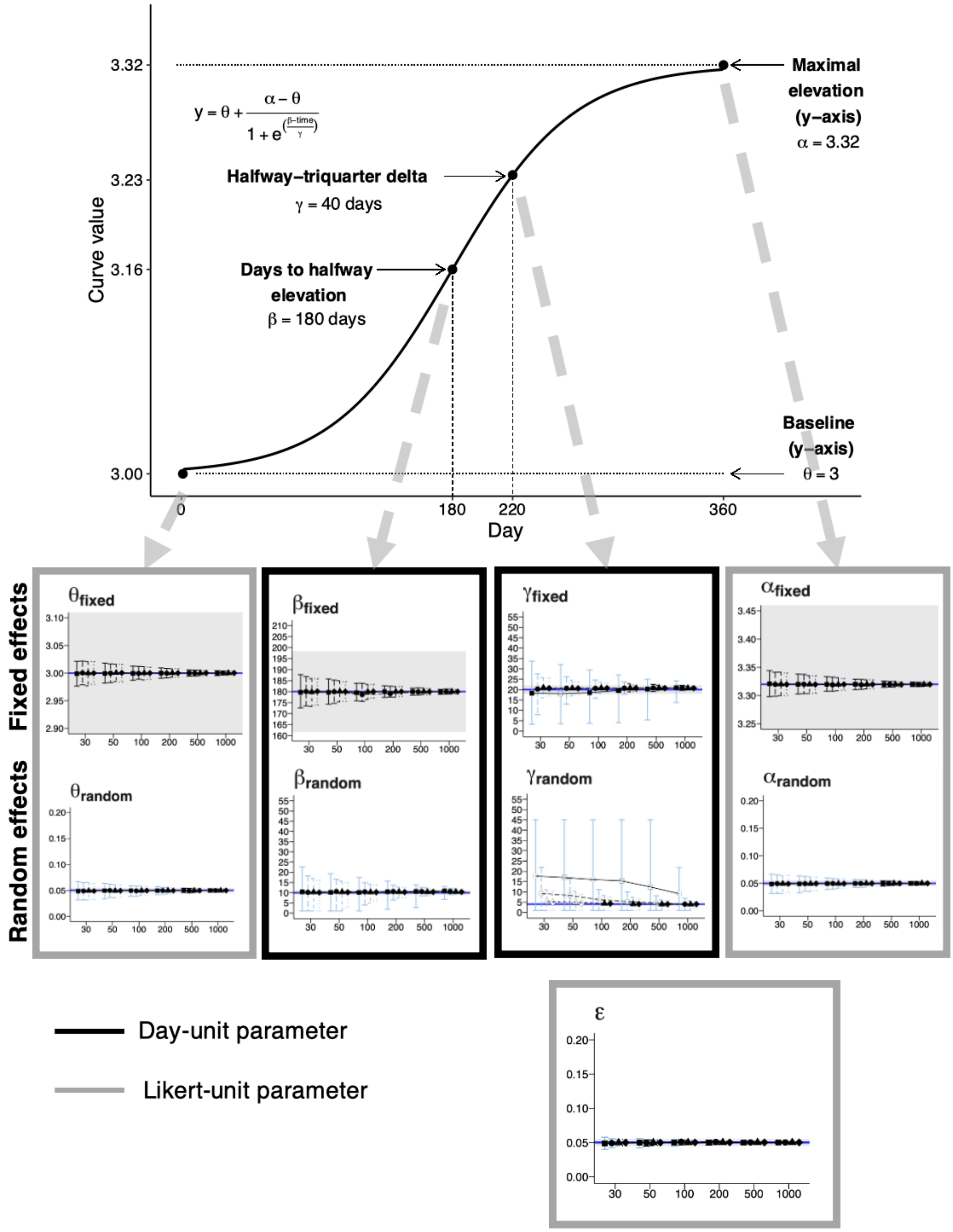
3.2.1 Framework for Interpreting Results

To conduct Experiment 2, the three variables of number of measurements (4 levels), spacing of measurements (4 levels), and sample size (9 levels) were manipulated, which yielded a total of 96 cells. Importantly, within each cell, bias and precision values were also computed for each of the nine parameters estimated by the structured latent growth curve models (for a review, see modelling of each generated data set). Thus, because the analysis of Experiment 2 computed values for many dependent variables, interpreting the results can become overwhelming. Therefore, I will provide a framework to help the reader efficiently navigate the results section.

Because I will present the results of Experiment 2 by each level of measurement spacing, the framework I will describe in Figure 4.3 shows a template for the bias/precision plots that I will present for each spacing schedule. The results of each spacing schedule contain a bias/precision plot for each of the nine estimated parameters. Each bias/precision plot shows the bias and precision for the estimation of one parameter across all measurement number and sample size levels. Within each bias/precision plot, dots indicate the average estimated value (which indicates bias) and error bars represent the middle 95% range of estimated values (which indicates precision). Bias/precision plots with black borders show the results for day-unit parameters and plots with gray border show the results for Likert-unit parameters. Importantly, only the results for the day-unit parameters will be presented (i.e., fixed- and random-effect days-to-halfway elevation and triquarter-halfway delta parameters [β_{fixed}, β_{random}, γ_{fixed}, γ_{random}, respectively]). The results for the Likert-unit parameters (i.e., fixed- and random-effect baseline and maximal elevation parameters [θ_{fixed}, θ_{random}, α_{fixed}, α_{random}, respectively]) were largely trivial and so are presented in Appendix F. Therefore, the results of each spacing schedule will only present the bias/precision plots for four parameters (i.e., the day-unit parameters).

Figure 3.1

Set of Bias/Precision Plots Constructed for Each Spacing Schedule in Experiment 2



Note. A bias/precision plot is constructed for each parameter of the logistic function (see Equation 2.1). Bias/precision plots with black borders show the results for day-unit parameters and plots with gray border show the results for Likert-unit parameters. For each parameter, bias and precision are shown across each combination of measurement number and sample size.

3.2.2 Pre-Processing of Data and Model Convergence

After collecting the output from the simulations, non-converged models (and their corresponding parameter estimates) were removed from subsequent analyses. Table G.2 in Appendix G provides the convergence success rates for each cell in Experiment 2. Model convergence was almost always above 90% and convergence rates, with rates only going below 90% in two cells (or instances) with five measurements.

3.2.3 Equal Spacing

For equal spacing, Table 3.1 provides a concise summary of the results for the day-unit parameters (see Figure 3.2 for the corresponding bias/precision plots). The sections that follow will present the results for each column of Table 3.1 and provide elaboration when necessary.

Before presenting the results for equal spacing, I provide a brief description of the concise summary table created for each spacing schedule and shown for equal spacing below in Table 3.1. Text in the 'Unbiased' and 'Precise' columns indicates the measurement number/sample size pairings that, respectively, resulted in unbiased and precise estimation. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the measurement number/sample size pairing that, respectively, resulted in unbiased estimation and the greatest improvements in bias and precision across all day-unit parameters (note that acceptable precision was not obtained in the estimation of all day-unit parameters with equal spacing). The 'Error Bar Length' column indicates the error bar length that resulted from using the lower-bounding measurement number/sample size pairing listed in the 'Qualitative Description' column (i.e., the maximum error bar

Table 3.1

Concise Summary of Results for Equal Spacing in Experiment 2

			Description	
Parameter	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 3.2A)	All cells	All cells	Unbiased and precise estimation in all cells	15.13
γ_{fixed} (Figure 3.2B)	All cells	NM $\geq$ 9 with $N \geq 500$	Largest improvements in precision using **NM = 7 with *N* $\geq$ 200** or **NM = 9 with *N* $\leq$ 100**	9.79
β_{random} (Figure 3.2C)	All cells	No cells	Largest improvements in precision with NM = 7	17.22
γ_{random} (Figure 3.2D)	**NM $\geq$ 7 with *N* = 1000** or **NM $\geq$ 9 with *N* $\geq$ 200** or **NM = 11 with *N* = 100**	No cells	Largest improvements in bias and precision using NM = 7 with *N* $\geq$ 100 or NM = 9 with *N* $\leq$ 50	10.08

Note. Text in the ‘Unbiased’ and ‘Precise’ columns indicates the measurement number/sample size pairings that, respectively, resulted in unbiased and precise estimation. Emboldened text in the ‘Unbiased’ and ‘Qualitative Description’ columns indicates the measurement number/sample size pairings that, respectively, resulted in unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision was not obtained in the estimation of all day-unit parameters with equal spacing). ‘Error Bar Length’ column indicates the maximum error bar length that resulted from using the measurement number/sample size recommendation listed in the ‘Qualitative Description’ column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect triquarter-halfway delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect triquarter-halfway delta parameter = 4. NM = number of measurements.

length).

3.2.3.1 Bias

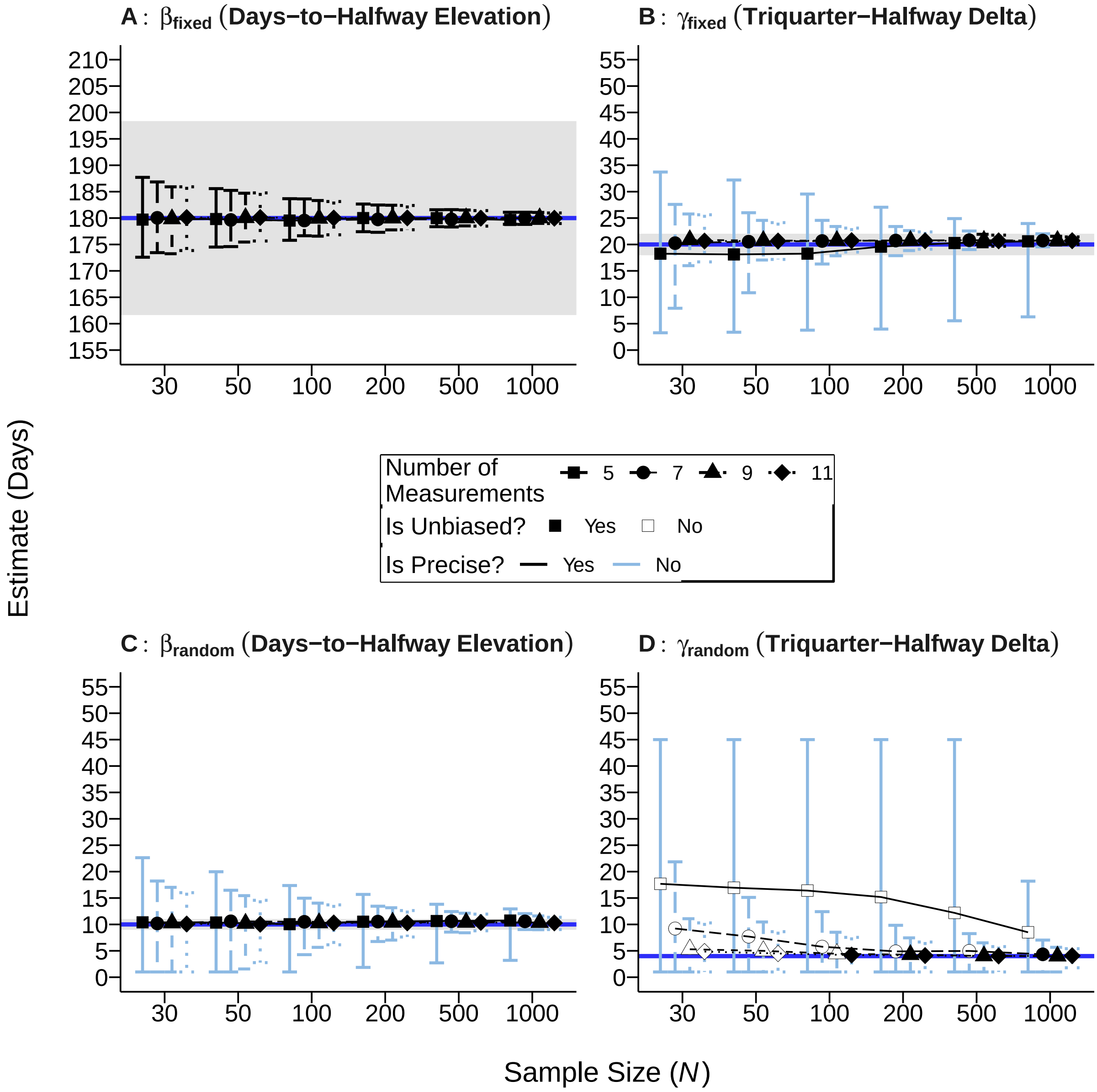
Before presenting the results for bias, I provide a description of the set of bias/precision plots shown in Figure 3.2 and in the results sections for the other spacing schedules in Experiment 2. Figure 3.2 shows the bias/precision plots for each day-unit parameter and Table 3.2 provides the partial ω^2 values for each independent variable of each day-unit parameter. In Figure 3.2, blue horizontal lines indicate the population values for each parameter (with population values of $\beta_{fixed} = 180.00$, $\beta_{random} = 10.00$, $\gamma_{fixed} = 20.00$, and $\gamma_{random} = 4.00$). Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Panels A–B show the bias/precision plots for the fixed- and random-effect days-to-halfway elevation parameters (β_{fixed} and β_{random}, respectively). Panels C–D show the bias/precision plots for the fixed- and random-effect triquarter-halfway delta parameters (γ_{fixed} and γ_{random}, respectively). Note that random-effect parameter units are in standard deviation units.

With respect to bias for equal spacing, estimates were biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.2A): no cells.

Figure 3.2

Bias/Precision Plots for Day-Unit Parameters With Equal Spacing in Experiment 2



Note. Panel A: Bias/precision plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Bias/precision plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Bias/precision plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Bias/precision plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect

parameter units are in standard deviation units. See Table H.2 for specific values estimated for each parameter and Table 3.2 for ω^2 effect size values.

Table 3.2
Partial ω^2 Values for Independent Variables With Equal Spacing in Experiment 2

	Effect		
Parameter	NM	S	NM x S
β_{fixed} (Figure 3.2A)	0.00	0.03	0.00
β_{random} (Figure 3.2B)	0.15	0.28	0.03
γ_{fixed} (Figure 3.2C)	0.31	0.15	0.09
γ_{random} (Figure 3.2D)	0.18	0.03	0.01

Note .NM = number of measurements (5, 7, 9, 11), S = sample size (30, 50, 100, 200, 500, 1000), NM x S = interaction between number of measurements and sample size, β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect triquarter-halfway delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect triquarter-halfway delta parameter.

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 3.2B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.2C): no cells.
- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 3.2D): five measurements across all sample sizes, seven measurements with $N \leq 500$, nine measurements with $N \leq 100$, and 11 measurements with $N \leq 50$.

In summary, with equal spacing, estimation of all the day-unit parameters across all manipulated nature-of-change values was unbiased using nine measurements with $N \geq 200$ or 11 measurements with $N = 100$, which is indicated by the emboldened text in the 'Unbiased' column of Table 3.1.

3.2.3.2 Precision

With respect to precision for equal spacing, estimates were imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.2A): all cells.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 3.2B): five and seven measurements across all sample sizes and nine and 11 measurements with $N \leq 200$.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.2C): all cells.
- random-effect triquarter-halfway delta parameter [γ_{random}] in Figure 3.2D): all cells.

In summary, with equal spacing, precise estimation was obtained for the fixed-effect day-unit parameters using at least nine measurements with $N \geq 500$, but no manipulated measurement number/sample size pairing resulted in precise estimation of the random-effect day-unit parameters (see the 'Precise' column of Table 3.1).

3.2.3.3 Qualitative Description

For equal spacing in Figure 3.2, although no manipulated measurement number resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) resulted from using moderate measurement number/sample size pairings. With respect to bias under equal spacing, the largest improvements resulted with the following measurement number/sample size pairings for the fixed- and random-effect triquarter-halfway delta parameters (γ_{fixed} and γ_{random}, respectively):

- fixed-effect triquarter-halfway delta parameters (γ_{fixed}): seven measurements with $N = 30$.
- random-effect triquarter-halfway delta parameters (γ_{random}): seven measurements

with $N \geq 100$ or nine measurements with $N \leq 50$.

With respect to precision under equal spacing, the largest improvements in the estimation of all day-unit parameters (except the fixed-effect days-to-halfway elevation parameter [β_{fixed}]) resulted from using the following measurement number/sample size pairings:

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which resulted in a maximum error bar length of 9.79 days.
- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements across all manipulated sample sizes, which which resulted in a error bar length of 17.22 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements with $N \geq 100$ or nine measurements with $N \leq 50$, which resulted in a maximum error bar length of 10.08 days.

For an applied researcher, one plausible question might be what measurement number/sample size pairing(s) results in the greatest improvements in bias and precision in the estimation of all day-unit parameters. In looking across the measurement number/sample size pairings in the above lists, it becomes apparent that the greatest improvements in bias and precision in the estimation of all day-unit parameters resulted from using seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see the emboldened text in the 'Qualitative Description' column of Table 3.1).

3.2.3.4 Summary of Results With Equal Spacing

In summarizing the results for equal spacing, estimation of all day-unit parameters was unbiased using nine measurements with $N \geq 200$ or 11 measurements with $N = 1000$

(see the emboldened text in in the 'Unbiased' column of Table 3.1). Precise estimation was never obtained in the estimation of all day-unit parameters with any manipulated measurement number/sample size pairing (see precision). Although it may be discouraging that no manipulated measurement number/sample size pairing under equal spacing resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters resulted with moderate measurement number/sample size pairings. With equal spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters resulted from using seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see the emboldened text in the 'Qualitative Description' column of Table 3.1).

3.2.4 Time-Interval Increasing Spacing

For time-interval increasing spacing, Table 3.3 provides a concise summary of the results for the day-unit parameters (see Figure 3.3 for the corresponding bias/precision plots). The sections that follow will present the results for each column of Table 3.3 and provide elaboration when necessary (for a description of Table 3.3, see concise summary table).

3.2.4.1 Bias

With respect to bias for time-interval increasing spacing, estimates were biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.3A): no cells.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 3.3B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.3C): NM = 5 with $N = 30$.

Table 3.3

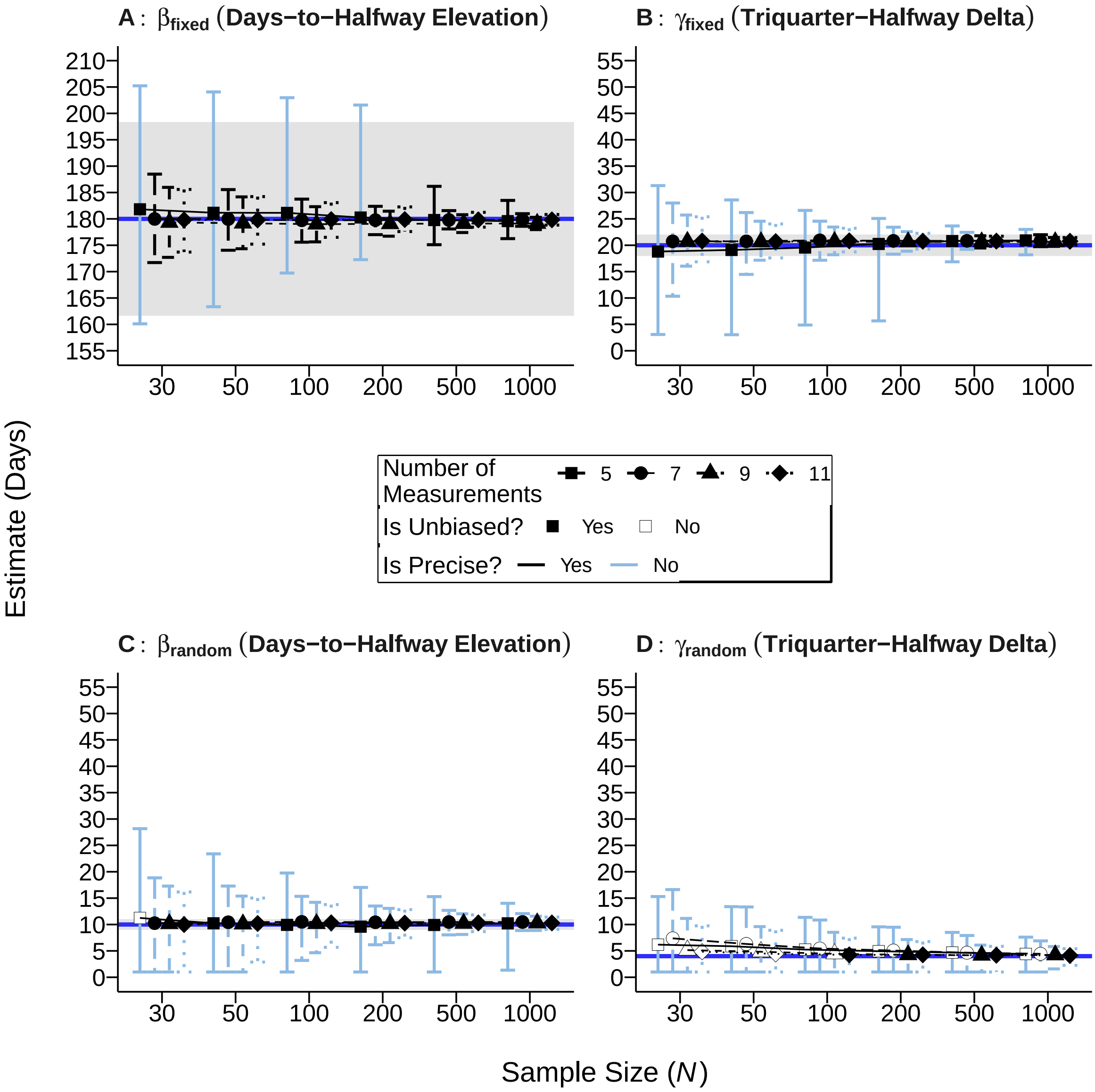
Concise Summary of Results for Time-Interval Increasing Spacing in Experiment 2

			Description	
Parameter	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 3.3A)	All cells	All cells except NM = 5 with $N \leq 200$	Largest improvements in precision using NM = 7 across all sample sizes	16.77
γ_{fixed} (Figure 3.3B)	All cells	NM $\geq$ 7 with N = 1000 or NM $\geq$ 9 with N = 1000	Largest improvements in precision using **NM = 7 with $N \geq 200$** or NM = 9 with $N \leq 100$	9.69
β_{random} (Figure 3.3C)	All cells except	No cells	Largest improvements in precision using NM = 7 across all sample sizes	17.85
γ_{random} (Figure 3.3D)	**NM $\geq$ 9 with $N \geq 200$** or **NM = 11 with N = 1000**	No cells	Largest improvements in bias and precision using **NM = 5 with $N \geq 500$** or **NM = 9 with $N \leq 200$**	10.15

Note. Text in the 'Unbiased' and 'Precise' columns indicates the measurement number/sample size pairings that, respectively, resulted in unbiased and precise estimation. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the measurement number/sample size pairings that, respectively, resulted in unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision was not obtained in the estimation of all day-unit parameters with time-interval increasing spacing). 'Error Bar Length' column indicates the maximum error bar length that resulted from using the measurement number/sample size recommendation listed in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect triquarter-halfway delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect triquarter-halfway delta parameter = 4. NM = number of measurements.

Figure 3.3

Bias/Precision Plots for Day-Unit Parameters With Time-Interval Increasing Spacing in Experiment 2



Note. Panel A: Bias/precision plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Bias/precision plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Bias/precision plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Bias/precision plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect

parameter units are in standard deviation units. See Table H.2 for specific values estimated for each parameter and Table 3.4 for ω^2 effect size values.

Table 3.4

Partial ω^2 Values for Independent Variables With Time-Interval Increasing Spacing in Experiment 2

	Effect		
Parameter	NM	S	NM x S
β_{fixed} (Figure 3.3A)	0.23	0.15	0.09
β_{random} (Figure 3.3B)	0.15	0.16	0.02
γ_{fixed} (Figure 3.3C)	0.17	0.16	0.07
γ_{random} (Figure 3.3D)	0.07	0.12	0.01

Note. NM = number of measurements (5, 7, 9, 11), S = sample size (30, 50, 100, 200, 500, 1000), NM x S = interaction between number of measurements and sample size, β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect triquarter-halfway delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect triquarter-halfway delta parameter.

- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 3.3D): five and seven measurements across all sample sizes, nine measurements with $N \leq 100$, and 11 measurements with $N \leq 50$.

In summary, with time-interval increasing spacing, estimation of all the day-unit parameters was unbiased using nine measurements with $N \geq 200$ or 11 measurements with $N = 100$, which is indicated by the emboldened text in the 'Unbiased' column of Table 3.3.

3.2.4.2 Precision

With respect to precision for time-interval increasing spacing, estimates were imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.3A): five measurements with $N \leq 100$.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 3.3B): five measurements across all sample sizes, seven measurements with $N \leq 500$, nine and 11 measurements with $N \leq 200$.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.3C): all cells.
- random-effect triquarter-halfway delta parameter [γ_{random}] in Figure 3.3D): all cells.

In summary, with time-interval increasing spacing, precise estimation for the fixed-effect day-unit parameters resulted from using at least nine measurements with $N \geq 500$, but no manipulated measurement number/sample size pairing resulted in precise estimation of the random-effect day-unit parameters (see the 'Precise' column of Table 3.3).

3.2.4.3 Qualitative Description

For time-interval increasing spacing in Figure 3.3, although no manipulated measurement number/sample size pairing resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) resulted from using moderate measurement number/sample size pairings. With respect to bias under time-interval increasing spacing, the largest improvements resulted with the following measurement number/sample size pairings for random-effect triquarter-halfway delta parameter (γ_{random}):

- random-effect triquarter-halfway delta parameters (γ_{random}): five measurements with $N \geq 100$ or nine measurements with $N \leq 50$.

With respect to precision under time-interval increasing spacing, the largest improvements in the estimation of each day-unit parameter resulted from using the following measurement number/sample size pairings:

- days-to-halfway elevation parameter (β_{fixed}): seven measurements with $N \geq 30$, which resulted in a maximum error bar length of 9.69 days.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which resulted in a maximum error bar length of 9.69 days.
- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements across all manipulated sample sizes, which resulted in a maximum error bar length of 17.85 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): five measurements with $N \geq 500$ or nine measurements with $N \leq 200$, which resulted in a maximum error bar length of 10.15 days.

For an applied researcher, one plausible question might be what measurement number/sample size pairing(s) results in the greatest improvements in bias and precision in the estimation of all day-unit parameters when using time-interval increasing spacing. In looking across the measurement number/sample size pairings in the above lists, it becomes apparent that the greatest improvements in bias and precision in the estimation of all day-unit parameters with time-interval increasing spacing resulted from using five measurements with $N \geq 500$ or nine measurements with $N \leq 200$ (see the emboldened text in the ‘Qualitative Description’ column of Table 3.3).

3.2.4.4 Summary of Results With Time-Interval Increasing Spacing

In summarizing the results for time-interval increasing spacing, estimation of all day-unit parameters was unbiased using nine measurements with $N \geq 200$ or 11 measurements with $N = 100$ (see bias). Precise estimation was never obtained in the estimation of all day-unit parameters with any manipulated measurement number/sample size pairing (see precision). Although it may be discouraging that no manipulated measurement number/sample size pairing under time-interval increasing spacing resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) across all the day-unit parameters were obtained with moderate measurement number/sample size pairings. With time-interval increasing spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters resulted from using five measurements with $N \geq 500$ or nine measurements with $N \leq 200$ (see qualitative description).

3.2.5 Time-Interval Decreasing Spacing

For time-interval decreasing spacing, Table 3.5 provides a concise summary of the results for the day-unit parameters (see Figure 3.4 for the corresponding bias/precision plots). The sections that follow will present the results for each column of Table 3.5 and provide elaboration when necessary (for a description of Table 3.5, see concise summary table).

3.2.5.1 Bias

With respect to bias for time-interval decreasing spacing, estimates were biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

Table 3.5

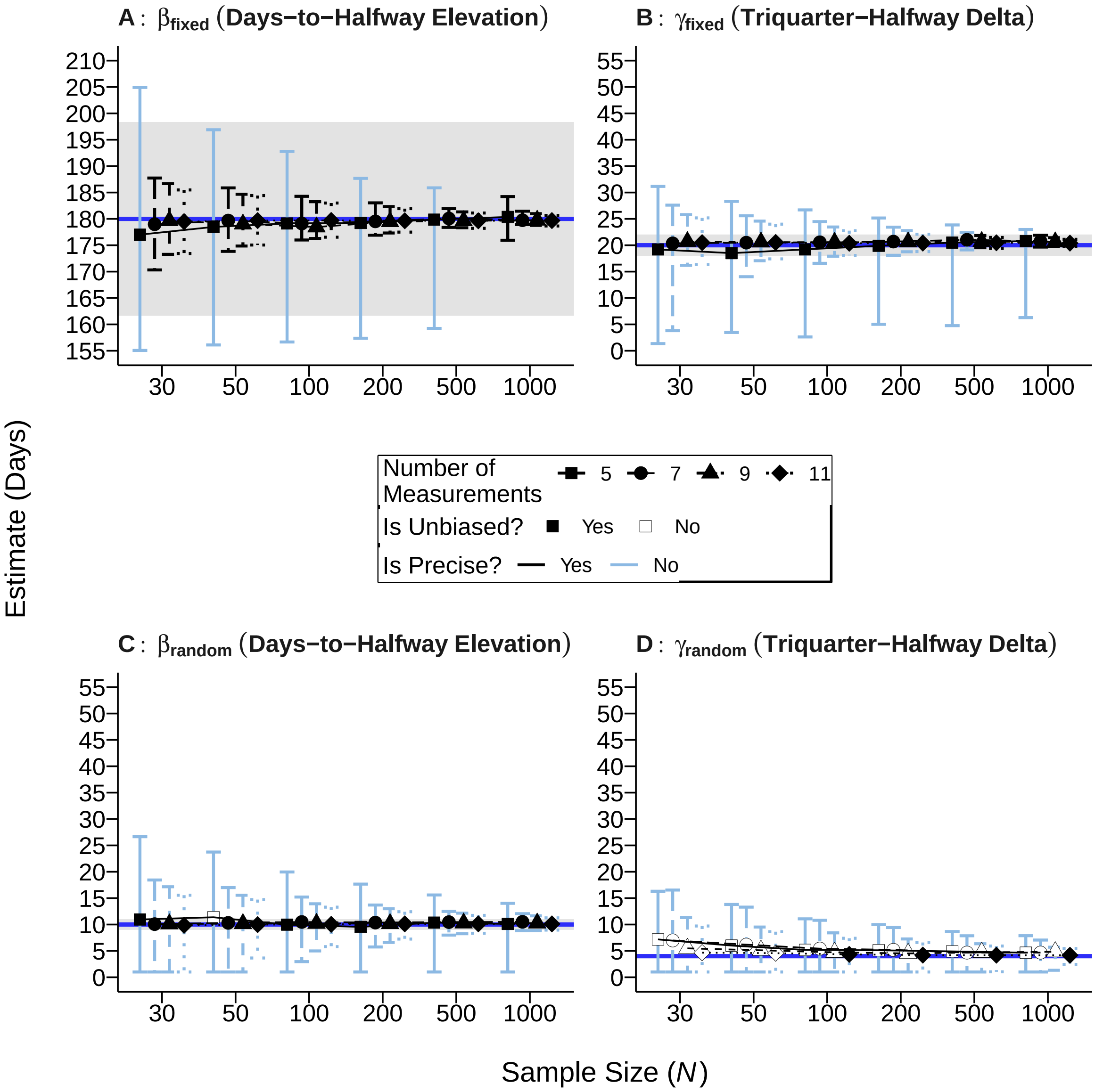
Concise Summary of Results for Time-Interval Decreasing Spacing in Experiment 2

			Description	
Parameter	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 3.4A)	All cells	All cells except NM = 5 with $N \leq 500$	Largest improvements in precision using NM = 7 across all sample sizes	17.42
γ_{fixed} (Figure 3.4B)	All cells	NM = 7 with N = 1000 or NM $\geq$ 9 with $N \geq 500$	Largest improvements in precision using **NM = 7 with $N \geq$ 200** or NM = 9 with $N \leq 100$	9.62
β_{random} (Figure 3.4C)	All cells except NM = 5 with N = 50	No cells	Largest improvements in precision using NM = 7 across all sample sizes	17.44
γ_{random} (Figure 3.4D)	**NM = 11 with $N \geq$ 100**	No cells	Largest improvements in bias and precision using **NM = 5 with $N \geq$ 500** or **NM = 9 with $N \leq$ 200**	10.32

Note. Text in the ‘Unbiased’ and ‘Precise’ columns indicates the measurement number/sample size pairings that, respectively, resulted in unbiased and precise estimation. Emboldened text in the ‘Unbiased’ and ‘Qualitative Description’ columns indicates the measurement number/sample size pairings that, respectively, resulted in unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision was not obtained in the estimation of all day-unit parameters with time-interval decreasing spacing). ‘Error Bar Length’ column indicates the maximum error bar length that resulted from using the measurement number/sample size recommendation listed in the ‘Qualitative Description’ column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect triquarter-halfway delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect triquarter-halfway delta parameter = 4. NM = number of measurements.

Figure 3.4

Bias/Precision Plots for Day-Unit Parameters With Time-Interval Decreasing Spacing in Experiment 2



1843 *Note.* Panel A: Bias/precision plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B:
1844 Bias/precision plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C:
1845 Bias/precision plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D:
1846 Bias/precision plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal
1847 lines in each panel represent the population value for each parameter. Population values for each day-unit
1848 parameter are as follows: β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands
1849 indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter
1850 estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated
1851 values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray
1852 bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or
1853 longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect

parameter units are in standard deviation units. See Table H.2 for specific values estimated for each parameter and Table 3.6 for ω^2 effect size values.

Table 3.6
Partial ω^2 Values for Independent Variables With Time-Interval Decreasing Spacing in Experiment 2

	Effect		
Parameter	NM	S	NM x S
β_{fixed} (Figure 3.4A)	0.05	0.03	0.01
β_{random} (Figure 3.4B)	0.14	0.12	0.01
γ_{fixed} (Figure 3.4C)	0.07	0.04	0.01
γ_{random} (Figure 3.4D)	0.05	0.09	0.00

Note. NM = number of measurements (5, 7, 9, 11), S = sample size (30, 50, 100, 200, 500, 1000), NM x S = interaction between number of measurements and sample size, β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect triquarter-halfway delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect triquarter-halfway delta parameter.

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.4A): no cells.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 3.4B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.4C): NM = 5 with $N = 30$.
- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 3.4D): five, seven, and nine measurements across all sample sizes an 11 measurements with $N \leq 50$, and 11 measurements with $N \leq 50$.

In summary, with time-interval decreasing spacing, estimation of all the day-unit parameters was unbiased using 11 measurements with $N \geq 100$, which is indicated by the emboldened text in the ‘Unbiased’ column of Table 3.5.

3.2.5.2 Precision

With respect to precision for time-interval decreasing spacing, estimates were imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter’s population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.4A): five measurements with $N \leq 500$.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 3.4B): five measurements across all sample sizes, seven measurements with $N \leq 500$, and nine and 11 measurements with $N \leq 200$.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.4C): all cells.
- random-effect triquarter-halfway delta parameter [γ_{random}] in Figure 3.4D): all cells.

In summary, with time-interval decreasing spacing, precise estimation for the fixed-effect day-unit parameters resulted from using at least seven measurements with $N = 1000$ or nine measurements $N \leq 500$. For the random-effect day-unit parameters, no manipulated measurement number/sample size pairing resulted in precise estimation (see the ‘Precise’ column of Table 3.5).

3.2.5.3 Qualitative Description

For time-interval decreasing spacing in Figure 3.4, although no manipulated measurement number resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) resulted from using moderate measurement number/sample size pairings. With respect to bias under time-interval decreasing spacing, the largest improvements resulted with the following measurement number/sample size pairings for the random-effect triquarter-halfway delta parameter (γ_{random}):

- random-effect triquarter-halfway delta parameters (γ_{random}): five measurements with $N \geq 100$ or nine measurements with $N \leq 50$.

With respect to precision under time-interval decreasing spacing, the largest improvements in the estimation of all day-unit parameters (except the fixed-effect days-to-halfway elevation parameter [β_{fixed}]) resulted from using the following measurement number/sample size pairings:

- days-to-halfway elevation parameter (β_{fixed}): seven measurements with $N \geq 30$, which resulted in a maximum error bar length of 9.62 days.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which resulted in a maximum error bar length of 9.62 days.
- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements across all manipulated sample sizes, which which resulted in a error bar length of 17.44 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): five measurements with $N \geq 500$ or nine measurements with $N \leq 200$, which resulted in a maximum error bar length of 10.32 days.

For an applied researcher, one plausible question might be what measurement number/sample size pairing(s) results in the greatest improvements in bias and precision in the estimation of all day-unit parameters with time-interval decreasing spacing. In looking across the measurement number/sample size pairings in the above lists, it becomes apparent that greatest improvements in bias and precision in the estimation of all day-unit parameters with time-interval decreasing spacing resulted with five measurements with $N \geq 500$, seven measurements with $N \geq 200$, or nine measurements with $N \leq 200$ (see the emboldened text in the 'Qualitative Description' column of Table 3.5).

3.2.5.4 Summary of Results Time-Interval Decreasing Spacing

In summarizing the results for time-interval decreasing spacing, estimation of all day-unit parameters was unbiased using 11 measurements with $N \geq 10$ (see bias). Precise estimation was never obtained in the estimation of all day-unit parameters with any manipulated measurement number/sample size pairing (see precision). Although it may be discouraging that no manipulated measurement number/sample size pairing under time-interval decreasing spacing resulted in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters were obtained with moderate measurement number/sample size pairings. With time-interval decreasing spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters resulted from using five measurements with $N \geq 500$ or nine measurements with $N \leq 200$ (see qualitative description).

3.2.6 Middle-and-Extreme Spacing

For middle-and-extreme spacing, Table 3.7 provides a concise summary of the results for the day-unit parameters (see Figure 3.5 for the corresponding bias/precision plots). The sections that follow will present the results for each column of Table 3.7 and provide elaboration when necessary (for a description of Table 3.7, see concise summary table).

Table 3.7

Concise Summary of Results for Middle-and-Extreme Spacing in Experiment 2

			Description	
Parameter	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 3.5A)	All cells	All cells	Largest improvements in precision using using **NM = 5**	14.96
γ_{fixed} (Figure 3.5B)	All cells	All number of measurements with $N \geq 500$	Largest improvements in precision using **NM = 5**	9.92
β_{random} (Figure 3.5C)	All cells	No cells	Largest improvements in precision using **NM = 5**	15.94
γ_{random} (Figure 3.5D)	**NM** $\in \{5, 9\}$ **with** $N \geq$ **100** or **NM** $\in \{7, 11\}$ **with** $N \leq$ **50**	No cells	Largest improvements in precision using **NM = 5**	10.13

Note. Text in the 'Unbiased' and 'Precise' columns indicates the measurement number/sample size pairings that, respectively, resulted in unbiased and precise estimation. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the measurement number/sample size pairings that, respectively, resulted in unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision was not obtained in the estimation of all day-unit parameters with middle-and-extreme spacing). 'Error Bar Length' column indicates the maximum error bar length that resulted from using the measurement number/sample size recommendation listed in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect triquarter-halfway delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect triquarter-halfway delta parameter = 4. NM = number of measurements.

3.2.6.1 Bias

With respect to bias for middle-and-extreme spacing, estimates were biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.4A): no cells.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 3.4B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.4C): no cells.
- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 3.4D): five and nine measurements with $N \leq 100$ and seven an 11 with $N \leq 50$.

In summary, with middle-and-extreme spacing, estimation of all the day-unit parameters was unbiased using five and nine measurements with $N \leq 100$ and seven an 11 with $N \leq 50$, which is indicated by the emboldened text in the 'Unbiased' column of Table 3.7.

3.2.6.2 Precision

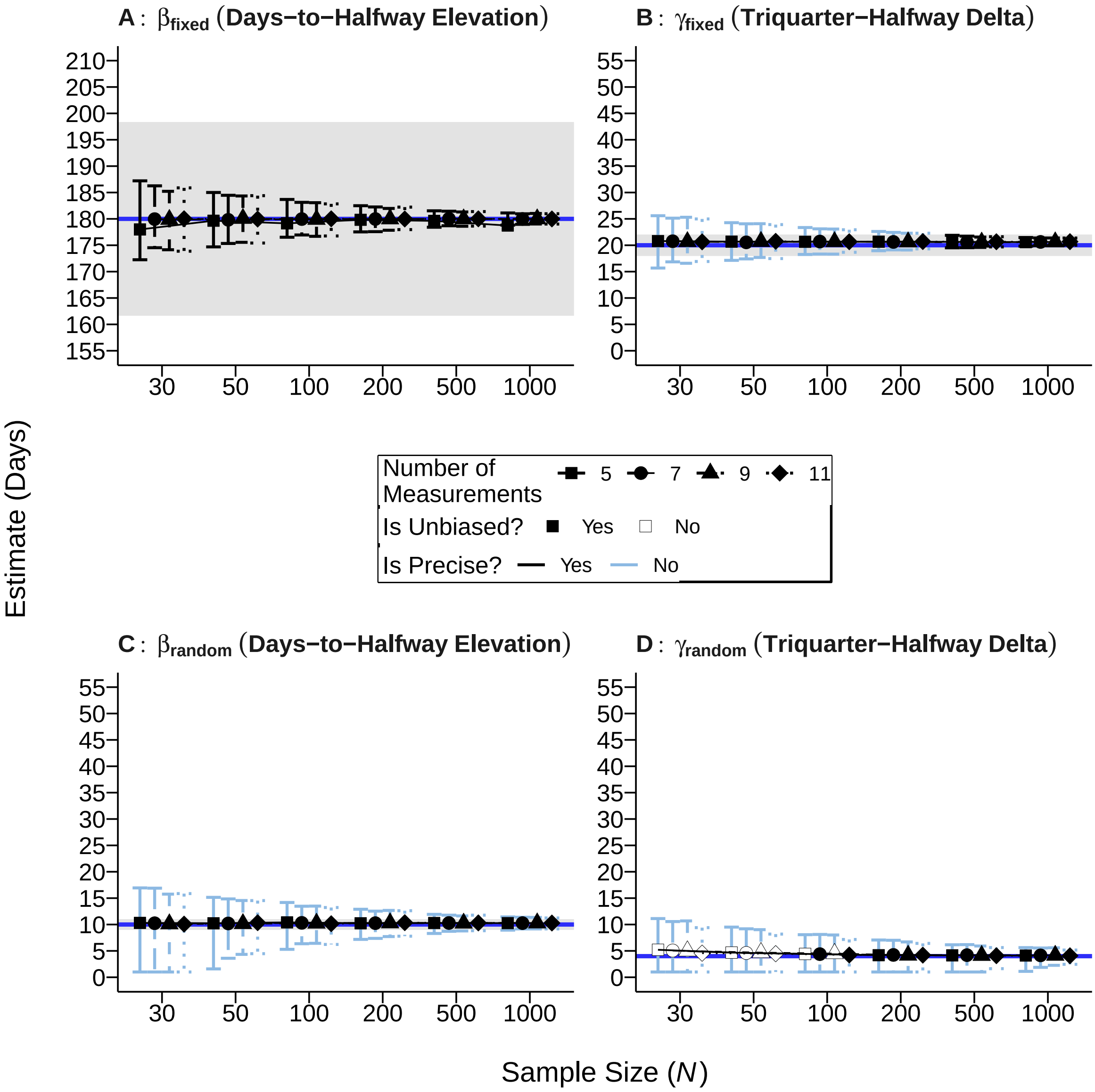
With respect to precision for middle-and-extreme spacing, estimates were imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.4A): no cells.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 3.4B): all measurements numbers with $N \geq 200$.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.4C): all cells.
- random-effect triquarter-halfway delta parameter (γ_{random}; Figure 3.4D): all cells.

In summary, with middle-and-extreme spacing, precise estimation for the fixed-effect day-unit parameters resulted from using at least five measurements with $N \geq 500$.

Figure 3.5

Bias/Precision Plots for Day-Unit Parameters With Middle-and-Extreme Spacing in Experiment 2



Note. Panel A: Bias/precision plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Bias/precision plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Bias/precision plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Bias/precision plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect

parameter units are in standard deviation units. See Table H.2 for specific values estimated for each parameter and Table 3.8 for ω^2 effect size values.

Table 3.8

Partial ω^2 Values for Independent Variables With Middle-and-Extreme Spacing in Experiment 2

	Effect		
Parameter	NM	S	NM x S
β_{fixed} (Figure 3.5A)	0.05	0.03	0.01
β_{random} (Figure 3.5B)	0.14	0.12	0.01
γ_{fixed} (Figure 3.5C)	0.07	0.04	0.01
γ_{random} (Figure 3.5D)	0.05	0.09	0.00

Note. NM = number of measurements (5, 7, 9, 11), S = sample size (30, 50, 100, 200, 500, 1000), NM x S = interaction between number of measurements and sample size, β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect triquarter-halfway delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect triquarter-halfway delta parameter.

For the random-effect day-unit parameters, no manipulated measurement number/sample size pairing resulted in precise estimation (see the 'Precise' column of Table 3.7).

3.2.6.3 Qualitative Description

For middle-and-extreme spacing in Figure 3.5, although no manipulated measurement number resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) resulted from using moderate measurement number/sample size pairings. With respect to bias under middle-and-extreme spacing, it was negligible under all manipulated measurement number/sample size pairings, and so there was little value in listing the pairings that resulted in the greatest improvements. With respect to precision under middle-and-extreme spacing, the largest improvements in the estimation of all day-unit parameters (except the fixed-effect days-to-halfway elevation parameter [β_{fixed}]) resulted from using the following measurement number/sample size pairings:

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): five measurements across all sample sizes, which resulted in a maximum error bar length of 9.92 days.
- random-effect days-to-halfway elevation parameter (β_{random}): five measurements across all sample sizes, which resulted in a maximum error bar length of 15.94 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): five measurements across all sample sizes, which resulted in a maximum error bar length of 10.13 days.

For an applied researcher, one plausible question might be what measurement number/sample size pairing(s) results in the greatest improvements in bias and precision in the estimation of all day-unit parameters with middle-and-extreme spacing. In looking across the measurement number/sample size pairings in the above lists, it becomes apparent that greatest improvements in bias and precision in the estimation of all day-unit parameters with middle-and-extreme spacing resulted from using five measurments with any sample size (see the emboldened text in the 'Qualitative Description' column of Table 3.7).

3.2.6.4 Summary of Results with Middle-and-Extreme Spacing

In summarizing the results for middle-and-extreme spacing, estimation of all day-unit parameters was unbiased using five or nine measurements with $N \leq 100$ and seven or 11 with $N \leq 50$ (see bias). Precise estimation was never obtained in the estimation of all day-unit parameters with any manipulated measurement number/sample size pairing (see precision). Although it may be discouraging that no manipulated measurement number/sample size pairing under time-interval decreasing spacing resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters resulted with moderate measurement number/sample size pairings. With middle-and-extreme spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters resulted from using five measurements with any sample size (see qualitative description).

3.3 What Measurement Number/Sample Size Pairings Should be Used With Each Spacing Schedule?

In Experiment 2, I was interested in determining the measurement number/sample size pairings that resulted in high model performance (unbiased and precise parameter estimation) for each spacing schedule. Table 3.9 summarizes the results for each spacing schedule in Experiment 2. Text within the 'Unbiased' and 'Precise' columns indicates the measurement number/sample size pairing needed to, respectively, obtain unbiased an precise estimation of all the day-unit parameters. The 'Error Bar Length' column indicates maximum error bar length that results in the estimation of each day-unit parameter from using the measurement number/sample size pairings listed in the 'Qualitative Description' column. Although no measurement number/sample size pairing results in high model performance for any spacing schedule, the greatest improvements in model performance result from using the following measurement number/sample size pairings for each spacing schedule (see Table 3.9):

- equal: seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$.

Table 3.9

Concise Summary of Results Across All Spacing Schedule Levels in Experiment 2

				Error Bar Summary			
Spacing Schedule	Unbiased	Precise	Qualitative Description	β_{fixed}	γ_{fixed}	β_{random}	γ_{random}
Equal (see Figure 3.2 and Table 3.1)	NM $\geq$ 7 with N = 1000 or NM $\geq$ 9 with $N \geq$ 100	No cells	Largest improvements in bias and precision using NM = 7 with $N \geq$ 200 or NM = 9 with $N \leq$ 100	12.67	9.79	16.02	10.08
Time-interval increasing (see Figure 3.3 and Table 3.3)	NM $\geq$ 9 with $N \geq$ 200 or NM = 11 with N = 1000	No cells	Largest improvements in bias and precision using NM = 7 with $N \geq$ 200 or NM = 9 with $N \leq$ 100	13.27	9.69	16.28	10.15
Time-interval decreasing (see Figure 3.4 and Table 3.5)	NM = 11 with $N \geq$ 1000	No cells	Largest improvements in bias and precision using NM = 7 with $N \geq$ 200 or NM = 9 with $N \leq$ 100	13.41	9.62	16.16	10.32
Middle and extreme (see Figure 3.5 and Table 3.7)	**NM $\geq$ 5 with $N \geq$ 200** or **NM $\in \{5, 7\}$ with N = 100**	No cells	Largest improvements in bias and precision with NM = 5	14.96	9.92	15.94	10.13

Note. 'Qualitative Description' column indicates the number of measurements that result in the greatest improvements in bias and precision across all day-unit parameters. 'Error Bar Summary' columns list the longest error bar lengths that result for each day-unit parameter using the measurement number/sample size pairing listed in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect triquarter-halfway delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect triquarter-halfway delta parameter = 4. N = sample size, NM = number of measurements.

- time-interval increasing: five measurements with $N \geq 500$, seven measurements with $N \geq 200$ or nine measurements with $N \leq 200$.
- time-interval decreasing: five measurements with $N \geq 500$, seven measurements with $N \geq 200$ or nine measurements with $N \leq 200$.
- middle-and-extreme: five measurements with any manipulated sample size.

Because each spacing schedule obtains comparable model performance as indicated by the similar error bar lengths, two statements can be made. First, using either seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ with any spacing schedule except middle-and-extreme spacing results in similar model performance. Second, given that only five measurements are needed with middle-and-extreme spacing to obtain model performance levels that the other spacing schedules obtained with at least seven measurements, middle-and-extreme spacing results in the highest model performance. Importantly, given that middle-and-extreme spacing results in the highest model performance in Experiment 1 with a midway halfway point (see section discussing measurement spacing), the result here that middle-and-extreme spacing leads to the highest model performance is an expected outcome because the nature-of-change was fixed to 180 (see constants).

The results of Experiment 2 are the first (to my knowledge) to provide measurement number and sample size guidelines for researchers interested in using nonlinear functions to model nonlinear change. Although previous simulation studies have investigated how to accurately model nonlinear change, three characteristics limit these results. First, some studies investigated the issue with fixed-effects models (e.g., Finch, 2017). Given that researchers often model effects as random, findings with fixed-effects effects models are limited in their application. Second, some studies used linear functions to model nonlinear change (e.g., Fine et al., 2019; J. Liu et al., 2022). Given that the parameters of linear functions become uninterpretable when modelling nonlinear change (with the intercept parameter being an exception), these models are less useful to practitioners. Third, some studies implemented unrealistic model fitting procedures by dropping a random-effect parameter from the model each time convergence failed (Finch, 2017). By dropping random-effect parameters when model convergence failed, estimation accuracy could not meaningfully evaluated for parameters because values could have been obtained with reduced or simplified models.

In summary, the results of Experiment 2 provide measurement number/sample size guidelines for researchers interested in modelling nonlinear change. Importantly, because no measurement number-sample pairing results in unbiased and precise estimation of all the day-unit parameters, the guidelines provided by this study are only suggestions to obtain the greatest improvements in model performance. Although researchers are encouraged to use larger measurement numbers and sample sizes than suggested in the current guidelines, the improvements in model performance are likely to be incommensurate with the efforts needed to increase measurement number and sample size.

3.4 Summary of Experiment 2

I designed Experiment 2 to investigate the measurement number and sample size combinations needed to obtain high model performance (i.e., unbiased and precise parameter estimation) under different spacing schedules. Although no measurement number/sample size pairing result in high model performance under any spacing schedule, the greatest improvements in model performance result from using modest measurement number/sample size pairings. Specifically, the greatest improvements in model performance can be obtained using either seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$.

4 Experiment 3

In Experiment 3, I was interested in examining how time structuredness affected model performance. Before presenting the results of Experiment 3, I present my design and analysis goals. For my design goals, I conducted a 3 (time structuredness: time-structured data, time-unstructured data resulting from a fast response rate, time-unstructured data resulting from a slow response rate) x 4 (number of measurements: 5, 7, 9, 11) x 6 (sample size: 30, 50, 100, 200, 500, 1000) study. For my analysis goals, I examined whether the number of measurements and sample sizes needed to obtain high model performance (i.e., low bias, high precision) increased as time structuredness decreased.

4.1 Methods

4.1.1 Variables Used in Simulation Experiment

4.1.1.1 Independent Variables

4.1.1.1.1 Number of Measurements

For the number of measurements, I used the same values as in Experiment 1 of 5, 7, 9, and 11 measurements (see number of measurements for more discussion).

4.1.1.1.2 Sample Size

For sample size, I used the same values as in Experiment 2 of 30, 50, 100, 200, 500, and 1000 (see sample size for more discussion).

4.1.1.1.3 Time Structuredness

Time structuredness describes the extent to which participants provide data over time with the same response pattern. That is, at each time point, do participants provide their data at the exact same time point. If one response pattern characterizes the way in which participants provide their data, then participants always provide data at the exact same moment, and the resulting data are *time structured*. If response patterns differ between participants, the resulting data lose their time structuredness and become *time unstructured*, with the extent of the time unstructuredness depending on the extent to which response patterns differ between participants. The manipulation of time structuredness was adopted from the manipulation used in Coulombe et al. (2016) with a slight modification. Below, I describe the original procedure used in Coulombe et al. (2016) and, following this explanation, I describe my improved procedure.

In Coulombe et al. (2016), time-unstructured data were generated according to an exponential pattern such that most data were obtained at the beginning of the response window, with a smaller amount of data being obtained towards the end of the response window. Importantly, Coulombe et al. (2016) employed a non-continuous function for generating time-unstructured data: A binning method was employed such that 80% of the data were obtained within a time period equivalent to 12% (fast response rate) or 30% (slow response rate) of the entire response window. Using a response window length of 10 days with a fast response rate, the procedure employed by Coulombe et al. (2016) for generating time-unstructured data would have generated the following percentages of data in each of the four bins (note that, using the data generation procedure for Coulombe et al. (2016), the effective response window length for a fast response rate would be 4 days in the current example instead of 10 days):[16]

1) Bin 1: 60% of the data would be generated in the initial 10% length of the response window (0–0.40 day).
2) Bin 2: 20% of the data would be generated in the next 20% length of the response response window (0.40–1.20 days).
3) Bin 3: 10% of the data would be generated in the next 30% length of the response window (1.20–2.40 days).
4) Bin 4: the remaining 10% of the data would be generated in the remaining 40% length of the response window (2.40–4.00 days).

Note that, summing the data percentages and time durations from the first two bins yields an 80% cumulative response rate that is obtained in the initial 12% length of the full-length response window of 10 days (i.e., $(\frac{1.2}{10})100\% = 12\%$). Also note that, in Coulombe et al. (2016), a data point in each bin was randomly assigned a measurement time within the bin's time range. In the current example where the full-length response window had a length of 10 days, a data point obtained in the first bin would be randomly assigned a measurement time between 0–0.40. Although Coulombe et al. (2016) generated time-unstructured data to resemble data collection conditions—response rates have been shown to follow an exponential pattern (Dillman et al., 2014; Pan, 2010)—the use of a pseudo-continuous binning function for generating time-unstructured data lacked ecological validity because response patterns are more likely to follow a continuous function.

[16]The data generation procedure in Coulombe et al. (2016) for a fast response rate assumed that all of the data were collected within the initial 40% length of the nominal response window length (i.e., 4 days in the current example).

To improve on the time structuredness manipulation of Coulombe et al. (2016), I developed a more ecologically valid manipulation by using a continuous function. Specifically, I used the exponential function shown below in Equation 4.1 to generate time-unstructured data:

$$y = M(1 - e^{-ax}), \tag{4.1}$$

where x stores the time delay for a measurement at a particular time point, y represents the cumulative response percentage achieved at a given x time delay, a sets the rate of growth of the cumulative response percentage over time, and M sets the range of possible y values. Two important points need to be made with respect to the M parameter (range of possible y values) and the response window length used in the current simulations. First, because the range of possible values for the cumulative response percentage (y) is 0–1 (data can be collected from a 0% to a maximum of 100% of respondents; $\{y : 0 \leq y \leq 1\}$), the M parameter had a value of 1 (M = 1). Second, the response window length in the current simulations was 36 days, and so the range of possible time delay values was between 0–36 ($\{x : 0 \leq x \leq 36\}$).[17]

To replicate the time structuredness manipulation in Coulombe et al. (2016) using the continuous exponential function of Equation 4.1, the growth rate parameter (a) had to be calibrated to achieve a cumulative response rate of 80% after either 12% or 30% of

[17] A value of 36 days was used because the generation of time-unstructured data had to remain independent of the manipulation of measurement number (i.e., the response window lengths used in generating time-unstructured data could not vary with the number of measurements). To ensure the manipulations of measurement number and time structuredness remained independent, the response window length had to remain constant for all measurement number conditions with equal spacing. Looking at Table 2.2, the longest possible response window that fit within all measurement number conditions with equal spacing was the interval length of the 11-measurement condition (i.e., 36 days).

the response window length of 36 days. The derivation below solves for a, with Equation 4.2 showing the equation for computing a.

$$
\begin{aligned}
y &= M(1 - e^{-ax}) \\
y &= M - Me^{-ax} \\
y &= 1 - e^{-ax} \\
e^{-ax} &= 1 - y \\
-ax \log(e) &= \log(1 - y) \\
a &= \frac{\log(1 - y)}{-x} \qquad (4.2)
\end{aligned}
$$

Because the target response rate was 80%, y took on a value of .80 ($y = .80$). Given that the response window length in the current simulations was 36 days, x took on a value of 4.32 (12% of 36) when time-unstructured data were defined by a fast response rate and 10.80 (30% of 36) when time-unstructured data were defined by a slow response rate. Using Equation 4.2 yielded the following growth rate parameter values for fast and slow response rates (a_{fast}, a_{slow}):

$$
\begin{aligned}
a_{fast} &= \frac{\log(1 - .80)}{-4.32} = 0.37 \\
a_{slow} &= \frac{\log(1 - .80)}{-10.80} = 0.15
\end{aligned}
$$

Therefore, to obtain 80% of the data with a fast response rate (i.e., in 4.32 days), the growth parameter (a) needed to have a value of 0.37 ($a_{fast} = 0.37$) and, to obtain 80% of the data with a slow response rate (i.e., in 10.80 days), the growth parameter (a) needed to have a value of 0.15 ($a_{slow} = 0.15$). Using the above growth rate values derived for the fast and slow response growth rate parameters (a_{fast}, a_{slow}), the following functions were generated for fast and slow response rates:

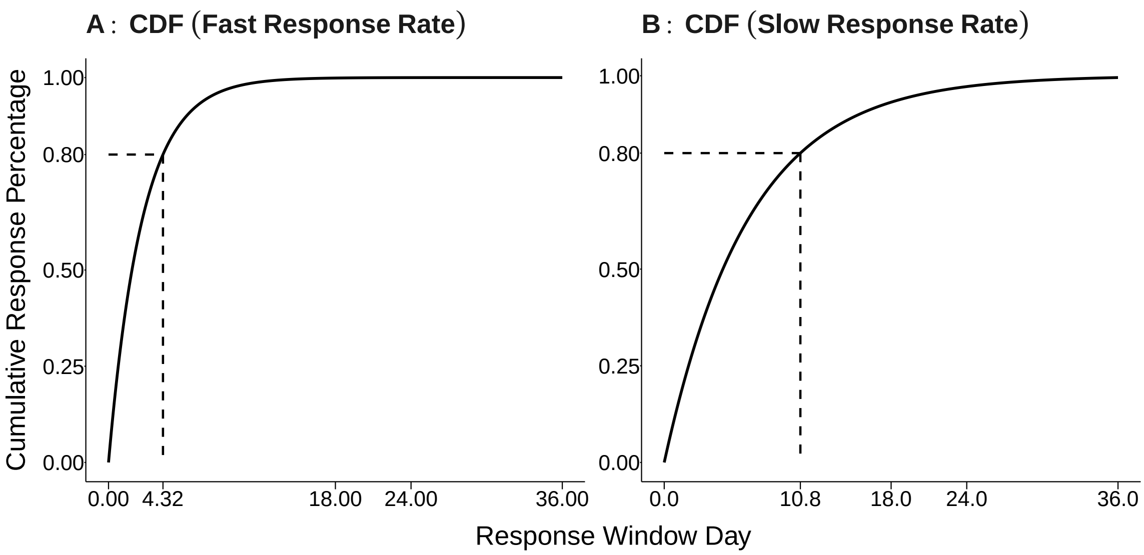
$$f_{fast}(x) = M(1 - e^{a_{fast}x}) = M(1 - e^{-0.37x}) \text{ and} \tag{4.3}$$

$$f_{slow}(x) = M(1 - e^{a_{slow}x}) = M(1 - e^{-0.15x}). \tag{4.4}$$

Using Equations 4.3–4.4, Figure 4.1 shows the resulting cumulative distribution functions (CDF) for time-unstructured data that show the cumulative response percentages as a function of time. Figure 4.1A shows the cumulative distribution function for a fast response rate (Equation 4.3), where an 80% response rate was obtained in 4.32 days. Figure 4.1B shows the cumulative distribution function for a slow response rate (Equation 4.4), where an 80% response rate was obtained in 10.80 days.

Figure 4.1

Cumulative Distribution Functions (CDF) With Fast and Slow Response Rates



Note. Panel A: Cumulative distribution function for a fast response rate (Equation 4.3), where an 80% response rate is obtained in 4.32 days. Panel B: Cumulative distribution function for a slow response rate

(Equation 4.4), where an 80% response rate is obtained in 10.80 days.

4.1.1.2 Constants

Given that each simulation experiment manipulated no more than three independent variables so that results could be readily interpreted (Halford et al., 2005), other variables had to be set to constant values. In Experiment 3, two important variables were set to constant values: nature of change and measurement spacing. For nature of change, I set the value for the fixed-effect days-to-halfway elevation parameter (β_{fixed}) across all cells to have a value of 180. For measurement spacing, I set the value across all cells to have equal spacing.

4.1.1.3 Dependent Variables

4.1.1.3.1 Convergence Success Rate

The proportion of iterations in a cell where models converged defined the *convergence success rate*.[18] Equation (4.5) below shows the calculation used to compute the convergence success rate:

$$\text{Convergence success rate} = \frac{\text{Number of models that successfully converged in a cell}}{n}, \tag{4.5}$$

where n represents the total number of models run in a cell.

4.1.1.3.2 Model Performance

Model performance was the combination of two metrics: bias and precision. More specifically, two questions were of importance in the estimation of a given logistic function

[18] Specifically, convergence was obtained if the convergence code returned by OpenMx was 0.

parameter: 1) How well was the parameter estimated on average (bias) and 2) what was a range of values that could be expected for an estimate from the output of a single model (precision). In the two sections that follow, I will discuss each metric of model performance and the cutoffs used to determine whether estimation was unbiased and precise.

4.1.1.3.2.1 Bias

Bias was calculated to evaluate the accuracy with which each logistic function parameter was estimated in each experimental cell. As shown below in Equation (4.6), *bias* was obtained by summing the differences between the population value set for a parameter and the value estimated for the parameter by each i converged model and then dividing the sum by the number of N converged models.

$$\text{Bias} = \frac{\sum_i^N (\text{Population value for parameter} - \text{Average estimated value}_i)}{N} \tag{4.6}$$

Bias was calculated for the fixed- and random-effect parameters of the baseline (θ_{fixed}, θ_{random}), maximal elevation (α_{fixed}, α_{random}), days-to-halfway elevation (β_{fixed}, β_{random}), and the halfway-triquarter delta parameters (γ_{fixed}, γ_{random}) and the error parameter (ϵ).

4.1.1.3.2.2 Precision

In addition to computing bias, precision was calculated to evaluate the variability with which each parameter was estimated. Importantly, metrics used to evaluate precision in previous studies assume estimates are normally distributed (e.g., mean-squared error and empirical standard error). Because some parameters in my simulations had skewed distributions, using a metric that assumed a normal distribution would likely yield inaccurate results. Correspondingly, I used a distribution-independent definition of precision. In my simulations, *precision* was defined as the range of values covered by the middle 95% of values estimated for a logistic parameter.

4.1.2 Overview of Data Generation

Data generation was computed the same way as in Experiment 1 (see data generation) with one addition to the procedure needed for time structuredness. The section that follows details how time structuredness was simulated. Note that the code used to run the simulations and create the data set can be found in Appendix B and the data file (`exp_3_data.csv`) can be found in the following GitHub repository: https://github.com/sciarraseb/dissertation.

4.1.2.0.1 Simulation Procedure for Time Structuredness

To simulate time-unstructured data, response rates at each collection point followed an exponential pattern described by either a fast or slow response rate (for a review, see time structuredness). Importantly, data generated for each person at each time point had to be sampled according to a probability density function defined by either the fast or slow response rate cumulative distribution function (respectively, see Equations 4.3–{eq:cdf-slow}). In the current context, a *probability density function* describes the probability of sampling any given time delay value x where the range of time delay values is 0–36 ($\{x : 0 \leq x \leq 36\}$). To obtain the probability density functions for fast and slow response rates, the response rate function shown in Equation (4.1) was differentiated with respect to x to obtain the function shown below in Equation 4.7:[19]

$$
\begin{aligned}
f' = \frac{\partial f(x)}{\partial x} &= \frac{\partial}{\partial x} M(1 - e^{-ax}). \\
&= M(e^{-ax}a)
\end{aligned}
\tag{4.7}
$$

To compute the probability density function for the fast response rate cumulative distribution function, the growth rate parameter a was set to 0.37 in Equation 4.7 to obtain the following function in Equation 4.8:

$$
f'_{fast}(x) = M(e^{-a_{fast}x}a_{fast}) = M(e^{-0.37x}0.37). \tag{4.8}
$$

To compute the probability density function for the slow response rate cumulative distribution function, the growth rate parameter a was set to 0.15 in Equation 4.7 to obtain the following function in Equation 4.9:

$$
f'_{slow}(x) = M(e^{-0.15}a_{slow}) = M(e^{-0.15}0.15). \tag{4.9}
$$

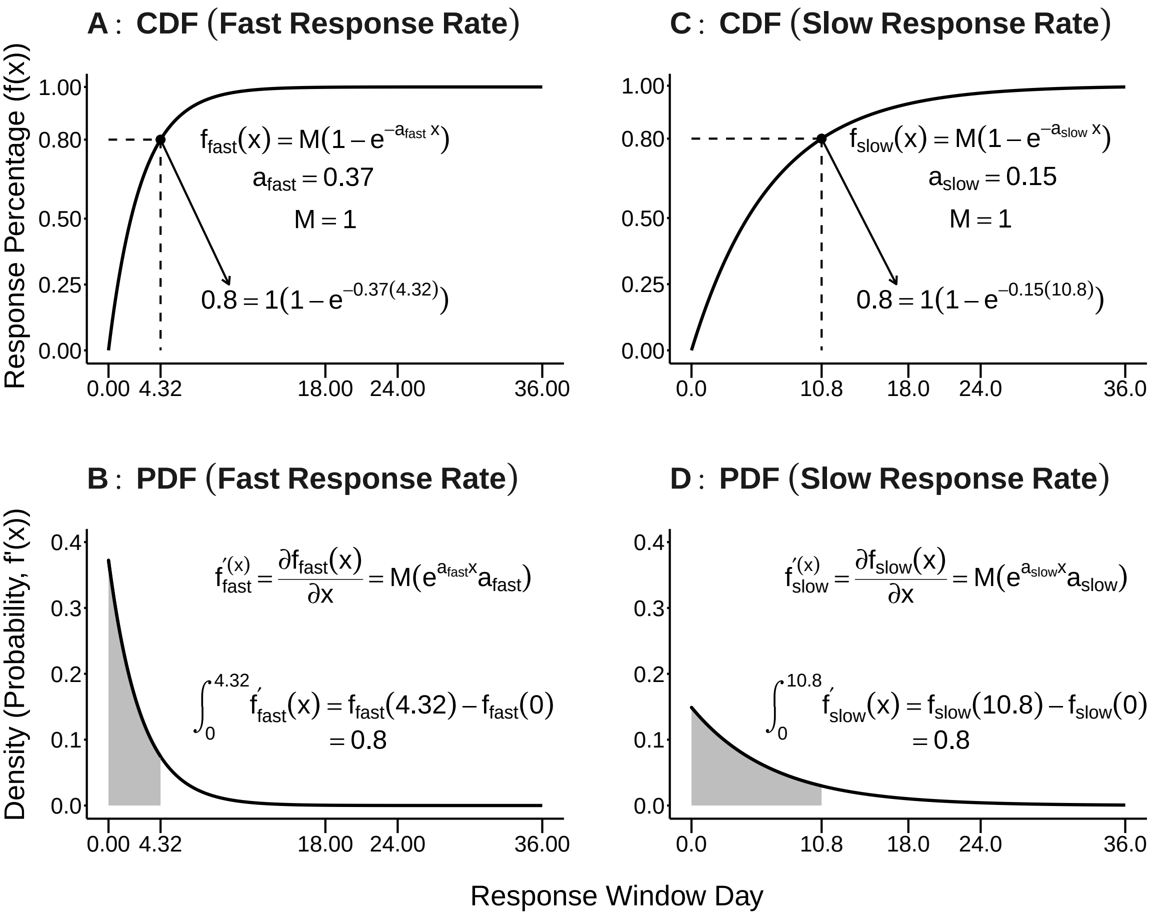
Figure 4.2 shows the fast and slow response cumulative distribution functions (CDF) and their corresponding probability density functions (PDF). Panel A shows the cumulative distribution function for the fast response rate (with a growth parameter value a set to 0.37; see Equation 4.3) and Panel B shows the probability density function that results from computing the derivative of the fast response rate cumulative distribution

[19]Euler's notation for differentiation is used to represent derivatives. In words, $\frac{\partial f(x)}{\partial x}$ means that the derivative of the function $f(x)$ is taken with respect to x.

function

Figure 4.2

Cumulative Distribution Functions (CDF) and Probability Density Functions (PDF) for Fast and Slow Response Rates



Note. Panel A: Cumulative distribution function for the fast response rate (with a growth parameter value a set to 0.37; see Equation 4.3). Panel B: Probability density function that results from computing the derivative of the fast response rate cumulative distribution function with respect to x (see Equation 4.8). Panel C: Cumulative distribution function for the slow response rate (with a growth parameter value a set to 0.15; see Equation 4.4). Panel D: Probability density function that results from computing the derivative of the slow response rate cumulative distribution function with respect to x (see Equation 4.9 and Time Structuredness for more discussion on time structuredness). For the fast response rate functions, an 80% response rate is obtained after 4.32 days or, equivalently, 80% of the area underneath the probability density function is obtained at 4.32 days ($\int_0^{4.32} f'_{fast}(x) = 0.80$). For the slow response rate functions, an 80% response rate is obtained after 10.80 days or, equivalently, 80% of the area underneath the probability density function is obtained at 10.80 days ($\int_0^{10.80} f'_{slow}(x) = 0.80$).

with respect to x (see Equation 4.8). Panel C shows the cumulative distribution function for the slow response rate (with a growth parameter value a set to 0.15; see Equation 4.4)) and Panel D shows the probability density function that results from computing the derivative of the slow response rate cumulative distribution function with respect to x (see Equation 4.9 and section on time structuredness for more discussion). For the fast response rate functions, an 80% response rate is obtained after 4.32 days or, equivalently, 80% of the area underneath the probability density function is obtained at 4.32 days ($\int_0^{4.32} f'_{fast}(x) = 0.80$; the integral from 0 to 4.32 of the probability density function for a fast response rate $f'(x)_{fast}$ is 0.80). For the slow response rate functions, an 80% response rate is obtained after 10.80 days or, equivalently, 80% of the area underneath the probability density function is obtained at 10.80 days ($\int_0^{10.80} f'_{slow}(x) = 0.80$; the integral from 0 to 10.80 of the probability density function for a slow response rate $f'(x)_{slow}$ is 0.80).

Having computed probability density functions for fast and slow response rates, time delays could be generated to create time-unstructured data. To generate time-unstructured data, a time delay was first generated by sampling values according to the probability density function defined by either a fast or slow response rate (Equations 4.8–4.9). The sampled time delay was then added to the value of the current measurement day for a person at a given time point. That is, if the collection window opened on day 60 and the generated time delay for a given person was 4.50 days, then their data would be generated by inserting a value of 64.50 for the $time_i$ parameter of the logistic function (Equation 2.1; along with the fixed-effect parameter values and the person-specific parameter values [or random-effects]).

4.1.3 Modelling of Each Generated Data Set

Each generated data set was modelled using the structured latent growth curve model outlined in Experiment 1 (see data modelling and explicated in Appendix D.

4.1.4 Analysis of Data Modelling Output and Accompanying Visualizations

Analysis and visualization was conducted as outlined in Experiment 1 (see analysis and visualization).

4.2 Results and Discussion

In the sections that follow, I organize the results by presenting them for each level of time structuredness (time-structured data, time-unstructured data resulting from a fast response rate, time-unstructured data resulting from a slow response rate). Importantly, only the results for the day-unit parameters will be presented (i.e., fixed- and random-effect days-to-halfway elevation and halfway-triquarter delta parameters [β_{fixed}, β_{random}, γ_{fixed}, γ_{random}, respectively]). The results for the likert-unit parameters (i.e., fixed- and random-effect baseline and maximal elevation parameters [θ_{fixed}, θ_{random}, α_{fixed}, α_{random}, respectively]) were largely trivial and so are presented in Appendix F.

For each level of time structuredness, I first provide a concise summary of the results and then provide a detailed report of the estimation accuracy of each day-unit parameter of the logistic function. Because the lengths of the detailed reports are considerable, I first provide concise summaries to establish a framework to interpret the detailed reports. The detailed report for the results of each time structuredness level will summarize the results of each (day-unit) parameter's bias/precision plot, report partial ω^2 values, and then provide a qualitative summary.

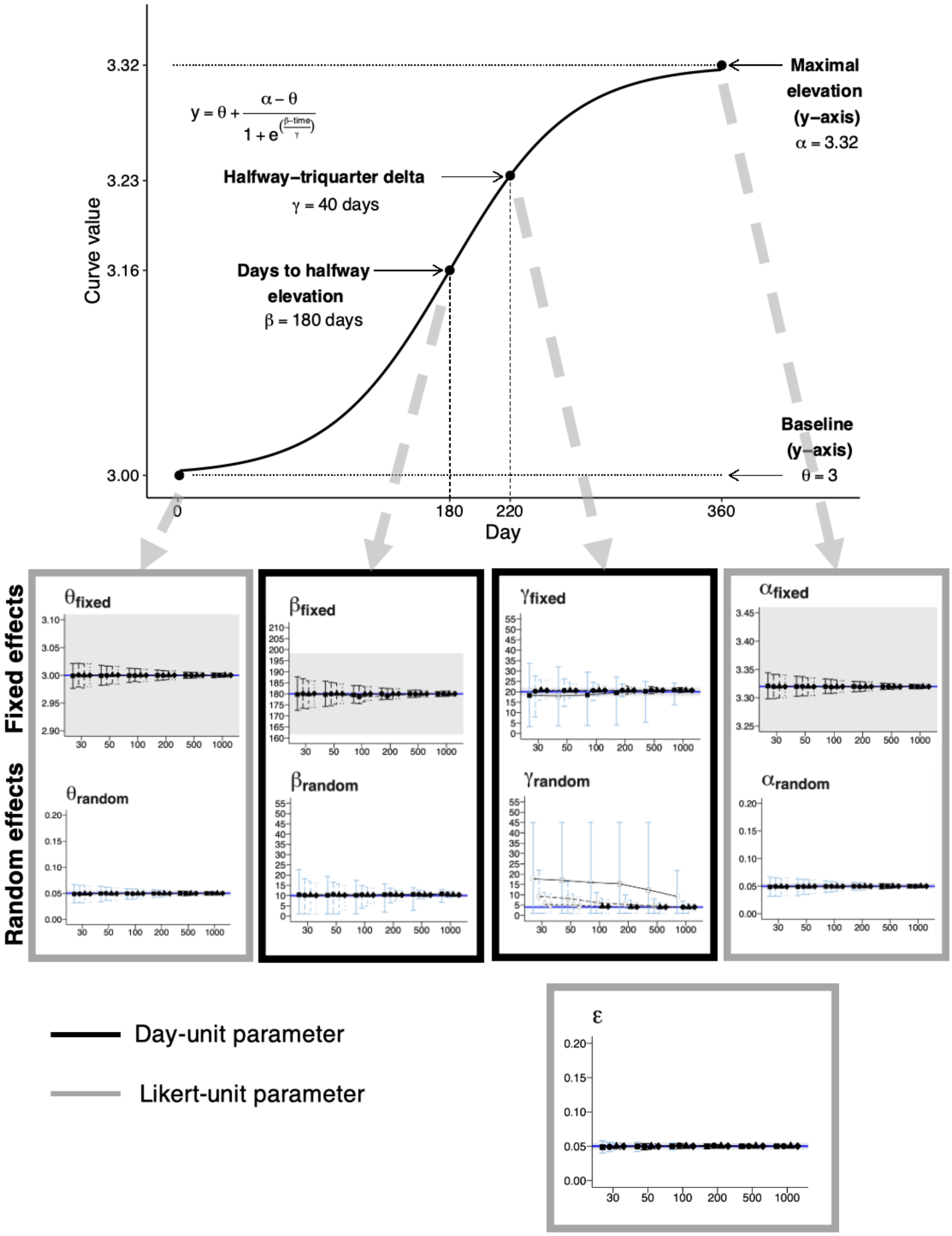
4.2.1 Framework for Interpreting Results

To conduct Experiment 3, the three variables of number of measurements (4 levels), sample size (6 levels), and time structuredness (3 levels) were manipulated, which yielded a total of 72 cells. Importantly, within each cell, bias and precision values were also computed for each of the nine parameters estimated by the structured latent growth curve model (for a review, see modelling of each generated data set). Thus, because the analysis of Experiment 3 computed values for many dependent variables, interpreting the results can become overwhelming. Therefore, I will provide a framework to help the reader efficiently navigate the results section.

Because I will present the results of Experiment 3 by each level of time structuredness, the framework I will describe in Figure 2.2 shows a template for the bias/precision plots that I will present for each level of time structuredness. The results presented for each time structuredness level contain a bias/precision plot for each of the nine estimated parameters. Each bias/precision plot shows the bias and precision for the estimation of one parameter across all measurement number and nature-of change levels. Within each bias/precision plot, dots indicate the average estimated value (which indicates bias) and error bars represent the middle 95% range of estimated values (which indicates precision). Bias/precision plots with black border show the results for day-unit parameters and plots with gray borders show the results for Likert-unit parameters. Importantly, only the results for the day-unit parameters will be presented (i.e., fixed- and random-effect days-to-halfway elevation and halfway-triquarter delta parameters [β_{fixed}, β_{random}, γ_{fixed}, γ_{random},

Figure 4.3

Set of Bias/Precision Plots Constructed for Each Spacing Schedule in Experiment 2



2321 *Note.* A bias/precision plot is constructed for each parameter of the logistic function (see Equation 2.1).
2322 Bias/precision plots with black borders show the results for day-unit parameters and plots with gray border
2323 show the results for Likert-unit parameters. For each parameter, bias and precision are shown across each

combination of measurement number and time structuredness.

respectively]). The results for the Likert-unit parameters (i.e., fixed- and random-effect baseline and maximal elevation parameters [θ_{fixed}, θ_{random}, α_{fixed}, α_{random}, respectively]) were largely trivial and so are presented in Appendix F. Therefore, the results of time structuredness level will only present the bias/precision plots for four parameters (i.e., the day-unit parameters).

4.2.2 Pre-Processing of Data and Model Convergence

After collecting the output from the simulations, non-converged models (and their corresponding parameter estimates) were removed from subsequent analyses. Table G.3 in Appendix G provides the convergence success rates for each cell in Experiment 3. Model convergence never goes below 90%.

4.2.3 Time-Structured Data

For time-structured data, Table 4.1 provides a concise summary of the results for the day-unit parameters (see Figure 4.4 for the corresponding bias/precision plots). The sections that follow will present the results for each column of Table 4.1 and provide elaboration when necessary.

Before presenting the results for equal spacing, I provide a brief description of the concise summary table created for each level of time structuredness and shown below for time-structured data in Table 4.1. Text in the 'Unbiased' and 'Precise' columns indicates the measurement number/sample size pairings that, respectively, resulted in unbiased and precise estimation. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the measurement number/sample size pairing that, respectively, resulted in unbiased estimation and the greatest improvements in bias and precision across

Table 4.1

Concise Summary of Results for Time-Structured Data in Experiment 3

			Description	
Parameter	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 4.4A)	All cells	All cells	Unbiased and precise estimation in all cells	15.13
γ_{fixed} (Figure 4.4B)	All cells	NM $\geq$ 9 with *N* = 500	Largest improvements in precision using **NM = 7 with *N* $\geq$ 200** or **NM = 9 with *N* $\leq$ 100**	9.79
β_{random} (Figure 4.4C)	All cells	No cells	Largest improvements in precision with NM = 7	17.22
γ_{random} (Figure 4.4D)	**NM $\geq$ 9 with *N* $\geq$ 200**	No cells	Largest improvements in precision using **NM = 7 with *N* $\geq$ 200** or **NM = 9 with *N* $\leq$ 100**	10.08

Note. Text in the 'Unbiased' and 'Precise' columns indicates the measurement number/sample size pairings that, respectively, resulted in unbiased and precise estimation. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the number of measurements that, respectively, resulted in unbiased estimation and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision was not obtained in the estimation of all day-unit parameters with time-structured data). 'Error Bar Length' column indicates the maximum error bar length that resulted from using the measurement number/sample size recommendation listed in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. NM = number of measurements.

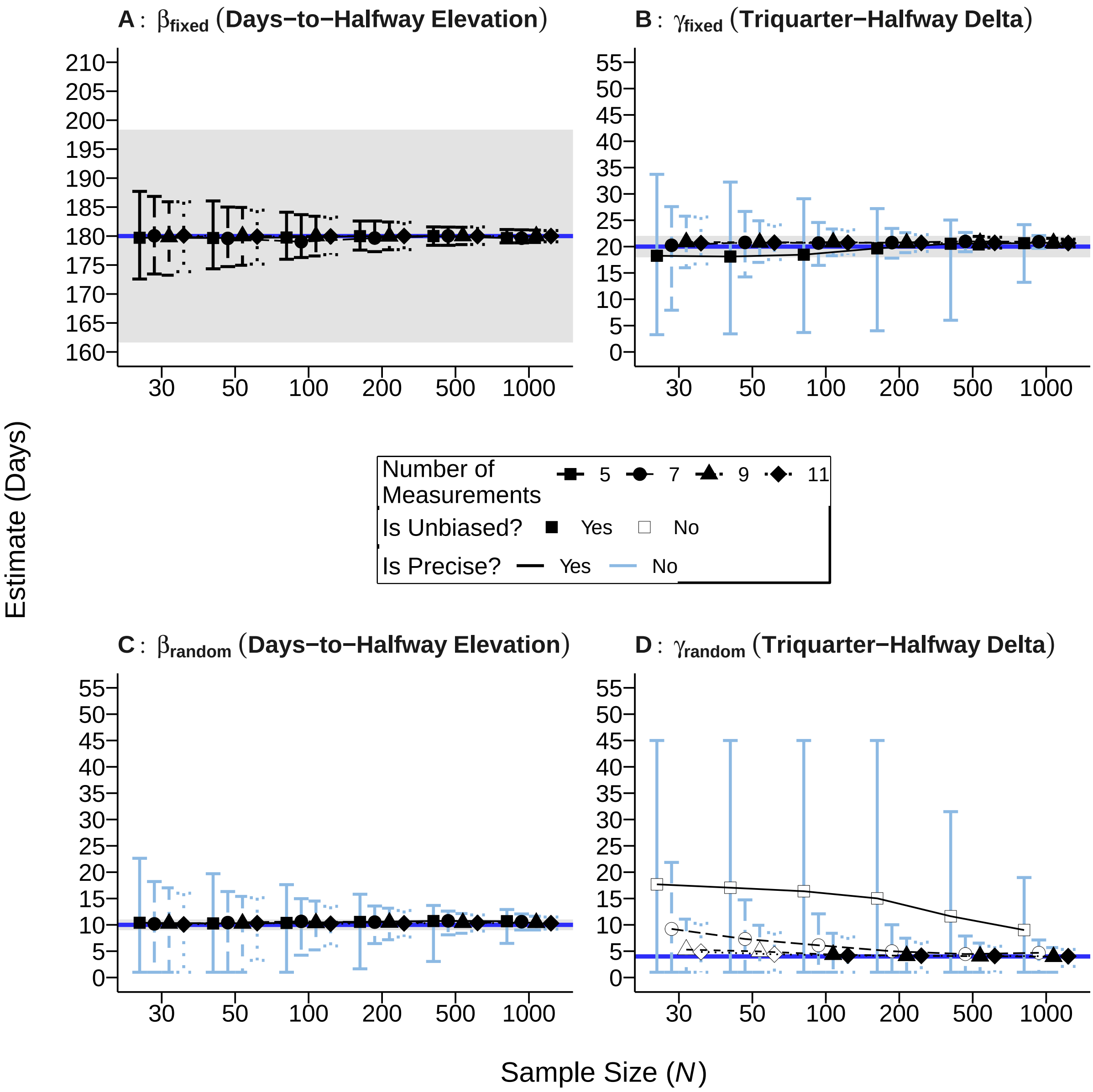
all day-unit parameters (acceptable precision was not obtained in the estimation of all day-unit parameters with time-structured data). The 'Error Bar Length' column indicates the error bar length that results from using the lower-bounding measurement number/sample size pairing listed in the 'Qualitative Description' column (i.e., the maximum error bar length).

4.2.3.0.1 Bias

Before presenting the results for bias, I provide a description of the set of bias/precision plots shown in Figure 4.4 and in the results sections for the other level of time structuredness in Experiment 3. Figure 4.4 shows the bias/precision plots for each day-unit parameter and Table 3.2 provides the partial ω^2 values for each independent variable of each day-unit parameter. In Figure 4.4, blue horizontal lines indicate the population values for each parameter (with population values of β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, and γ_{random} = 4.00). Gray bands indicate the ±10% margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Panels A–B show the bias/precision plots for the fixed- and random-effect days-to-halfway elevation parameters (β_{fixed} and β_{random}, respectively). Panels C–D show the bias/precision plots for the fixed- and random-effect triquarter-halfway delta parameters (γ_{fixed} and γ_{random}, respectively). Note that random-effect parameter units are in standard deviation units.

Figure 4.4

Bias/Precision Plots for Day-Unit Parameters With Time-Structured Data in Experiment 3



Note. Panel A: Bias/precision plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Bias/precision plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Bias/precision plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Bias/precision plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect

parameter units are in standard deviation units. See Table H.3 for specific values estimated for each parameter and Table 4.2 for ω^2 effect size values.

Table 4.2

Partial ω^2 Values for Manipulated Variables With Time-Structured Data in Experiment 3

	Effect		
Parameter	NM	S	NM x S
β_{fixed} (Figure 4.4A)	0.00	0.02	0.00
β_{random} (Figure 4.4B)	0.14	0.27	0.03
γ_{fixed} (Figure 4.4C)	0.25	0.12	0.07
γ_{random} (Figure 4.4D)	0.18	0.03	0.01

NM = number of measurements (5, 7, 9, 11), S = sample size (30, 50, 100, 200, 500, 100), NM x S = interaction between number of measurements and sample size.

With respect to bias for time-structured data, estimates were biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 4.4A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 4.4B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 4.4C): no cells.
- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 4.4D): five and seven measurements across all sample sizes and nine and 11 measurements with $N \leq 100$.

In summary, with time-structured data, estimation of all the day-unit parameters across all manipulated nature-of-change values were unbiased using at least nine measurements with $N \geq 200$, which is indicated by the emboldened text in the 'Unbiased' column of Table 4.1.

4.2.3.0.2 Precision

With respect to precision for time-structured data, estimates were imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 4.4A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 4.4B): five and seven measurements across all sample sizes and nine and 11 measurements with $N \leq 200$.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 4.4C): all cells.
- random-effect halfway-triquarter delta parameter [γ_{random}] in Figure 4.4D): all cells.

In summary, with time-structured data, precise estimation for the fixed-effect day-unit parameters resulted from using at least nine measurements with $N \geq 500$, but no manipulated measurement number/sample size pairing resulted in precise estimation of the random-effect day-unit parameters (see the 'Precise' column of Table 4.1).

4.2.3.0.3 Qualitative Description

For time-structured data in Figure 4.4, although no manipulated measurement number resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) resulted from using moderate measurement number/sample size pairings. With respect to bias under time-structured data, the largest improvements resulted with the following measurement number/sample size pairing(s) for the random-effect triquarter-halfway delta parameter (γ_{fixed}):

- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements with $N \geq 100$ or nine measurements with $N \leq 50$.

With respect to precision under time-structured data, the largest improvements in the estimation of all the day-unit parameters (except the fixed-effect days-to-halfway elevation parameter [β_{fixed}]) resulted from using the following measurement number/sample size pairings:

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which resulted in a maximum error bar length of 9.79 days.
- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements across all manipulated sample sizes, which resulted in a error bar length of 17.22 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which resulted in a maximum error bar length of 10.08 days.

For an applied researcher, one plausible question might be what measurement number/sample size pairing(s) results in the greatest improvements in bias and precision in the estimation of all day-unit parameters with time-structured data. In looking across the measurement number/sample size pairings in the above lists, it becomes apparent that the greatest improvements in bias and precision in the estimation of all day-unit parameters resulted from using seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see the emboldened text in the 'Qualitative Description' column of Table 4.1).

4.2.3.1 Summary of Results for Time-Structured Data

In summarizing the results for time-structured data, estimation of all the day-unit parameters was unbiased using at least nine measurements with $N \geq 200$ (see bias).

Precise estimation was never obtained in the estimation of all day-unit parameters with any manipulated measurement number/sample size pairing (see precision). Although it may be discouraging that no manipulated measurement number/sample size pairing under equal spacing resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) across all the day-unit parameters resulted from using moderate measurement number/sample size pairings. With time-structured data, the largest improvements in bias and precision in the estimation of all the day-unit parameters resulted from using seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see qualitiative description).

4.2.4 Time-Unstructured Data Characterized by a Fast Response Rate

For time-unstructured data characterized by a fast response rate, Table 4.3 provides a concise summary of the results for the day-unit parameters (see Figure 4.5 for the corresponding bias/precision plots). The sections that follow will present the results for each column of Table 4.3 and provide elaboration when necessary (for a description of Table 4.3, see concise summary).

4.2.4.1 Bias

With respect to bias for time-unstructured data characterized by a fast response rate, estimates were biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 4.5A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 4.5B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 4.5C): no cells.

Table 4.3

Concise Summary of Results for Time-Unstructured Data (Fast Response Rate) in Experiment 3

			Description	
Parameter	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 4.5A)	All cells	All cells	Unbiased and precise estimation in all cells	15.35
γ_{fixed} (Figure 4.5B)	All cells	NM $\geq$ 9 with $N \geq$ 500	Largest improvements in precision using **NM = 7 with $N \geq$ 200** or **NM = 9 with $N \leq$ 100**	10.25
β_{random} (Figure 4.5C)	All cells	No cells	Largest improvements in precision with NM = 7	17.47
γ_{random} (Figure 4.5D)	**NM $\geq$ 7 with N = 1000** or **NM $\geq$ 9 with $N \geq$ 200** or **NM = 11 with N = 100**	No cells	Largest improvements in precision using **NM = 7 with $N \geq$ 200** or **NM = 9 with $N \leq$ 100**	10.51

Note. Text in the 'Unbiased' and 'Precise' columns indicates the measurement number/sample size pairings that, respectively, resulted in unbiased and precise estimation. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the number of measurements that, respectively, resulted in unbiased estimation and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision was not obtained in the estimation of all day-unit parameters with time-unstructured data characterized by a fast response rate). 'Error Bar Length' column indicates the maximum error bar length that resulted from using the measurement number/sample size recommendation listed in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. NM = number of measurements.

- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 4.5D): five measurements across all sample sizes, seven measurements with $N \leq 500$, nine measurements with $N \geq 100$, and 11 measurements with $N \leq 50$.

Importantly, for the fixed-effect days-to-halfway elevation parameter (β_{fixed}), although bias was still within the acceptable margin of error, bias appeared to be constant across all manipulated measurement number/sample size pairings. In comparing the bias/precision plots between time-unstructured data characterized by a fast response rate (Figure 4.5A) and time-structured data (Figure 4.4A), the systematic decline in bias observed for fixed-effect days-to-halfway elevation parameter (β_{fixed}) appeared to result from thedecrease in time structuredness.

In summary, with time-unstructured data characterized by a fast response rate, estimation of all the day-unit parameters across all manipulated nature-of-change values was unbiased using at least seven measurements with $N = 1000$, nine measurements with $N \geq 200$, or 11 measurements with $N \geq 100$, which is indicated by the emboldened text in the 'Unbiased' column of Table 4.3.

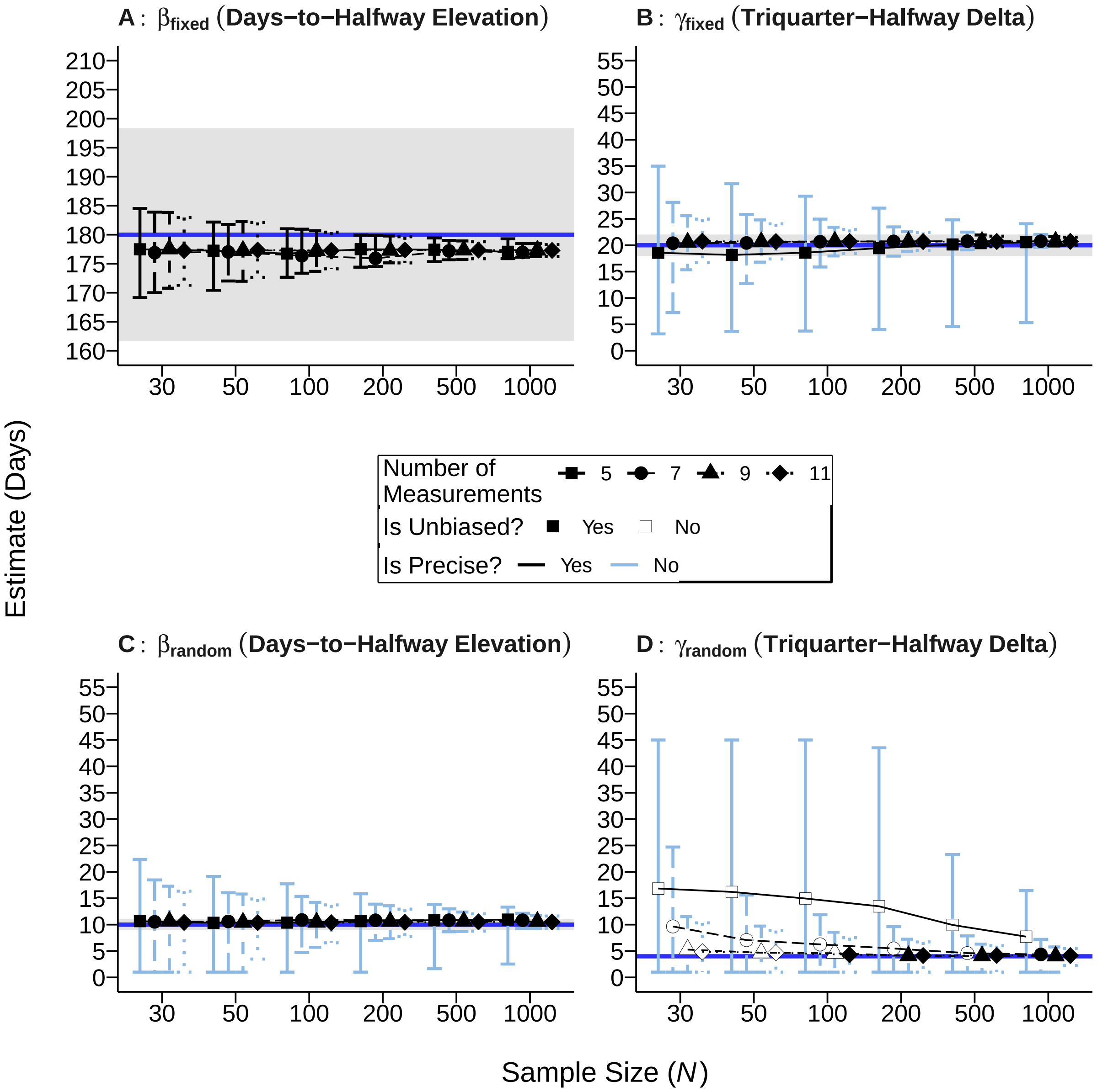
4.2.4.2 Precision

With respect to precision for time-unstructured data characterized by a fast response rate, estimates were imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 4.5A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 4.5B): five and seven measurements across all sample sizes and nine and 11 measurements with $N \leq 200$.

Figure 4.5

Bias/Precision Plots for Day-Unit Parameters With Time-Unstructured Data Characterized by a Fast Response Rate in Experiment 3



Note. Panel A: Bias/precision plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Bias/precision plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Bias/precision plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Bias/precision plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or

longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table H.3 for specific values estimated for each parameter and Table 4.4 for ω^2 effect size values.

Table 4.4

Partial ω^2 Values for Manipulated Variables With Time-Unstructured Data Characterized by a Fast Response Rate in Experiment 3

	Effect		
Parameter	NM	S	NM x S
β_{fixed} (Figure 4.5A)	0.00	0.02	0.00
β_{random} (Figure 4.5B)	0.15	0.27	0.03
γ_{fixed} (Figure 4.5C)	0.29	0.14	0.08
γ_{random} (Figure 4.5D)	0.17	0.04	0.01

NM = number of measurements (5, 7, 9, 11), S = sample size (30, 50, 100, 200, 500, 100), NM x S = interaction between number of measurements and sample size.

- random-effect days-to-halfway elevation parameter (β_{random}; Figure 4.5C): all cells.
- random-effect halfway-triquarter delta parameter [γ_{random}] in Figure 4.5D): all cells.

In summary, with time-unstructured data characterized by a fast response rate, precise estimation for the fixed-effect day-unit parameters resulted from using at least nine measurements with $N \geq 500$, but no manipulated measurement number/sample size pairing resulted in precise estimation of the random-effect day-unit parameters (see the 'Precise' column of Table 4.3).

4.2.4.3 Qualitative Description

For time-unstructured data characterized by a fast response rate (see Figure 4.5), although no manipulated measurement number resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) resulted from using moderate measurement number/sample size pairings. With respect to bias under time-unstructured data characterized by a fast response rate, the largest improvements in bias resulted with the following measurement number/sample size pairing(s) for the random-effect triquarter-halfway delta parameter (γ_{fixed}):

- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements with $N \geq 100$ or nine measurements with $N \leq 50$.

With respect to precision under time-unstructured data characterized by a fast response rate, the largest improvements in the estimation of all the day-unit parameters (except the fixed-effect days-to-halfway elevation parameter [β_{fixed}]) resulted from using the following measurement number/sample size pairings:

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which resulted in a maximum error bar length of 10.25 days.
- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements across all manipulated sample sizes, which resulted in a maximum error bar length of 17.47 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which resulted in a maximum error bar length of 10.51 days.

For an applied researcher, one plausible question might be what measurement number/sample size pairing(s) results in the greatest improvements in bias and precision in the estimation of all day-unit parameters with time-unstructured data characterized by a fast response rate. In looking across the measurement number/sample size pairings in the above lists, it becomes apparent that greatest improvements in bias and precision in the estimation of all day-unit parameters resulted from using seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see the emboldened text in the 'Qualitative Description' column of Table 4.3).

4.2.4.4 Summary of Results for Time-Unstructured Characterized by a Fast Response Rate

In summarizing the results for time-unstructured data characterized by a fast response rate, estimation of all the day-unit parameters was unbiased using least seven measurements with $N = 1000$, nine measurements with $N \geq 200$, or 11 measurements with $N \geq 100$ (see bias). Importantly, bias for some day-unit parameters was constant across manipulated measurement number/sample size pairings. Precise estimation was never obtained in the estimation of all day-unit parameters with any manipulated measurement number/sample size pairing (see precision). Although it may be discouraging that no manipulated measurement number/sample size pairing resulted in precise estimation of all the day-unit parameters with time-unstructured data characterized by a fast response rate, the largest improvements in precision (and bias) across all day-unit parameters resulted with moderate measurement number/sample size pairings. With time-unstructured data characterized by a fast response rate, the largest improvements in bias and precision in the estimation of all day-unit parameters resulted from using seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see qualitiative description).

4.2.5 Time-Unstructured Data Characterized by a Slow Response Rate

For time-unstructured data characterized by a slow response rate, Table 4.5 provides a concise summary of the results for the day-unit parameters (see Figure 4.6 for the corresponding bias/precision plots). The sections that follow will present the results for each column of Table 4.5 and provide elaboration when necessary (for a description of Table 4.5, see concise summary).

4.2.5.1 Bias

With respect to bias for time-unstructured data characterized by a slow response rate, estimates were biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 4.6A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 4.6B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 4.6C): no cells.
- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 4.6D): five measurements across all sample sizes, seven measurements with $N \leq 500$, nine measurements with $N \geq 100$, and 11 measurements with $N \leq 50$.

Note that, for all parameters except the halfway-triquarter delta parameter (γ_{fixed}), bias appeared to be constant across all manipulated measurement number/sample size pairings.

In summary, with time-unstructured data characterized by a slow response rate, estimation of all the day-unit parameters across all manipulated nature-of-change values was unbiased using at least seven measurements with $N = 1000$, nine measurements with $N \geq 200$, or 11 measurements with $N \geq 100$, which is indicated by the emboldened text

Table 4.5

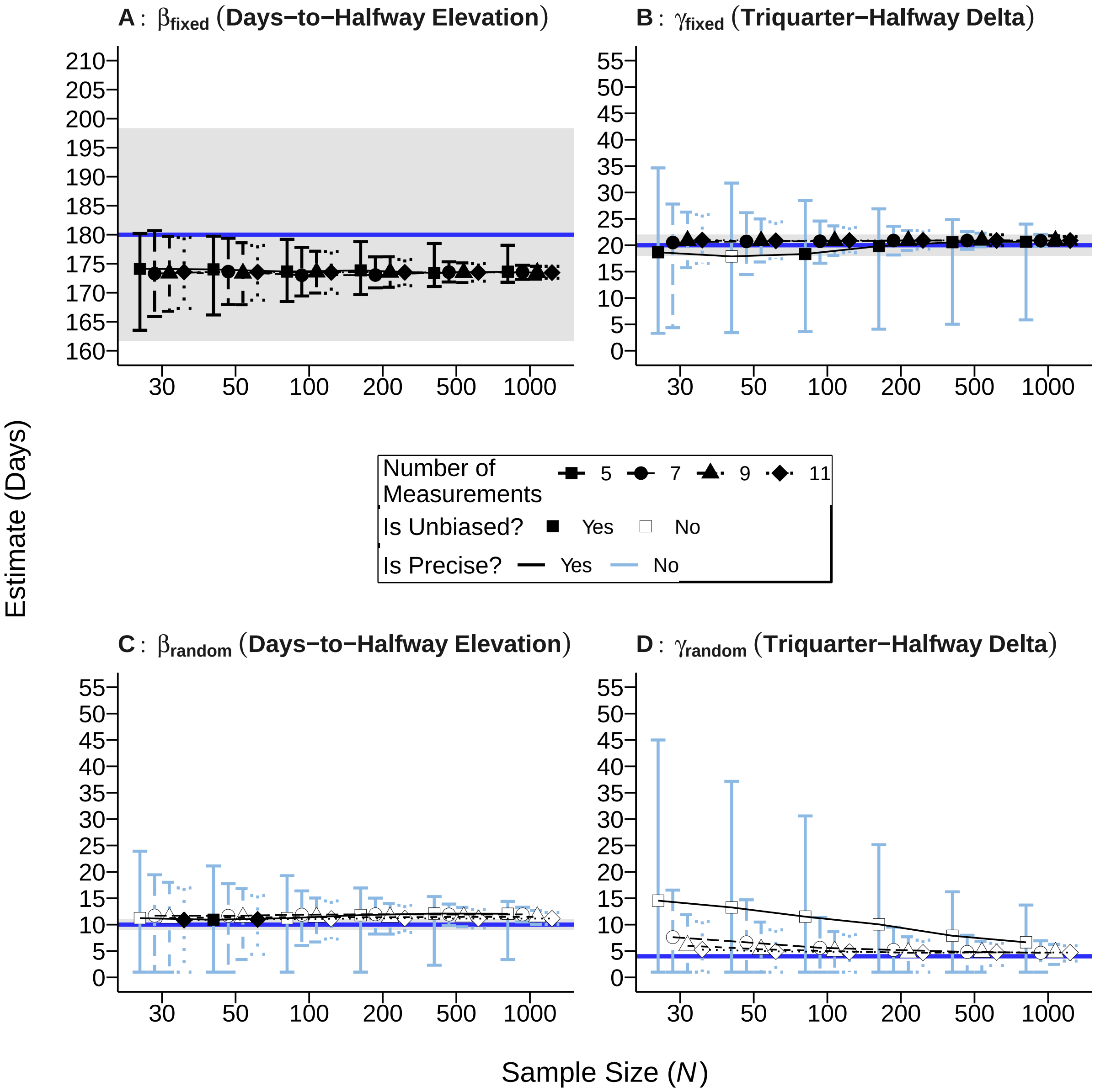
Concise Summary of Results for Time-Unstructured Data (Slow Response Rate) in Experiment 3

			Summary	
			Description	
Parameter	Unbiased	Precise	Qualitative Summary	Error Bar Length
β_{fixed} (Figure 4.6A)	All cells	All cells	Low bias and high precision in all cells	16.68
γ_{fixed} (Figure 4.6B)	All cells except NM = 5 with *N* = 50	NM = 7 with *N* = 200 or NM = 9 with $N \leq 500$	Largest improvements in precision using **NM = 7 with $N \geq$ 200** or **NM = 9 with $N \leq$ 100**	10.53
β_{random} (Figure 4.6C)	No cells except NM = 5 with *N* = 30 and NM = 11 with $N \leq 50$	No cells	Largest improvements in precision with NM = 7	18.44
γ_{random} (Figure 4.6D)	No cells	No cells	Largest improvements in bias and precision using NM = 7 with $N \geq 200$ or M = 9 with $N \leq 100$	10.9

Note. Text in the 'Unbiased' and 'Precise' columns indicates the measurement number/sample size pairings that, respectively, resulted in unbiased and precise estimation. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the number of measurements that, respectively, resulted in unbiased estimation and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision was not obtained in the estimation of all day-unit parameters with time-unstructured data characterized by a slow response rate). 'Error Bar Length' column indicates the maximum error bar length that resulted from using the measurement number/sample size recommendation listed in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. NM = number of measurements.

Figure 4.6

Bias/Precision Plots for Day-Unit Parameters With Time-Unstructured Data Characterized by a Slow Response Rate in Experiment 3



Note. Panel A: Bias/precision plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Bias/precision plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Bias/precision plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Bias/precision plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or

longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table H.3 for specific values estimated for each parameter and Table 4.6 for ω^2 effect size values.

Table 4.6

Partial ω^2 Values for Manipulated Variables With Time-Unstructured Data Characterized by a Slow Response Rate in Experiment 3

	Effect		
Parameter	NM	S	NM x S
β_{fixed} (Figure 4.6A)	0.00	0.02	0.00
β_{random} (Figure 4.6B)	0.15	0.27	0.03
γ_{fixed} (Figure 4.6C)	0.29	0.14	0.08
γ_{random} (Figure 4.6D)	0.17	0.04	0.01

NM = number of measurements (5, 7, 9, 11), S = sample size (30, 50, 100, 200, 500, 100), NM x S = interaction between number of measurements and sample size.

in the 'Unbiased' column of Table 4.5.

4.2.5.2 Precision

With respect to precision for time-unstructured data characterized by a slow response rate, estimates were imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 4.6A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 4.6B): five and seven measurements across all sample sizes and nine and 11 measurements with $N \leq 200$.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 4.6C): all cells.
- random-effect halfway-triquarter delta parameter [γ_{random}] in Figure 4.6D): all cells.

In summary, with time-unstructured data characterized by a slow response rate, precise estimation for the fixed-effect day-unit parameters resulted from using at least nine measurements with $N \geq 500$, but no manipulated measurement number/sample size pairing resulted in precise estimation of the random-effect day-unit parameters (see the 'Precise' column of Table 4.5).

4.2.5.3 Qualitative Description

For time-unstructured data characterized by a slow response rate (see Figure 4.6), although no manipulated measurement number resulted in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) resulted from using moderate measurement number/sample size pairings. With respect to bias under time-unstructured data characterized by a slow response rate, the largest improvements resulted with the following measurement number/sample size pairings for the random-effect triquarter-halfway delta parameter (γ_{fixed}):

- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements with $N \geq 100$ or nine measurements with $N \leq 50$.

With respect to precision under time-unstructured data characterized by a slow response rate, the largest improvements in the estimation of all the day-unit parameters (except the fixed-effect days-to-halfway elevation parameter [β_{fixed}]) resulted from using the following measurement number/sample size pairings:

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which resulted in a maximum error bar length of 10.53 days.
- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements

across all manipulated sample sizes, which resulted in a maximum error bar length of 18.44 days.

- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which resulted in a maximum error bar length of 10.9 days.

For an applied researcher, one plausible question might be what measurement number/sample size pairing(s) results in the greatest improvements in bias and precision in the estimation of all day-unit parameters with time-unstructured data characterized by a fast response rate. In looking across the measurement number/sample size pairings in the above lists, it becomes apparent that the greatest improvements in bias and precision in the estimation of all day-unit parameters resulted from using seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see the emboldened text in the 'Qualitative Description' column of Table 4.5).

4.2.5.4 Summary of Results Time-Unstructured Characterized by a Slow Response Rate

In summarizing the results for time-unstructured data characterized by a slow response rate, estimation of all the day-unit parameters was unbiased using least seven measurements with $N = 1000$, nine measurements with $N \geq 200$, or 11 measurements with $N \geq 100$ (see bias). Importantly, bias for most day-unit parameters was constant across manipulated measurement number/sample size pairings. Precise estimation was never obtained in the estimation of all day-unit parameters with any manipulated measurement number/sample size pairing (see precision). Although it may be discouraging that no manipulated measurement number/sample size pairing resulted in precise estimation of all the day-unit parameters with time-unstructured data characterized by a slow response rate, the largest improvements in precision (and bias) across all day-unit parameters resulted with moderate measurement number/sample size pairings. With time-unstructured data characterized by a slow response rate, the largest improvements in bias and precision in the estimation of all day-unit parameters resulted from using seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see qualitiative description).

4.2.6 How Does Time Structuredness Affect Model Performance?

In Experiment 3, I was interested in how decreasing time structuredness affected model performance. Table 4.7 summarizes the results for each spacing schedule in Experiment 3. Text within the 'Unbiased' and 'Precise' columns indicates the measurement number/sample size pairing needed to, respectively, obtain unbiased an precise estimation for all the day-unit parameters. The 'Error Bar Length' column indicates longest error bar lengths that result in the estimation of each day-unit parameter from using the measurement number/sample size pairings listed in the 'Qualitative Description' column. In looking at the 'Qualitative Description' column, the greatest improvements in bias and precision for all time structuredness levels result from using either seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$.

Although the same measurement number/sample size pairing can be used to obtain the greatest improvements in model performance under any time structuredness level, two results suggest that model performance decreases as the time structuredness decreases. First, the error bar lengths in Table 4.7 increase as time structuredness decreases. As an

Table 4.7

Concise Summary of Results Across All Time Structuredness Levels in Experiment 3

				Error Bar Summary			
Time Structuredness	Unbiased	Precise	Qualitative Description	β_{fixed}	γ_{fixed}	β_{random}	γ_{random}
Time structured (see Figure 4.4 and Table 4.1)	NM $\geq$ 9 with $N \geq$ 200	No cells	Largest improvements in precision using **NM = 7 with $N \geq$ 200** or **NM = 9 with $N \leq$ 100**	15.13	9.79	17.22	10.08
Time unstructured (fast response rate; see Figure 4.5 and Table 4.3)	NM $\geq$ 7 with N = 1000 or NM $\geq$ 9 with $N \geq$ 200 or NM = 11 with N = 100	No cells	Largest improvements in precision using **NM = 7 with $N \geq$ 200** or **NM = 9 with $N \leq$ 100**	15.35	10.25	17.47	10.51
Time unstructured (slow response rate; see Figure 4.6 and Table 4.5)	No cells	No cells	Largest improvements in precision using **NM = 7 with $N \geq$ 200** or **NM = 9 with $N \leq$ 100**	16.68	10.53	18.44	10.90

Note. 'Qualitative Description' column indicates the number of measurements that obtains the greatest improvements in bias and precision across all day-unit parameters. 'Error Bar Summary' columns list the error bar lengths that result for each day-unit parameter using the measurement number listed in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter $\in \{80, 180, 280\}$; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. NM = number of measurements.

example, the error bar length of the fixed-effect days-to-halfway elevation parameter is 15.13 days with time-structured data and increases to 16.68 days with time-unstructured data characterized by a slow response rate. Second, and more alarming, the bias incurred as time structuredness decreases is constant across all measurement number/sample size pairings (see Figure 4.6). That is, the increase in bias that results from time-unstructured data cannot be reduced by increasing the number of measurements or sample size. An an example, the fixed-effect days-to-halfway elevation parameter is underestimated by approximately 6 days across all measurement number/sample size pairings (β_{fixed}; see Figure 4.6A).

To understand why bias is systematic as time structure decreases, it is important to first understand latent growth curve models more deeply. By default, latent growth curve models assume time-structured data. As a reminder, data are time structured when participants provide data at the exact same moment at each time point (e.g., if a study collects data on the first day of each month for a year, then time-structured data would only be obtained if participants all provide their data at the exact same moment each time data are collected). In other words, one response schedule characterized the response patterns of all participants. Consider a random-intercept-random-slope model shown in Figure 4.7 that is used to model stress ratings collected on the first day of each month over the course of five months from j people. Stress ratings at each i time point for each j person are predicted by person-specific intercepts (b_{0j}) and slopes (b_{1j}; in addition to a residual term [ϵ_{ij}]) as shown below in Equation 4.10 (which is often called Level-1 equation):

$$Stress_{ij} = b_{0j} + b_{1j}(Stress_{ij}) + \epsilon_{ij}. \tag{4.10}$$

The person-specific intercepts and slopes are the sum of a fixed-effect parameter whose value is constant across all people (γ_{00} and γ_{10}) and a random-effect parameter that represents the variance of the person-specific variables (i.e., σ_{00} and σ_{10}). The fixed-effect intercept and slope, respectively, represent the mean starting stress value (i.e., average stress value at Time = 0) and the average slope value. Importantly, by estimating a random-effect parameter (in addition to the fixed-effect parameters), deviations from the mean intercept an slope values can be obtained for each j person (σ_{0j} and σ_{1j}) and these values can be used to compute person-specific intercepts and slopes as shown in Equations 4.11–4.12 (which are often called Level-2 equations):

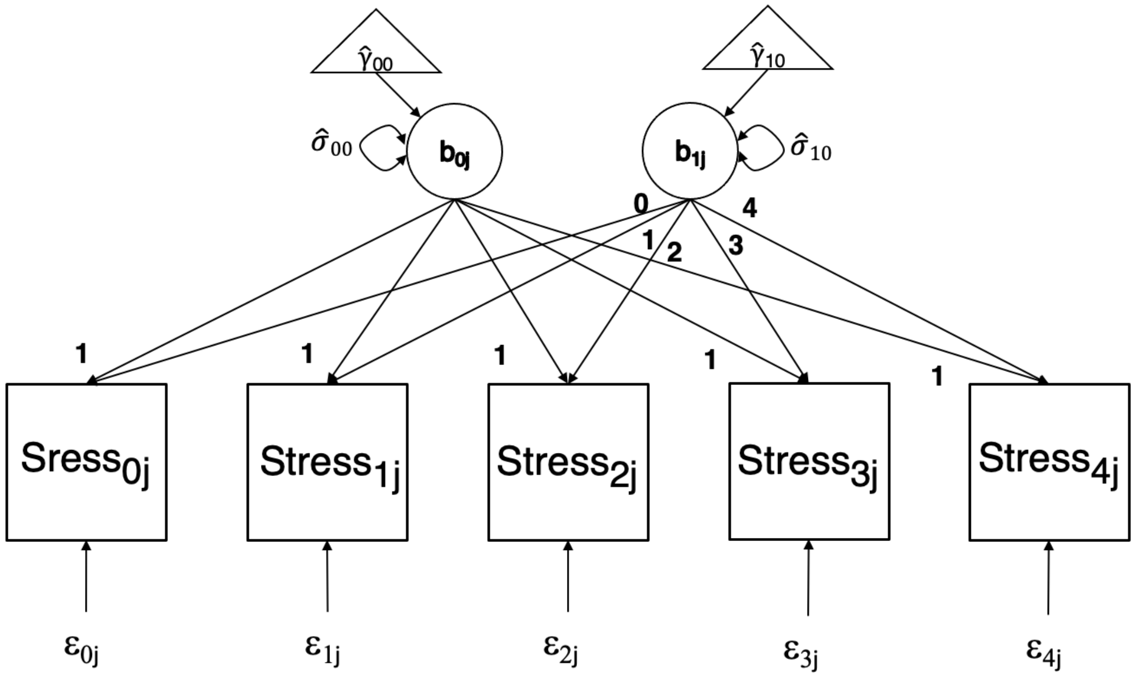
$$b_{0j} = \hat{\gamma_{00}} + \sigma_{0j} \tag{4.11}$$

$$b_{1j} = \hat{\gamma_{10}} + \sigma_{1j} \tag{4.12}$$

Note that the fixed- and random-effect parameters in Figure 4.7 are superscribed with a caret ($\hat{}$) to indicate that the values of these parameters are estimated by the latent growth curve model. Also note that, in Figure 4.7, circles indicate latent variables, triangles indicate constants, and squares indicate observed (or manifest variables).

Figure 4.7

Path Diagram for a Random-Intercept-Random-Slope Latent Growth Curve Model



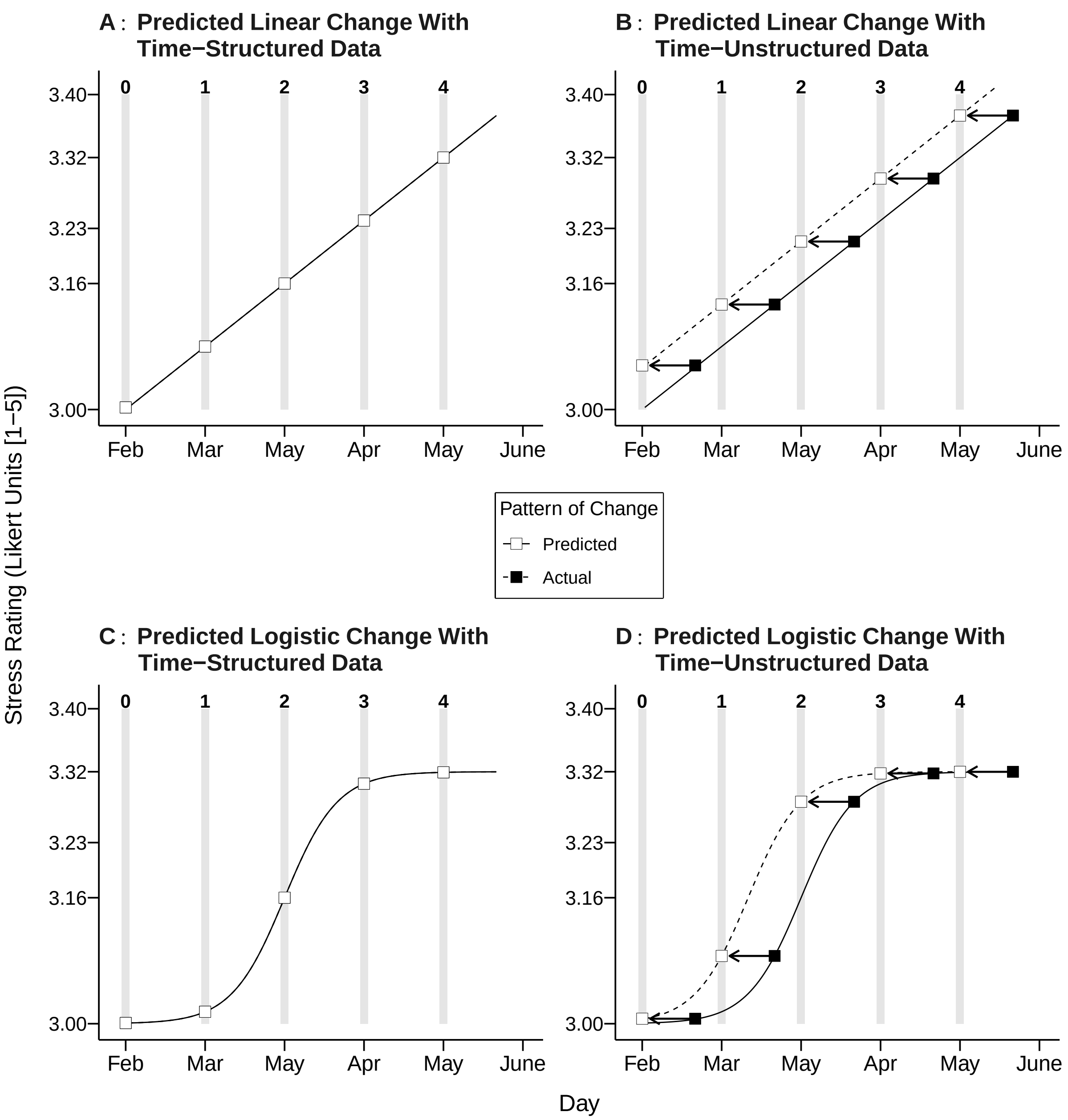
Note. Stress at each i time point for each j person is predicted by a person-specific slope (b_{0j}), person-specific intercept (b_{1j}), and residual (ϵ_{ij}; see Equation 4.10 [Level-1 equation]). The person-specific effects are also called *random effects* and each is the sum of a fixed-effect parameter whose value is constant across all people (γ_{00} and γ_{10}) and a random-effect parameter that represents the variance of the person-specific variables (i.e., σ_{00} and σ_{10}; see Equations 4.11–4.12 [Level-2 equations]). Note that the fixed- and random-effect parameters are superscribed with a caret ($\hat{}$) to indicate that the values of these parameters are estimated by the latent growth curve model. Also note that circles indicate latent variables, triangles indicate constants, and squares indicate observed (or manifest variables).

To understand why bias in parameter estimation increases as time structuredness decreases, it is important to discuss one component of the latent growth curve model not yet discussed: loadings. In latent variable models, *loadings* comprise numbers that indicate how a latent variable should be modelled. The numbers in loadings satisfy two needs of latent variables. First, loadings give latent variables a unit; latent variables are inherently unitless, and so require a unit so that they can be meaningfully interpreted. By fixing at least one pathway between a latent variable and observed variable with a loading, the latent variable takes on the units of the observed variable. In the current example, the intercept and slope latent variables take on the units of the stress ratings (e.g., Likert units). Second, in latent growth curve models, latent variables need their effect to be specified, and loadings satisfy this need. In the current example, the intercept has a constant effect at each time point, and this is represented by setting its loadings at each time point to 1. The slope represents linearly increasing change over time, and so its loadings are set to increase by an integer value of 1 after each time point.

Although loadings allow latent variables to model change over time, their values are constant across participants and it is this characteristic that causes model performance to decrease as time structuredness decreases. In focusing on the slope variable in Figure 4.7, the loadings of 0, 1, 2, 3, and 4 assume that only one response pattern describes how each participant provides their data over some period of time. If the period of time is assume to be five months, then the loadings assume that each participant provides data on the first day of each month, which is indicated by the gray rectangles (along with the loading number above each gray rectangle) in each panel of Figure 4.8. With time-structured data, constant loadings do not decrease model performance because each participant provides their data on the first day of each month. As examples of model performance with time-structured data, panels A and C of Figure 4.8 show the predicted and actual patterns for individual participants with linear and logistic patterns of change, respectively. Because each individual participant displays a response pattern identical to the one specified by the loadings, the predicted and actual patterns of change are identical. With time-unstructured data however, the predicted and actual patterns of change no longer overlap because response patterns in participants differ from the one assumed by the

Figure 4.8

Model Performance Decreases as Time Structuredness Decreases



Note. Panel A: Predicted and actual linear patterns of change are identical because of time-structured data. Panel B: Predicted and actual linear patterns of change are different because of time-untructured data. Panel C: Predicted and actual logistic patterns of change are identical because of time-structured data. Panel D: Predicted and actual logistic patterns of change differ because model because of time-unstructured data. Predicted patterns of change are based on empty dots and actual patterns of change are based on filled dots. Shaded vertical rectangles indicate the response pattern expected across all participants by the loadings set in the latent growth curve model depicted in Figure 4.7.

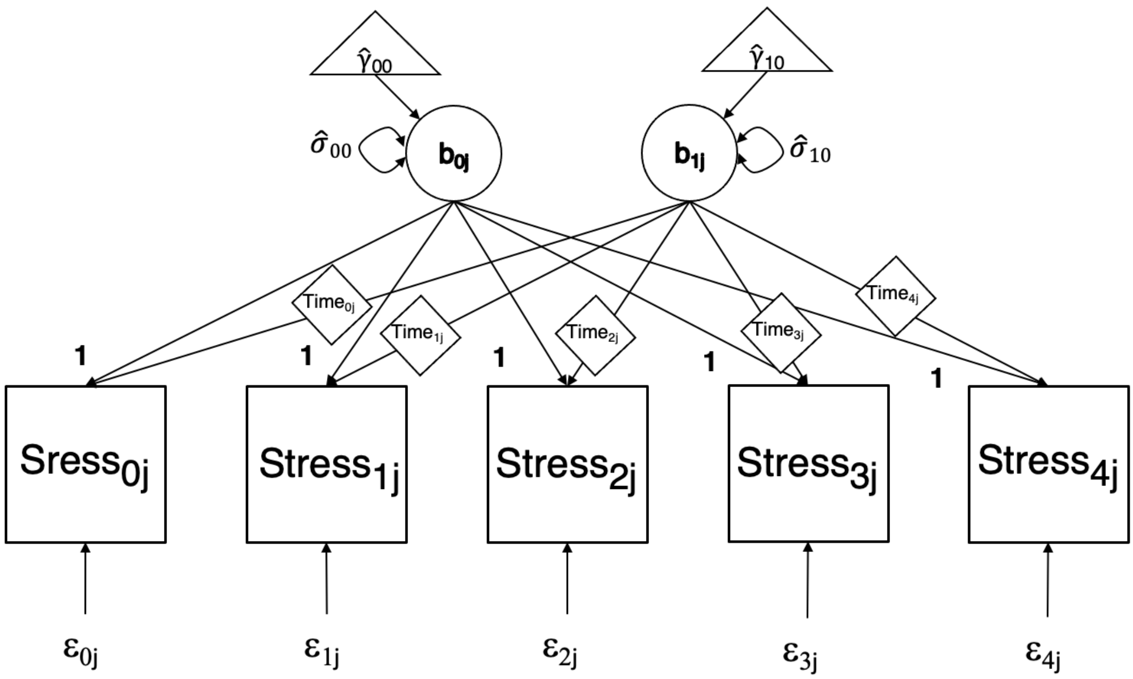
loadings. As examples of model performance with time-unstructured data, panels B and D of Figure 4.8 show the predicted and actual patterns for individual participants with linear and logistic patterns of change, respectively. Although each participant provides data many days after the first day of each month, the constant loadings set in the model lead it to assume that data were collected on the first day of each month. Because the model misattributes the time at which data are recorded, the predicted patterns of change are shifted leftward, leading to a decrease in model performance. In Figure 4.8B, the intercept parameter value (b_{0j}) increases due to time-unstructured data. In Figure 4.8D, the value for the fixed-effect days-to-halfway elevation parameter (β_{fixed}) decreases due to time-unstructured data. Therefore, the loading structured specified by default in latent growth curve model causes model performance to decrease when data are time unstructured.

4.2.7 Eliminating the Bias Caused by Time Unstructuredness: Using Definition Variables

In examining the effects of time structuredness, the results show that model performance decreases as time structuredness decreases. Importantly, increasing the number of measurements and/or sample size has no effect on eliminating the decline in model performance. Because data are likely to be time unstructured under realistic conditions, the resulting decline in model performance seems inevitable and this can be disconcerting. Fortunately, the error incurred when time unstructuredness is overlooked can be prevented by allowing loadings to vary across people by using *definition variables*: Observed variables are placed in parameter matrices so that values in the matrix (specifically, the loadings) are constrained to person-specific values (Blozis & Cho, 2008; Mehta & Neale, 2005; Mehta & West, 2000; Sterba, 2014). In the current example, definition variables are used to set loadings to the specific time points at which each participant provides their data. Thus, the observed variable is the specific i time point at which a j person provides a datum and this value is inserted into the λ matrix (for details of this matrix, see Appendix D). Figure 4.9 shows a path diagram for a random-intercept-random-slope latent variable model with definition variables. In comparing it to the latent growth curve model in Figure 4.7, there is only one difference. Instead of setting the loadings to be constant across all participants, definition variables (indicated by diamonds) are used so that loadings for each j person are set to the specific i time point at which a datum was provided.

Figure 4.9

Path Diagram for a Random-Intercept-Random-Slope Latent Growth Curve Model With Definition Variables



Note. Stress at each i time point for each j person is predicted by a person-specific slope (b_{0j}), person-specific intercept (b_{1j}), and residual (ϵ_{ij}; see Equation 4.10 [Level-1 equation]). The person-specific effects are also called *random effects* and each is the sum of a fixed-effect parameter whose value is constant across all people (γ_{00} and γ_{10}) and a random-effect parameter that represents the variance of the person-specific variables (i.e., σ_{00} and σ_{10}; see Equations 4.11–4.12 [Level-2 equations]). Note that the fixed- and random-effect parameters are superscribed with a caret ($\hat{}$) to indicate that the values of these parameters are estimated by the latent growth curve model. To account for time-unstructured data, loadings are allowed to vary using definition variables (diamonds). Specifically, loadings for each j person are set to the specific i time point at which a datum was provided. Also note that circles indicate latent variables, triangles indicate constants, and squares indicate observed (or manifest variables).

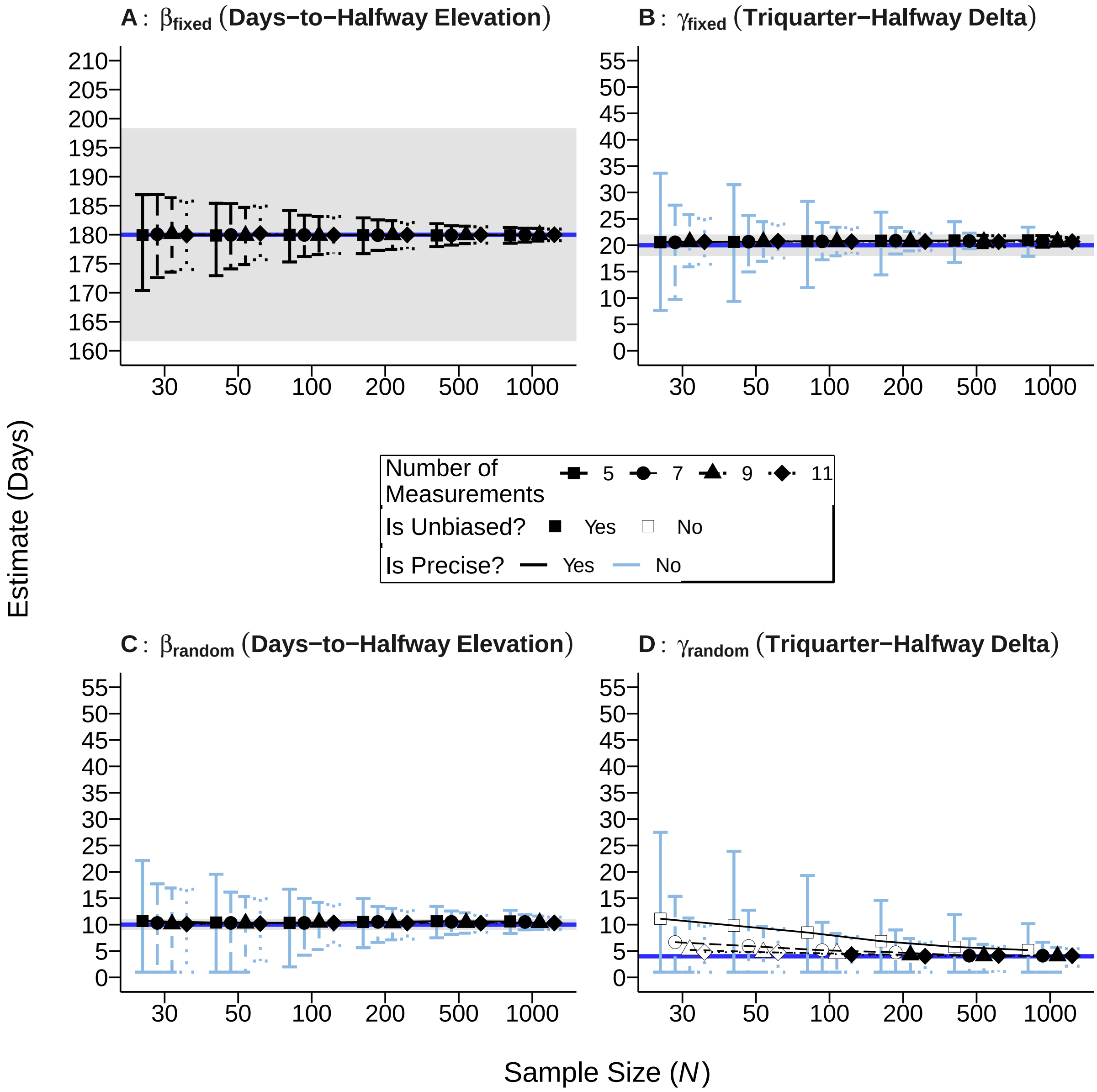
To show that definition variables can eliminate the error incurred by time-unstructured data, I ran an additional set of simulations. In these simulations, time-unstructured data characterized by a slow response rate were analyzed with a structured latent growth curve model equipped with definition variables (see Appendix I for the corresponding code). Number of measurements and sample size were manipulated as in Experiment 3, thus yielding 24 cells (i.e., 4[number of measurements: 5, 7, 9, 11] x 6[sample size: 30, 50, 100, 200, 500, 1000]). As in all previous simulation experiments, I only present the results for the day-unit parameters because the results for the Likert-unit parameters were largely negligible (for Likert-unit bias/precision plots, see Appendix F). Similar to the results for convergence success rates obtained in all other simulation experiments, convergence success rates across all cells were always above 90%, with the specific values presented in Table G.4.[20]

Figure 4.10 shows the bias/precision plots that result from using definition variables to model time-unstructured data characterized by a slow response rate. In comparing the bias/precision plot of Figure 4.10 to that of Figure 4.6, model performance improves in the following four ways:

[20]It should be noted that convergence times increased by approximately eightfold when definition variables were used.

Figure 4.10

Bias/Precision Plots for Day-Unit Parameters When Using Definition Variables To Model Time-Unstructured Data Characterized by a Slow Response Rate



Note. Panel A: Bias/precision plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Bias/precision plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Bias/precision plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Bias/precision plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or

longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table H.3 for specific values estimated for each parameter and Table 4.8 for ω^2 effect size values.

Table 4.8

Partial ω^2 Values for Manipulated Variables With Time-Unstructured Data Characterized by a Slow Response Rate With a Model Using Definition Variables in Experiment 3

	Effect		
Parameter	NM	S	NM x S
β_{fixed} (Figure 4.10A)	0.00	0.02	0.00
β_{random} (Figure 4.10B)	0.14	0.27	0.03
γ_{fixed} (Figure 4.10C)	0.25	0.12	0.07
γ_{random} (Figure 4.10D)	0.18	0.03	0.01

NM = number of measurements (5, 7, 9, 11), S = sample size (30, 50, 100, 200, 500, 100), NM x S = interaction between number of measurements and sample size.

1) Bias in the estimation of the fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 4.6A) almost entirely disappears when using definition variables (Figure 4.10A).

2) Bias in the estimation of the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}; Figure 4.6B) almost entirely disappears when using definition variables (Figure 4.10B).

3) Bias in the estimation of the random-effect days-to-halfway elevation parameter (β_{random}; Figure 4.6C) almost entirely disappears when using definition variables (Figure 4.10C).

4) Bias in the estimation of the random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 4.6D) returns to levels observed with time-structured data (see Figure 4.4A) with definition variables. Precision also decreases (especially with five

measurements) when using definition variables (Figure 4.10C).

Therefore, given the improvements in the estimation of each day-unit parameter that follow from using definition variables, latent variable models, by default, should use definition variables to improve model performance when data are time unstructured.

4.3 Summary of Experiment 3

I designed Experiment 3 to investigate whether model performance decreased as time structuredness decreased. Across all manipulated levels of time structuredness, the greatest improvements in model performance result from using either seven measurements with $N \geq 200$ and nine measurements with $N \leq 100$. Importantly, although the measurement number/sample size pairings that result in the greatest improvements in model performance do not change as time structuredness decreases, the absolute level of model performance itself decreases. In using the same measurement number/sample size pairing across all levels of time structuredness, precision slightly increases and, more importantly, bias decreases such that it is constant; that is, the decrease in bias cannot be avoided by using increasing measurement number and/or sample size. Given that data are unlikely to be time structured, then the decrease in model performance seems inevitable. Fortunately, the decrease in model performance that results from time-unstructured data can be avoided by using definition variables in latent growth curve models, which I show to be the case by in an additional set of simulations. Therefore, the greatest improvements in model performance result from using either seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ and, definition variables should be used to prevent model performance from decreasing as time structuredness decreases.

5 General Discussion

In systematically reviewing the simulation literature, I found that studies rarely conducted comprehensive investigations into the effects of longitudinal design and analysis factors on model performance with nonlinear patterns of change. Specifically, few studies examined three-way interactions between any of the following four variables: 1) measurement spacing, 2) number of measurements, 3) sample size, and 4) time structuredness. Given that longitudinal designs are necessary for understanding the temporal dynamics of psychological processes (for a more detailed explanation, see Appendix A), it is important that researchers understand how longitudinal design and analysis factors affect the performance of longitudinal analyses. Therefore, to address these gaps in the literature, I designed three simulation experiments.

In each simulation experiment, a logistic pattern of change (i.e., s-shaped change pattern) was modelled under conditions that varied in nature of change (i.e., shape of the logistic curve), measurement number, sample size, and time structuredness.[21] To fit a logistic function where each parameter could be meaningfully interpreted, each simulation experiment used a structured latent growth model to estimate nonlinear change (for a detailed explanation, see Appendix D).

To investigate the effects of longitudinal design and analysis factors on model performance, my simulation experiments examined the accuracy with which each logistic function parameter was estimated. In computing the estimation accuracy of each parameter, two questions were of importance: 1) How well was the parameter estimated on

[21] Importantly, no simulation experiment manipulated more than three variables at once so that results would not be too difficult to understand (Halford et al., 2005).

average (bias) and 2) what was a range of values that could be expected for an estimate from the output of a single model (precision). Thus, model performance was the combination of bias and precision, and these two metrics were computed for each logistic function parameter. To succinctly summarize each experiment, I have created Table 5.1. Each row of Table 5.1 contains a summary of a simulation experiment.

In Experiment 1, I was interested in answering two questions: 1) Does placing measurements near periods of change increase model performance and 2) how should measurements be spaced when the nature of change is unknown. To answer these two questions, I manipulated measurement spacing, number of measurements, and nature of change (i.e., shape of the s-shaped curve). With respect to the first question, the results of Experiment 1 suggest that model performance increases when measurements are placed closer to periods of change (see section discussing measurement spacing). With respect to the second question, the results of Experiment 1 suggest that measurements should be spaced equally over time when the nature of change is unknown (see section discussing measurement spacing when the nature of change is unknown).

In Experiment 2, I was interested in the measurement number/sample size pairings needed to obtain high model performance (i.e., low bias, high precision) under different spacing schedules. To answer this question, I manipulated measurement spacing, measurement number, and sample size. Although no manipulated measurement number/sample size pairing results in high model performance (low bias, high precision) of all parameters, moderate measurement numbers and sample sizes often yield low bias and the largest improvements in model performance. For all spacing schedules (except middle-and-extreme spacing), the largest improvements in model performance result from using either either seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$. The results for middle-and-extreme spacing are largely a byproduct of the nature of change used in Experiment 2, and so are of little value to emphasize.

Table 5.1
Summary of Each Simulation Experiment

Simulation Exeriment	Independent Variables	Main Results
Experiment 1	Spacing of measurements Number of measurements Nature of change	• Model performance is higher when measurements are placed closer to periods of change • Measurements should be spaced equally when the nature of change is unknown
Experiment 2	Spacing of measurements Number of measurements Sample size	• The greatest improvements in model performance result from using either seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$
Experiment 3	Number of Measurements Sample size Time structuredness	• The greatest improvements in model performance across all time structuredness levels result from using either seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ • Use definition variables to prevent model performance from decreasing as time structuredness decreases

In Experiment 3, I was interested in examining how time structuredness affected model performance. To answer this question, I manipulated measurement spacing, measurement number, and time structuredness. Although the measurement number/sample size pairings that result in the greatest improvements in model performance are the same as in Experiment 2, two results suggest that model performance decreases as time structuredness decreases. First, precision decreases as time structuredness decreases. That is, precision decreases as response patterns of participants become increasingly dissimilar. Second, and more concerning, bias decreases as time structuredness decreases regardless of the measurement number or sample size. That is, as response patterns of participants become increasingly dissimilar, bias increases across all measurement number/sample size pairings.

Importantly, the decrease in model performance that results as time structuredness decreases can be prevented by using a latent growth curve model with definition variables. By default. latent growth curve models assume an identical response pattern for all participants (i.e., time-structured data). Definition variables can be used in latent growth curve models to allow individual response patterns to be modelled (Mehta & Neale, 2005; Mehta & West, 2000). In an additional set of simulations (see section on definition variables), I generate time-unstructured data and analyze the data with a structured latent growth curve model that has definition variables. When definition variables are used, the decrease in model performance that results from a decrease in time structuredness disappears. Therefore, to obtain the largest improvements in model performance, either seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ must be used and, importantly, the latent growth curve model must use definition variables.

In summary, the results of my simulation experiments are the first (to my knowledge) to provide specific measurement number and sample size recommendations needed to accurately model nonlinear change over time. Importantly, although previous studies have investigated the effects of some longitudinal design and analysis factors on model performance with nonlinear patterns of change, the results of these studies are limited because they either use unrealistic fixed-effects models (e.g., Finch, 2017), use models with with non-meaningful parameter interpretations (e.g., Fine et al., 2019; J. Liu et al., 2022), or use unrealistic model fitting procedures (Finch, 2017). Additionally, I developed novel and replicable procedures for creating spacing schedules (see Appendix C) and simulating time-unstructured data (see time structuredness).

The sections that follow will discuss the limitations of the current simulation experiments and avenues for future research. The scope of the discussion will then expand to include issues concerning the nature of longitudinal designs, the importance of modelling nonlinear change, and suggestions for modelling such change.

5.1 Limitations and Future Directions

Recall that in designing each simulation experiment, I decided to manipulate no more than three variables so that results could be readily understood (Halford et al., 2005). Although limiting the number of independent variables has its advantages, there are a number of non-manipulated variables could have influenced the results. In the sections that follow, I review the possible impact of not manipulating these variables.

5.1.1 Cutoff Values for Bias and Precision

In simulation research, cutoff values for parameters are often set to a percentage of a parameter's population value (e.g., Muthén et al., 1997) for two reasons. First, cutoff values are needed to allow bias and precision to be categorized so that results can be clearly presented. In the current set of simulation experiments, cutoff values for bias and precision were set to 10% of the parameter's population value (Muthén et al., 1997). If a parameter estimate was outside a 10% error margin, then estimation was considered biased. If an error bar whisker length was longer than 10% of the parameter's population value, then estimation was considered imprecise. Therefore, using cutoff values allows categorical decisions to be made modelling performance.

Second, cutoff values are needed to allow results from different simulation studies to be meaningfully compared. If another study uses a cutoff value of 15%, then the results of this study become difficult to compare with the results of the current simulation experiments because each study uses different cutoff standards. Therefore, it is important that simulation studies use a common standard of 10% (Muthén et al., 1997)—as I have done in my simulation experiments. Although simulation studies use cutoff values to simplify results and allow meaningful comparisons of results, it is also important that cutoff values themselves represent meaningful boundary values.

Given the need for using cutoff values in simulation research, it was necessary to do so in my experiments. Although several methods exist for setting cutoff values that each have their advantages and disadvantages, I decided to choose a method that aligned with the conventions of simulation research. Thus, I used a percentage-based cutoff rule (Muthén et al., 1997). Like other methods for setting cutoff values, the percentage-based cutoff method has limitations and I discuss these limitations in the paragraphs that follow.

In simply defining cutoff values as a percentage of a population value, cutoff values can lead to problematic conclusions. As a simple example, consider a scenario where a beverage company wants to produce a caffeinated drink that can only increase heart rate and body temperature by a certain amount. Specifically, neither heart rate nor body temperature can increase by 10% of their resting values. Given that, for males and females, any value below 70 and 80, respectively, constitutes a healthy resting heart rate (Nanchen, 2018), a 10% increase would translate to an increase of 7 and 8 beats per minute, which is arguably less than the increase in heart rate caused from walking (e.g., Whitley & Schoene, 1987). Thus, requiring that a caffeinated drink not increase resting heart rate by a value equal to or greater than 10% appears to be a responsible stipulation. Unfortunately, setting a 10%-cutoff rule for body temperature allows for far less desirable outcomes than a 10% cutoff for heart rate. Using a typical body temperature of 37 °C for resting body temperature, a 10%-cutoff would allow for a change in body temperature of 3.7 °C. Given that deviations of less than 3.7 °C from resting body temperature can lead to physiological impairments and even death (Moran & Mendal, 2002), restricting the caffeinated drink to not increase body temperature by 10% of its resting value is unwise. Therefore, a percentage cutoff rule can fail to create useful cutoff values by overlooking the underlying nature of the variable in question.

In the current simulation experiments, the percentage-cutoff rule may have led to overly pessimistic conclusions about model performance. As an example, consider the estimation of the random effect parameters. In each simulation experiment, no measurement number/sample size pairing resulted in high model performance (low bias, high precision) of any random-effect parameter[22] Specifically, the random-effect day-unit parameters were never modelled precisely with any measurement number/sample size pairing. Although the lack of precise estimation for the random-effect day-unit parameters is concerning, the result may be a byproduct of having used conventional standards for precision. For a given parameter, the cutoff value used to deem estimation precise was proportional to the population value set for that parameter. Specifically, the cutoff values for precision (and bias) were set to 10% of the parameter's population value (Muthén et al., 1997)—as is

[22]It should be mentioned that low bias was obtained from using moderate measurement number/sample size pairings.

suggested by the literature. In setting the cutoff value to a percentage of the parameter's population value, the margin of error becomes a function of the population value: Large population values have large margins of error and small population values have small margins of error. Given that the random-effect parameters had the smallest population values (e.g., 10.00, 4.00, and 0.05) and that even the largest measurement number/sample size pairing of 11 measurements with $N = 1000$ did not model with high precision, it is conceivable that the associated 10%-error margins (e.g., 1.00, 0.04, and 0.005) may have been too small.

Future research could consider using more useful cutoff values. One way to set useful cutoff values in simulation experiments is to contextualize cutoff values with respect to a real-world phenomenon. Using smallest effect sizes of interest offers one way to contextualize cutoff values (Lakens, 2017; Lakens et al., 2018). Introduced to improve null-hypothesis significance testing, a smallest effect size of interest constitutes the smallest effect size above which a researcher considers an observed effect meaningful (Lakens, 2017). Instead of testing the typical zero-effect null hypothesis, a researcher can specify a smallest effect size of interest as the null hypothesis. Using a smallest effect size of interest (in tandem with equivalence testing), a researcher can more definitively conclude whether an effect is trivially small or not and, consequently, be less likely to incorrectly dismiss an effect as nonexistent. Thus, smallest effect sizes of interest allow researchers to make more meaningful conclusions. Although the current simulation experiments did not employ significance testing, the cutoff values used to determine whether estimation was biased and precise could be improved in future research by treating them as smallest cutoff values of interest. By replacing the current percentage-based cutoff values with smallest cutoff values of interest for each parameter, conclusions are likely to become more meaningful because cutoff values are contextualized with respect to real-world phenomena.

One effective way to determine smallest cutoff values of interest in future research would be to use anchor-based methods (Anvari & Lakens, 2021). As an example, I detail a two-step procedure for how an anchor-based method could be used to determine a cutoff value for a the Likert-unit parameter of the fixed-effect baseline parameter (θ_{fixed}). First, a survey for some Likert-unit variable such as job satisfaction could be given at two time points to employees. Importantly, after completing the survey at the second time point, employees would also indicate how much job satisfaction changed by answering an anchor question (e.g., "Job satisfaction increased/decreased by a little, increased/decreased a lot, or did not change."). Second, a smallest cutoff value of interest would need to be computed. Given that the fixed-effect baseline parameter (θ_{fixed}) represents the starting value, then employees that indicated no change in job satisfaction could be said to still be at baseline and their data could be used to compute a smallest effect size of interest for the baseline parameter(θ_{fixed}). Specifically, the difference in job satisfaction between the two time points could be calculated for employees that indicated no change. Therefore, using the anchor-based method, the smallest cutoff value of interest for the fixed-effect baseline parameter (θ_{fixed}) is the mean change in some Likert-unit variable—job satisfaction in the current example—from respondents that indicate no change.[23]

[23]If the mean observed change in job satisfaction from employees that indicate no change is a near-zero value, using this value as a smallest effect-size of interest for the fixed-effect baseline parameter (θ_{fixed}) would likely be too conservative. In such situations, the smallest effect-size of interest for the fixed-effect baseline parameter (θ_{fixed}) could be determined by computing the mean change in job satisfaction from employees that indicate a small change (i.e., 'little increase/decrease), as it could be said that these employees have slightly moved away from baseline.

5.1.2 External Validity of Simulation Experiments

In the current set of simulation experiments, data were were generated under ideal conditions in three ways. First, the current simulation experiments always assumed complete data (i.e., 100% response rate). Unfortunately, researchers rarely obtain complete data and, instead, have some amount of data that are missing. One investigation estimated that, using a sample of 300 articles published over a period of three years, 90% of articles had missing data, with each study estimated to have over 30% of data points missing (McKnight et al., 2007, Chapter 1). Perhaps even more concerning, missing data often compound over time (Newman, 2003).[24] Future research could simulate more realistic conditions for response rates in longitudinal designs, missing data could be set to increase—either linearly or nonlinearly—over time under three types of commonly simulated missing data mechanisms: 1) missing data are random, 2) missing data depend on the value of another variable, and 3) missing data depend on their own values (Newman, 2009).

Second, the current simulation experiments assumed measurement invariance over time. That is, at each time point, the manifest variable was assumed to be measured with the same measurement model—specifically, aspects of the measurement model such as factor loadings, intercepts, and error variances were assumed to remain constant over time (Mellenbergh, 1989; Vandenberg & Lance, 2000). For a longitudinal design, it is important that the measurement of a latent variable meet the conditions for invariance so that change over time can be meaningfully interpreted. As an example, consider a situation

[24] It should be noted that great recommendations exist on increasing response rate. In fact, an entire book of recommendations exists on this issue (see Dillman et al., 2014).

where a researcher measures some latent variable over time such as job satisfaction using a four-item survey where each item measures some component of job satisfaction on a Likert scale (range of 1–5). If the loadings of a specific item change over time, then the response values from participants cannot be meaningfully interpreted. For example, if a participant gives the same answers to each item across two time points but factor loadings of any item(s) change between the two time points, then their job satisfaction scores between the time points will, counterintuitively, be different. Thus, even though job satisfaction did not change over time, changes in the measurement model of job satisfaction caused the observed scores to be different. Unfortunately, measurement invariance is seldom observed (Van De Schoot et al., 2015; Vandenberg & Lance, 2000) because measurement model components often change over time (e.g., E. I. Fried et al., 2016). Thus, it can be argued that it is more realistic to assume measurement non-invariance. To simulate measurement non-invariance, future research could generate data such that aspects of measurement models change over time (e.g., Kim & Willson, 2014b).

Third, the current simulations assumed error variances in the observed variables to be constant and uncorrelated over time. Unfortunately, error variances over time are likely to correlate with each other and be nonconstant or heterogeneous (Bliese & Ployhart, 2002; Blozis & Harring, 2018; Braun et al., 2013; DeShon et al., 1998; Ding et al., 2016; Goldstein et al., 1994; Lester et al., 2019). Future research could simulate more realistic error variance structures by generating errors to correlate with each other and to decrease over time—as observed in a longitudinal analysis of fatigue (Lang et al., 2018).

5.1.3 Simulations With Other Longitudinal Analyses

Given that researchers are often interested in investigating questions outside of modelling a nonlinear pattern of change, longitudinal analyses outside of the structured latent growth curve model used in the current simulation experiments may be used in other circumstances. Although the structured latent growth curve modelling framework used in the current simulations allows nonlinear change to be meaningfully modelled (see Appendix E), the framework cannot be used to understand all meaningful components of change. As an example, if a researcher is interested in modelling different response patterns in some variable in response to some organizational event—for instance, work engagement patterns after mergers (Seppälä et al., 2018)—a structured latent growth curve model could not meaningfully model such data because it assumes one pattern of responding. Therefore, to develop a comprehensive understanding of change over time, a variety of longitudinal analyses may be considered and it is important that future simulation research investigate the performance of these analyses. I outline four longitudinal analyses below that future simulation experiments should consider investigating.

First, discontinuous growth models are needed to model punctuated change (Bliese et al., 2020; Bliese & Lang, 2016).[25] Given that change in organizations often results from discrete events, the pattern of change is often punctuated or discontinuous (Morgeson et al., 2015). Examples of punctuated change in organizations have been observed in

[25]In the multilevel framework, discontinuous growth modelling is also referred to as piecewise hierarchical linear modelling (Raudenbush & Bryk, 2002) and multiphase mixed-effects models (Cudeck & Klebe, 2002). In the latent variable or structural equation modelling framework, discontinuous growth modelling is also referred to as piecewise growth modelling (Chou et al., 2004; Kohli & Harring, 2013). Note that spline models are technically different from discontinuous growth models because spline models cannot model vertical displacements at knot points and, thus, are models for continuous change (for a review, see Edwards & Parry, 2017).

life satisfaction after unemployment (Lucas et al., 2004), trust after betrayal (Fulmer & Gelfand, 2015), and firm performance after an economic recession (Kim & Willson, 2014a; for more examples, see Bliese & Lang, 2016). Discontinuous growth models can model punctuated change by selectively activating and deactivating growth factors—that is, assigning nonzero- and zero-value weights, respectively—after certain time points (Bliese & Lang, 2016). Therefore, given that punctuated change merits the need for discontinuous growth modelling in organizational research, future simulation studies should investigate the effects of longitudinal design and analysis factors on the performance of such models.

Second, time series models are needed to model cyclical patterns (Pickup, 2014). Technological advances such as smartphones and wearable sensors have allowed researchers to collect intensive longitudinal data sets where data are collected over at least 20 time points (Collins, 2006) with the experience sampling method (Larson & Csikszentmihalyi, 2014). With intensive longitudinal data sets, researchers are often interested in modelling cyclical patterns such as those with affect and performance (Dalal et al., 2014) and stress (Fuller et al., 2003). Time series models allow researchers to model cyclical patterns through a variety of methods (e.g., decomposition, autoregressive integrated moving average, etc.). Therefore, the rise of intensive longitudinal data made possible by technological advances merits the use of time series models, and future simulation studies should investigate the effects of longitudinal design and analysis factors affect the performance of such models.

Third, second-order growth models are needed to model measurement invariance (Hancock et al., 2001; Sayer & Cumsille, 2001). In organizational research, many variables are latent—that is, they cannot be directly observed (e.g., job satisfaction, organizational commitment, trust). Because latent variables cannot be directly measured, nomological networks[26]—correlation matrices specifying relations between the target latent variable and other variables—are constructed to develop valid measures of latent variables (Cronbach & Meehl, 1955). As discussed previously, an unfortunate phenomenon with surveys is that the accuracy with which they measure a latent variable is seldom invariant over time—that is, measurement accuracy is often non-invariant (Van De Schoot et al., 2015; Vandenberg & Lance, 2000). If measurement non-invariance is overlooked, model performance decreases (Jeon & Kim, 2020; Kim & Willson, 2014a). Fortunately, second-order latent growth curve models allow researchers to include measurement models and, thus, test for measurement invariance and estimate parameters with greater accuracy (e.g., Kim & Willson, 2014b). Therefore, given that the common occurrence of measurement non-invariance in organizational research merits the use of second-order latent growth models, future simulation studies should investigate the effects of longitudinal design and analysis factors on the performance of such models.

Fourth, growth mixture models are needed to model heterogeneous response patterns (van der Nest et al., 2020; M. Wang & Bodner, 2007). In organizations, employees are likely to respond to changes in different ways, thus exhibiting heterogeneous response

[26] Although a nomological network gives meaning to a latent variable by specifying relations with other variables, it should be noted that nomological networks have limitations in establishing validity—whether a survey measures what is purports to measure. In psychology, almost all variables psychology are correlated with each other (Meehl, 1978), and so using the correlations specified in a nomological network to establish validity is imprecise because many latent variables are likely to satisfy the network of relations. One potentially more effective method to establish validity is to first assume the existence of the latent variable and then develop theory that specifies processes by which changes in the latent variable manifest themselves in reality. Surveys can the be constructed by causatively testing whether the theorized manifestations that follow from changes in the latent variable actually emerge (for a review, see Borsboom et al., 2004).

patterns. Examples of heterogeneous response patterns have been observed in job performance patterns during organizational restructuring (Miraglia et al., 2015), work engagement patterns after mergers (Seppälä et al., 2018), and leadership development throughout training (Day & Sin, 2011). Growth mixture models allow heterogeneity in response patterns to be modelled by including a latent categorical variable that allows participants to be placed into different response category patterns (cf. Bauer, 2007). Therefore, given that heterogeneous response patterns in organizations merit the use of interest for modelling cyclical patterns with intensive longitudinal data merits the use of time series models, future simulation studies should investigate the effects of longitudinal design and analysis factors on the performance of such models.

5.2 Nonlinear Patterns and Longitudinal Research

5.2.1 A New Perpective on Longitudinal Designs for Modelling Change

The results of the current simulation experiments suggest that previous measurement number recommendations for longitudinal research need to be modified when modelling nonlinear patterns of change. Previous suggestions for conducting longitudinal research recommend that at least three measurements be used (Chan, 1998; Ployhart & Vandenberg, 2010). The requirement that a longitudinal study use at least three measurements is largely to obtain an estimate of change that is not confounded by measurement error (Rogosa et al., 1982) and allow a nonlinear pattern of change to be modelled. Unfortunately, although using at least three measurements allows a nonlinear pattern of change to be modelled, doing so provides no guarantee that a nonlinear pattern of change will be accurately modelled. The results of the current simulation experiments suggest that, at the very least, five measurements are needed to accurately model a nonlinear pattern of change. Importantly, five measurements only results in adequate model performance if the measurements are placed near periods of change. Given that organizational theories seldom delineate nonlinear patterns of change (for a rare example, see Methot et al., 2017), it is unlikely that researchers will place measurements near periods of change. In situations where researchers have little insight into the pattern of nonlinear change, the current simulation experiments suggest that at least seven measurements be used. Therefore, when researchers do not have strong theory to suggest a nonlinear pattern of change, the current simulations suggest that at least seven measurements are needed.

Although the current results suggest that seven measurements are needed to model nonlinear change, these results by no means imply that longitudinal designs with fewer measurements are of no value. Studies measuring a variable at two time points (i.e., pre-post designs) can be used to estimate meaningful anchors (Anvari & Lakens, 2021). Studies measuring change between three and seven time points can, for instance, be used to investigate causality by determining whether reverse causality occurs (Leszczensky & Wolbring, 2019). As a last point, it should be noted that studies using fewer than seven measurements may be able to provide accurate parameter estimates for nonlinear models that estimate fewer parameters than the nine parameters estimated by the model in the current simulations. If a latent variable model estimates fewer parameters, the optimization problem becomes less complex, and so it is conceivable that the convergence algorithm can find accurate parameter estimates with fewer than seven measurements.

5.2.2 Why is it Important to Model Nonlinear Patterns of Change?

For at least 30 years, research in organizational psychology has had a minimal effect on practitioners and their practices (Daft & Lewin, 1990; for a review, see Lawler

& Benson, 2022). Few practitioners–specifically, an estimated 1%—read journal articles (Rynes, Colbert, et al., 2002), which is accompanied by a poor understanding by managers of fundamental principles in organizational psychology, which has been observed across multiple countries including the Netherlands (Sanders et al., 2008), the United States (Rynes, Colbert, et al., 2002), Finland, South Korea, and Spain (Tenhiälä et al., 2014). Perhaps most unfortunate, a poor understanding of organizational psychology by managers is associated with large effects on financial and individual performance (for a review, see Rynes, Brown, et al., 2002). Additionally, an estimated 55% of practitioners are skeptical that evidence-based human resource practices can affect any positive change (Spears & Bolton, 2015). With the gap between academics and practitioners being so patently wide, some academics have cast doubt on the possibility of academic-practitioner research collaborations (Kieser & Leiner, 2009).

One factor that may contribute to the academic-practitioner gap is that research seldom provides specific recommendations to practitioners. When considering the typical organizational theory, propositions often lack any degree of specificity: They often specify non-zero linear relations between variables (Edwards & Berry, 2010). Because it is difficult to develop specific recommendations from non-zero relations, it becomes unsurprising that reviews of the organizational literature estimate 3% of human resource articles address problems faced by practitioners (Sackett & Larson, 1990) and, in reviewing of 5780 articles from 1963–2007, concluded that research is often late to address practitioner issues (Cascio & Aguinis, 2008). Thus, with organizational theories often providing vague predictions, it becomes difficult to develop specific recommendations for practitioners.

Organizational research can provide specific recommendations to practitioners by modelling nonlinear patterns of change. In modelling nonlinear change, organizational researcher can understand how processes unfold over time and when specific psychological phenomena emerge (T. R. Mitchell & James, 2001; Navarro et al., 2020). As an example of the usefulness of modelling nonlinear change, Vancouver et al. (2020) uses computational modelling to predict specific nonlinear patterns of self-efficacy and performance in response to different events over time. In predicting nonlinear patterns, the theory provides specific insight into how much specific events affect performance and self-efficacy, how long such effects last, and how performance and self-efficacy affect each other. Given that change over time is likely to be nonlinear (Cudeck & Harring, 2007), it is likely that many opportunities exist for organizational research to provide specific recommendations for solving problems faced by practitioners.

In summary, a concerning gap exists between academics and practitioners in organizational research whereby academics seldom address the problems faced by practitioners (e.g., Sackett & Larson, 1990) and practitioners rarely consult research when making decisions (Rynes, Brown, et al., 2002). One cause for the academic-practitioner gap is the paucity of specific recommendations provided by academics. One way that academics can reduce the gap from practitioners is to model nonlinear patterns of change over time. In modelling a nonlinear patterns of change, organizational research can develop an understanding o how processes evolve over time and when psychological phenomena emerge (T. R. Mitchell & James, 2001; Navarro et al., 2020). With an understanding of the temporal dynamics of psychological processes, organizational research can then provide specific recommendations to practitioners.

5.2.3 Suggestions for Modelling Nonlinear Change

In modelling nonlinear change, researchers can either do so using the multilevel or latent growth curve framework. Although the multilevel and latent growth curve frameworks return identical results under many conditions (e.g., Bauer, 2003), researchers should consider using the latent growth curve framework over the multilevel framework for two reasons. First, the multilevel framework encounters convergence problems when specifying nonlinear models, and the frequency of convergence problems increases with the number of random-effect parameters (for a review, see McNeish & Bauer, 2020). Second, the latent variable framework allows data to be more realistically modelled than the multilevel approach thanks to, in large part, its ability to include measurement models to investigate phenomena such as measurement invariance (Hancock et al., 2001; Sayer & Cumsille, 2001).

In modelling nonlinear change, researchers should prioritize the interpretability of their models so that results can be more easily applied. As an example, the structured latent growth curve model used in the current simulation experiments provides a meaningful representation of logistic pattern of change. In the current simulations, the number of days needed to reach the halfway- and triquarter-halfway elevation points (among other parameters) were estimated.[27] To add another level of meaning, a latent categorical variable can be added to the model to create a growth mixture model (van der Nest et al., 2020). Using a growth mixture model, not only can nonlinear change be defined in a meaningful way, but response groups can be modelled and people can be categorized into

[27] Note that parameters of nonlinear functions can be reparameterized to estimate other meaningful aspects of a curve (K. J. Preacher & Hancock, 2015).

the groups based on their individual pattern of change. Thus, in prioritizing the meaning of statistical models, the current example shows how heterogeneous logistic response patterns can be meaningfully modelled and how frequently each pattern occurs.

5.3 Conclusion

Investigating nonlinear patterns of change is a growing area of organizational research. By understanding nonlinear patterns of change, organizational research can develop a more nuanced understanding of temporal dynamics and provide practitioners with more specific recommendations. The simulation experiments conducted in my dissertation contribute to this goal by providing boundary conditions for model performance.

References

Adolf, J. K., & Fried, E. I. (2019). Ergodicity is sufficient but not necessary for group-to-individual generalizability. *Proceedings of the National Academy of Sciences*, *116*(14), 6540–6541. https://doi.org/10.1073/pnas.1818675116 (cited on p. 237)

Aguinis, H., & Bakker, R. M. (2021). Time is of the essence: Improving the conceptualization and measurement of time. *Human Resource Management Review*, *31*(2), 100763. https://doi.org/10.1016/j.hrmr.2020.100763 (cited on p. 1)

Anvari, F., & Lakens, D. (2021). Using anchor-based methods to determine the smallest effect size of interest. *Journal of Experimental Social Psychology*, *96*, 104159. https://doi.org/10.1016/j.jesp.2021.104159 (cited on pp. 191, 198)

Auguie, B. (2019). *egg: Extensions for ggplot2: Custom geom, custom themes, plot alignment, labelled panels, symmetric scales, and fixed panel size* [R package version 0.4.5]. https://CRAN.R-project.org/package=egg

Aydin, B., Leite, W. L., & Algina, J. (2014). The consequences of ignoring variability in measurement occasions within data collection waves in latent growth ,odels. *Multivariate Behavioral Research*, *49*(2), 149–160. https://doi.org/10.1080/00273171.2014.887901 (cited on pp. 21, 22)

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173 (cited on p. 3)

Bauer, D. J. (2003). Estimating multilevel linear models as structural equation models. *Journal of Educational and Behavioral Statistics*, *28*(2), 135–167. https://doi.org/10.3102/10769986028002135 (cited on p. 201)

Bauer, D. J. (2007). Observations on the use of growth mixture models in psychological research. *Multivariate Behavioral Research*, *42*(4), 757–786. https://doi.org/10.1080/00273170701710338 (cited on p. 197)

Beal, D. J. (2015). Esm 2.0: State of the art and future potential of experience sampling methods in organizational research. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(1), 383–407. https://doi.org/10.1146/annurev-orgpsych-032414-111335 (cited on p. 5)

Bergman, L., & Magnusson, D. (1990). General issues about data quality in longitudinal research. In L. Bergman & D. Magnusson (Eds.), *Data quality in longitudinal research* (pp. 1–31). Cambridge University Press. https://bit.ly/3i1CGny. (Cited on p. 7)

Birkhoff, G. D. (1931). Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, *17*(12), 656–660. https://doi.org/10.1073/pnas.17.2.656 (cited on p. 235)

Bliese, P. D., Kautz, J., & Lang, J. W. B. (2020). Discontinuous growth models: Illustrations, recommendations, and an R function for generating the design matrix. In Y. Griep & S. D. Hansen (Eds.), *Handbook on the temporal dynamics of organizational behavior* (pp. 319–350). Edward Elgar Publishing. https://doi.org/10.4337/9781788974387.00030. (Cited on p. 194)

Bliese, P. D., & Lang, J. W. B. (2016). Understanding relative and absolute change in discontinuous growth models. *Organizational Research Methods*, *19*(4), 562–592. https://doi.org/10.1177/1094428116633502 (cited on pp. 194, 195)

Bliese, P. D., & Ployhart, R. E. (2002). Growth modeling using random coefficient models: Model building, testing, and illustrations. *Organizational Research Methods*, *5*(4), 362–387. https://doi.org/10.1177/109442802237116 (cited on p. 193)

Blozis, S. A. (2004). Structured latent curve models for the study of change in multivariate repeated measures. *Psychological Methods*, *9*(3), 334–353. https://doi.org/10.1037/1082-989x.9.3.334 (cited on pp. 253, 265)

Blozis, S. A., & Cho, Y. I. (2008). Coding and centering of time in latent curve models in the presence of interindividual time heterogeneity. *Structural Equation Modeling: A Multidisciplinary Journal*, *15*(3), 413–433. https://doi.org/10.1080/10705510802154299 (cited on p. 177)

Blozis, S. A., & Harring, J. R. (2018). Fitting nonlinear mixed-effects models with alternative residual covariance structures. *Sociological Methods & Research*, *50*(2), 531–566. https://doi.org/10.1177/0049124118789718 (cited on p. 193)

Bodenmann, G., Atkins, D. C., Schär, M., & Poffet, V. (2010). The association between daily stress and sexual activity. *Journal of Family Psychology*, *24*(3), 271–279. https://doi.org/10.1037/a0019365 (cited on p. 6)

Boker, S. M., Neale, M. C., Maes, H. H., Wilde, M. J., Spiegel, M., Brick, T. R., Estabrook, R., Bates, T. C., Mehta, P., von Oertzen, T., Gore, R. J., Hunter, M. D., Hackett, D. C., Karch, J., Brandmaier, A., Pritikin, J. N., Zahery, M., & Kirkpatrick, R. M. (2020, August). *Openmx user guide release 2.18.1*. (Cited on p. 268).

Boneau, C. A. (1960). The effects of violations of assumptions underlying the *t* test. *Psychological Bulletin*, *57*(1), 49–64. https://doi.org/10.1037/h0041412 (cited on p. 14)

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). John Wiley & Sons. https://doi.org/10.1002/9780470743386.ch7. (Cited on p. 33)

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111*(4), 1061–1071. https://doi.org/10.1037/0033-295x.111.4.1061 (cited on p. 196)

Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology*, *100*(2), 431–449. https://doi.org/10.1037/a0038047 (cited on pp. 32, 40)

Braun, M. T., Kuljanin, G., & DeShon, R. P. (2013). Spurious results in the analysis of longitudinal data in organizational research. *Organizational Research Methods*, *16*(2), 302–330. https://doi.org/10.1177/1094428112469668 (cited on p. 193)

Brown, M. B., & Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association*, *69*(346), 364–367. https://doi.org/10.1080/01621459.1974.10482955 (cited on p. 47)

Browne, M. W. (1993). Structured latent curve models. In C. M. Cuadras & C. R. Rao (Eds.), *Multivariate analysis: Future directions 2* (1st ed., pp. 171–197). Elsevier. https://doi.org/10.1016/B978-0-444-81531-6.50016-7. (Cited on pp. 253, 254, 260, 265)

Browne, M. W., & du Toit, S. H. C. (1991). Models for learning data. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change: Recent advances,* *unanswered questions, future directions* (pp. 47–68). American Psychological Association. https://doi.org/10.1037/10099-004. (Cited on pp. 254, 260)

Burchinal, M., & Appelbaum, M. I. (1991). Estimating individual developmental functions: Methods and their assumptions. *Child Development*, *62*(1), 23–42. https://doi.org/10.2307/1130702 (cited on p. 7)

Cascio, W. F., & Aguinis, H. (2008). Research in industrial and organizational psychology from 1963 to 2007: Changes, choices, and trends. *Journal of Applied Psychology*, *93*(5), 1062–1081. https://doi.org/10.1037/0021-9010.93.5.1062 (cited on p. 199)

Chan, D. (1998). The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal mean and covariance structures analysis (lmacs) and multiple indicator latent growth modeling (mlgm). *Organizational Research Methods*, *1*(4), 421–483. https://doi.org/10.1177/109442819814004 (cited on p. 197)

Chen, C. X., Martin, M., & Merchant, K. A. (2014). The effect of measurement timing on the information content of customer satisfaction measures. *Management Accounting Research*, *25*(3), 187–205. https://doi.org/10.1016/j.mar.2013.12.003 (cited on p. 9)

Choe, C. H. (2005). *Computational ergodic theory*. Springer-Verlag. https://doi.org/10.1007/b138894. (Cited on p. 235)

Chou, C.-P., Yang, D., Pentz, M. A., & Hser, Y.-I. (2004). Piecewise growth curve modeling approach for longitudinal prevention study. *Computational Statistics & Data Analysis*, *46*(2), 213–225. https://doi.org/10.1016/s0167-9473(03)00149-x (cited on p. 194)

Cohen, A. (1993). Organizational commitment and turnover: A meta-analysis. *Academy of Management Journal*, *36*(5), 1140–1157. https://doi.org/10.2307/256650 (cited on p. 6)

Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, *112*(4), 558–577. https://doi.org/10.1037/0021-843x.112.4.558 (cited on pp. 3, 5, 8)

Cole, D. A., & Maxwell, S. E. (2009). Statistical methods for risk-outcome research: Being sensitive to longitudinal structure. *Annual Review of Clinical Psychology*, *5*(1), 71–96. https://doi.org/10.1146/annurev-clinpsy-060508-130357 (cited on p. 9)

Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology*, *57*(1), 505–528. https://doi.org/10.1146/annurev.psych.57.102904.190146 (cited on pp. 8, 195)

Coulombe, P. (2016). *Partially and fully time-unstructured residual variance-covariance matrices in growth curve modeling: Consequences of ignoring variability in times of assessment* (No. 10155460) [Doctoral dissertation, University of New Mexico]. ProQuest Dissertations and Theses Global. https://bit.ly/3WDWzA4. (Cited on pp. 21, 22)

Coulombe, P., Selig, J. P., & Delaney, H. D. (2016). Ignoring individual differences in times of assessment in growth curve modeling. *International Journal of Behavioral*

Development, *40*(1), 76–86. https://doi.org/10.1177/0165025415577684 (cited on pp. 8, 11, 18, 21, 22, 33, 35, 37, 100, 136–138)

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. https://doi.org/10.1037/h0040957 (cited on p. 196)

Cudeck, R., & Harring, J. R. (2007). Analysis of nonlinear patterns of change with random coefficient models. *Annual Review of Psychology*, *58*(1), 615–637. https://doi.org/10.1146/annurev.psych.58.110405.085520 (cited on pp. 6, 23, 200)

Cudeck, R., & Klebe, K. J. (2002). Multiphase mixed-effects models for repeated measures data. *Psychological Methods*, *7*(1), 41–63. https://doi.org/10.1037/1082-989x.7.1.41 (cited on p. 194)

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, *62*(1), 583–619. https://doi.org/10.1146/annurev.psych.093008.100356 (cited on pp. 4, 238)

Daft, R. L., & Lewin, A. Y. (1990). Can organization studies begin to break out of the normal science straitjacket? an editorial essay. *Organization Science*, *1*(1), 1–9. https://doi.org/10.1287/orsc.1.1.1 (cited on p. 198)

Dalal, R. S., Bhave, D. P., & Fiset, J. (2014). Within-person variability in job performance. *Journal of Management*, *40*(5), 1396–1436. https://doi.org/10.1177/0149206314532691 (cited on pp. 1, 195)

Day, D. V., & Sin, H.-P. (2011). Longitudinal tests of an integrative model of leader development: Charting and understanding developmental trajectories. *The Leadership Quarterly*, *22*(3), 545–560. https://doi.org/10.1016/j.leaqua.2011.04.011 (cited on pp. 10, 197, 238)

DeShon, R. P., Ployhart, R. E., & Sacco, J. M. (1998). The estimation of reliability in longitudinal models. *International Journal of Behavioral Development*, *22*(3), 493–515. https://doi.org/10.1080/016502598384243 (cited on p. 193)

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. John Wiley & Sons. https://bit.ly/3i20CHB. (Cited on pp. 7, 137, 192)

Ding, C. G., Jane, T.-D., Wu, C.-H., Lin, H.-R., & Shen, C.-K. (2016). A systematic approach for identifying level-1 error covariance structures in latent growth modeling. *International Journal of Behavioral Development*, *41*(3), 444–455. https://doi.org/10.1177/0165025416647800 (cited on p. 193)

Dormann, C., & Griffin, M. A. (2015). Optimal time lags in panel studies. *Psychological Methods*, *20*(4), 489–505. https://doi.org/10.1037/met0000041 (cited on p. 8)

Dormann, C., & van de Ven, B. (2014). Timing in methods for studying psychosocial factors at work. In M. F. Dollard, A. Shimazu, R. B. Nordin, P. Brough, & M. R. Tuckey (Eds.), *Psychosocial factors at work in the asia pacific* (pp. 89–116). Springer Dordrecht. https://doi.org/10.1007/978-94-017-8975-2_4. (Cited on pp. 8, 9)

Dowle, M., & Srinivasan, A. (2022). *data.table: Extension of ‘data.frame‘* [R package version 1.14.6]. https://CRAN.R-project.org/package=data.table

Edwards, J. R., & Berry, J. W. (2010). The presence of something or the absence of nothing: Increasing theoretical precision in management research. *Organizational Research Methods*, *13*(4), 668–689. https://doi.org/10.1177/1094428110380467 (cited on p. 199)

Edwards, J. R., & Parry, M. E. (2017). On the use of spline regression in the study of congruence in organizational research. *Organizational Research Methods*, *21*(1), 68–110. https://doi.org/10.1177/1094428117715067 (cited on p. 194)

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*(2), 121–138. https://doi.org/10.1037/1082-989x.12.2.121 (cited on p. 7)

Finch, W. (2017). Investigation of parameter estimation accuracy for growth curve modeling with categorical indicators. *Methodology*, *13*(3), 98–112. https://doi.org/10.1027/1614-2241/a000134 (cited on pp. 8, 21, 22, 133, 134, 186, 187)

Fine, K., & Grimm, K. (2020). Examination of nonlinear and functional mixed-effects models with nonparametrically generated data. *Multivariate Behavioral Research*, 1–18. https://doi.org/10.1080/00273171.2020.1754746 (cited on pp. 18, 21–23)

Fine, K., Suk, H., & Grimm, K. (2019). An examination of a functional mixed-effects modeling approach to the analysis of longitudinal data. *Multivariate Behavioral Research*, *54*(4), 475–491. https://doi.org/10.1080/00273171.2018.1520626 (cited on pp. 8, 18, 21, 22, 134, 187)

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National*

Academy of Sciences, *115*(27). https://doi.org/10.1073/pnas.1711978115 (cited on pp. 5, 237, 238)

Fisher, C. D. (2008). What if we took within-person variability seriously? *Industrial and Organizational Psychology*, *1*(2), 185–189. https://doi.org/10.1111/j.1754-9434.2008.00036.x (cited on p. 1)

Fournier, M., d'Arripe-Longueville, F., Rovere, C., Easthope, C. S., Schwabe, L., El Methni, J., & Radel, R. (2017). Effects of circadian cortisol on the development of a health habit. *Health Psychology*, *36*(11), 1059–1064. https://doi.org/10.1037/hea0000510 (cited on p. 7)

Fried, E. I., van Borkulo, C. D., Epskamp, S., Schoevers, R. A., Tuerlinckx, F., & Borsboom, D. (2016). Measuring depression over time . . . or not? Lack of unidimensionality and longitudinal measurement invariance in four common rating scales of depression. *Psychological Assessment*, *28*(11), 1354–1367. https://doi.org/10.1037/pas0000275 (cited on p. 193)

Fried, Y., & Slowik, L. H. (2004). Enriching goal-setting theory with time: An integrated approach. *Academy of Management Review*, *29*(3), 404–422. https://doi.org/10.5465/amr.2004.13670973 (cited on p. 1)

Fuller, J. A., Stanton, J. M., Fisher, G. G., Spitzmüller, C., Russell, S. S., & Smith, P. C. (2003). A lengthy look at the daily grind: Time series analysis of events, mood, stress, and satisfaction. *Journal of Applied Psychology*, *88*(6), 1019–1033. https://doi.org/10.1037/0021-9010.88.6.1019 (cited on p. 195)

Fulmer, C. A., & Gelfand, M. J. (2015). Trust after violations: Are collectivists more or less forgiving? *Journal of Trust Research*, *5*(2), 109–131. https://doi.org/10.1080/21515581.2015.1051050 (cited on p. 195)

Gasimova, F., Robitzsch, A., Wilhelm, O., & Hülür, G. (2014). A hierarchical bayesian model with correlated residuals for investigating stability and change in intensive longitudinal data settings. *Methodology*, *10*(4), 126–137. https://doi.org/10.1027/1614-2241/a000083 (cited on pp. 18, 21, 22)

George, J. M., & Jones, G. R. (2000). The role of time in theory and theory building. *Journal of Management*, *26*(4), 657–684. https://doi.org/10.1177/014920630002600404 (cited on p. 1)

Goldstein, H., Healy, M. J. R., & Rasbash, J. (1994). Multilevel time series models with applications to repeated measures data. *Statistics in Medicine*, *13*(16), 1643–1655. https://doi.org/10.1002/sim.4780131605 (cited on p. 193)

Griffeth, R. W., Hom, P. W., & Gaertner, S. (2000). A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. *Journal of Management*, *26*(3), 463–488. https://doi.org/10.1177/014920630002600305 (cited on p. 6)

Grimm, K., & Widaman, K. (2010). Residual structures in latent growth curve modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(3), 424–442. https://doi.org/10.1080/10705511.2010.489006 (cited on p. 7)

Halford, G. S., Baker, R., McCredden, J. E., & Bain, J. D. (2005). How many variables can humans process? *Psychological Science*, *16*(1), 70–76. https://doi.org/10.1111/j.0956-7976.2005.00782.x (cited on pp. 30, 40, 101, 141, 183, 187)

Hamaker, E. L. (2012). Why researchers should think “within-person”: A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of studying research methods for daily life* (pp. 43–61). The Guilford Press. https://bit.ly/3jxfq14. (Cited on p. 238)

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2005). Statistical modeling of the individual: Rationale and application of multivariate stationary time series analysis. *Multivariate Behavioral Research*, *40*(2), 207–233. https://doi.org/10.1207/s15327906mbr4002_3 (cited on p. 5)

Hancock, G., Kuo, W.-L., & Lawrence, F. (2001). An illustration of second-order latent growth models. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(3), 470–489. https://doi.org/10.1207/s15328007sem0803_7 (cited on pp. 195, 201)

Hom, P. W., Caranikas-Walker, F., Prussia, G. E., & Griffeth, R. W. (1992). A meta-analytical structural equations analysis of a model of employee turnover. *Journal of Applied Psychology*, *77*(6), 890–909. https://doi.org/10.1037/0021-9010.77.6.890 (cited on p. 6)

Howell, D. C. (2009). *Statistical methods for psychology* (7th ed.). Wadsworth Publishing. (Cited on p. 46).

Huber, N. (2021). *ggbrace: Curly braces in ggplot* [R package version 0.1.0]. https://github.com/NicolasH2/ggbrace

Huh, D., Kaysen, D. L., & Atkins, D. C. (2015). Modeling cyclical patterns in daily college drinking data with many zeroes. *Multivariate Behavioral Research*, *50*(2), 184–196. https://doi.org/10.1080/00273171.2014.977433 (cited on p. 6)

Igic, I., Keller, A. C., Elfering, A., Tschan, F., Kälin, W., & Semmer, N. K. (2017). Ten-year trajectories of stressors and resources at work: Cumulative and chronic effects on health and well-being. *Journal of Applied Psychology*, *102*(9), 1317–1343. https://doi.org/10.1037/apl0000225 (cited on p. 238)

Jeon, M.-J., & Kim, S.-Y. (2020). Performance of second-order latent growth model under partial longitudinal measurement invariance: A comparison of two scaling approaches. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(2), 261–277. https://doi.org/10.1080/10705511.2020.1783270 (cited on p. 196)

Kenny, D. A., & Judd, C. M. (2014). Power anomalies in testing mediation. *Psychological Science*, *25*(2), 334–339. https://doi.org/10.1177/0956797613502676 (cited on p. 14)

Kieser, A., & Leiner, L. (2009). Why the rigour-relevance gap in nanagement research is unbridgeable. *Journal of Management Studies*, *46*(3), 516–533. https://doi.org/10.1111/j.1467-6486.2009.00831.x (cited on p. 199)

Kim, E. S., & Willson, V. L. (2014a). Measurement invariance across groups in latent growth modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 408–424. https://doi.org/10.1080/10705511.2014.915374 (cited on pp. 195, 196)

Kim, E. S., & Willson, V. L. (2014b). Testing measurement invariance across groups in longitudinal data: Multigroup second-order latent growth model. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(4), 566–576. https://doi.org/10.1080/10705511.2014.919821 (cited on pp. 193, 196)

Kohli, N., & Harring, J. R. (2013). Modeling growth in latent variables using a piecewise function. *Multivariate Behavioral Research*, *48*(3), 370–397. https://doi.org/10.1080/00273171.2013.778191 (cited on p. 194)

Kraft, D. (1994). Algorithm 733: Tomp–fortran modules for optimal control calculations. *ACM Transactions on Mathematical Software*, *20*(3), 262–281. https://doi.org/10.1145/192115.192124 (cited on p. 269)

Kunisch, S., Bartunek, J. M., Mueller, J., & Huy, Q. N. (2017). Time in strategic change research. *Academy of Management Annals*, *11*(2), 1005–1064. https://doi.org/10.5465/annals.2015.0133 (cited on p. 1)

Lakens, D. (2017). Equivalence tests: A practical primer for *t* tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, *8*(4), 355–362. https://doi.org/10.1177/1948550617697177 (cited on p. 190)

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269. https://doi.org/10.1177/2515245918770963 (cited on p. 190)

Lang, J. W. B., Bliese, P. D., & de Voogt, A. (2018). Modeling consensus emergence in groups using longitudinal multilevel methods. *Personnel Psychology*, *71*(2), 255–281. https://doi.org/10.1111/peps.12260 (cited on p. 193)

Larsen, R. J., & Kasimatis, M. (1990). Individual differences in entrainment of mood to the weekly calendar. *Journal of Personality and Social Psychology*, *58*(1), 164–171. https://doi.org/10.1037/0022-3514.58.1.164 (cited on p. 6)

Larson, R., & Csikszentmihalyi, M. (2014). The experience sampling method. In M. Csikszentmihalyi & R. Larson (Eds.), *Flow and the foudations of positive psychology:*

The collected works of mihaly csikszentmihalyi (pp. 21–34). Springer Netherlands. https://doi.org/10.1007/978-94-017-9088-8_2. (Cited on p. 195)

Last, F. (2022). *nonlinSims: Generates longitudinal data that follows logistic pattern* [R package version 0.0.0.9000]. https://github.com/sciarraseb/nonlinSims

Lawler, E. E., & Benson, G. S. (2022). The practitioner-academic gap: A view from the middle. *Human Resource Management Review*, *32*(1), 100748. https://doi.org/10.1016/j.hrmr.2020.100748 (cited on p. 198)

Lawrence, T. B., Winn, M. I., & Jennings, P. D. (2001). The temporal dynamics of institutionalization. *Academy of Management Review*, *26*(4), 624–644. https://doi.org/10.5465/amr.2001.5393901 (cited on pp. 1, 32)

Lester, H. F., Cullen-Lester, K. L., & Walters, R. W. (2019). From nuisance to novel research questions: Using multilevel models to predict heterogeneous variances. *Organizational Research Methods*, *24*(2), 342–388. https://doi.org/10.1177/1094428119887434 (cited on p. 193)

Leszczensky, L., & Wolbring, T. (2019). How to deal with reverse causality using panel data? recommendations for researchers based on a simulation study. *Sociological Methods & Research*, *51*(2), 837–865. https://doi.org/10.1177/0049124119882473 (cited on p. 198)

Liu, J., & Perera, R. (2022). Estimating knots and their association in parallel bilinear spline growth curve models in the framework of individual measurement occasions. *Psychological Methods*, *27*(5), 703–729. https://doi.org/10.1037/met0000309 (cited on pp. 18, 21, 22)

Liu, J., Perera, R. A., Kang, L., Sabo, R. T., & Kirkpatrick, R. M. (2022). Obtaining interpretable parameters from reparameterized longitudinal models: Transformation matrices between growth factors in two parameter spaces. *Journal of Educational and Behavioral Statistics*, *47*(2), 167–201. https://doi.org/10.3102/10769986211052009 (cited on pp. 21, 22, 134, 187)

Liu, Y., Liu, H., Li, H., & Zhao, Q. (2015). The effects of individually varying times of observations on growth parameter estimations in piecewise growth model. *Journal of Applied Statistics*, *42*(9), 1843–1860. https://doi.org/10.1080/02664763.2015.1014884 (cited on pp. 18, 21, 22)

Loehlin, J. C., & Beaujean, A. A. (2017). *Latent variable models*. Routledge. https://doi.org/10.4324/9781315643199. (Cited on p. 37)

Lucas, R. E., Clark, A. E., Georgellis, Y., & Diener, E. (2004). Unemployment alters the set point for life satisfaction. *Psychological Science*, *15*(1), 8–13. https://doi.org/10.1111/j.0963-7214.2004.01501002.x (cited on p. 195)

Magee, C. A., Caputi, P., & Lee, J. K. (2016). Distinct longitudinal patterns of absenteeism and their antecedents in full-time australian employees. *Journal of Occupational Health Psychology*, *21*(1), 24–36. https://doi.org/10.1037/a0039138 (cited on p. 238)

Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods*, *12*(1), 23–44. https://doi.org/10.1037/1082-989x.12.1.23 (cited on pp. 2, 3, 5)

Maxwell, S. E., Cole, D. A., & Mitchell, M. A. (2011). Bias in cross-sectional analyses of longitudinal mediation: Partial and complete mediation under an autoregressive model. *Multivariate Behavioral Research*, *46*(5), 816–841. https://doi.org/10.1080/00273171.2011.606716 (cited on pp. 3, 5)

McKnight, P., McKnight, P. K., Sidani, S., & Figueredo, A. J. (2007). *Missing data: A gentle introduction*. Guilford Press. https://bit.ly/3i5of1W. (Cited on p. 192)

McNeish, D., & Bauer, D. J. (2020). Reducing incidence of nonpositive definite covariance matrices in mixed effect models. *Multivariate Behavioral Research*, *57*(2-3), 318–340. https://doi.org/10.1080/00273171.2020.1830019 (cited on pp. 28, 201)

McNeish, D., & Matta, T. (2017). Differentiating between mixed-effects and latent-curve approaches to growth modeling. *Behavior Research Methods*, *50*(4), 1398–1414. https://doi.org/10.3758/s13428-017-0976-5 (cited on p. 28)

Medaglia, J. D., Jeronimus, B. F., & Fisher, A. J. (2019). Reply to adolf and fried: Conditional equivalence and imperatives for person-level science. *Proceedings of the National Academy of Sciences*, *116*(14), 6542–6543. https://doi.org/10.1073/pnas.1820221116 (cited on p. 237)

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir karl, sir ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, *46*(4), 806–834. https://doi.org/10.1037/0022-006x.46.4.806 (cited on p. 196)

Mehta, P. D., & Neale, M. C. (2005). People are variables too: Multilevel structural equations modeling. *Psychological Methods*, *10*(3), 259–284. https://doi.org/10.1037/1082-989x.10.3.259 (cited on pp. 10, 11, 177, 186)

Mehta, P. D., & West, S. G. (2000). Putting the individual back into individual growth curves. *Psychological Methods*, *5*(1), 23–43. https://doi.org/10.1037/1082-989x.5.1.23 (cited on pp. 10, 11, 177, 186)

Mellenbergh, G. J. (1989). Item bias and item response theory. *International Journal of Educational Research*, *13*(2), 127–143. https://doi.org/10.1016/0883-0355(89)90002-5 (cited on pp. 7, 192)

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*(1), 107–122. https://doi.org/10.1007/bf02294746 (cited on pp. 252, 253)

Methot, J. R., Lepak, D., Shipp, A. J., & Boswell, W. R. (2017). Good citizen interrupted: Calibrating a temporal theory of citizenship behavior. *Academy of Management Review*, *42*(1), 10–31. https://doi.org/10.5465/amr.2014.0415 (cited on p. 198)

Mill, J. S. (2011). Of the law of universal causation. In *A system of logic, tatiocinative and inductive: Being a connected view of the principles of evidence, and the methods of scientific investigation* (pp. 392–424). Cambridge University Press. (Original work published in 1843). https://doi.org/10.1017/cbo9781139149839.021. (Cited on p. 2)

Miller, M. L., & Ferrer, E. (2017). The effect of sampling-time variation on latent growth curve models. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(6), 831–854. https://doi.org/10.1080/10705511.2017.1346476 (cited on pp. 18, 21, 22)

Miraglia, M., Alessandri, G., & Borgogni, L. (2015). Trajectory classes of job performance. *Career Development International*, *20*(4), 424–442. https://doi.org/10.1108/cdi-03-2015-0032 (cited on pp. 197, 238)

Mitchell, M. A., & Maxwell, S. E. (2013). A comparison of the cross-sectional and sequential designs when assessing longitudinal mediation. *Multivariate Behavioral*

Research, *48*(3), 301–339. https://doi.org/10.1080/00273171.2013.784696 (cited on pp. 2–4, 6)

Mitchell, T. R., & James, L. R. (2001). Building better theory: Time and the specification of when things happen. *Academy of Management Review*, *26*(4), 530–547. https://doi.org/10.5465/amr.2001.5393889 (cited on pp. 1, 8, 9, 200)

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research & Perspective*, *2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1 (cited on pp. 4, 23, 235, 237, 238)

Molenaar, P. C. M. (2008). Consequences of the ergodic theorems for classical test theory, factor analysis, and the analysis of developmental processes. SAGE Publications. https://doi.org/10.4135/9781412976589.n5. (Cited on p. 238)

Molenaar, P. C., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science*, *18*(2), 112–117. https://doi.org/10.1111/j.1467-8721.2009.01619.x (cited on pp. 4, 235, 238)

Moran, D. S., & Mendal, L. (2002). Core temperature measurement. *Sports Medicine*, *32*(14), 879–885. https://doi.org/10.2165/00007256-200232140-00001 (cited on p. 189)

Morgeson, F. P., Mitchell, T. R., & Liu, D. (2015). Event system theory: An event-oriented approach to the organizational sciences. *Academy of Management Review*, *40*(4), 515–537. https://doi.org/10.5465/amr.2012.0099 (cited on p. 194)

Murphy, D., Beretvas, S., & Pituch, K. (2011). The effects of autocorrelation on the curve-of-factors growth model. *Structural Equation Modeling: A Multidisciplinary*

Journal, *18*(3), 430–448. https://doi.org/10.1080/10705511.2011.582399 (cited on pp. 18, 21, 22)

Murre, J. M. J., & Dros, J. (2015). Replication and analysis of Ebbinghaus' forgetting curve. *PLOS ONE*, *10*(7), e0120644. https://doi.org/10.1371/journal.pone.0120644 (cited on p. 7)

Muthén, B. O., du Toit, S. H. C., & Spisic, D. (1997). *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes* [Unpublished technical report]. https://bit.ly/3Q3Yyew. (Cited on pp. 43, 187–189)

Nanchen, D. (2018). Resting heart rate: What is normal? *Heart*, *104*(13), 1048–1049. https://doi.org/10.1136/heartjnl-2017-312731 (cited on p. 188)

Navarro, J., Roe, R. A., & Artiles, M. (2015). Taking time seriously: Changing practices and perspectives in work/organizational psychology. *Journal of Work and Organizational Psychology*, *31*(3), 135–145. https://doi.org/10.1016/j.rpto.2015.07.002 (cited on p. 1)

Navarro, J., Rueff-Lopes, R., & Rico, R. (2020). New nonlinear and dynamic avenues for the study of work and organizational psychology: An introduction to the special issue. *European Journal of Work and Organizational Psychology*, *29*(4), 477–482. https://doi.org/10.1080/1359432x.2020.1794952 (cited on p. 200)

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, *15*(3), 263–280. https://doi.org/10.1002/ejsp.2420150303 (cited on p. 7)

Neuwirth, E. (2022). *RColorBrewer: ColorBrewer palettes* [R package version 1.1-3]. https://CRAN.R-project.org/package=RColorBrewer

Newman, D. A. (2009). Missing data techniques and low response rates: The role of systematic nonresponse parameters. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences* (pp. 7–36). Routledge. https://bit.ly/3VzJvuj. (Cited on p. 192)

Newman, D. A. (2003). Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods*, *6*(3), 328–362. https://doi.org/10.1177/1094428103254673 (cited on p. 192)

Newman, D. A. (2008). Missing data techniques and low response rates: The role of systematic nonresponse parameters. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences* (pp. 7–36). Routledge. https://doi.org/10.4324/9780203867266. (Cited on p. 7)

Newsom, J., & Smith, N. (2020). Performance of latent growth curve models with binary variables. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(6), 888–907. https://doi.org/10.1080/10705511.2019.1705825 (cited on pp. 21, 22)

Nixon, A. E., Mazzola, J. J., Bauer, J., Krueger, J. R., & Spector, P. E. (2011). Can work make you sick? a meta-analysis of the relationships between job stressors and physical symptoms. *Work & Stress*, *25*(1), 1–22. https://doi.org/10.1080/02678373.2011.569175 (cited on p. 5)

Okada, K. (2013). Is omega squared less biased? a ccmparison of three major effect size indices in one-way anova. *Behaviormetrika*, *40*(2), 129–147. https://doi.org/10.2333/bhmk.40.129 (cited on p. 46)

O'Laughlin, K. D., Martin, M. J., & Ferrer, E. (2018). Cross-sectional analysis of longitudinal mediation processes. *Multivariate Behavioral Research*, *53*(3), 375–402. https://doi.org/10.1080/00273171.2018.1454822 (cited on pp. 4–6)

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*(11), 776–783. https://doi.org/10.1037/h0043424 (cited on p. 7)

O'Rourke, H. P., Fine, K. L., Grimm, K. J., & MacKinnon, D. P. (2022). The importance of time metric precision when implementing bivariate latent change score models. *Multivariate Behavioral Research*, *57*(4), 1–19. https://doi.org/10.1080/00273171.2021.1874261 (cited on pp. 21–23)

Ostroff, C., Kinicki, A. J., & Clark, M. A. (2002). Substantive and operational issues of response bias across levels of analysis: An example of climate-satisfaction relationships. *Journal of Applied Psychology*, *87*(2), 355–368. https://doi.org/10.1037/0021-9010.87.2.355 (cited on p. 7)

Pan, B. (2010). Online travel surveys and response patterns. *Journal of Travel Research*, *49*(1), 121–135. https://doi.org/10.1177/0047287509336467 (cited on p. 137)

Petersen, K. E. (1983, May 12). *Ergodic theory*. Cambridge University Press. https://doi.org/10.1017/cbo9780511608728. (Cited on p. 235)

Pickup, M. (2014). *Introduction to time series analysis*. Sage Publications. https://bit.ly/3WDtdll. (Cited on p. 195)

Ployhart, R. E., & Vandenberg, R. J. (2010). Longitudinal research: The theory, design, and analysis of change. *Journal of Management*, *36*(1), 94–120. https://doi.org/10.1177/0149206309352110 (cited on pp. 1, 8, 94, 197)

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879 (cited on p. 7)

Preacher, K., Wichman, A., MacCallum, R., & Briggs, N. (2008). *Latent growth curve modeling*. SAGE Publications. https://doi.org/10.4135/9781412984737. (Cited on p. 252)

Preacher, K. J., & Hancock, G. R. (2015). Meaningful aspects of change as novel random coefficients: A general method for reparameterizing longitudinal models. *Psychological Methods*, *20*(1), 84–101. https://doi.org/10.1037/met0000028 (cited on pp. 34, 201, 254, 265)

Ram, N., Brose, A., & Molenaar, P. C. M. (2013). Dynamic factor analysis: Modeling person-specific process. In T. D. Little (Ed.), *The Oxford handbook of quantitative methods in psychology* (pp. 441–457). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199934898.013.0021. (Cited on p. 5)

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed., Vol. 1). SAGE Publications. https://bit.ly/3FZbOww. (Cited on pp. 5, 194)

Riketta, M. (2008). The causal relation between job attitudes and performance: A meta-analysis of panel studies. *Journal of Applied Psychology*, *93*(2), 472–481. https://doi.org/10.1037/0021-9010.93.2.472 (cited on p. 6)

Robert, C., & Casella, G. (2010). *Introducing monte carlo methods with r*. Springer New York. https://doi.org/10.1007/978-1-4419-1576-4. (Cited on p. 14)

Roe, R. A. (2008). Time in applied psychology. *European Psychologist*, *13*(1), 37–52. https://doi.org/10.1027/1016-9040.13.1.37 (cited on pp. 1, 2, 6)

Roe, R. A. (2014, September 22–26). The representation of time in psychological research on air pilots and air traffic controllers. In A. Droog (Chair), *Aviation psychology: Facilitating change(s)* [Symposium]. Proceedings of the 31st EAAP Conference, Valetta, Malta (cited on p. 2).

Roe, R. A., Gockel, C., & Meyer, B. (2012). Time and change in teams: Where we are and where we are moving. *European Journal of Work and Organizational Psychology*, *21*(5), 629–656. https://doi.org/10.1080/1359432x.2012.729821 (cited on p. 1)

Rogosa, D., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin*, *92*(3), 726–748. https://doi.org/10.1037/0033-2909.92.3.726 (cited on p. 197)

RStudio Team. (2020). *Rstudio: Integrated development environment for r*. RStudio, PBC. Boston, MA. http://www.rstudio.com/. (Cited on pp. 31, 36)

Rynes, S. L., Brown, K. G., & Colbert, A. E. (2002). Seven common misconceptions about human resource practices: Research findings versus practitioner beliefs. *Academy of Management Perspectives*, *16*(3), 92–103. https://doi.org/10.5465/ame.2002.8540341 (cited on pp. 199, 200)

Rynes, S. L., Colbert, A. E., & Brown, K. G. (2002). HR professionals' beliefs about effective human resource practices: Correspondence between research and practice. *Human Resource Management*, *41*(2), 149–174. https://doi.org/10.1002/hrm.10029 (cited on p. 199)

Sackett, P. R., & Larson, J., J. R. (1990). Research strategies and tactics in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (pp. 419–489). Consulting Psychologists Press. https://bit.ly/3GymBiR. (Cited on pp. 199, 200)

Sanders, K., van Riemsdijk, M., & Groen, B. (2008). The gap between research and practice: A replication study on the hr professionals' beliefs about effective human resource practices. *The International Journal of Human Resource Management*, *19*(10), 1976–1988. https://doi.org/10.1080/09585190802324304 (cited on p. 199)

Sayer, A. G., & Cumsille, P. E. (2001). Second-order latent growth models. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 179–200). American Psychological Association. https://doi.org/10.1037/10409-006. (Cited on pp. 195, 201)

Sciarra, S. (2022). *nonlinSimsAnalysis: Analyzes data simulated from nonlinSims package* [R package version 0.1.0]. https://github.com/sciarraseb/nonlinSimsAnalysis

Seppälä, P., Hakanen, J. J., Tolvanen, A., & Demerouti, E. (2018). A job resources-based intervention to boost work engagement and team innovativeness during organizational restructuring. *Journal of Organizational Change Management*, *31*(7), 1419–1437. https://doi.org/10.1108/jocm-11-2017-0448 (cited on pp. 194, 197)

Shapiro, A., & Browne, M. W. (1987). Analysis of covariance structures under elliptical distributions. *Journal of the American Statistical Association*, *82*(400), 1092–1097. https://doi.org/10.1080/01621459.1987.10478544 (cited on p. 265)

Shipp, A. J., & Cole, M. S. (2015). Time in individual-level organizational studies: What is it, how is it used, and why isn't it exploited more often? *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(1), 237–260. https://doi.org/10.1146/annurev-orgpsych-032414-111245 (cited on p. 1)

Siegler, R. (2006). Microgenetic analyses of learning. In D. Kuuhn, R. S. Siegler, W. Damon, & R. Lerner (Eds.), *Handbook of child psychology: Cognition, perception, and language* (6th ed., pp. 464–510). John Wiley & Sons. https://bit.ly/3IeZNFT. (Cited on p. 94)

Sonnentag, S. (2012). Time in organizational research: Catching up on a long neglected topic in order to improve theory. *Organizational Psychology Review*, *2*(4), 361–368. https://doi.org/10.1177/2041386612442079 (cited on p. 1)

Spears, M., & Bolton, R. (2015). *Evidence-based hr: The bridge between your people and delivering business strategy*. KPMG International. (Cited on p. 199).

Spector, P. E. (2019). Do not cross me: Optimizing the use of cross-sectional designs. *Journal of Business and Psychology*, *34*(2), 125–137. https://doi.org/10.1007/s10869-018-09613-8 (cited on p. 237)

Stanley, D. J., & Spence, J. R. (2014). Expectations for eeplications. *Perspectives on Psychological Science*, *9*(3), 305–318. https://doi.org/10.1177/1745691614528518 (cited on p. 14)

Steel, R. P., Hendrix, W. H., & Balogh, S. P. (1990). Confounding effects of the turnover base rate on relations between time lag and turnover study outcomes: An extension of meta-analysis findings and conclusions. *Journal of Organizational Behavior*, *11*(3), 237–242. https://doi.org/10.1002/job.4030110306 (cited on p. 6)

Steel, R. P., & Ovalle, N. K. (1984). A review and meta-analysis of research on the relationship between behavioral intentions and employee turnover. *Journal of Applied Psychology*, *69*(4), 673–686. https://doi.org/10.1037/0021-9010.69.4.673 (cited on p. 6)

Sterba, S. K. (2014). Fitting nonlinear latent growth curve models with individually varyingt time points. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(4), 630–647. https://doi.org/10.1080/10705511.2014.919828 (cited on p. 177)

Stockdale, G. D. (2007). *Factors affecting goodness of fit of the quasi-simplex, linear growth curve, and latent difference score models to oppositive data structures: A simulation study* (Doctoral dissertation, University of California No. 3303209). ProQuest Dissertations and Theses Global. https://bit.ly/3CBSzIH. (Cited on p. 17)

Tenhiälä, A., Giluk, T. L., Kepes, S., Simón, C., Oh, I.-S., & Kim, S. (2014). The research-practice gap in human resource management: A cross-cultural study. *Human Resource Management*, *55*(2), 179–200. https://doi.org/10.1002/hrm.21656 (cited on p. 199)

Tiberio, S. S. (2008). *The effects of misspecified measurement intervals in multivariate latent differential equation models* (Doctoral dissertation, University of Notre Dame

No. 3441759). ProQuest Dissertations and Theses Global. https://bit.ly/3Z3EjSe. (Cited on p. 17)

Timmons, A. C., & Preacher, K. J. (2015). The importance of temporal design: How do measurement intervals affect the accuracy and efficiency of parameter estimates in longitudinal research? *Multivariate Behavioral Research*, *50*(1), 41–55. https://doi.org/10.1080/00273171.2014.961056 (cited on pp. 8, 21, 36, 94)

Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenburg, M. (2015). Editorial: Measurement invariance. *Frontiers in Psychology*, *6*, 1064. https://doi.org/10.3389/fpsyg.2015.01064 (cited on pp. 193, 196)

van der Nest, G., Lima Passos, V., Candel, M. J., & van Breukelen, G. J. (2020). An overview of mixture modelling for latent evolutions in longitudinal data: Modelling approaches, fit statistics and software. *Advances in Life Course Research*, *43*, 100323. https://doi.org/10.1016/j.alcr.2019.100323 (cited on pp. 196, 201)

Vancouver, J. B., Wang, M., & Li, X. (2020). Translating informal theories into formal theories: The case of the dynamic computational model of the integrated model of work motivation. *Organizational Research Methods*, *23*(2), 238–274. https://doi.org/10.1177/1094428118780308 (cited on p. 200)

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–70. https://doi.org/10.1177/109442810031002 (cited on pp. 192, 193, 196)

Vantilborgh, T., Hofmans, J., & Judge, T. A. (2018). The time has come to study dynamics at work. *Journal of Organizational Behavior*, *39*(9), 1045–1049. https://doi.org/10.1002/job.2327 (cited on p. 1)

Voelkle, M. C., Brose, A., Schmiedek, F., & Lindenberger, U. (2014). Toward a unified framework for the study of between-person and within-person structures: Building a bridge between two research paradigms. *Multivariate Behavioral Research*, *49*(3), 193–213. https://doi.org/10.1080/00273171.2014.889593 (cited on p. 237)

Wang, L., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods*, *20*(1), 63–83. https://doi.org/10.1037/met0000030 (cited on pp. 7, 238)

Wang, M., & Bodner, T. E. (2007). Growth mixture modeling. *Organizational Research Methods*, *10*(4), 635–656. https://doi.org/10.1177/1094428106289397 (cited on pp. 5, 196)

Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, *14*(4), 490–495. https://doi.org/10.5465/amr.1989.4308371 (cited on p. 1)

Whitley, J. D., & Schoene, L. L. (1987). Comparison of heart rate responses. *Physical Therapy*, *67*(10), 1501–1504. https://doi.org/10.1093/ptj/67.10.1501 (cited on p. 188)

Wickham, H. (2022). *tidyverse: Easily install and load the tidyverse* [R package version 1.3.2]. https://CRAN.R-project.org/package=tidyverse

Wickham, H., Hester, J., & Chang, W. (2021). *devtools: Tools to make developing r packages easier* [R package version 2.4.1]. https://CRAN.R-project.org/package=devtools

Wilke, C. O. (2020). *cowplot: Streamlined plot theme and plot annotations for ggplot2* [R package version 1.1.1]. https://wilkelab.org/cowplot/

Wilke, C. O., & Wiernik, B. M. (2022). *ggtext: Improved text rendering support for ggplot2* [R package version 0.1.2]. https://wilkelab.org/ggtext/

Wu, J.-Y., Kwok, O.-M., & Willson, V. (2014). Using design-based latent growth curve modeling with cluster-level predictor to address dependency. *The Journal of Experimental Education*, *82*(4), 431–454. https://doi.org/10.1080/00220973.2013.876226 (cited on pp. 21, 22)

Xia, W., Ye, M., Liu, J., Cao, M., & Sun, X.-M. (2020). Analysis of a nonlinear opinion dynamics model with biased assimilation. *Automatica*, *120*, 109113. https://doi.org/10.1016/j.automatica.2020.109113 (cited on p. 7)

Xie, Y. (2022). *knitr: A general-purpose package for dynamic report generation in r* [R package version 1.41]. https://yihui.org/knitr/

Ye, F. (2016). Latent growth curve analysis with dichotomous items: Comparing four approaches. *British Journal of Mathematical and Statistical Psychology*, *69*(1), 43–61. https://doi.org/10.1111/bmsp.12058 (cited on pp. 21, 22)

Yigit, S., & Mendes, M. (2018). Which effect size measure is appropriate for one-way and two-way anova models? a Monte Carlo simulation study. *Statistical Journal*, *16*(3), 295–313. https://doi.org/10.2333/bhmk.40.129 (cited on p. 46)

Zaheer, S., Albert, S., & Zaheer, A. (1999). Time scales and organizational theory. *Academy of Management Review*, *24*(4), 725–741. https://doi.org/10.5465/amr.1999.2553250 (cited on p. 1)

Zhu, H. (2021). *kableExtra: Construct complex table with kable and pipe syntax* [R package version 1.3.4]. https://CRAN.R-project.org/package=kableExtra

Appendix A: Ergodicity and the Need to Conduct Longitudinal Research

To understand why cross-sectional results are unlikely to agree with longitudinal results for any given analysis, a discussion of data structures is apropos. Consider an example where a researcher obtains data from 50 people measured over 100 time points such that each row contains a p person's data over the 100 time points and each column contains data from 50 people at a t time point. For didactic purposes, all data are assumed to be sampled from a normal distribution. To understand whether findings in any given cross-sectional data set yield the same findings in any given longitudinal data set, the researcher randomly samples one cross-sectional and one longitudinal data set and computes the mean and variance in each set. To conduct a cross-sectional analysis, the researcher randomly samples the data across the 50 people at a given time point and computes a mean of the scores at the sampled time point ($\bar{X}_t$) using Equation A.1 shown below:

$$\bar{X}_t = \frac{1}{P}\sum_{p=1}^{P} x_p, \tag{A.1}$$

where the scores of all P people are summed (x_p) and then divided by the number of people (P). To compute the variance of the scores at the sampled time point (S_t^2), the researcher uses Equation A.2 shown below:

$$S_t^2 = \frac{1}{P}\sum_{p=1}^{P}(x_p - \bar{X}_t)^2, \tag{A.2}$$

where the sum of the squared differences between each person's score (x_p) and the average value at the given t time point ($\bar{X}_t$) is computed and then divided by the number of people (P). To conduct a longitudinal analysis, the researcher randomly samples one person's data across the 100 time points and also computes a mean and variance of the scores. To compute the mean across the t time points of the longitudinal data set ($\bar{X}_p$), the researcher uses Equation A.3 shown below:

$$\bar{X}_p = \frac{1}{T}\sum_{t=1}^{T} x_t, \tag{A.3}$$

where the scores at each t time point are summed (x_t) and then divided by the number of time points (T). The researcher also computes a variance of the sampled person's scores across all time points (S_p^2) using Equation A.4 shown below:

$$S_p^2 = \frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{X}_p)^2, \tag{A.4}$$

where the sum of squared differences between the score at each time point (x_t) and the average value of the p person's scores ($\bar{X}_p$) is computed and then divided by the number of time points (T).

If the researcher wants to treat the mean and variance from the cross-sectional and longitudinal data sets as interchangeable, then two conditions outlined by ergodic theory must be satisfied (Molenaar, 2004; Molenaar & Campbell, 2009).[28] First, a given cross-sectional mean and variance can only closely estimate the mean and variance of any

[28]Note that ergodic theory is an entire mathematical discipline (for an introduction, see Petersen, 1983). In the current context, the most important ergodic theorems are those proven by Birkhoff (1931, for a review, see Choe, 2005, Chapter 3)

given person's data (i.e., a longitudinal data set) to the extent that each person's data originate from a normal distribution with the same mean and variance. If each person's data originate from a different normal distribution, then computing the mean and variance at a given time point would, at best, describe the values of one person. When each person's data are generated from the same normal distribution, the condition of *homogeneity* is met. Importantly, satisfying the condition of homogeneity does not guarantee that the mean and variance obtained from another cross-sectional data set will closely estimate the mean and variance of any given person (i.e., any given longitudinal data set). The mean and variance values computed from any given cross-sectional data set can only closely estimate the values of any given person to the extent that the cross-sectional mean and variance remain constant over time. If the mean and variance of observations remain constant over time, then the second condition of *stationarity* is satisfied. Therefore, the researcher can only treat means and variances from cross-sectional and longitudinal data sets as interchangeable if each person's data are generated from the same normal distribution (homogeneity) and if the mean and variance remain constant over time (stationarity). When the conditions of homogeneity and stationarity are satisfied, a process is said to be *ergodic*: Analyses of cross-sectional data sets will return the same values as analyses on longitudinal data sets.

Given that psychological studies almost never collect data from only one person, one potential reservation may be that the conditions required for ergodicity only hold when a longitudinal data set contains the data of one person. That is, if the researcher uses the full data set containing the data of 100 people sampled over 100 time points and computes 100 cross-sectional means and variances (Equation A.1 and Equation A.2, respectively) and 100 longitudinal means and variances (Equation A.3 and Equation A.4, respectively), wouldn't the average of the cross-sectional means and variances be the same as the average of the longitudinal means and variances? Although averaging the cross-sectional means returns the same value as averaging the longitudinal means, the average longitudinal variance remains different from the average cross-sectional variance (for several empirical examples, see A. J. Fisher et al., 2018). Therefore, the conditions of ergodicity apply even with larger longitudinal and cross-sectional sample sizes.

The guaranteed differences in cross-sectional and longitudinal variance values that result from non-ergodic processes have far-reaching implications. Almost every analysis employed in organizational research—whether it be correlation, regression, factor analysis, mediation, etc.—analyzes variability, and so, when a process is non-ergodic, cross-sectional variability will differ from longitudinal variability, and the results obtained from applying any given analysis on each of the variabilities will differ as a consequence. Because variability is central to so many analyses, the non-equivalence of longitudinal and cross-sectional variances that results from a non-ergodic process explains why discussions of ergodicity often point out that "for non-ergodic processes, an analysis of the structure of IEV [interindividual variability] will yield results that differ from results obtained in an analogous analysis of IAV [intraindividual variability]"(Molenaar, 2004, p. 202).[29]

[29]It is important to note that a violation of one or both ergodic conditions (homogeneity and stationarity) does not mean that an analysis of cross-sectional variability yields results that have no relation to the results gained from applying the analysis on longitudinal variability (i.e., the causes of cross-sectional variability are independent from the causes of longitudinal variability). An analysis of cross-sectional variability can still give insight into temporal dynamics if the causes of non-ergodicity can be identified (Voelkle et al., 2014; for similar discussion, see Spector, 2019). Thus, conceptualizing ergodicity on a continuum with non-erdogicity and ergodicity on opposite ends provides a more accurate perspective for understanding ergodicity (Adolf & Fried, 2019; Medaglia et al., 2019).

With an understanding of the conditions required for ergodicity, a brief review of organizational phenomena finds that these conditions are regularly violated. Focusing only on homogeneity (each person's data are generated from the same distribution), several instances in organizational research violate this condition. As examples of homogeneity violations, employees show different patterns of absenteeism over five years (Magee et al., 2016), leadership development over the course of a seminar (Day & Sin, 2011), career stress over the course of 10 years (Igic et al., 2017), and job performance in response to organizational restructuring (Miraglia et al., 2015). With respect to stationarity (constant values for statistical parameters across people over time), several examples can be generated by realizing how calendar events affect psychological processes and behaviours throughout the year. As examples of stationarity violations, consider how salespeople, on average, undoubtedly sell more products during holidays, how employees, on average, take more sick days during the winter months, and how accountants, on average, experience more stress during tax season. With violations of ergodic conditions commonly occurring in organizational psychology, it becomes fitting to echo the commonly held sentiment that few, if any, psychological processes are ergodic (Curran & Bauer, 2011; A. J. Fisher et al., 2018; Hamaker, 2012; Molenaar, 2004, 2008; Molenaar & Campbell, 2009; L. Wang & Maxwell, 2015). As a result, longitudinal research is necessary for understanding psychological processes.

Appendix B: Code Used to Run Monte Carlo Simulations for all Experiments

The code used to compute the simulations of each experiment are shown in Code Block B.1. Note that the cell size is 1000 (i.e., `num_iterations = 1000`).

Code Block B.1

Code Use to Run Monte Carlo Simulations for Each Simulation Experiment

```r
devtools::install_github(repo = 'sciarraseb/nonlinSims', force=T)

library(easypackages)
packages <- c('devtools', 'nonlinSims', 'parallel', 'tidyverse', "OpenMx", "data.table", 'progress', 'tictoc')
libraries(packages)

time_period <- 360

#Population values for parameters
#fixed effects
sd_scale <- 1
common_effect_size <- 0.32
theta_fixed <- 3
alpha_fixed <- theta_fixed + common_effect_size
beta_fixed <- 180
gamma_fixed <- 20

#random effects
sd_theta <- 0.05
sd_alpha <- 0.05
sd_beta <- 10
sd_gamma <- 4
sd_error <- 0.05

#List containing population parameter values
pop_params_4l <- generate_four_param_pop_curve(
  theta_fixed =  theta_fixed, alpha_fixed = alpha_fixed,
   beta_fixed = beta_fixed, gamma_fixed = gamma_fixed,
   sd_theta = sd_theta, sd_alpha = sd_alpha,
   sd_beta = sd_beta, sd_gamma = sd_gamma, sd_error = sd_error



num_iterations <- 1e3 #n=1000 (cell size)
seed <- 27 #ensures replicability

# Experiment 1 (number measurements,  spacing, midpoint) --------
factor_list_exp_1 <- list('num_measurements' = seq(from = 5, to = 11, by = 2),
                          'time_structuredness' = c('time_structured'),
                          'spacing' = c('equal', 'time_inc', 'time_dec', 'mid_ext'),
                          'midpoint' = c(80, 180, 280),
                          'sample_size' = c(225))

tic()
exp_1_data <- run_exp_simulation(factor_list = factor_list_exp_1, num_iterations = num_iterations, pop_params = pop_params_4l,
                                 num_cores = detectCores()-1, seed = seed)
toc()

#Average computation time is 1 iteration per second. As an example, Experiment has 48 cells x 1000 iterations/cell = 48 000 iterations and seconds/3600s/hour ~ 13.33 hours (simulations computed with 15 cores)
write_csv(x = exp_1_data, file = '~/Desktop/exp_1_data.csv')

# Experiment 2 (number measurements, spacing,  sample size) ---
factor_list_exp_2 <- list('num_measurements' = seq(from = 5, to = 11, by = 2),
                          'time_structuredness' = c('time_structured'),
                          'spacing' = c('equal', 'time_inc', 'time_dec', 'mid_ext'),
                          'midpoint' = 180,
                          'sample_size' = c(30, 50, 100, 200, 500, 1000))

tic()
exp_2_data <- run_exp_simulation(factor_list = factor_list_exp_2, num_iterations = num_iterations, pop_params = pop_params_4l,
```

```
                                   num_cores = detectCores(), seed = seed)
toc()

write_csv(x = exp_2_data, file = 'Desktop/exp_2_data.csv')

# Experiment 3 (number measurements, sample size, time structuredness) -----
factor_list_exp_3 <- list('num_measurements' = seq(from = 5, to = 11, by = 2),
                           'time_structuredness' = c('time_structured', 'fast_response',
                           'slow_response'),
                           'spacing' = c('equal'),
                           'midpoint' = 180,
                           'sample_size' = c(30, 50, 100, 200, 500, 1000))
tic()
exp_3_data <- run_exp_simulation(factor_list = factor_list_exp_3, num_iterations =
num_iterations, pop_params = pop_params_41,
                                   num_cores = detectCores(), seed = seed)
toc()

write_csv(x = exp_3_data, file = '~/Desktop/exp_3_data.csv')



# Experiment 3 (definition variables with slow response rate ) ------
factor_list_exp_def <- list('num_measurements' = seq(from = 5, to = 11, by = 2),
                             'time_structuredness' = c('slow_response'),
                             'spacing' = c('equal'),
                             'midpoint' = 180,
                             'sample_size' = c(30, 50, 100, 200, 500, 1000))
tic()
exp_3_def_data <- run_exp_simulation(factor_list = factor_list_exp_def, num_iterations =
num_iterations, pop_params = pop_params_41,
                                       num_cores = detectCores() - 1, seed = seed,
                                       definition = T)
toc()
#240734.993 sec elapsed (7 cores used; simulation time increased by roughly a
magnitude of 8).
write_csv(x = exp_3_def_data, file = 'exp_3_def.csv')
```

Appendix C: Procedure for Generating Measurement Schedules Measurement Schedules

Given that no procedure existed (to my knowledge) for creating measurement schedules, I devised a method for generating measurement schedules for the four spacing conditions (equal, time-interval increasing, time-interval decreasing, and middle-and-extreme spacing). The code I used to automate the generation of these schedules can be found within the code for the `compute_measurement_schedules()` function for the `nonlinSims` package (see https://github.com/sciarraseb/nonlinSims). For each measurement spacing conditions across all measurement number levels, a two-step procedure was employed to generate measurement schedules in Experiments 1 and 2. At a broad level, the first step computes values for setup variables and the second step computes the interval lengths.

C.1 Procedure for Constructing Measurement Schedules With Equal Spacing

Figure C.1 shows how the two-step procedure is implemented to construct a measurement schedule with equal spacing and five measurements. In the first step, the number of intervals (NI) is computed by subtracting one from the number of measurements (NM). With five measurements ($NM = 5$), there are four intervals ($NI = 4$). In the second step, interval lengths are calculated by dividing the length of the measurement period (MP) by the number of intervals (NI), yielding an interval length of 90 days ($\frac{MP}{NI} = \frac{360}{4} = 90$) for each interval and the following measurement days:

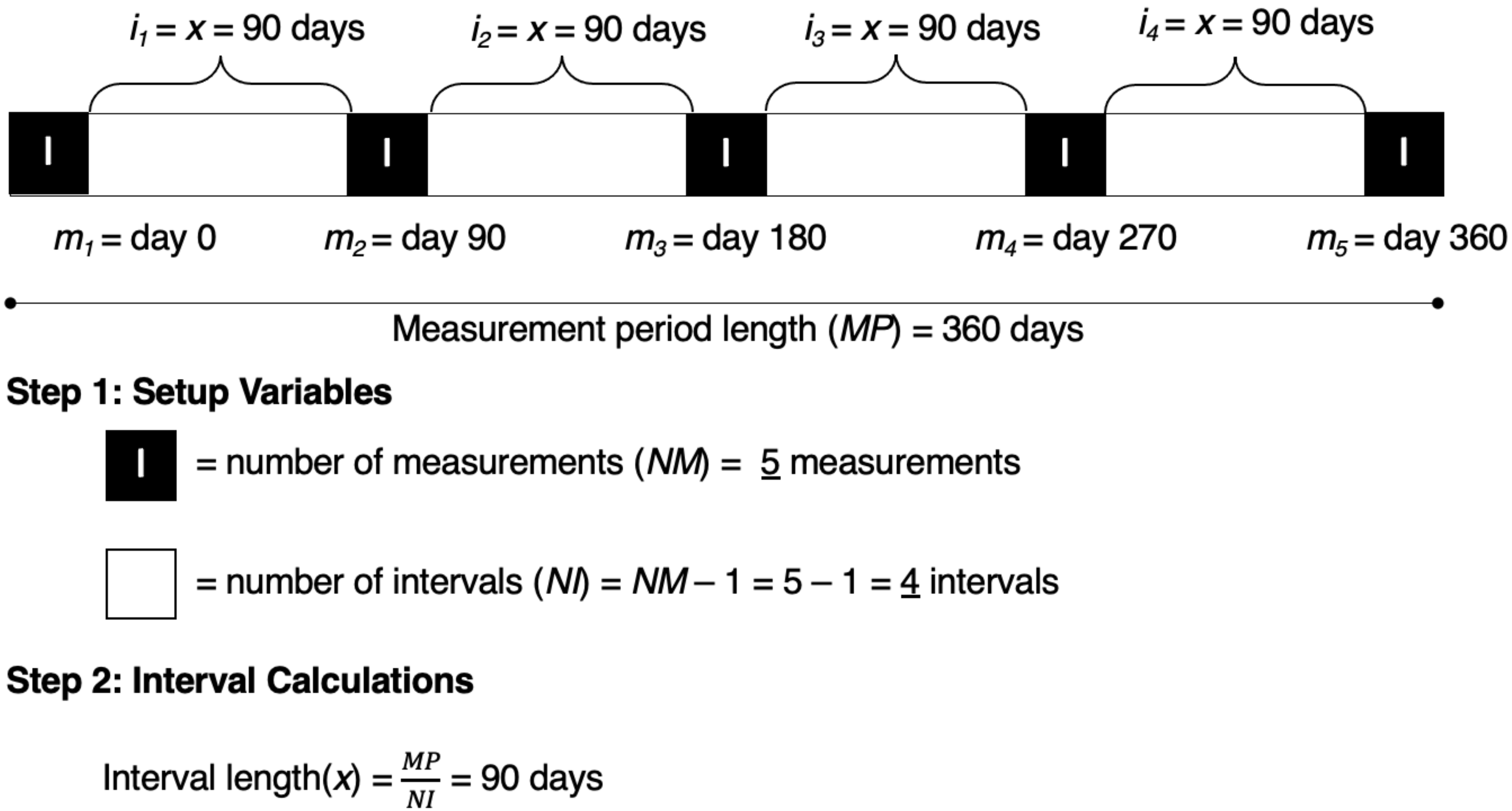
- m_1 = day 0
- m_2 = day 90
- m_3 = day 180
- m_4 = day 270
- m_5 = day 360.

C.2 Procedure for Constructing Measurement Schedules With Time-Interval Increasing Spacing

Figure C.2 shows how the two-step procedure is implemented to construct a measurement schedule with time-interval increasing spacing and five measurements. In the first step, the number of intervals (NI) is computed by subtracting one from the number of measurements (NM). With five measurements ($NM = 5$), there are four intervals

Figure C.1

Procedure for Computing Measurement Schedules With Equal Spacing



Note. In Step 1, setup variables are calculated. With five measurements $(NM = 5)$, there are four intervals $(NI = 4)$. In Step 2, interval lengths are calculated by dividing the length of the measurement period (MP) by the number of intervals (NI), yielding an interval length of 90 days ($\frac{MP}{NI} = \frac{360}{4} = 90$) for each interval.

($NI = 4$). Because interval lengths increase over time, I decided that intervals would increase by an integer multiple of a constant length (c) after each measurement day (m_i) according to the function shown below in Equation C.1:

$$\text{Constant-length increment} = \sum_{x=0}^{NI-1} xc, \tag{C.1}$$

where x represents the integer multiple that increases by 1 after each measurement day. Importantly, to calculate the constant-length increment (c) by which interval lengths increase over time, it is important to realize that two terms contribute to the length of any interval: A shortest-interval length (s) and a constant-length value (c), as shown below in Equation C.2:

$$\text{Interval length} = s + \sum_{x=0}^{NI-1} xc. \tag{C.2}$$

Because the shortest-interval length (s) contributes to the length of each interval–in this example, four intervals—then the sum of these lengths can be subtracted from the measurement period length of 360 days ($MP = 360$). In the current example with five measurements, 240 days remain ($r = 240$) after subtracting the days needed for the shortest-interval lengths (see Equation C.3).

$$\text{Remaining days } (r) = MP - (NI)s = 360 - (30)4 = 240 \text{ days} \tag{C.3}$$

Having computed the number of remaining days, the constant-length value (c) can then be obtained by dividing the number of remaining days by the number of constant-value interval lengths (c_i), as shown below in Equation C.4:

$$\text{Constant-value interval length}(c) = \frac{r}{\sum_{i=0}^{NI-2} i} = \frac{240}{3+2+1} = 40 \text{ days} \tag{C.4}$$

Therefore, having computed the value for c, the following interval lengths are obtained:

- $i_1 = s + 0(c) = 30 + 0(30) = 30$ days
- $i_2 = s + 1(c) = 30 + 1(40) = 70$ days
- $i_3 = s + 0(c) = 30 + 2(40) = 110$ days
- $i_4 = s + 0(c) = 30 + 3(40) = 150$ days

and the following measurement days are obtained:

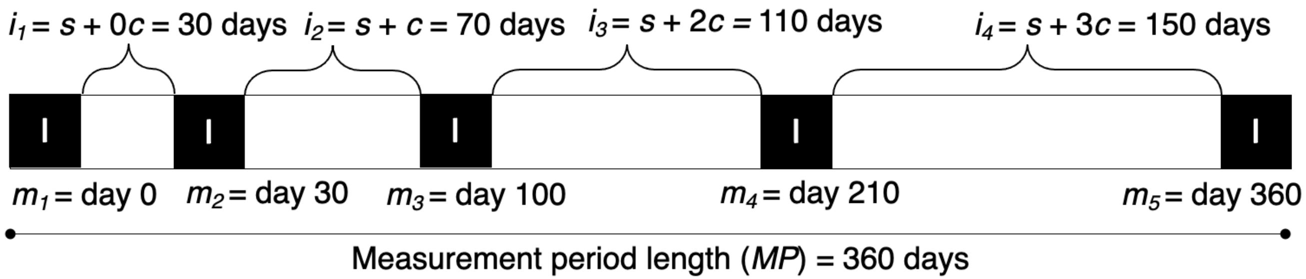
- $m_1 =$ day 0
- $m_2 =$ day 30
- $m_3 =$ day 100
- $m_4 =$ day 210
- $m_5 =$ day 360.

C.3 Procedure for Constructing Measurement Schedules With Time-Interval Decreasing Spacing

Figure C.3 shows how the two-step procedure is implemented to construct a measurement schedule with time-interval decreasing spacing and five measurements. Because the procedure for calculating time-decreasing intervals simply requires that the order of time-interval increasing intervals are reversed, the procedure is, thus, essentially identical to the procedure shown in the previous section. Therefore, with five measurements, time-interval decreasing spacing produces the following intervals:

Figure C.2

Procedure for Computing Measurement Schedules With Time-Interval Increasing Spacing



Step 1: Setup Variables

■ = number of measurements (NM) = 5 measurements

□ = number of intervals (NI) = $NM - 1 = 5 - 1 = 4$ intervals

Step 2: Interval Calculations

s = shortest-interval length = 30 days

Remaining days(r) = $MP - (NI)s = 360 - 4(30) =$ 240 days

$$\text{Constant length}(c) = \frac{r}{\sum_{i=0}^{NI-1} c_i} = \frac{240}{0+1+2+3} = \underline{40 \text{ days}}$$

Note. In Step 1, setup variables are calculated. With five measurements ($NM = 5$), there are four intervals ($NI = 4$). In Step 2, two components contribute to each interval length: A shortest-interval length (s) and a constant-length value (c), as shown in Equation C.2. Because the shortest-interval length (s) contributes to each interval, the sum of these lengths can be subtracted from the measurement period length of 360 days ($MP = 360$). In the current example with five measurements, 240 days remain ($r = 240$) after subtracting the days needed for the shortest-interval lengths (see Equation C.3). To calculate the constant-length value (c), the remaining days (r) are divided by the number of constant-value interval lengths (c_i), as shown in Equation C.4.

- $i_1 = s + 0(c) = 30 + 3(40) = 150$ days
- $i_2 = s + 0(c) = 30 + 2(40) = 110$ days
- $i_3 = s + 1(c) = 30 + 1(40) = 70$ days
- $i_4 = s + 0(c) = 30 + 0(30) = 30$ days

and the following measurement days are obtained:

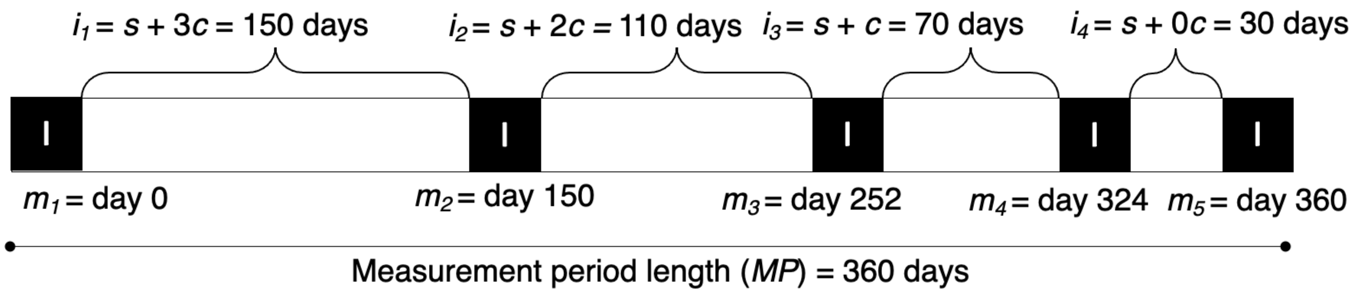
- $m_1 =$ day 0
- $m_2 =$ day 150
- $m_3 =$ day 260
- $m_4 =$ day 330
- $m_5 =$ day 360.

C.4 Procedure for Constructing Measurement Schedules With Middle-and-Extreme Spacing

Figure C.4 shows how the two-step procedure is implemented to construct a measurement schedule with middle-and-extreme spacing and five measurements. In the first step, the number of intervals (NI) is computed by subtracting one from the number of measurements (NM). With five measurements ($NM = 5$), there are four intervals ($NI = 4$). Importantly, because middle-and-extreme spacing places measurements near the extremities and the middle of the measurement window, the number of measurements in both these sections must also be calculated. The number of extreme measurements is first calculated by dividing the number of measurements by 3 and taking the floor (i.e., rounded-down value [$\lfloor x \rfloor$]) of this value and multiplying it by 2, as shown below in Equation C.5:

Figure C.3

Procedure for Computing Measurement Schedules With Time-Interval Decreasing Spacing



Step 1: Setup Variables

I = number of measurements (NM) = 5 measurements

= number of intervals (NI) = NM – 1 = 5 – 1 = 4 intervals

Step 2: Interval Calculations

s = shortest-interval length = 30 days

Remaining days(r) = MP – (NI)s = 360 – 4(30) = 240 days

Constant length(c) = $\frac{r}{\sum_{i=0}^{NI-1} c_i} = \frac{240}{0+1+2+3}$ = 40 days

Note. In Step 1, setup variables are calculated. With five measurements ($NM = 5$), there are four intervals ($NI = 4$). In Step 2, two components contribute to each interval length: A shortest-interval length (s) and a constant-length value (c), as shown in Equation C.2. Because the shortest-interval length (s) contributes to each interval, the sum of these lengths can be subtracted from the measurement period length of 360 days ($MP = 360$). In the current example with five measurements, 240 days remain ($r = 240$) after subtracting the days needed for the shortest-interval lengths (see Equation C.3). To calculate the constant-length value (c), the remaining days (r) are divided by the number of constant-value interval lengths (c_i), as shown in Equation C.4.

$$\text{Number of extreme measurements}(ex) = 2\lfloor\frac{NM}{3}\rfloor = 2\lfloor\frac{5}{3}\rfloor = 2. \tag{C.5}$$

The number of middle measurements can then be calculated by subtracting the number of extreme measurements (ex) from the number of measurements (NM), as shown below in Equation C.7:

$$\text{Number of middle measurements}(mi) = NM - ex = 5 - 2 = 3. \tag{C.6}$$

In Step 2, interval lengths are calculated. For middle-and-extreme spacing, there are two types of interval lengths: 1) Intervals separating either two middle or two extreme measurements and 2) intervals separating one middle and one extreme measurement. Intervals separating two middle or two extreme measurements (w_i) are set to the shortest-interval length (s), which I set to be 30 days ($w_i = s = 30$). Intervals separating one middle and one extreme measurement (b_i) are set to the sum of two components: 1) A shortest-interval length (s) and a 2) constant-value interval length (c), as shown below in Equation C.7:

$$b_i = s + c. \tag{C.7}$$

To obtain the constant-value interval length (c), the sum of shortest-value interval lengths (s) is subtracted from the measurement period of 360 days ($MP = 360$). In the current example with five measurements, 240 days remain ($r = 240$) after subtracting the days needed for the shortest-interval lengths (see Equation C.8).

$$\text{Remaining days } (r) = MP - (NI)s = 360 - (30)4 = 240 \text{ days} \tag{C.8}$$

Having computed the number of remaining days, the constant-length value (c) can then be obtained by dividing the number of remaining days by the number of intervals separating middle and extreme measurements, which will always be 2, as shown below in Equation C.9:

$$\text{Constant-value interval length}(c) = \frac{r}{2} = \frac{240}{2} = 120 \text{ days} \tag{C.9}$$

Therefore, having computed the value for c, the following interval lengths are obtained:

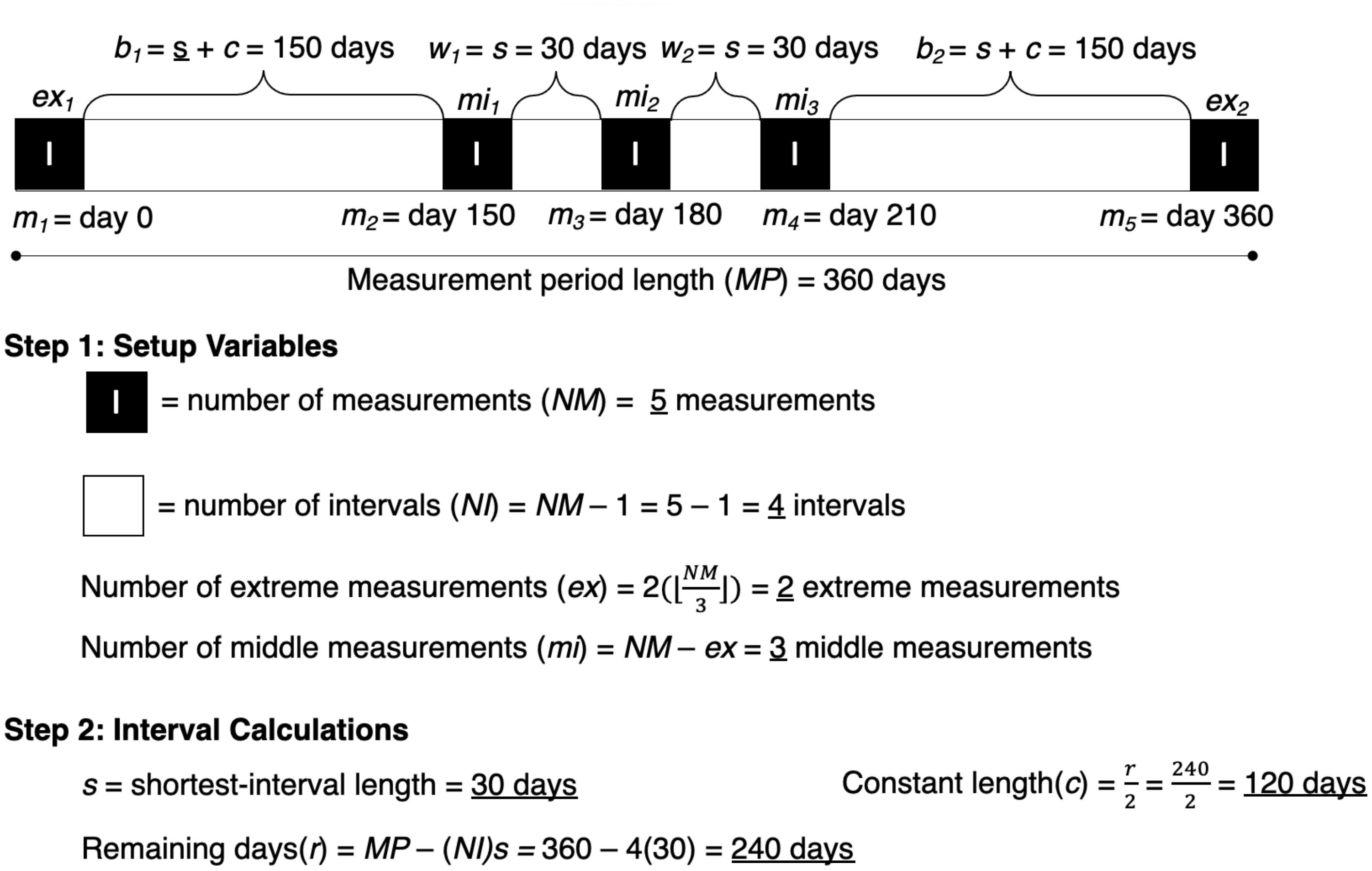
- $b_1 = s + c = 30 + 120 = 150$ days
- $w_1 = s = 30 = 30$ days
- $w_2 = s = 30 = 30$ days
- $b_2 = s + c = 30 + 120 = 150$ days

and the following measurement days are obtained:

- $m_1 =$ day 0
- $m_2 =$ day 150
- $m_3 =$ day 180
- $m_4 =$ day 21
- $m_5 =$ day 360.

Figure C.4

Procedure for Computing Measurement Schedules With Middle-and-Extreme Spacing



Note. In Step 1, setup variables are calculated. With five measurements ($NM = 5$), there are four intervals ($NI = 4$). Importantly, because middle-and-extreme spacing places measurements near the extremities and the middle of the measurement window, the number of measurements in both these sections must also be calculated. The number of extreme measurements is first calculated by dividing the number of measurements by 3 and taking the taking the floor (i.e., rounded-down value [$\lfloor x \rfloor$]) of this value and multiplying it by 2 (see Equation C.5). The number of middle measurements can then be calculated by subtracting the number of extreme measurements (ex) from the number of measurements (NM; see Equation C.7). In Step 2, interval lengths are calculated. For middle-and-extreme spacing, there are two types of interval lengths: 1) Intervals separating either two middle or two extreme measurements and 2) intervals separating one middle and one extreme measurement. Intervals separating two middle or two extreme measurements are set to the shortest-interval length (s), which I set to be 30 days ($w_i = s = 30$). Intervals separating one middle and one extreme measurement are set to the sum of two components: 1) A shortest-interval length (s) and a 2) constant-value interval length (c; see Equation C.9). To obtain the constant-value interval length (c), the sum of shortest-value interval lengths (s) is subtracted from the measurement period of 360 days ($MP = 360$). In the current example with five measurements, 240 days remain ($r = 240$) after subtracting the days needed for the shortest-interval lengths (see Equation C.8). Having computed the number of remaining days, the constant-length value (c) can then be obtained by dividing the number of remaining days by the number of intervals separating middle and extreme measurements, which will always be 2 (see Equation C.9).

Appendix D: Using Nonlinear Function in the Structural Equation Modelling Framework

D.1 Nonlinear Latent Growth Curve Model Used to Analyze Each Generated Data Set

The sections that follow will first review the framework used to build latent growth curve models and then explain how nonlinear functions can be modified to fit into this framework.

D.1.1 Brief Review of the Latent Growth Curve Model

The latent growth curve model proposed by Meredith and Tisak (1990) is briefly reviewed here (for a review, see K. Preacher et al., 2008). Consider an example where data are collected at five time points ($T = 5$) from p people ($\mathbf{y_p} = [y_1, y_2, y_3, y_4, y_5]$). A simple model to fit is one where change over time is defined by a straight line and each person's pattern of change is some variation of this straight line. In modelling parlance, an intercept-slope model is fit where both the intercept and slope are random effects whose values are allowed to vary for each person.

To fit a random-effect intercept-slope model, a general linear pattern can first be specified in the Λ matrix shown below in Equation D.1:

$$\Lambda = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}. \tag{D.1}$$

In each column of Λ, the effect a parameter is specified over the five time points; that is, Λ is a matrix with two columns (one for the intercept and one for the slope parameter) and five rows (one for each time point).[30] The first column of Λ specifies the intercept parameter. Because the effect of the intercept parameter is constant over time, a column of 1s is used to represent its effect. The second column of Λ specifies the slope parameter. Because a linear pattern of growth is assumed, the second column contains a series of monotonically increasing integer numbers across the time points and begins with 0. [31]

To specify the intercept and slope parameters as random effects that vary across people, a weight can be applied to each column of Λ and each weight can vary across people. That is, a p person's pattern of change can be reproduced with a unique set of weights in $\boldsymbol{\iota}_\mathbf{p}$ that determines the extent to which each basis column of Λ contributes to the person's observed change over time. By allowing the weights for the intercept and slope parameters to vary across people, variability can be estimated in these parameters. Discrepancies between the values predicted by $\Lambda\boldsymbol{\iota}_\mathbf{p}$ and a person's observed scores across all five time points are stored in an error vector $\mathcal{E}_\mathbf{p}$. Thus, a person's observed data ($\mathbf{y_p}$) is reproduced using the function shown below in Equation D.2:

$$y_p = \Lambda\boldsymbol{\iota}_\mathbf{p} + \mathcal{E}_\mathbf{p}. \tag{D.2}$$

Note that Equation D.2 defines the general structural equation modelling framework.

[30]The columns of Λ are often called basis curves (Blozis, 2004) or basis functions (Meredith & Tisak, 1990; Browne, 1993) because each column specifies a particular component of change.

[31]The set of numbers specified for the slope starts at zero because there is presumably no effect of any variable at the first time point.

D.1.2 Fitting a Nonlinear Function in the Structural Equation Modelling Framework

Unfortunately, the logistic function of Equation 2.1—where each parameter is estimated as a fixed- and random-effect—cannot be directly used in a latent growth curve model because it violates the linear nature of the structural equation modelling framework (Equation D.2). Structural equation models only permit linear combinations—specifically, the products of matrix-vector and/or matrix-matrix multiplication—and so directly fitting a nonlinear function such as the logistic function in Equation 2.1 is not possible.

One solution to fitting the logistic function within the structural equation modelling framework is to implement the structured latent curve modelling approach (Browne, 1993; Browne & du Toit, 1991; for an excellent review, see K. J. Preacher & Hancock, 2015). Briefly, the structured latent curve modelling approach constructs a Taylor series approximation of a nonlinear function so that the nonlinear function can be fit into the structural equation modelling framework (Equation D.2). The sections that follow will present the structured latent curve modelling approach in four parts such that 1) Taylor series approximations will first be reviewed, 2) a Taylor series approximation will then be constructed for the logistic function, 3) the logistic Taylor series approximation will be modified and fit into the structural equation modelling framework, and 4) the process of parameter estimation will be reviewed.

D.1.2.1 Taylor Series': Approximations of Linear Functions

A Taylor series uses derivative information of a nonlinear function to construct a linear function that is an approximation of the nonlinear function.[32] Equation D.3 shows the general formula for a Taylor series such that

$$P^N(f(x), a) = \sum_{n=0}^{N} \frac{f^n a}{n!}(x-a)^n, \tag{D.3}$$

where N is the highest derivative order of the function $f(a)$ that is taken beginning from a zero-value derivative order ($n = 0$), a is the point where the Taylor series is derived (i.e., the point of derivation), and x is the point where the Taylor series is evaluated (i.e., the point of evaluation). As an example of a Taylor series, consider the second-order Taylor series of $f(x) = \cos(x)$. Note that, across the continuum of x values (i.e., from $-\infty$ to ∞), $\cos(x)$ returns values between -1 and 1 in an oscillatory manner. Computing the second-order Taylor series of $f(x) = \cos(x)$ yields the following function shown in

[32]Linear functions are defined as functions where no parameter exists within its own partial derivative (at any order). For example, none of the parameters in the polynomial equation of $y = a + bt + ct^2 + dt^3$ exist within their own partial derivative: $\frac{\partial y}{\partial a} = 1$, $\frac{\partial y}{\partial b} = t$, $\frac{\partial y}{\partial c} = t^2$, and $\frac{\partial y}{\partial d} = t^3$. Conversely, the logistic function is nonlinear because β and γ exist in their own partial derivatives. For example, the derivative of the logistic function $y = \theta + \frac{\alpha - \theta}{1 + e^{\frac{\beta - t}{\gamma}}}$ with respect to β is $\frac{(\theta - \alpha)(e^{\frac{\beta - t}{\gamma}})(\frac{1}{\gamma})}{1 + (e^{\frac{\beta - t}{\gamma}})^2}$ and so is nonlinear because it contains β.

Equation D.4:

$$\begin{aligned}
P^2(\cos(x), a) &= \frac{\frac{\partial^0 \cos(a)}{\partial a^0}}{0!}(x-a)^0 + \frac{\frac{\partial^1 \cos(a)}{\partial a^1}}{1!}(x-a)^1 + \frac{\frac{\partial^2 \cos(a)}{\partial a^2}}{2!}(x-a)^2 \\
&= \frac{\cos(0)}{0!}(x-0)^0 - \frac{\sin(0)}{1!}(x-0)^1 - \frac{\cos(0)}{2!}(x-0)^2 \\
&= \frac{1}{1}1 - \frac{0}{1}x - \frac{1}{2}x^2 \\
P^2(\cos(x), 0) &= 1 - \frac{1}{2}x^2. \qquad \text{(D.4)}
\end{aligned}$$

Importantly, the second-order Taylor series of $\cos(x)$ shown in Equation D.4 is linear, whereas the function $\cos(x)$ is not linear. To show that the second-order Taylor series of $1 - \frac{1}{2}x^2$ is linear, we can reformulate it by adding placeholder parameters in front of each term (b and c), resulting in the following modified equation of Equation D.5:

$$P^2_{reform}(\cos(x), a) = b1 - c\frac{1}{2}x^2. \qquad \text{(D.5)}$$

If the partial derivative of $P^2(\cos(x), a)$ is taken with respect to b and c, no parameter exists within its own partial derivative, meaning the function is linear (see Equations D.6–D.7 below).

$$\frac{\partial P^2_{reform}(\cos(x), a)}{\partial b} = 1 \text{ and} \qquad \text{(D.6)}$$

$$\frac{\partial P^2_{reform}(\cos(x), c)}{\partial c} = -\frac{1}{2}x^2. \qquad \text{(D.7)}$$

Conversely, the fourth-order partial derivative of $\cos(x)$ contains itself (see Equation D.8), and so is a nonlinear function.

$$\frac{\partial^4 \cos(x)}{\partial x^4} = \cos(x). \tag{D.8}$$

Therefore, Taylor series' can generate linear versions of nonlinear functions by using local derivative information.

Although Taylor series' provide linear versions of nonlinear functions, it is important to emphasize that the linear versions are approximations. More specifically, the second-order Taylor series of $\cos(x)$ perfectly estimates $\cos(x)$ when the point of evaluation x is set equal to the point of derivation a, but estimates $\cos(x)$ with an increasing amount of error as the difference between x and a increases (see Example D.1). Thus, Taylor series are approximations because they are only locally accurate (i.e., near the point of derivation).

Example D.1. *Estimates of Taylor series approximation of $f(x) = \cos(x)$ as the difference between the point of evaluation* x *and the point of derivation* a *increases.*

Taylor series approximation of $\cos(x)$ (specifically, the second-order Taylor series; $P^2[\cos(x), a]$) estimates values that are exactly equal to the values returned by $\cos(x)$ when the point of evaluation (x) is set to the point of derivation (a). The example below computes the value predicted by the Taylor series approximation of $P^2[\cos(x), a]$ and by

$\cos(x)$ when $x = a = 0$.

$$
\begin{aligned}
P^2(\cos(x=0), a=0) &= \cos(x=0) \\
1 - \frac{1}{2}x^2 &= \cos(0) \\
1 - \frac{1}{2}0^2 &= 1 \\
1 - 0 &= 1 \\
1 &= 1
\end{aligned}
$$

Taylor series approximation of $\cos(x)$ (specifically, the second-order Taylor series; $P^2[\cos(x), a]$) estimates a value that is approximately equal ($\approx$) to the value returned by $f\cos(x)$ when the difference between the point of evaluation x and the point of derivation a is small. The example below computes the value predicted by the Taylor series approximation of $P^2[\cos(x), a]$ and by $\cos(x)$ when $x = 1$ and $a = 0$.

$$
\begin{aligned}
P^2(\cos(x=1), 0) &\approx \cos(x=1) \\
1 - \frac{1}{2}x^2 &\approx \cos(1) \\
1 - \frac{1}{2}1^2 &\approx 0.54 \\
1 - 0.5 &\approx 0.54 \\
0.5 &\approx 0.54
\end{aligned}
$$

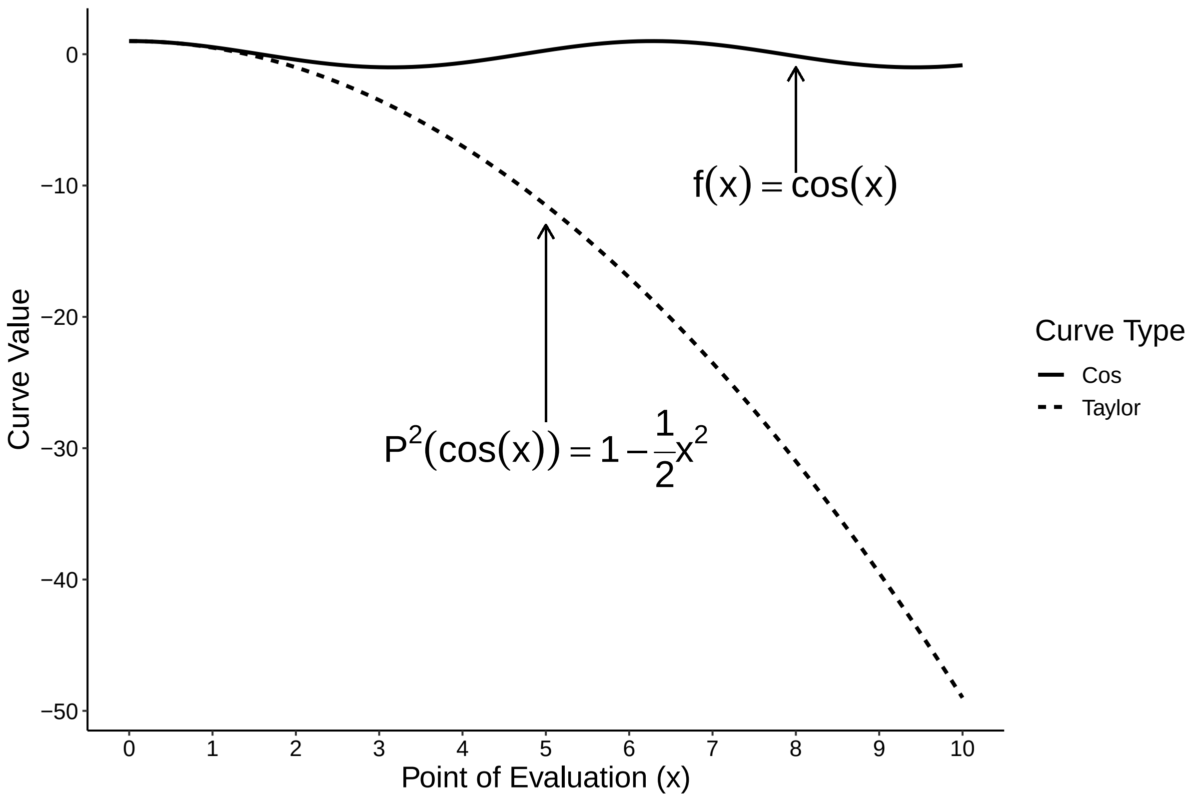
Taylor series approximation of $f\cos(x)$ (specifically, the second-order Taylor series; $P^2[\cos(x), a]$) estimates a a value that is clearly not equal ($\neq$) to the value returned by $f\cos(x)$ when the difference between the point of evaluation x and the point of derivation a is large. The example below computes the value predicted by the Taylor series approximation of $P^2[\cos(x), a]$ and by $\cos(x)$ when $x = 4$ and $a = 0$.

$$
\begin{aligned}
P^2(\cos(x=4), 0) &\neq \cos(x=4) \\
1 - \frac{1}{2}x^2 &\neq \cos(4) \\
1 - \frac{1}{2}4^2 &\neq -0.65 \\
1 - 16 &\neq -0.65 \\
0.5 &\neq -0.65
\end{aligned}
$$

Figure D.1 provides a comprehensive visualization of the of the point conveyed in Example D.1 about the accuracy of Taylor series approximations. In Figure D.1, the values returned by the nonlinear function of $\cos(x)$ and its second-order Taylor series $P^2[\cos(x)] = 1 - \frac{1}{2}x^2$ are shown. The second order Taylor series perfectly estimates $\cos(x)$ when the point of evaluation (x) equals the point of derivation (a; $x = a = 0$), but incurs an increasingly large amount of error as the difference between the point of evaluation and the point of derivation increases. For example, at $x = 10$, $\cos(10) = -0.84$, but the Taylor series outputs a value of -49.50 ($P^2[cos(50)] = 1 - \frac{1}{2}10^2 = -49.50$).

Figure D.1

Estimation Accuracy of Taylor Series Approximation of Nonlinear Function (cos(x))



Note. The second order Taylor series perfectly estimates $\cos(x)$ when the point of evaluation (x) equals the point of derivation (a; $x = a = 0$), but incurs an increasingly large amount of error as the difference between the point of evaluation and the point of derivation increases. For example, at $x = 10$, $\cos(x) = -0.84$, but the Taylor series outputs a value of -49.50 ($P^2[cos(50)] = 1 - \frac{1}{2}10^2 = -49.50$).

D.1.2.2 Taylor Series of the Logistic Function

Given that the Taylor series provides a linear version of a nonlinear function, the structured latent curve modelling approach uses Taylor series' to fit nonlinear functions into the linear nature of the structural equation modelling framework (Browne, 1993; Browne & du Toit, 1991). In the current simulations, the logistic function was used to generate data (see Equation D.10), and so a Taylor series approximation was constructed for the logistic function in the analysis. Note that, because the logistic function had four parameters (θ, α, β, γ), partial derivatives were computed with respect to each of the parameters. Using a derivative order set to one ($n = 1$), the following Taylor series was constructed for the logistic function (Equation D.9):

$$P^1(L(\Theta, t)) = L + \frac{\partial L}{\partial \theta}(x_\theta - a_\theta)^1 + \frac{\partial L}{\partial \alpha}(x_\alpha - a_\alpha)^1 + \frac{\partial L}{\partial \beta}(x_\beta - a_\beta)^1 + \frac{\partial L}{\partial \gamma_\gamma}(x_\gamma - a_\gamma)^1, \tag{D.9}$$

where $\mathbf{L}(\Theta, \mathbf{t})$ represents the logistic function shown below in Equation D.10:

$$\mathbf{L}(\Theta, \mathbf{t}) = \theta + \frac{\alpha - \theta}{1 + e^{\frac{\beta - t}{\gamma}}} + \epsilon, \tag{D.10}$$

with $\Theta = [\theta, \alpha, \beta, \gamma]$ and $\mathbf{L}(\Theta, \mathbf{t})$ being a vector of scores across all $\mathbf{t}$ time points. Because each parameter of the logistic function has a unique meaning (see section on data generation), they are unlikely to have the same population value, and so the derivation (a) will, therefore, differ for each parameter. To set the derivation values (a), the mean values estimated by the structured latent growth curve model for each parameter (i.e., fixed-effect values) are used, meaning that each derivation value in Equation D.9 is replaced with a model estimate as shown below:

- $a_\theta = \hat{\theta}$
- $a_\alpha = \hat{\alpha}$
- $a_\beta = \hat{\beta}$
- $a_\gamma = \hat{\gamma}$

where a caret ($\hat{}$) denotes a parameter value that is estimated by the analysis. In order to compute curves for each p person, evaluation points for each parameter (x_θ, x_α, x_β, x_γ) are set to the value computed for a given person (θ_p, α_p, β_p, γ_p). Thus, each evaluation value in Equation D.9 is replaced with a person-specfic value as shown below:

- $x_\theta = \theta_p$
- $x_\alpha = \alpha_p$
- $x_\beta = \beta_p$
- $x_\gamma = \gamma_p$.

Substituting the above values for the derivation and evaluation values of x and a in the initial logistic Taylor series (Equation D.9) yields the following function (Equation D.11):

$$P^1(L(\Theta,t)) = L(\Theta,t) + \frac{\partial L}{\partial \theta}(\theta_p - \hat{\theta})^1 + \frac{\partial L}{\partial \alpha}(\alpha_p - \hat{\alpha})^1 + \frac{\partial L}{\partial \beta}(\beta_p - \hat{\beta})^1 + \frac{\partial L}{\partial \gamma}(\gamma_p - \hat{\beta})^1. \tag{D.11}$$

Two important points about Equation D.9 deserve mentioning. First, the average population logistic curve (i.e., the fixed-effect parameter values) will have a perfect logistic function shape. In estimating the average population logistic curve, the evaluation values (x) are set equal to the derivation value counterparts (a); that is, each mean value estimated for a parameter ($\hat{\theta}$, $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$) replaces the corresponding derivation-evaluation pair in Equation D.9. Second, it is possible that estimates of random-effect parameters (i.e., variability observed in a parameter's value across people) may be misleading. To compute the values for the random-effect parameters, the evaluation values (a) are set to the logistic function values needed to compute each p person's observed curve (θ_p, α_p, β_p, γ_p). Because Taylor series approximations are only locally accurate, the curves computed for individuals can accommodate shapes that do not resemble a logistic (i.e., s-shaped) pattern (see Example D.1). Thus, estimates of random-effect parameters (i.e., variability observed in a parameter's value across people) can be influenced by curves that do not have a logistic shape and, therefore, may be misleading.

D.1.2.3 Fitting the Logistic Taylor Series Into the Structual Equation Modelling Framework

With the logistic Taylor series computed in Equation D.11, it can be fit into the structural equation modelling framework by transforming it from its scalar form (Equation D.11) into its matrix form (see Equation D.16). In transforming the scalar form of the logistic Taylor series into a matrix form, three steps will be completed, with each step transforming a component of the scalar form into a matrix representation. The paragraphs that follow detail each of these three steps.

First, the partial derivative information must be transformed into their matrix form. The matrix Λ shown below contains the partial derivative information presented in the scalar Taylor series function (see Equation D.11):[33]

$$\Lambda = \begin{bmatrix} \frac{\partial L(\Theta,t_1)}{\partial \theta} & \frac{\partial L(\Theta,t_1)}{\partial \alpha} & \frac{\partial L(\Theta,t_1)}{\partial \beta} & \frac{\partial L(\Theta,t_1)}{\partial \gamma} \\ \frac{\partial L(\Theta,t_2)}{\partial \theta} & \frac{\partial L(\Theta,t_2)}{\partial \alpha} & \frac{\partial L(\Theta,t_2)}{\partial \beta} & \frac{\partial L(\Theta,t_2)}{\partial \gamma} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial L(\Theta,t_n)}{\partial \theta} & \frac{\partial L(\Theta,t_n)}{\partial \alpha} & \frac{\partial L(\Theta,t_n)}{\partial \beta} & \frac{\partial L(\Theta,t_n)}{\partial \gamma} \end{bmatrix}.$$

As in the structural equation modelling framework (see Equation D.2) where each column of Λ specifies a basis curve (i.e., loadings of a growth parameter onto all time points that specify the effect of the parameter over time), each column of Λ in the structured latent curve modelling approach similarly contains the loadings of a logistic function parameter across all the n time points, but the loading values are now determined by the partial

[33] This is also known as a Jacobian matrix.

derivative of the logistic function with respect to that parameter.

Second, the difference between the evaluation and derivation values $(x - a)$ must be transformed into their matrix form. As a reminder, the difference between the evaluation and derivation values is needed so that person-specific curves can be computed. Thus, the difference between the evaluation and derivation values can be conceptualized as person-specific deviation. The vector $\boldsymbol{\iota}_\mathbf{p}$ contains the person-specific deviations (e.g., $\hat{\theta} - \theta_p$) from each mean estimated parameter value as shown below:

$$\boldsymbol{\iota}_\mathbf{p} = \begin{bmatrix} \hat{\theta} - \theta_p \\ \hat{\alpha} - \alpha_p \\ \hat{\beta} - \beta_p \\ \hat{\gamma}_i - \gamma_p \end{bmatrix},$$

where a caret ($\hat{}$) denotes the mean value estimated for a given parameter and a subscript p indicates a parameter value computed for a person.

With a matrix of logistic function loadings (Λ) and the vector of person-specific deviations ($\boldsymbol{\iota}_\mathbf{p}$), person-specific deviations can be computed for each basis column of Λ. Specifically, the person-specific basis column deviations can be computed by post-multiplying the matrix of loadings (Λ) by the vector of person-specific deviations ($\boldsymbol{\iota}_\mathbf{p}$), as shown below in Equation D.12:

$$\text{Basis column deviations}_p = \Lambda \boldsymbol{\iota}_\mathbf{p}. \tag{D.12}$$

Because $\Lambda \boldsymbol{\iota}_\mathbf{p}$ only provides the extent to which each person's curve deviates from the average curve ($\mathbf{L}(\Theta, \mathbf{t})$), it cannot alone be used to compute person-specfic curves. To compute person-specific curves ($\mathbf{y_p}$), the average logistic curve must be added to Equation D.12, as shown below in Equation D.13:

$$\mathbf{y_p} = \mathbf{L}(\Theta, \mathbf{t}) + \Lambda\iota_{\mathbf{p}} + \mathcal{E}_{\mathbf{p}}. \tag{D.13}$$

Unfortunately, the logistic function ($\mathbf{L}(\Theta, \mathbf{t})$) in the above expression (Equation D.13) is simply the original logistic function (see Equation D.10), and so Equation D.13 above is nonlinear. Because Equation D.13 is nonlinear, it cannot be inserted in the structural equation modelling framework, which requires a linear function (see Equation D.2). Thus, the logistic function term in Equation D.13 ($\mathbf{L}(\Theta, \mathbf{t})$) must be linearized so that the logistic Taylor series can be used in the structural equation modelling framework.

Third, and last, the logistic function component ($\mathbf{L}(\Theta, \mathbf{t})$) must be linearized. By taking advantage of some clever linear algebra, the logistic function component can be rewritten as the product of the partial derivative matrix (Λ) and a mean vector (τ; Browne, 1993; Shapiro & Browne, 1987) as shown below in Equation D.14:

$$\mathbf{L}(\Theta, \mathbf{t}) = \boldsymbol{\Lambda}\tau. \tag{D.14}$$

Importantly, the values of the mean vector $\boldsymbol{\tau}$ need to be determined so that a linear representation of the logistic function can be created. Example D.2 below solves for the mean vector (τ) and shows that the values obtained for the linear parameters (i.e., θ and α) constitute the mean values estimated by the analysis (i.e., the fixed-effect values) and zeroes are obtained for the nonlinear parameters (i.e., θ and α). Given that the vector $\boldsymbol{\tau}$ contains mean estimated values, it is often called the mean vector (Blozis, 2004; K. J.

Preacher & Hancock, 2015).

Example D.2. *Computation of mean vector* τ.

Given the parameter estimates of $\hat{\theta} = 3.00$, $\hat{\alpha} = 3.32$, $\hat{\beta} = 180.00$, and $\hat{\gamma} = 20.00$ and $\mathbf{t} = [0, 1, 2, 3]$, $\tau = [3.00, 3.32, 0, 0]$, then

$$\mathbf{L}(\Theta, \mathbf{t}) = \boldsymbol{\Lambda}\tau$$

$$[3.00, 3.02, 3.30, 3.32] = \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 \\ 0.95 & 0.05 & -0.00 & 0.00 \\ 0.05 & 0.95 & -0.00 & -0.00 \\ 0.00 & 1.00 & 0.00 & 0.00 \end{bmatrix} \tau$$

$$\begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 \\ 0.95 & 0.05 & -0.00 & 0.00 \\ 0.05 & 0.95 & -0.00 & -0.00 \\ 0.00 & 1.00 & 0.00 & 0.00 \end{bmatrix}^{-1} \begin{bmatrix} 3.00 \\ 3.02 \\ 3.30 \\ 3.32 \end{bmatrix} = \boldsymbol{\Lambda}\tau$$

$$\tau = [3.00, 3.32, 0, 0]$$

With $\mathbf{L}(\Theta, \mathbf{t}) = \Lambda\tau$, Equation D.13 can be rewritten in a linear equation as shown below in Equation D.15:

$$\mathbf{y_p} = \Lambda\tau + \Lambda\iota_\mathbf{p} + \mathcal{E}_\mathbf{p}. \tag{D.15}$$

Two important points should be made about Equation D.15. First, with some algebraic modification, it can be shown to have the exact same form as the general structural equation modelling framework (see Equation D.2) that expresses a person's score (y_p) as the sum of a loading matrix (Λ) post-multiplied by a vector of person-specific deviations (ι_p) and an error vector ($\mathcal{E}_\mathbf{p}$). To show the equivalence between Equation D.15 and Equation D.2, the mean vector $\boldsymbol{\tau}$ and vector of person-specific deviations $\boldsymbol{\iota}_\mathbf{p}$ can be combined into a new vector $\mathbf{s_p}$ that, like the product of $\boldsymbol{\Lambda\tau}$ (see Equation D.14), also represents the person-specific weights applied to the basis curves in Λ such that

$$\mathbf{s_p} = \tau + \iota_\mathbf{p} = \begin{bmatrix} \hat{\theta} + \hat{\theta} - \theta_p \\ \hat{\alpha} + \hat{\alpha} - \alpha_p \\ 0 + \hat{\beta} - \beta_p \\ 0 + \hat{\gamma} - \gamma_p \end{bmatrix},$$

which allows Equation D.15 to be reexpressed in Equation D.16 below and, thus, take on the exact same form as the general structural equation modelling framework (see Equation D.2)

$$\mathbf{y_p} = \Lambda \mathbf{s_p} + \mathcal{E}_\mathbf{p}. \tag{D.16}$$

Second, the logistic Taylor series shown in Equation D.15 reproduces the nonlinear logistic function. Because the expected value of the person-specific weights ($\mathbf{s_p}$) is the mean vector (τ; $\mathbb{E}[\mathbf{s_p}] = \tau$), the expected set of scores predicted across all people ($\mathbb{E}[\mathbf{y_p}]$) gives back the original expression for the logistic function matrix-vector product in Equation D.14

as shown below in Equation D.17:

$$\mathbb{E}[\mathbf{y_p}] = \Lambda\tau = \mathbf{L}(\Theta, \mathbf{t}). \tag{D.17}$$

Therefore, the structured latent curve modelling approach successfully reproduces the output of the nonlinear logistic function (Equation D.10) with the linear function of Equation D.16. Note that that no error term exists in Equation D.17 because the expected value of the error values is zero ($\mathbb{E}[\mathcal{E}_\mathbf{p}] = 0$).

D.1.2.4 Estimating Parameters in the Structured Latent Curve Modelling Approach

To estimate the parameter values, the full-information maximum likelihood shown in Equation D.18 is computed for each person (i.e., likelihood of observing a p person's data given the estimated parameter values):

$$\mathcal{L}_p = k_p \ln(2\pi) + \ln(|\mathbf{\Sigma_p}| + (\mathbf{y_p} - \mu_\mathbf{p})^\top \mathbf{\Sigma_p}^{-1}(\mathbf{y_p} - \mu_\mathbf{p}), \tag{D.18}$$

where k_p is the number of non-missing values for a given p person, $\mathbf{\Sigma_p}$ is the model-implied covariance matrix with rows and columns filtered at time points where person p has missing data, $\mathbf{y_p}$ is a vector containing the data points collected for a p person (i.e., filtered data), and $\mu_\mathbf{p}$ is the model-implied mean vector that is filtered at time points where person p has missing data. Note that, because all my simulations assumed complete data across all times points, no filtering procedures were executed (for a review of the filtering procedure, see Boker et al., 2020, Chapter 5). Thus, computing the above full-information maximum likelihood in Equation D.18 is equivalent to computing the below likelihood function in Equation D.19:

$$\mathcal{L}_p = k_p \ln(2\pi) + \ln(|\boldsymbol{\Sigma}| + (\mathbf{y_p} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y_p} - \boldsymbol{\mu}), \tag{D.19}$$

where $\boldsymbol{\Sigma}$ is the model-implied covariance matrix, $\mathbf{y_p}$ contains the data collected from a p person, and $\boldsymbol{\mu}$ is the model-implied mean vector. The model-implied covariance matrix $\boldsymbol{\Sigma}$ is computed using Equation D.20 below:

$$\boldsymbol{\Sigma} = \Lambda \Psi \Lambda + \Omega_{\mathcal{E}}, \tag{D.20}$$

where Ψ is the random-effect covariance matrix and $\Omega_{\mathcal{E}}$ contains the error variances at each time point. The mean vector $\boldsymbol{\mu}$ is computed using Equation D.21 shown below:

$$\boldsymbol{\mu} = \Lambda \boldsymbol{\tau}. \tag{D.21}$$

Parameter estimation is conducted by finding values for the model-implied covariance matrix $\boldsymbol{\Sigma}$ and the model-implied mean vector $\boldsymbol{\mu}$ that maximizes the sum of log-likelihoods across all P people (see Equation D.22 below):

$$\mathcal{L} = \underset{\boldsymbol{\Sigma}, \boldsymbol{\mu}}{\arg\max} \sum_{p=1}^{P} \mathcal{L}_p. \tag{D.22}$$

In OpenMx, the above problem is solved using the sequential least squares quadratic program (for a review, see Kraft, 1994).

Appendix E: OpenMx Code for Structured Latent Growth Curve Model Used in Simulation Experiments

The code that I used to model logistic pattern of change (see data generation) is shown in Code Block E.1. Note that, the code is largely excerpted from the `run_exp_simulations()` and `create_logistic_model_ns()` functions from the `nonlinSims` package, and so readers interested in obtaining more information should consult the source code of this package. One important point to mention is that the model specified in Code Block E.1 assumes time-structured data.

Code Block E.1

OpenMx Code for Structured Latent Growth Curve Model That Assumes Time-Structured Data

```r
#Days on which measurements are assumed to be taken (note that model assumes
time-structured data; that is, at each time point, participants provide data at the
exact same moment). The measurement days obtained by finding the unique values in the
`measurement_day` column of the generated data set.
measurement_days <- unique(data$measurement_day)

#Manifest variable names (i.e., names of columns containing data at each time point,
manifest_vars <- nonlinSims:::extract_manifest_var_names(data_wide = data_wide)

#Now convert data to wide format (needed for OpenMx)
data_wide <- data[ , c(1:3, 5)] %>%
    pivot_wider(names_from = measurement_day, values_from = c(obs_score,
    actual_measurement_day))

#Remove . from column names so that OpenMx does not run into error (this occurs
because, with some spacing schedules, measurement days are not integer values.)
names(data_wide) <- str_replace(string = names(data_wide), pattern = '\\.', replacement
= '_')

#Latent variable names (theta = baseline, alpha = maximal elevation, beta =
days-to-halfway elevation, gamma = triquarter-haflway elevation)
latent_vars <- c('theta', 'alpha', 'beta', 'gamma')

latent_growth_curve_model <- mxModel(
  model = model_name,
  type = 'RAM', independent = T,
  mxData(observed = data_wide, type = 'raw'),

  manifestVars = manifest_vars,
  latentVars = latent_vars,

  #Residual variances; by using one label, they are assumed to all be equal
  (homogeneity of variance). That is, there is no complex error structure.
  mxPath(from = manifest_vars,
         arrows=2, free=TRUE,  labels='epsilon', values = 1, lbound = 0),

  #Latent variable covariances and variances (note that only the variances are
  estimated. )
```

```
mxPath(from = latent_vars,
       connect='unique.pairs', arrows=2,
       free = c(TRUE,FALSE, FALSE, FALSE,
                TRUE, FALSE, FALSE,
                TRUE, FALSE,
                TRUE),
       values=c(1, NA, NA, NA,
                1, NA, NA,
                1, NA,
                1),
       labels=c('theta_rand', 'NA(cov_theta_alpha)', 'NA(cov_theta_beta)',
                'NA(cov_theta_gamma)',
                'alpha_rand','NA(cov_alpha_beta)', 'NA(cov_alpha_gamma)',
                'beta_rand', 'NA(cov_beta_gamma)',
                'gamma_rand'),
       lbound = c(1e-3, NA, NA, NA,
                  1e-3, NA, NA,
                  1, NA,
                  1),
       ubound = c(2, NA, NA, NA,
                  2, NA, NA,
                  90^2, NA,
                  45^2)),

# Latent variable means (linear parameters). Note that the parameters of beta and gamma do not have estimated means because they are nonlinear parameters (i.e., the logistic function's first-order partial derivative with respect to each of those two parameters contains those two parameters. )
mxPath(from = 'one', to = c('theta', 'alpha'), free = c(TRUE, TRUE), arrows = 1,
       labels = c('theta_fixed', 'alpha_fixed'), lbound = 0, ubound = 7,
       values = c(1, 1)),

#Functional constraints (needed to estimate mean values of fixed-effect parameters)
mxMatrix(type = 'Full', nrow = length(manifest_vars), ncol = 1, free = TRUE,
         labels = 'theta_fixed', name = 't', values = 1, lbound = 0,  ubound = 7),
mxMatrix(type = 'Full', nrow = length(manifest_vars), ncol = 1, free = TRUE,
         labels = 'alpha_fixed', name = 'a', values = 1, lbound = 0,  ubound = 7),
mxMatrix(type = 'Full', nrow = length(manifest_vars), ncol = 1, free = TRUE,
         labels = 'beta_fixed', name = 'b', values = 1, lbound = 1, ubound = 360),
mxMatrix(type = 'Full', nrow = length(manifest_vars), ncol = 1, free = TRUE,
         labels = 'gamma_fixed', name = 'g', values = 1, lbound = 1, ubound = 360),

mxMatrix(type = 'Full', nrow = length(manifest_vars), ncol = 1, free = FALSE,
         values = measurement_days, name = 'time'),

#Algebra specifying first-order partial derivatives;
mxAlgebra(expression = 1 - 1/(1 + exp((b - time)/g)), name="Tl"),
mxAlgebra(expression = 1/(1 + exp((b - time)/g)), name = 'Al'),

mxAlgebra(expression = -((a - t) * (exp((b - time)/g) * (1/g))/(1 + exp((b - time)/g))^2), name = 'Bl'),
mxAlgebra(expression =  (a - t) * (exp((b - time)/g) * ((b - time)/g^2))/(1 + exp((b -time)/g))^2, name = 'Gl'),

#Factor loadings; all fixed and, importantly, constrained to change according to their partial derivatives (i.e., nonlinear functions)
mxPath(from = 'theta', to = manifest_vars, arrows=1, free=FALSE,
       labels = sprintf(fmt = 'Tl[%d,1]', 1:length(manifest_vars))),
mxPath(from = 'alpha', to = manifest_vars, arrows=1, free=FALSE,
       labels = sprintf(fmt = 'Al[%d,1]', 1:length(manifest_vars))),
mxPath(from='beta', to = manifest_vars, arrows=1,  free=FALSE,
       labels =  sprintf(fmt = 'Bl[%d,1]', 1:length(manifest_vars))),
mxPath(from='gamma', to = manifest_vars, arrows=1,  free=FALSE,
       labels =  sprintf(fmt = 'Gl[%d,1]', 1:length(manifest_vars))),

#Fit function used to estimate free parameter values.
```

```
  mxFitFunctionML(vector = FALSE)
)

#Use starting value function from OpenMx to generate good starting values (uses
weighted least squares)
latent_growth_model <- mxAutoStart(model = latent_growth_model)

#Fit model using mxTryHard(). Increases probability of convergence by attempting model
convergence by randomly shifting starting values.
model_results <- mxTryHard(latent_growth_model)
```

Appendix F: Complete Versions of Bias/Precision Plots (Day- and Likert-Unit Parameters)

F.1 Experiment 1

F.1.1 Equal Spacing

Figure F.1

Bias/Precision Plots for Day- and Likert-Unit Parameters With Equal Spacing in Experiment 1

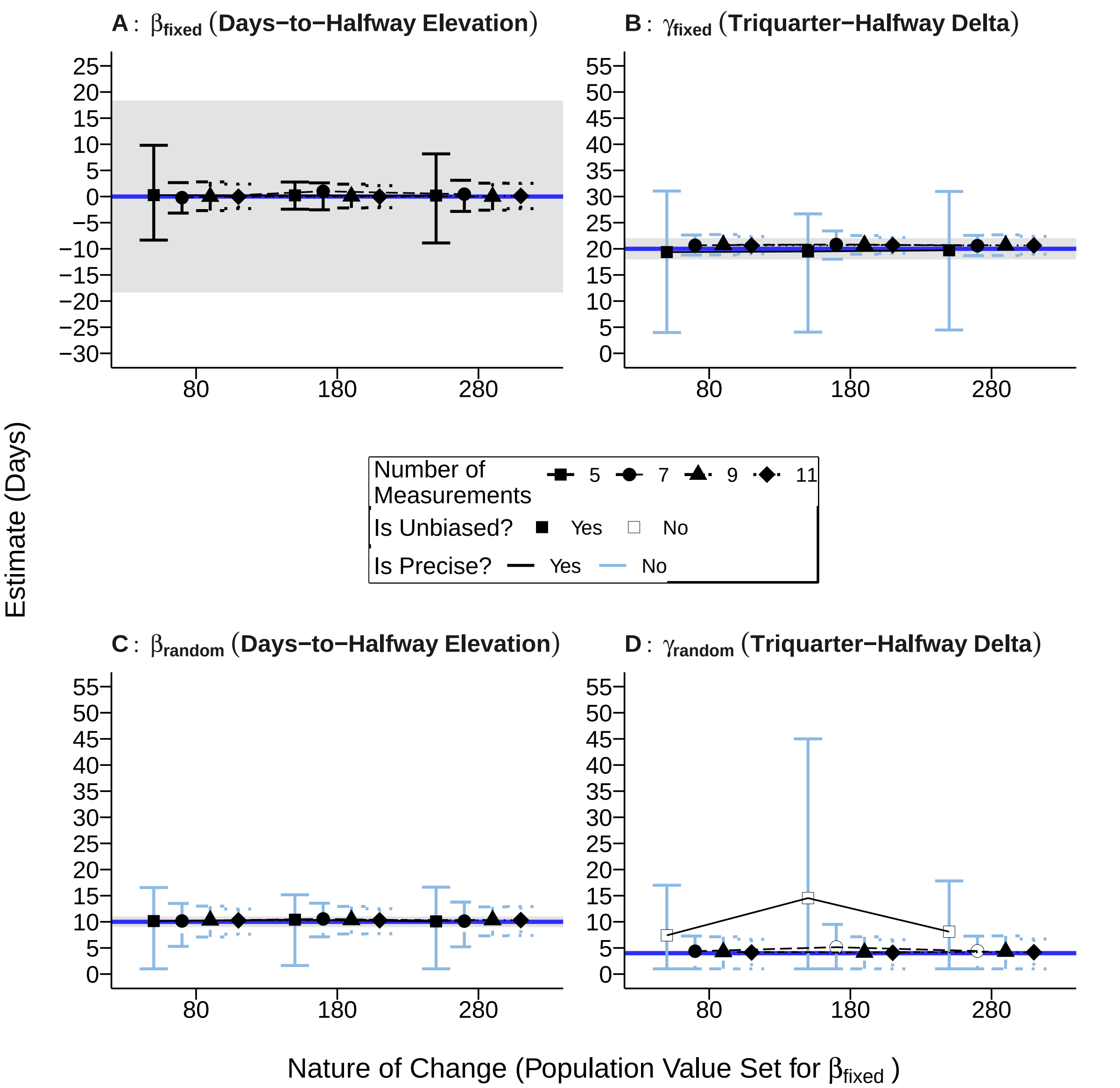
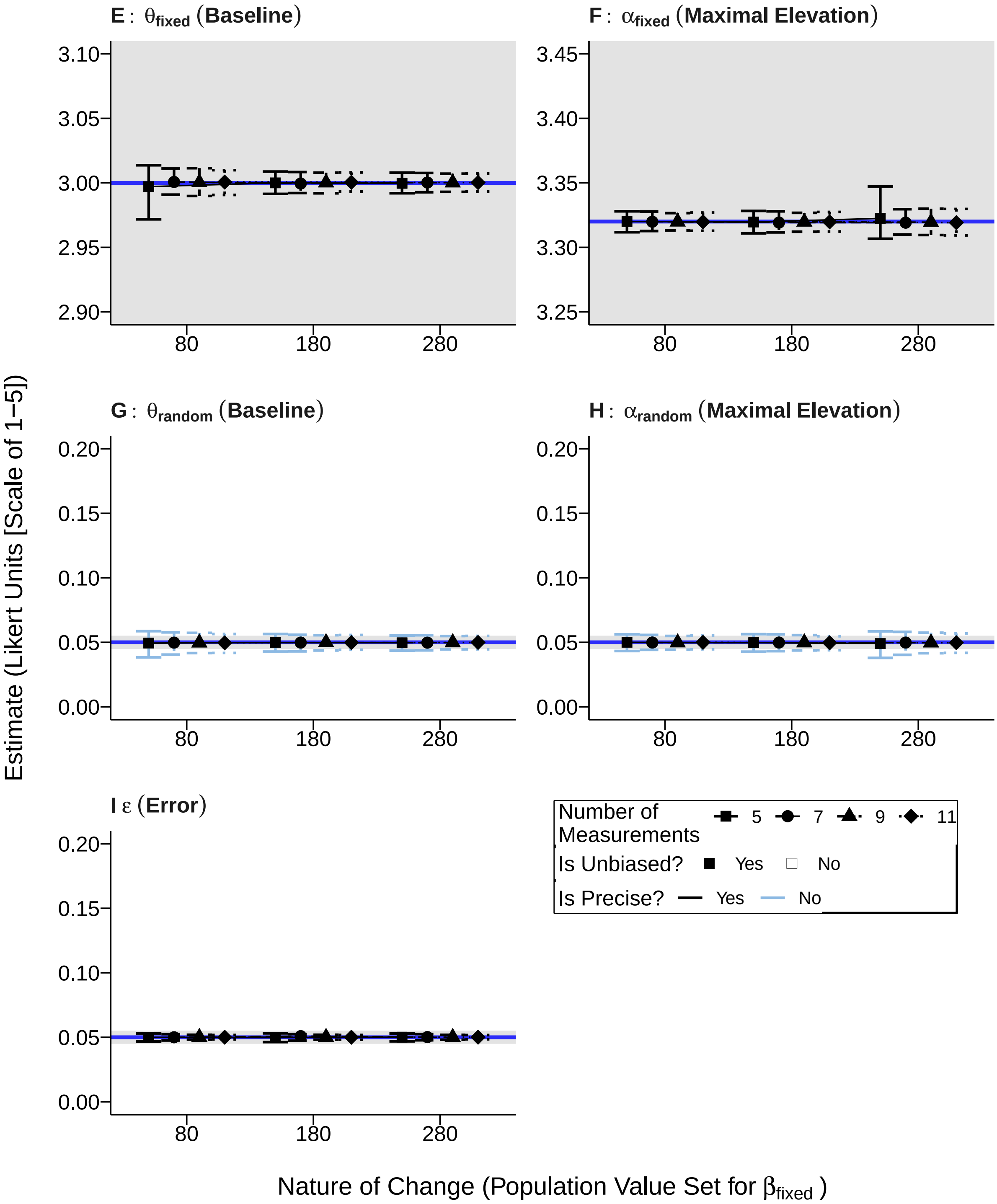


Figure F.1

Bias/Precision Plots for Day- and Likert-Unit Parameters With Equal Spacing in Experiment 1 (continued)



4449 *Note.* Panels A–B: Bias/precision plots for the fixed- and random-effect days-to-halfway elevation
4450 parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Bias/precision plots for the fixed- and
4451 random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F:
4452 Bias/precision plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and θ_{random}).

Panels G–H: Bias/precision plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table H.1 for specific values estimated for each parameter.

F.1.2 Time-Interval Increasing Spacing

Figure F.2

Bias/Precision Plots for Day- and Likert-Unit Parameters With Time-Interval Increasing Spacing in Experiment 1

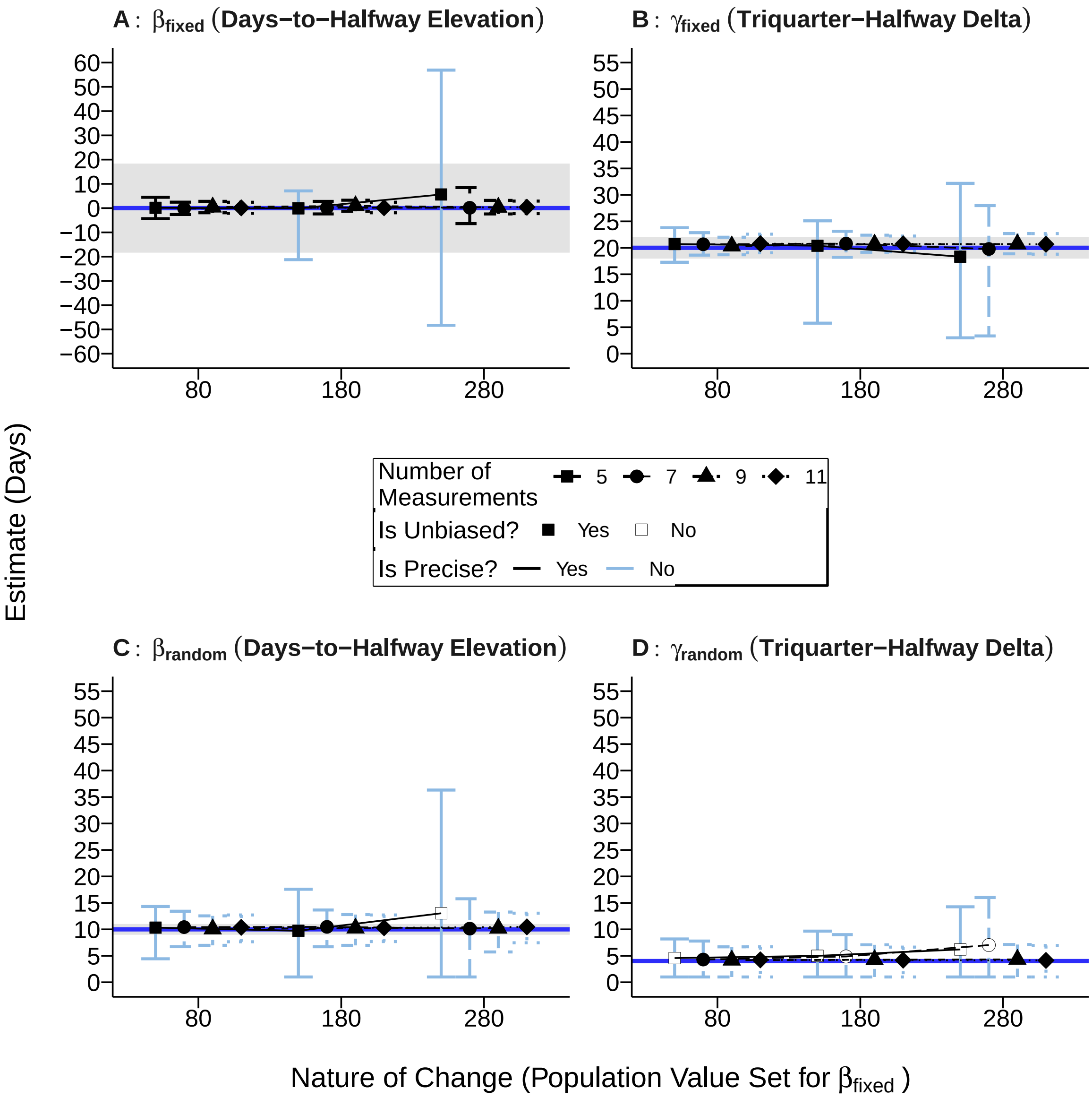
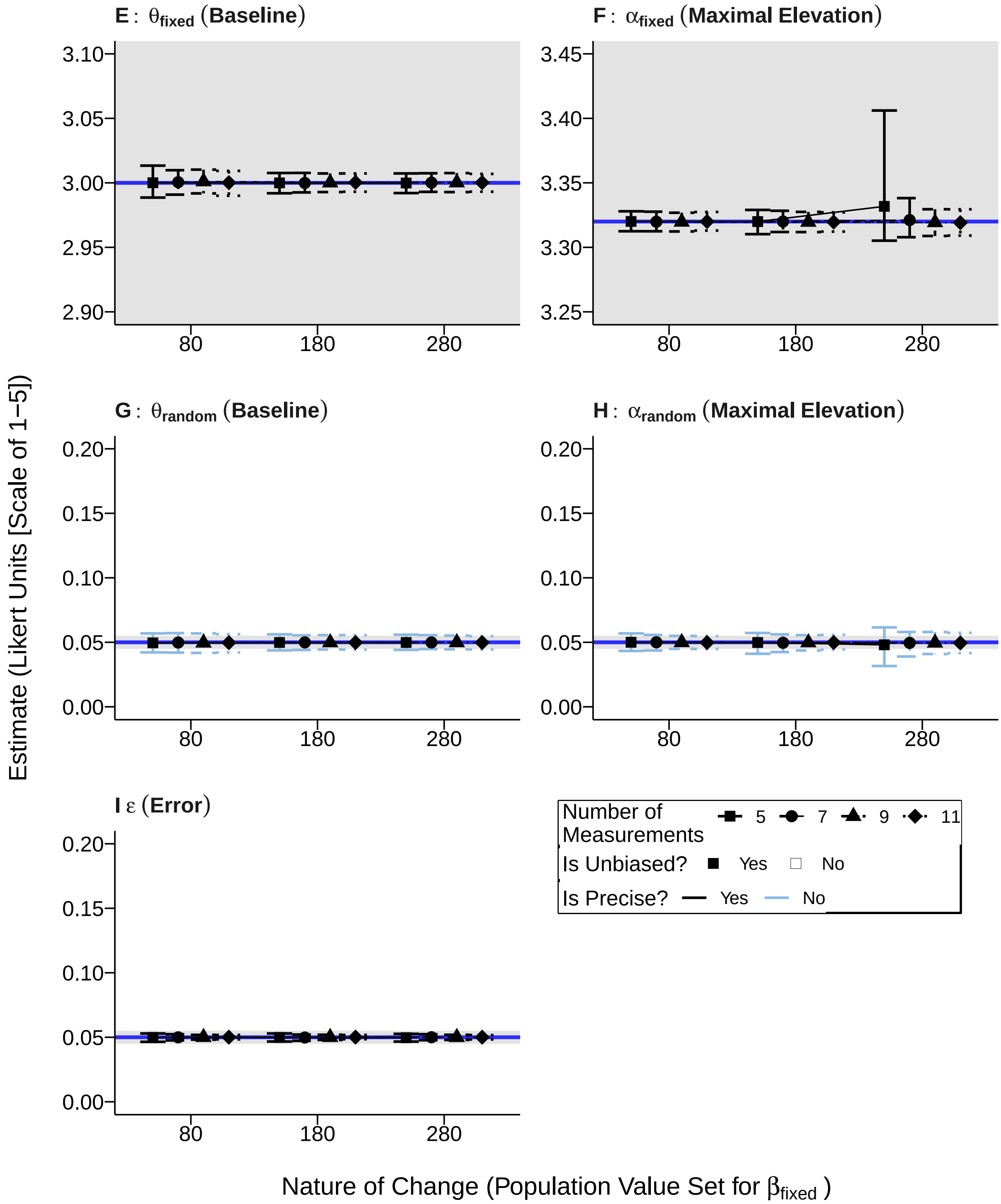


Figure F.2

Bias/Precision Plots for Day- and Likert-Unit Parameters With Time-Interval Increasing Spacing in Experiment 1 (continued)



Note. Panels A–B: Bias/precision plots for the fixed- and random-effect days-to-halfway elevation parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Bias/precision plots for the fixed- and random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F: Bias/precision plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and θ_{random}).

Panels G–H: Bias/precision plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table H.1 for specific values estimated for each parameter.

F.1.3 Time-Interval Decreasing Spacing

Figure F.3

Bias/Precision Plots for Day- and Likert-Unit Parameters With Time-Interval Decreasing Spacing in Experiment 1

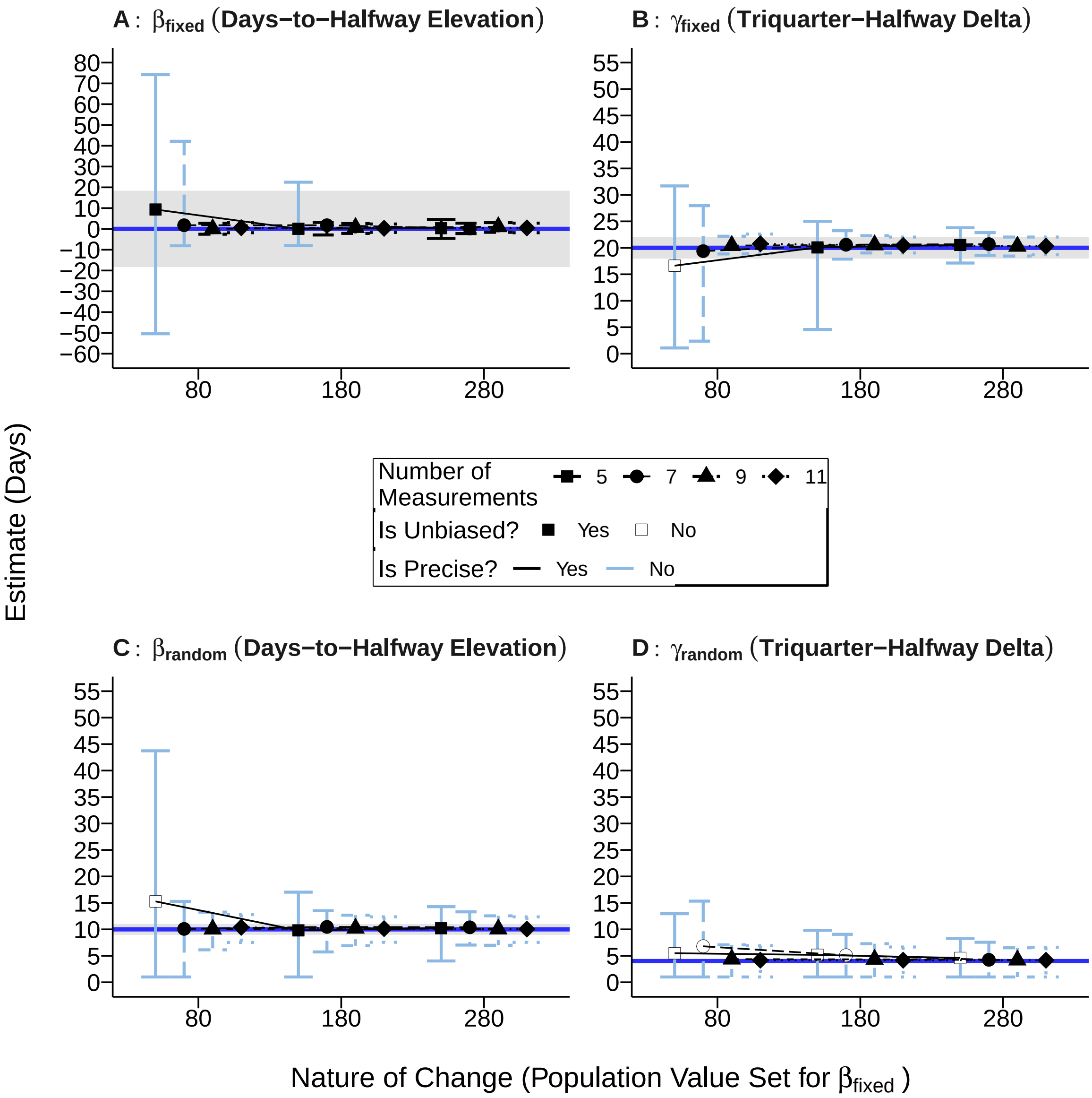
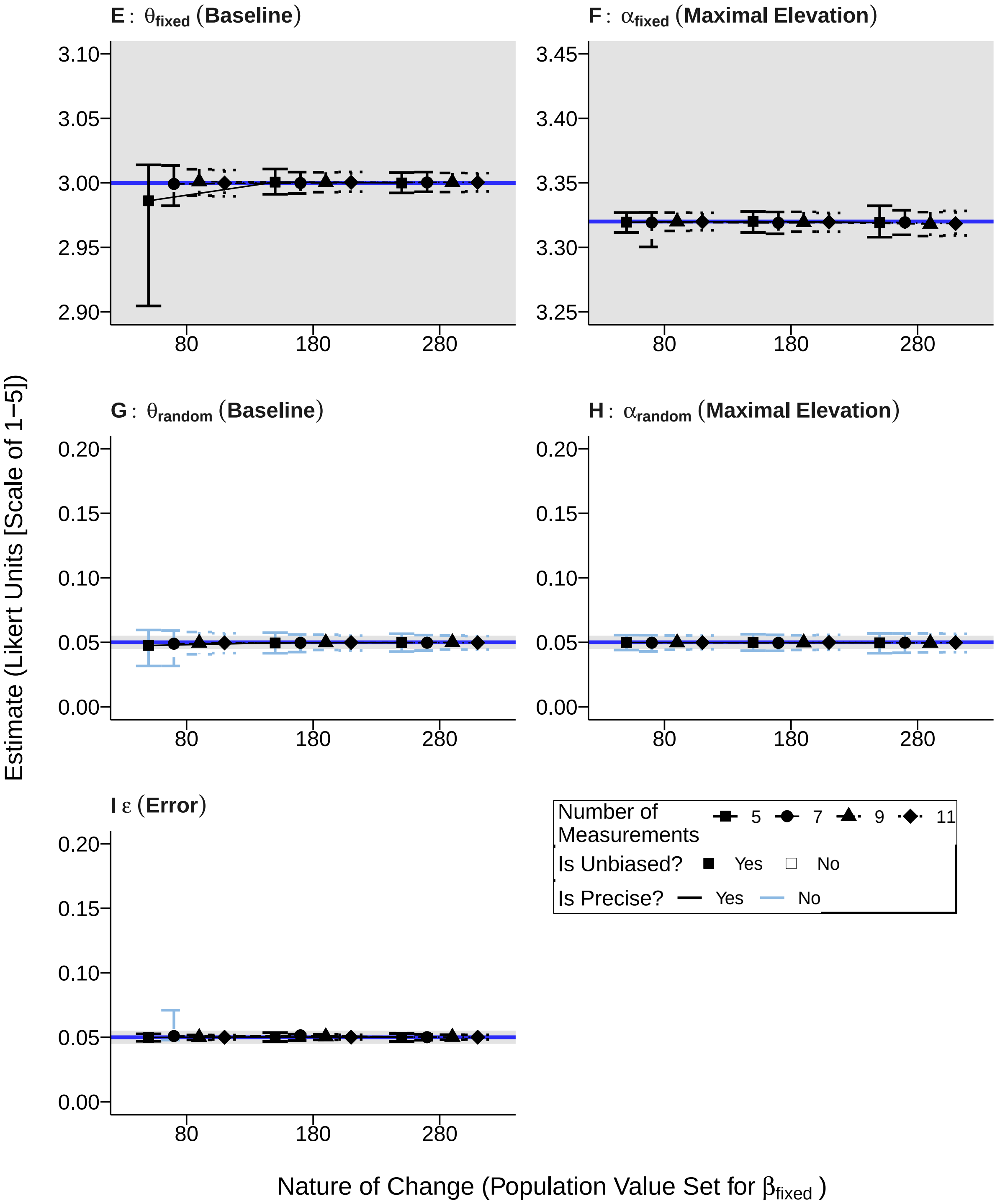


Figure F.3

Bias/Precision Plots for Day- and Likert-Unit Parameters With Time-Interval Decreasing Spacing in Experiment 1 (continued)



Note. Panels A–B: Bias/precision plots for the fixed- and random-effect days-to-halfway elevation parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Bias/precision plots for the fixed- and random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F: Bias/precision plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and θ_{random}).

Panels G–H: Bias/precision plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table H.1 for specific values estimated for each parameter.

F.1.4 Middle-and-Extreme Spacing

Figure F.4

Bias/Precision Plots for Day- and Likert-Unit Parameters With Middle-and-Extreme Spacing in Experiment 1

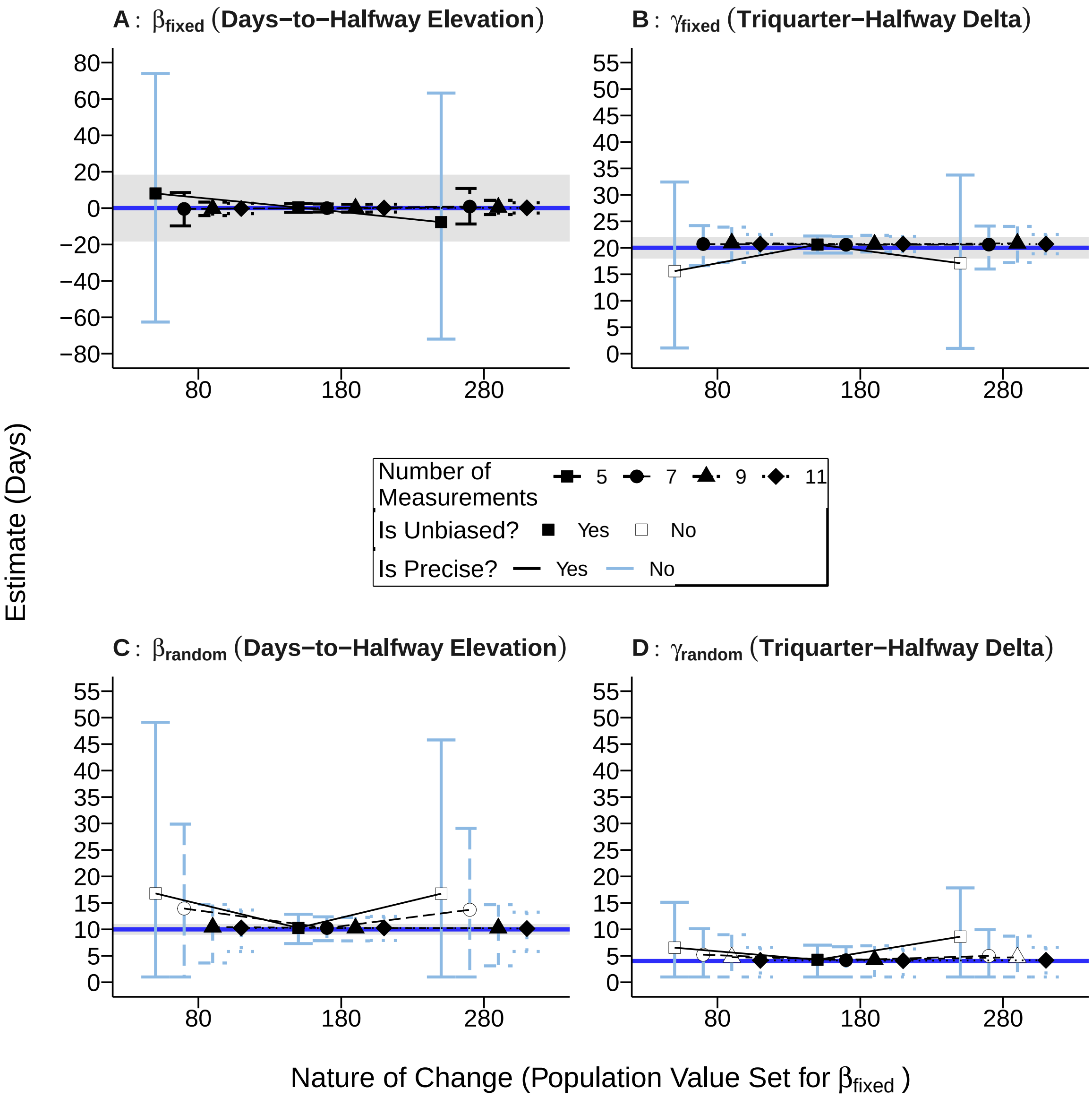
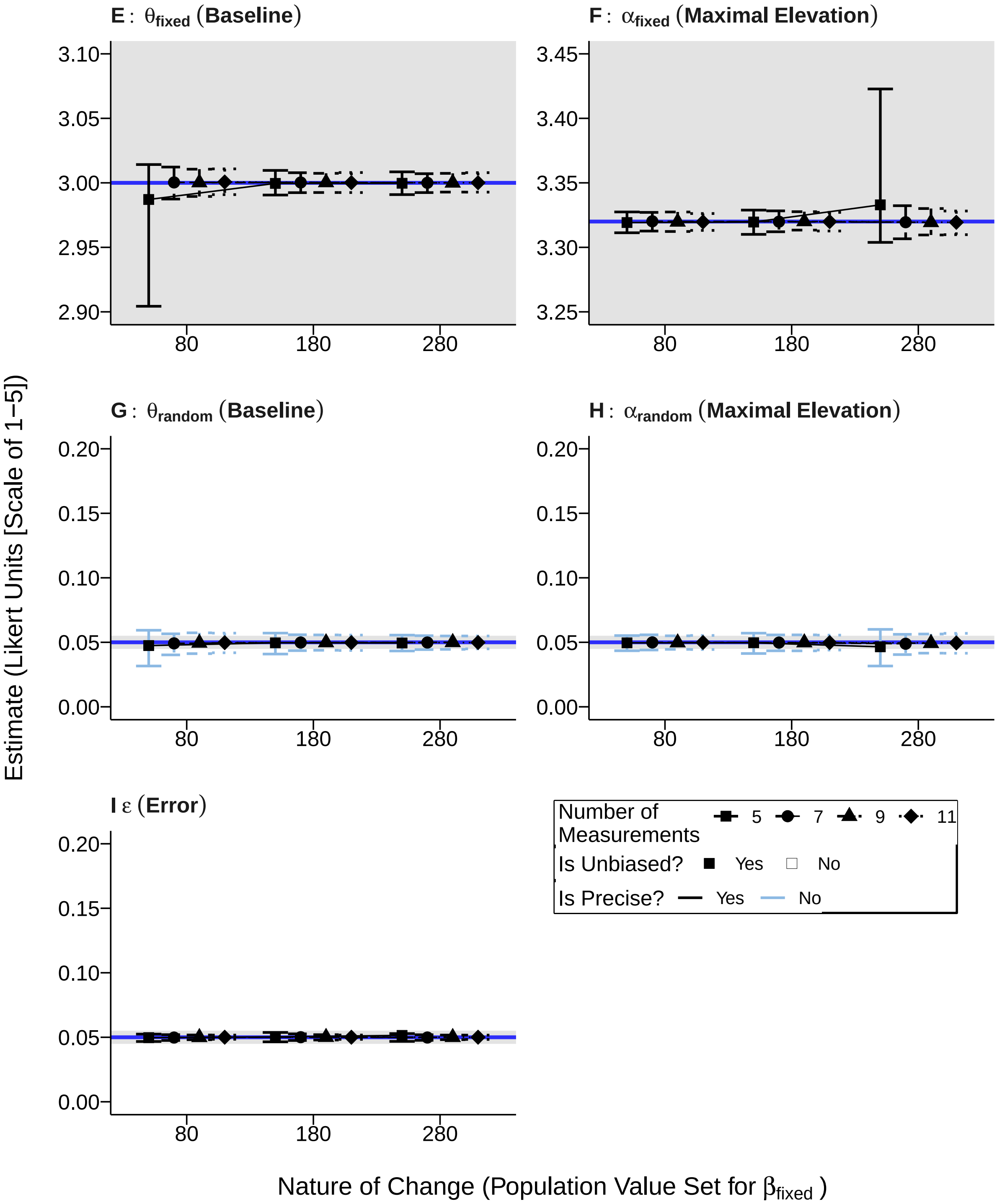


Figure F.4

Bias/Precision Plots for Day- and Likert-Unit Parameters With Middle-and-Extreme Spacing in Experiment 1 (continued)



Note. Panels A–B: Bias/precision plots for the fixed- and random-effect days-to-halfway elevation parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Bias/precision plots for the fixed- and random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F: Bias/precision plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and θ_{random}).

Panels G–H: Bias/precision plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table H.1 for specific values estimated for each parameter.

F.2 Experiment 2

F.2.5 Equal Spacing

Figure F.5

Bias/Precision Plots for Day- and Likert-Unit Parameters With Equal Spacing in Experiment 2

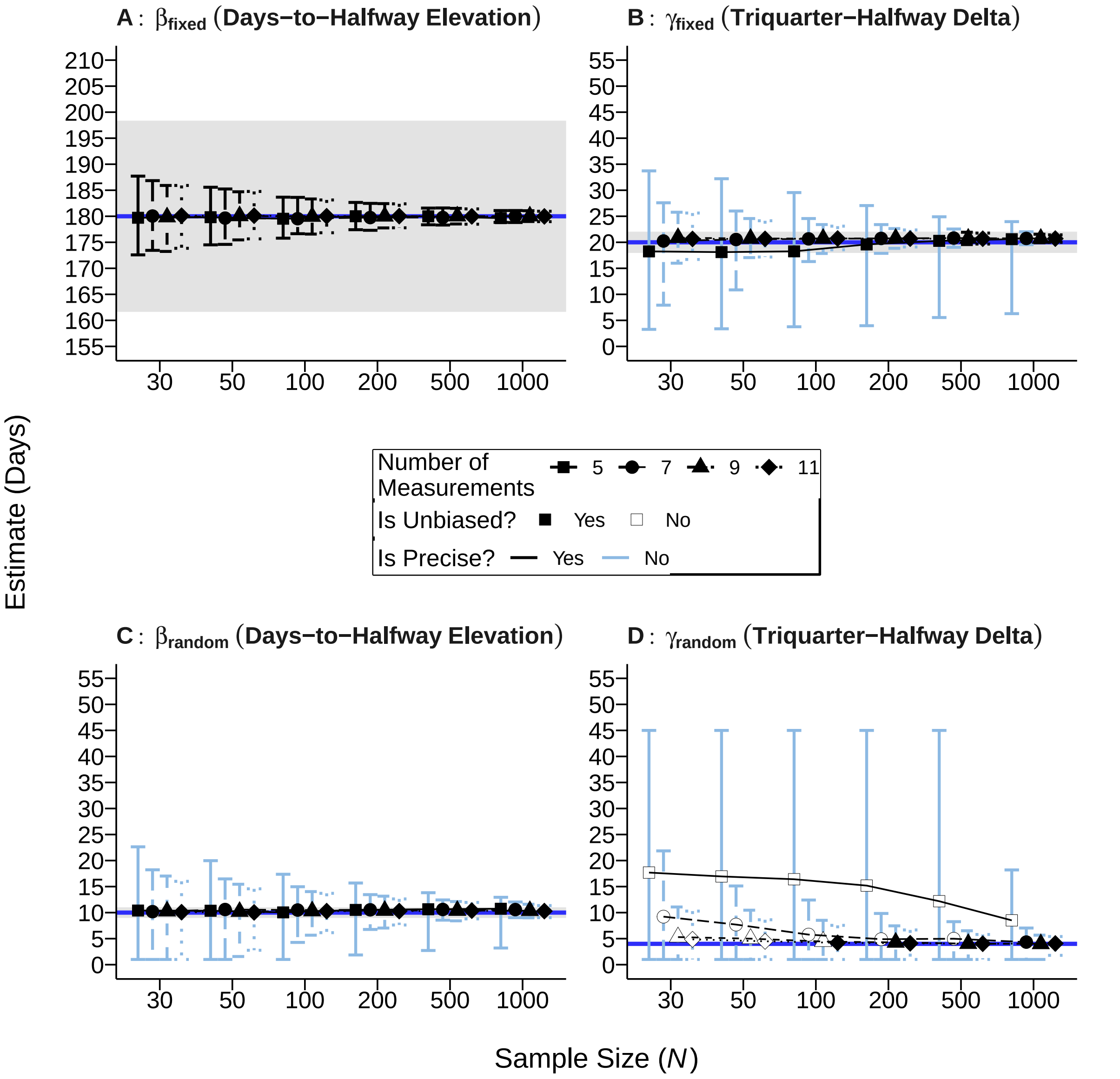
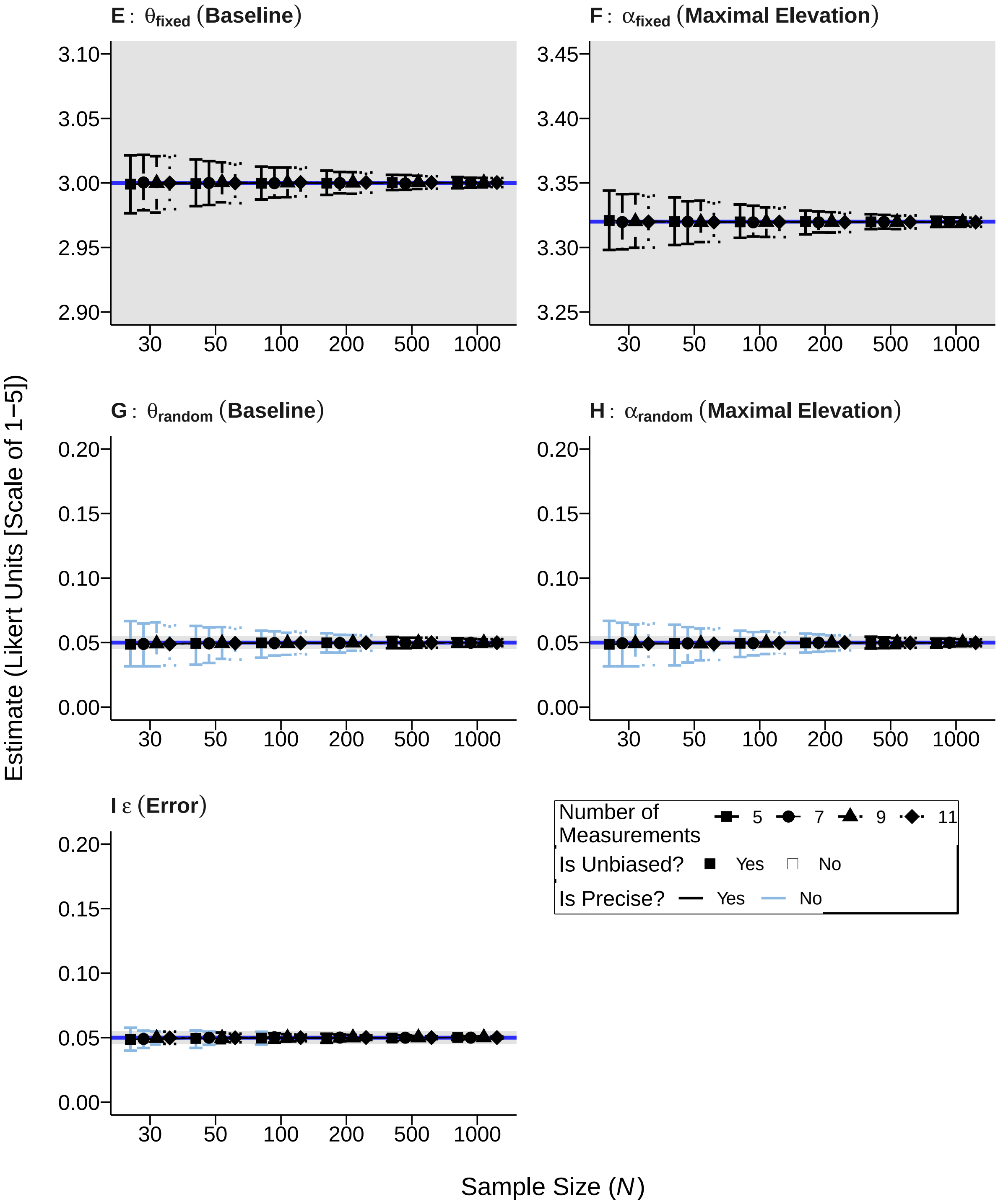


Figure F.5

Bias/Precision Plots for Day- and Likert-Unit Parameters With Equal Spacing in Experiment 2 (continued)



Note. Panels A–B: Bias/precision plots for the fixed- and random-effect days-to-halfway elevation parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Bias/precision plots for the fixed- and random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F: Bias/precision plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and θ_{random}).

Panels G–H: Bias/precision plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table H.2 for specific values estimated for each parameter.

F.2.6 Time-Interval Increasing Spacing

Figure F.6

Bias/Precision Plots for Day- and Likert-Unit Parameters With Time-Interval Increasing Spacing in Experiment 2

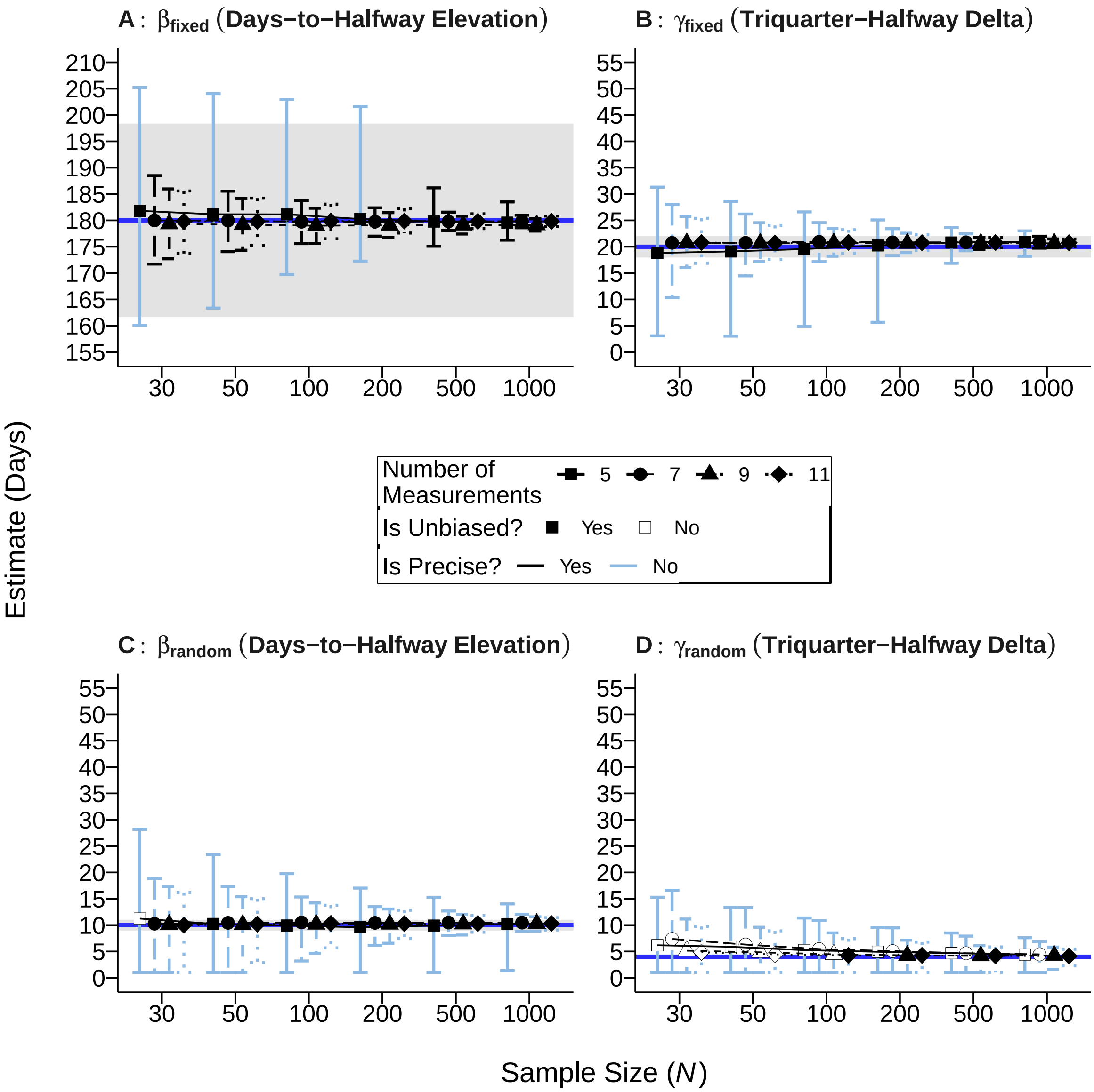
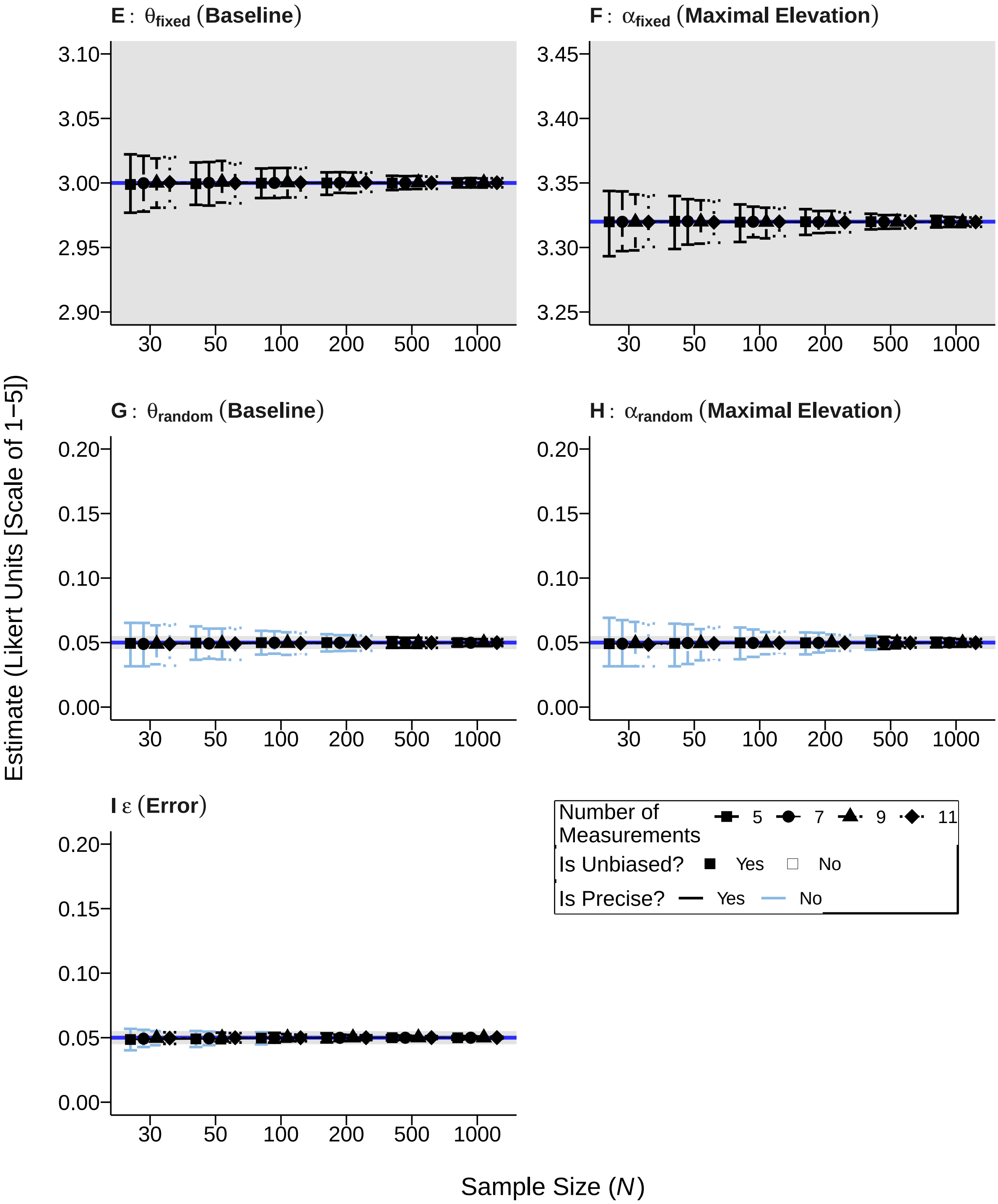


Figure F.6

Bias/Precision Plots for Day- and Likert-Unit Parameters With Time-Interval Increasing Spacing in Experiment 2 (continued)



4534 *Note.* Panels A–B: Bias/precision plots for the fixed- and random-effect days-to-halfway elevation
4535 parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Bias/precision plots for the fixed- and
4536 random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F:
4537 Bias/precision plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and θ_{random}).

Panels G–H: Bias/precision plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table H.2 for specific values estimated for each parameter.

4549 ## F.2.7 Time-Interval Decreasing Spacing

Figure F.7

Bias/Precision Plots for Day- and Likert-Unit Parameters With Time-Interval Decreasing Spacing in Experiment 2

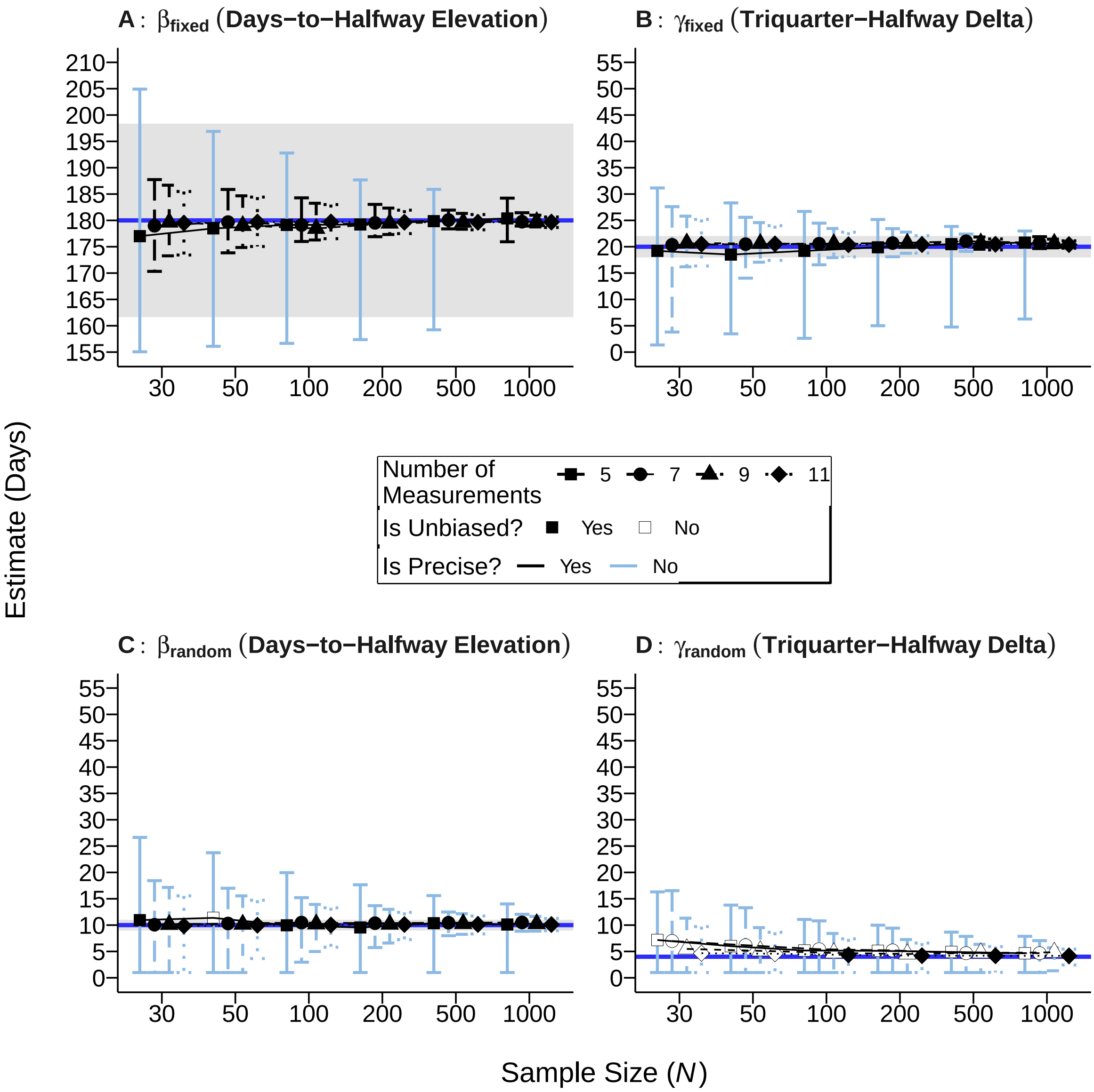
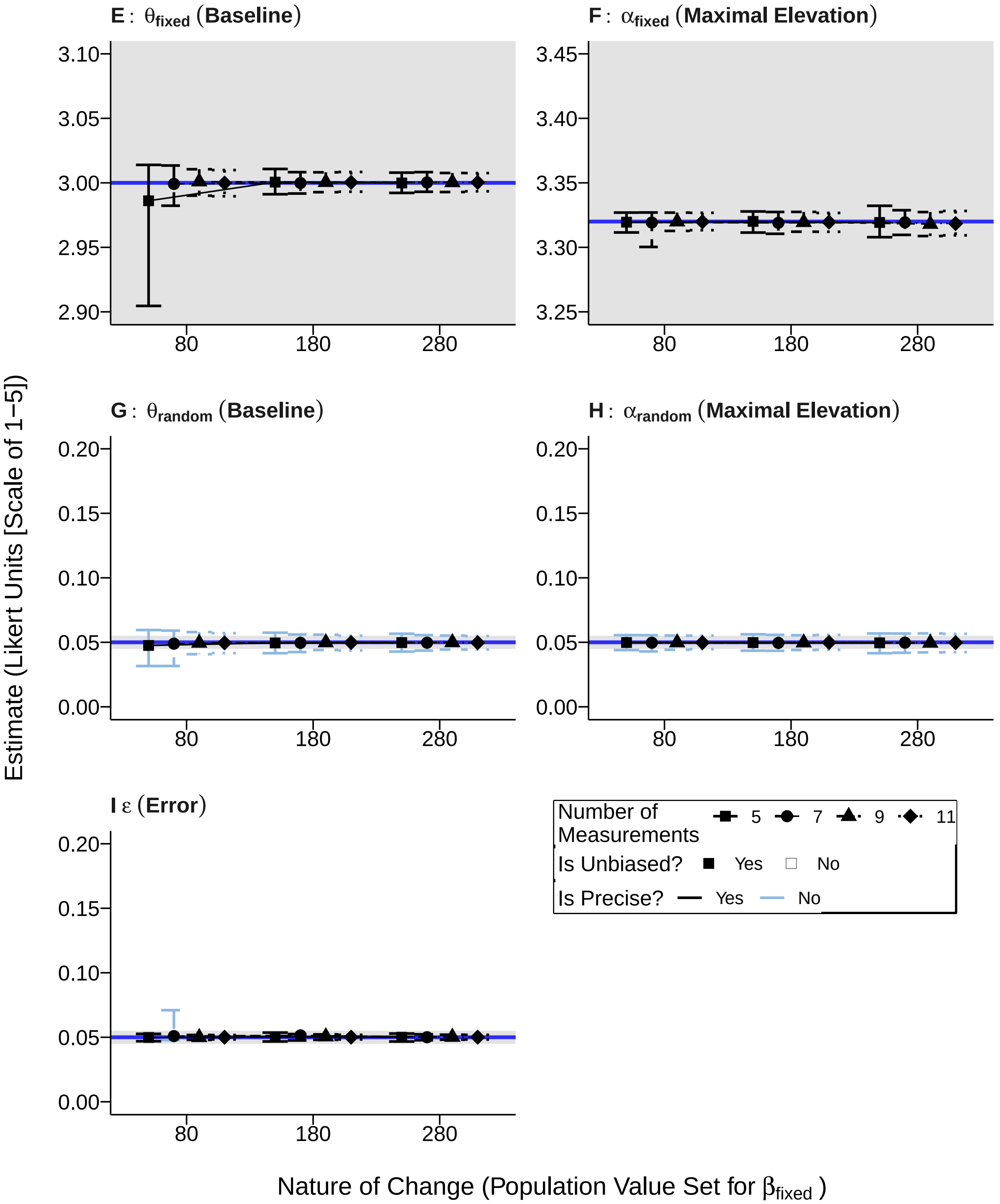


Figure F.7

Bias/Precision Plots for Day- and Likert-Unit Parameters With Time-Interval Decreasing Spacing in Experiment 2 (continued)



4550 *Note.* Panels A–B: Bias/precision plots for the fixed- and random-effect days-to-halfway elevation
4551 parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Bias/precision plots for the fixed- and
4552 random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F:
4553 Bias/precision plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and θ_{random}).

Panels G–H: Bias/precision plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table H.2 for specific values estimated for each parameter.

F.2.8 Middle-and-Extreme Spacing

Figure F.8

Bias/Precision Plots for Day- and Likert-Unit Parameters With Middle-and-Extreme Spacing in Experiment 2

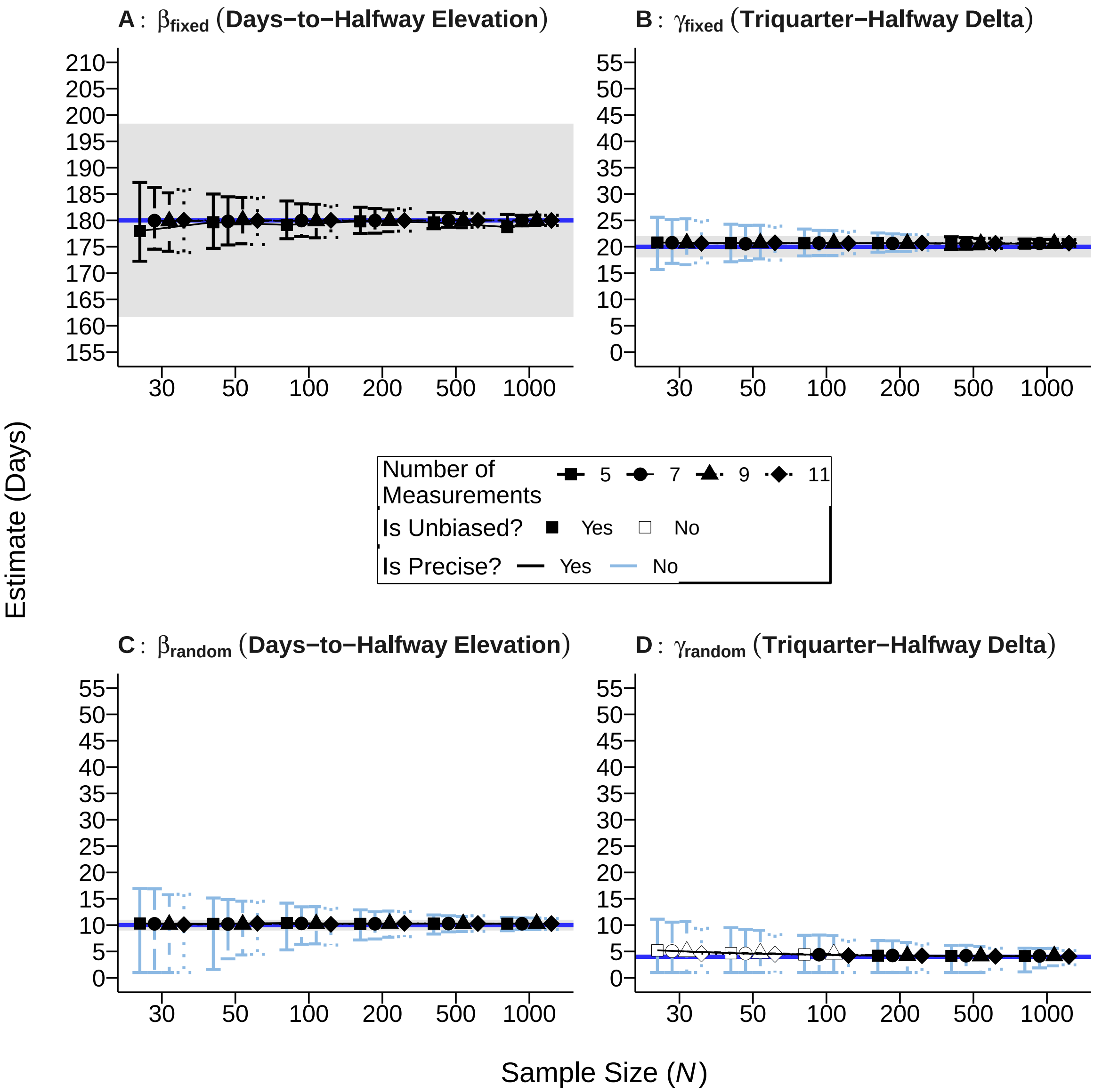
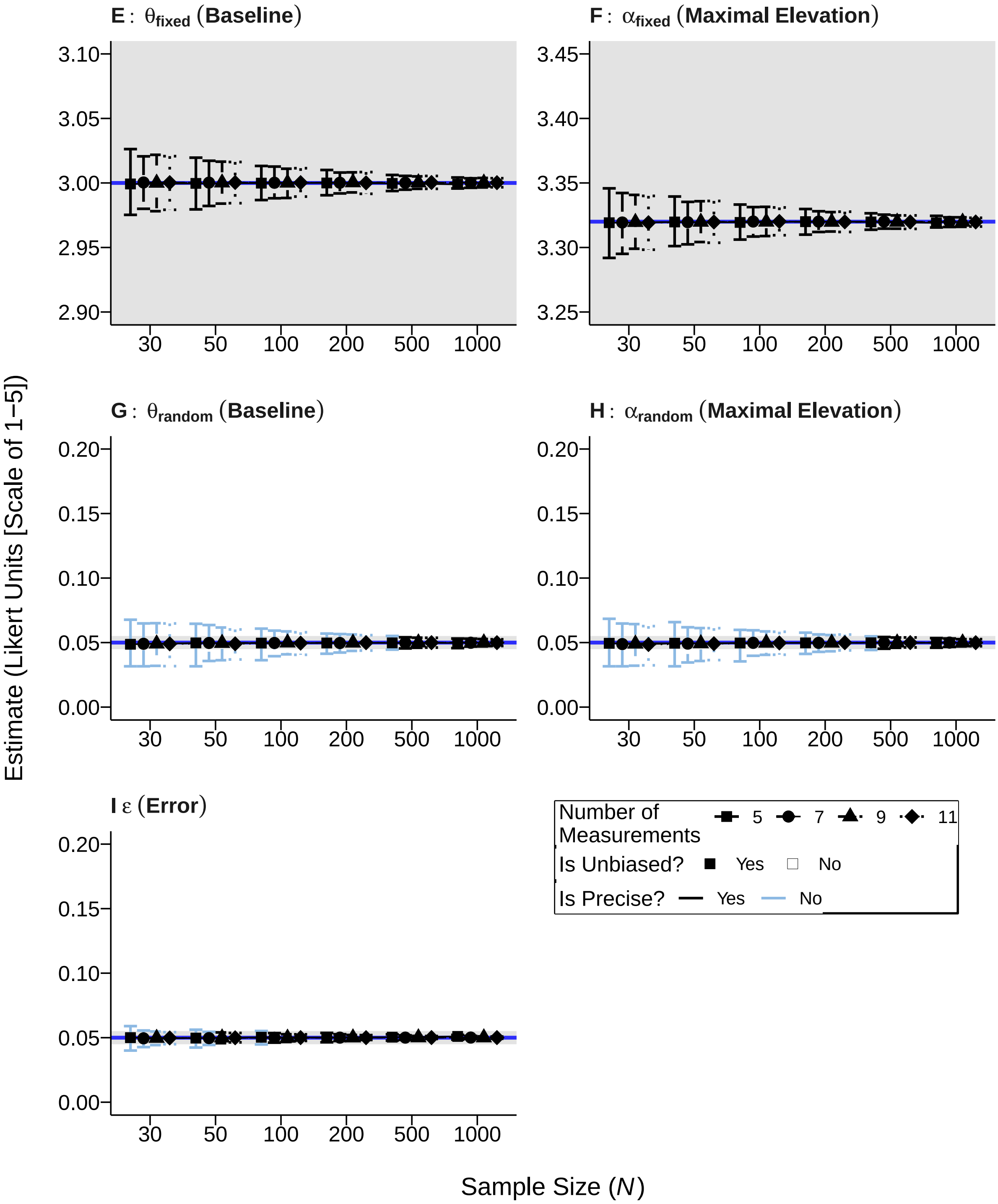


Figure F.8

Bias/Precision Plots for Day- and Likert-Unit Parameters With Middle-and-Extreme Spacing in Experiment 2 (continued)



4566 *Note.* Panels A–B: Bias/precision plots for the fixed- and random-effect days-to-halfway elevation
4567 parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Bias/precision plots for the fixed- and
4568 random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F:
4569 Bias/precision plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and θ_{random}).

Panels G–H: Bias/precision plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table H.2 for specific values estimated for each parameter.

F.3 Experiment 3

F.3.9 Time-Structured Data

Figure F.9

Bias/Precision Plots for Day- and Likert-Unit Parameters With Time-Structured Data in Experiment 3

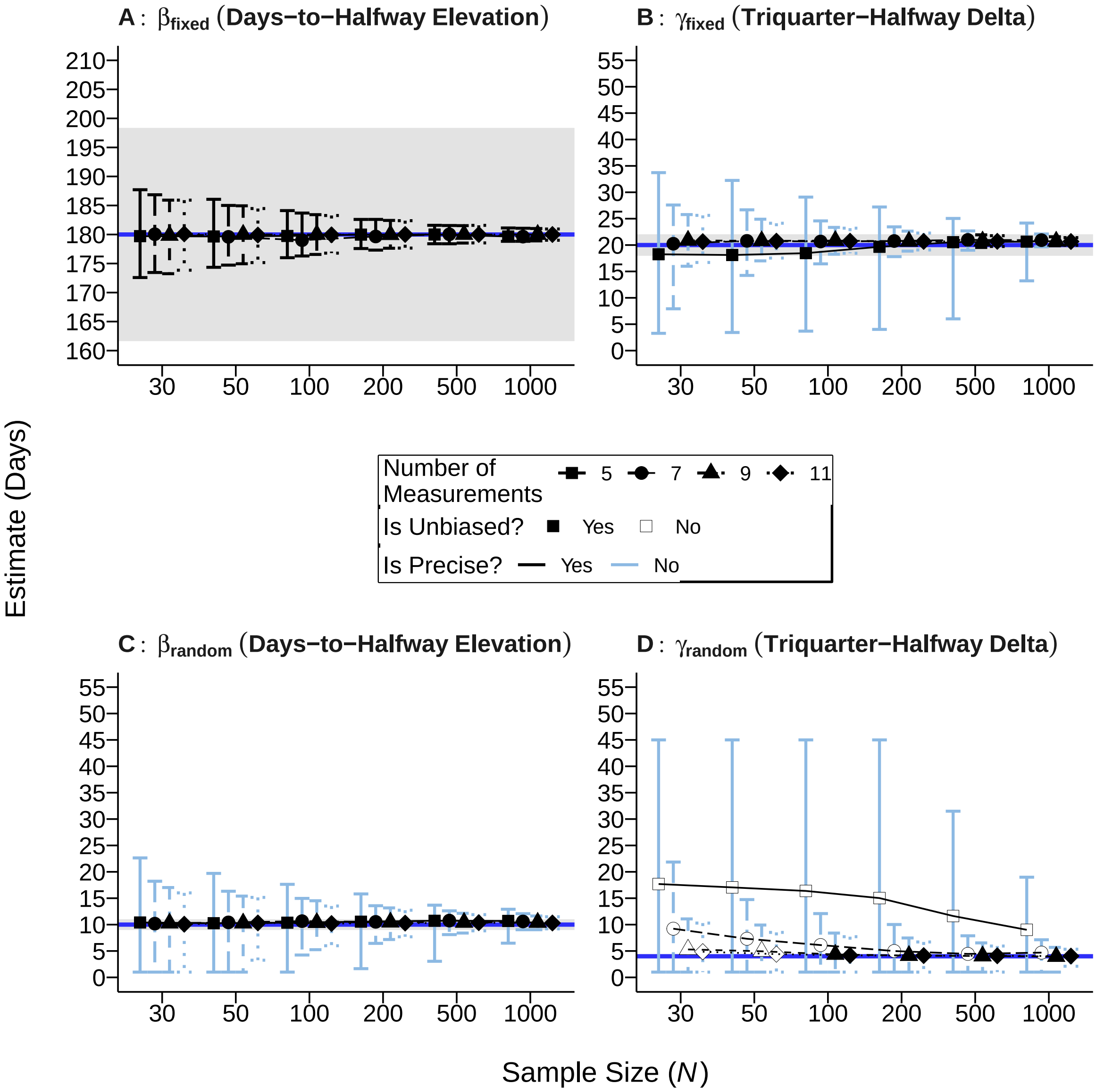
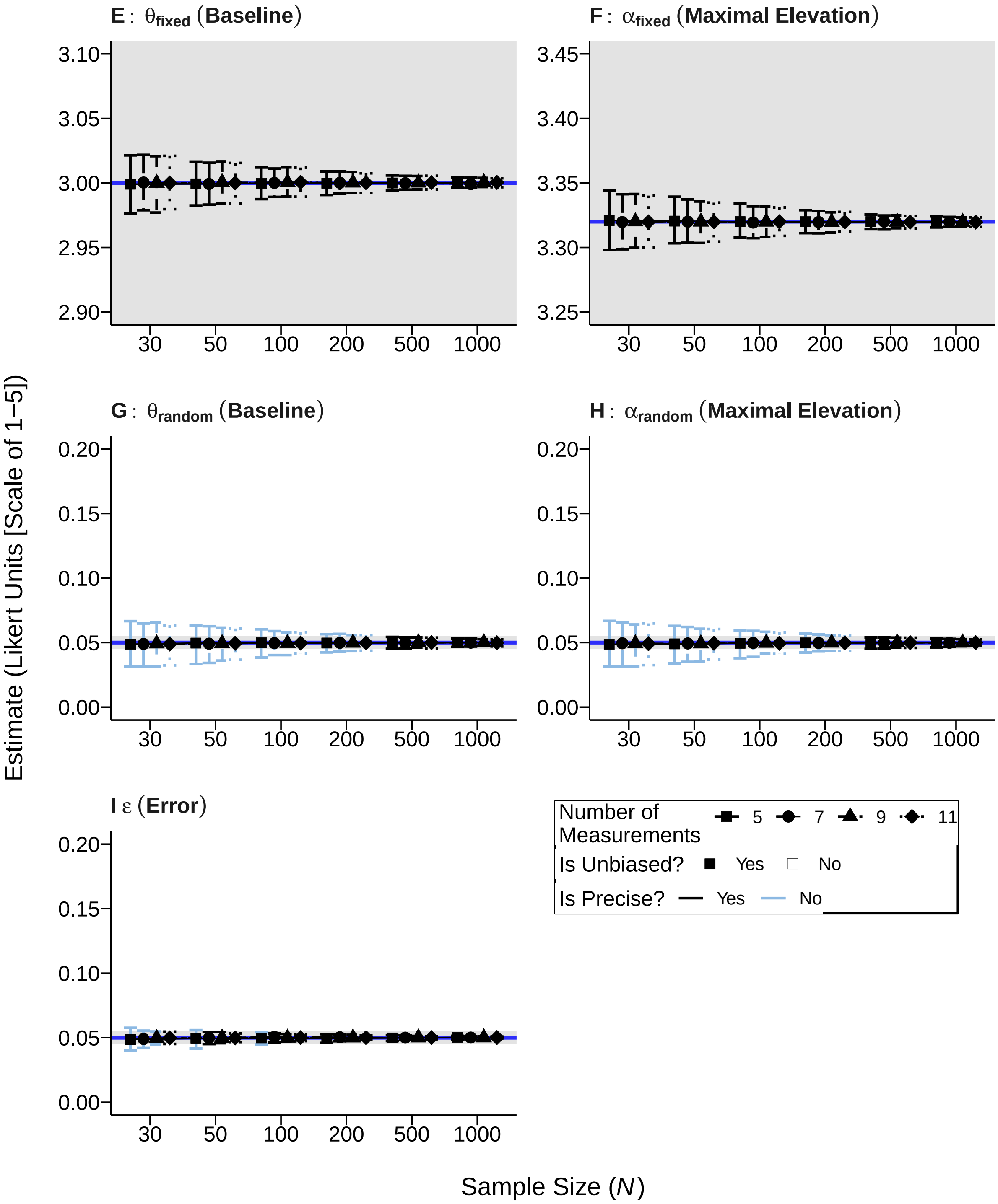


Figure F.9

Bias/Precision Plots for Day- and Likert-Unit Parameters With Time-Structured Data in Experiment 3 (continued)



4583 *Note.* Panels A–B: Bias/precision plots for the fixed- and random-effect days-to-halfway elevation
4584 parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Bias/precision plots for the fixed- and
4585 random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F:
4586 Bias/precision plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and θ_{random}).

Panels G–H: Bias/precision plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table H.3 for specific values estimated for each parameter.

F.3.10 Time-Unstructured Data Characterized by a Fast Response Rate

Figure F.10

Bias/Precision Plots for Day- and Likert-Unit Parameters With Time-Unstructured Data Characterized by a Fast Response Rate in Experiment 3

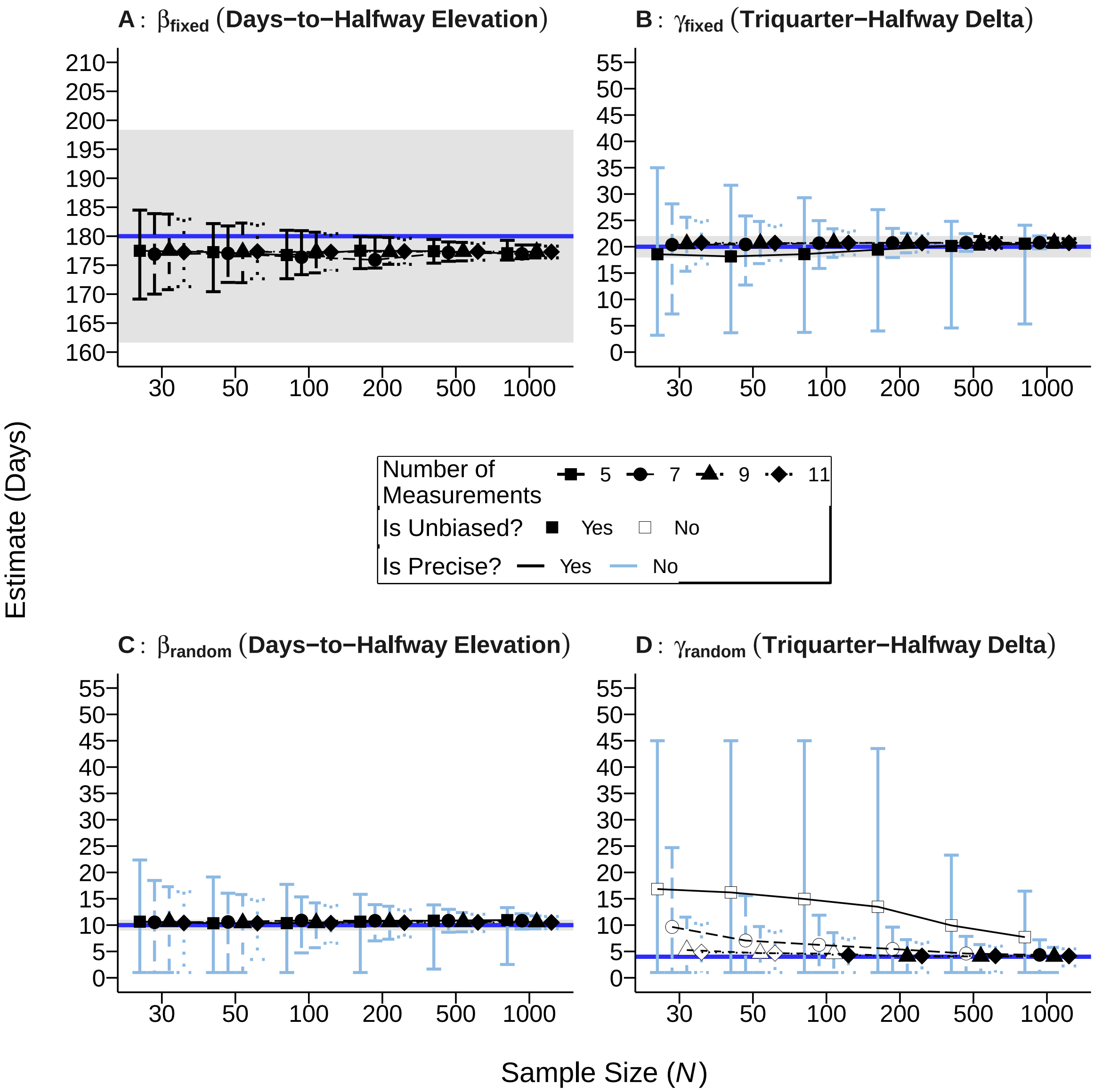
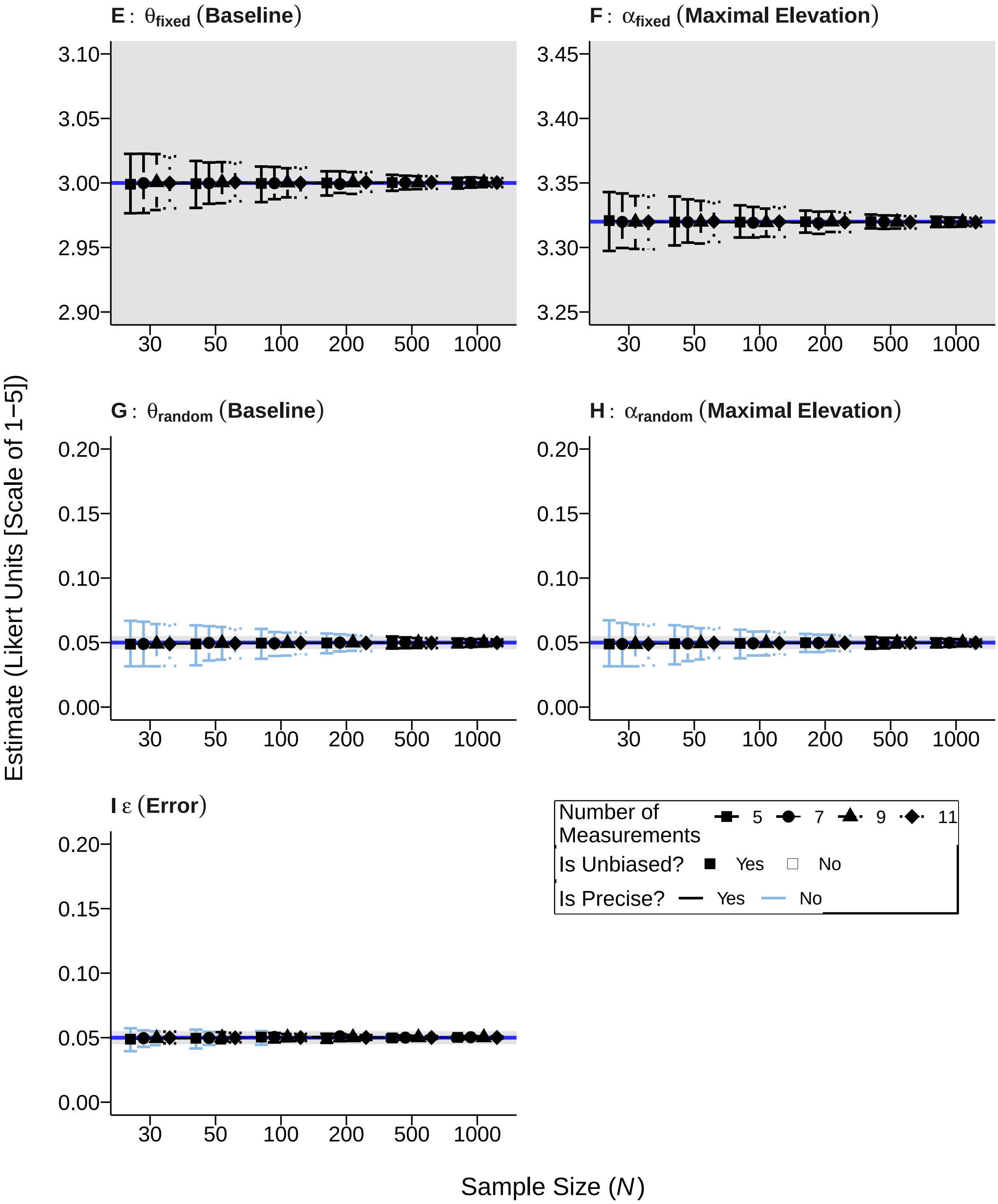


Figure F.10

Bias/Precision Plots for Day- and Likert-Unit Parameters With Time-Unstructured Data Characterized by a Fast Response Rate in Experiment 3 (continued)



Note. Panels A–B: Bias/precision plots for the fixed- and random-effect days-to-halfway elevation parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Bias/precision plots for the fixed- and random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F: Bias/precision plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and θ_{random}).

Panels G–H: Bias/precision plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table H.3 for specific values estimated for each parameter.

F.3.11 Time-Unstructured Data Characterized by a Slow Response Rate

Figure F.11

Bias/Precision Plots for Day- and Likert-Unit Parameters With Time-Unstructured Data Characterized by a Slow Response Rate in Experiment 3

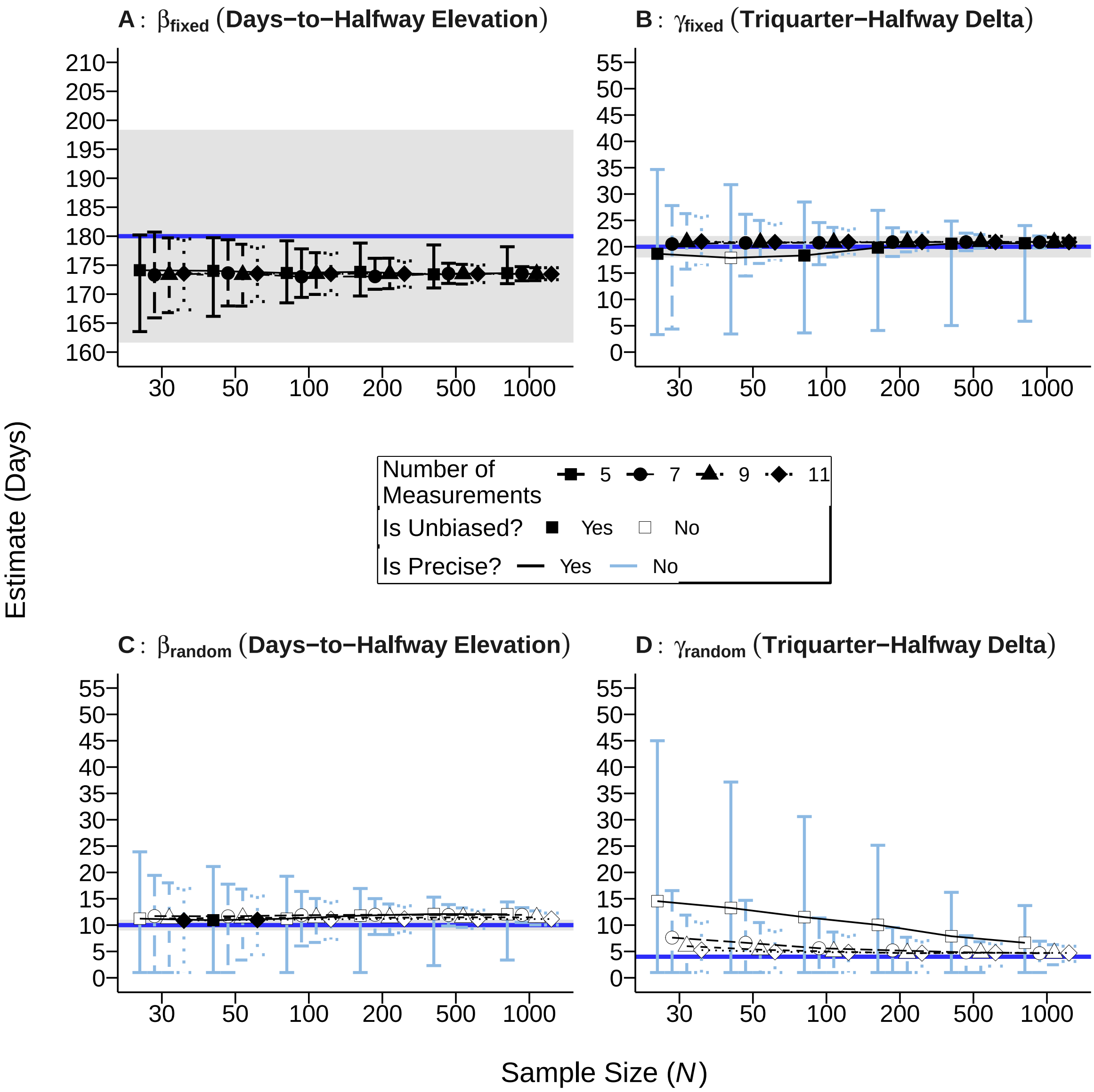
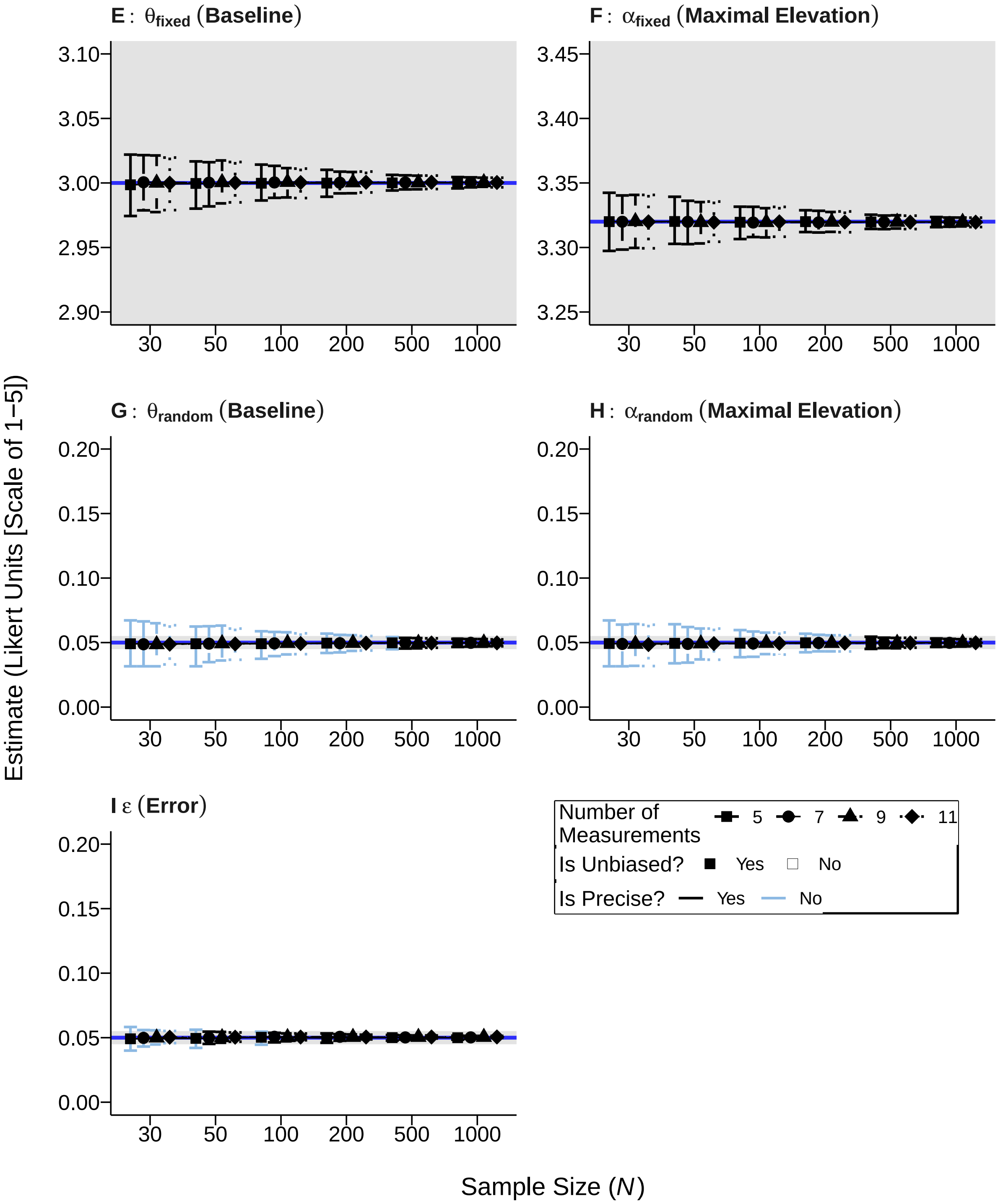


Figure F.11

Bias/Precision Plots for Day- and Likert-Unit Parameters With Time-Unstructured Data Characterized by a Slow Response Rate in Experiment 3 (continued)



4615 *Note.* Panels A–B: Bias/precision plots for the fixed- and random-effect days-to-halfway elevation
4616 parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Bias/precision plots for the fixed- and
4617 random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F:
4618 Bias/precision plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and θ_{random}).

Panels G–H: Bias/precision plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table H.3 for specific values estimated for each parameter.

F.3.12 Time-Unstructured Data Characterized by a Slow Response Rate and Modelled with Definition Variables

Figure F.12

Bias/Precision Plots for Day- and Likert-Unit Parameters When Using Definition Variables To Model Time-Unstructured Data Characterized by a Slow Response Rate

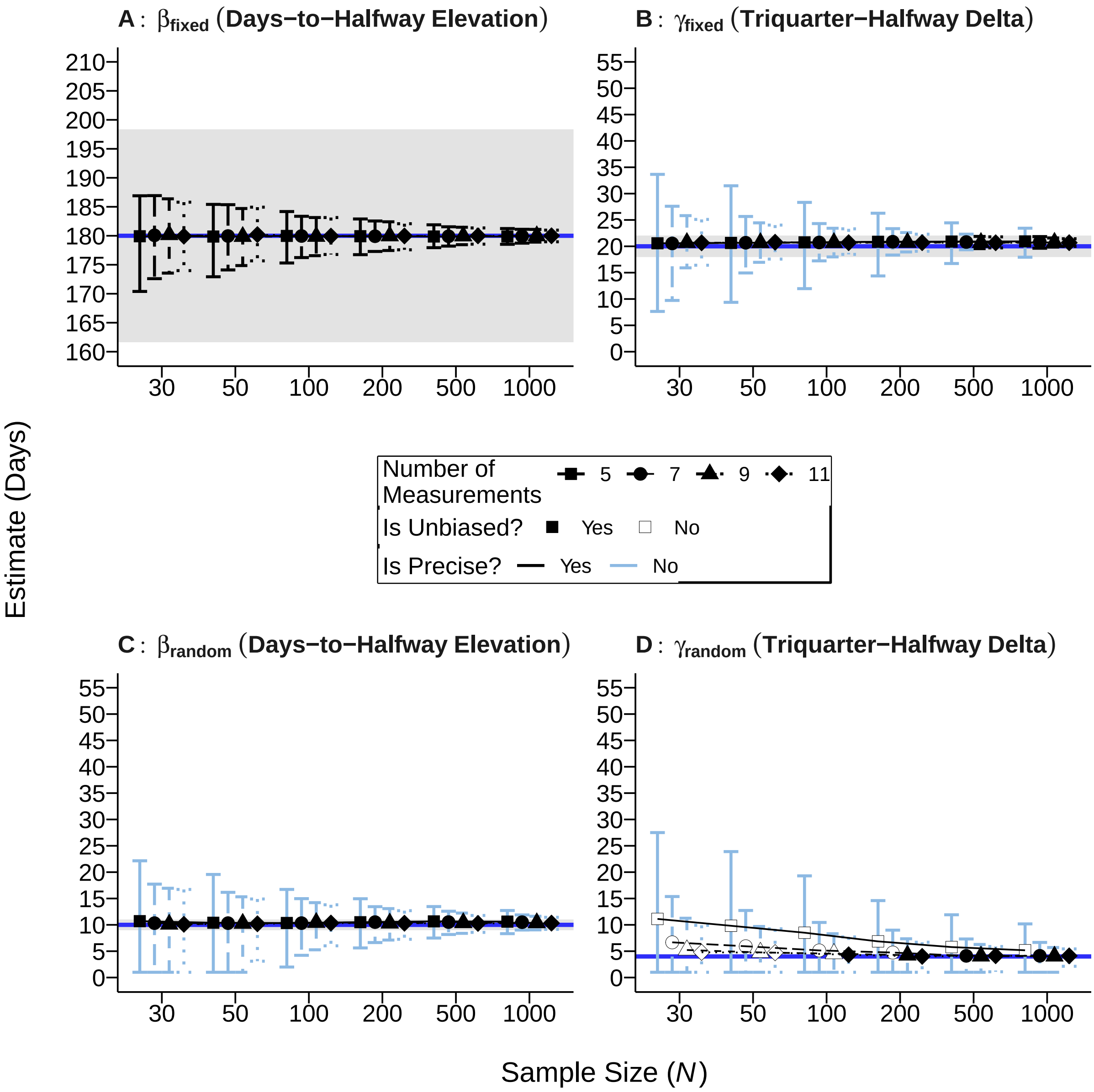
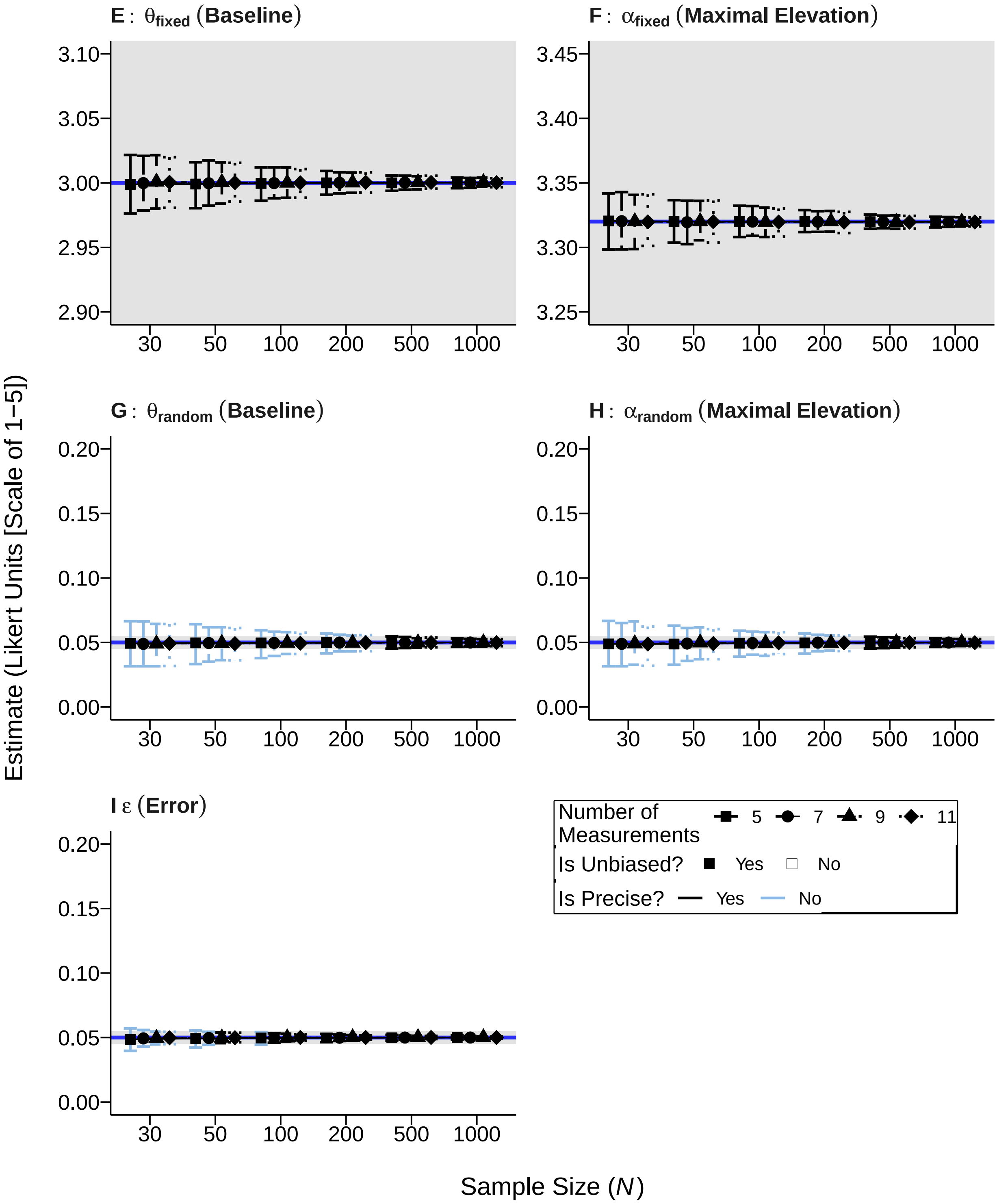


Figure F.12

Bias/Precision Plots for Day- and Likert-Unit Parameters When Using Definition Variables To Model Time-Unstructured Data Characterized by a Slow Response Rate (continued)



4632 *Note.* Panels A–B: Bias/precision plots for the fixed- and random-effect days-to-halfway elevation
4633 parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Bias/precision plots for the fixed- and
4634 random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F:
4635 Bias/precision plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and θ_{random}).

Panels G–H: Bias/precision plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table H.3 for specific values estimated for each parameter.

Appendix G: Convergence Success Rates

G.1 Experiment 1

Table G.1
Convergence Success Rates in Experiment 1

		Days to Halfway Elevation		
Measurement Spacing	Number of Measurements	80	180	280
Equal	5	1.00	0.98	0.95
	7	1.00	1.00	0.99
	9	1.00	1.00	1.00
	11	1.00	1.00	1.00
Time-interval increasing	5	1.00	1.00	1.00
	7	1.00	1.00	1.00
	9	1.00	1.00	1.00
	11	1.00	1.00	1.00
Time-interval decreasing	5	1.00	0.96	0.82
	7	1.00	0.99	0.98
	9	1.00	1.00	1.00
	11	1.00	1.00	1.00

	5	1.00	0.96	0.86
Middle-and-extreme	7	1.00	1.00	1.00
	9	1.00	1.00	1.00
	11	1.00	1.00	1.00

Note. Cells shaded in gray indicate conditions where less than 90% of models converged.

G.2 Experiment 2

Table G.2
Convergence Success Rates in Experiment 2

		Sample Size (*N*)					
Measurement Spacing	Number of Measurements	30	50	100	200	500	1000
Equal	5	1.00	1.00	0.99	0.98	0.95	0.92
	7	1.00	1.00	1.00	1.00	0.99	0.98
	9	1.00	1.00	1.00	1.00	1.00	1.00
	11	1.00	1.00	1.00	1.00	1.00	1.00
Time-interval increasing	5	1.00	1.00	1.00	1.00	1.00	1.00
	7	1.00	1.00	1.00	1.00	1.00	1.00
	9	1.00	1.00	1.00	1.00	1.00	1.00
	11	1.00	1.00	1.00	1.00	1.00	1.00
Time-interval decreasing	5	1.00	0.99	0.98	0.95	0.93	0.88
	7	1.00	1.00	0.99	0.99	0.98	0.95
	9	1.00	1.00	1.00	1.00	1.00	0.99
	11	1.00	1.00	1.00	1.00	1.00	1.00
Middle-and-extreme	5	1.00	0.99	0.98	0.96	0.90	0.81
	7	1.00	1.00	1.00	1.00	1.00	1.00
	9	1.00	1.00	1.00	1.00	1.00	1.00
	11	1.00	1.00	1.00	1.00	1.00	1.00

Note. Cells shaded in gray indicate conditions where less than 90% of models converged.

G.3 Experiment 3

Table G.3
Convergence Success Rates in Experiment 3

Time Structuredness	Number of Measurements	Sample Size (*N*) 30	50	100	200	500	1000
Time structured	5	1.00	0.99	0.99	0.98	0.96	0.90
	7	1.00	1.00	1.00	1.00	0.99	0.98
	9	1.00	1.00	1.00	1.00	1.00	1.00
	11	1.00	1.00	1.00	1.00	1.00	1.00
Time unstructured (fast response)	5	1.00	1.00	0.98	0.99	0.96	0.90
	7	1.00	1.00	1.00	0.99	0.98	0.99
	9	1.00	1.00	1.00	1.00	1.00	1.00
	11	1.00	1.00	1.00	1.00	1.00	1.00
Time unstructured (slow response)	5	1.00	1.00	0.99	1.00	0.95	0.92
	7	1.00	1.00	1.00	0.99	0.99	0.98
	9	1.00	1.00	1.00	1.00	1.00	1.00
	11	1.00	1.00	1.00	1.00	1.00	1.00
Time unstructured (slow response) with definition variables	5	1.00	1.00	1.00	1.00	0.99	0.98
	7	1.00	1.00	1.00	1.00	1.00	0.99
	9	1.00	1.00	1.00	1.00	1.00	1.00
	11	1.00	1.00	1.00	1.00	1.00	1.00

Note. Cells shaded in gray indicate conditions where less than 90% of models converged.

Table G.4
Convergence Success in Experiment 3 With Definition Variables

Time Structuredness	Number of Measurements	Sample size (*N*) 30	50	100	200	500	1000

Time unstructured	5	1.00	1.00	1.00	1.00	0.99	0.98
(slow response)	7	1.00	1.00	1.00	1.00	1.00	0.99
with definition	9	1.00	1.00	1.00	1.00	1.00	1.00
variables	11	1.00	1.00	1.00	1.00	1.00	1.00

Note. Cells shaded in gray indicate conditions where less than 90% of models converged.

Appendix H: Parameter Estimate Tables

H.1 Experiment 1

Table H.1

Parameter Values Estimated for Day- and Likert-Unit Parameters in Experiment 1

		β_{fixed} (Days to halfway elevation)			β_{random} (Days to halfway elevation) Pop value = 10.00			γ_{fixed} (Triquarter-halfway delta) Pop value = 20.00			γ_{random} (Triquarter-halfway delta) Pop value = 4.00		
Measurement Spacing	Number of Measurements	80	180	280	80	180	280	80	180	280	80	180	280
Equal spacing	5	79.73	179.78	279.81$^{\square}$	10.14	10.40	10.08	19.37	19.49	19.71	7.41$^{\square}$	14.53$^{\square}$	8.11$^{\square}$
	7	80.21	178.99	279.55$^{\square}$	10.16	10.55	10.13	20.67	20.83	20.60	4.37	5.14$^{\square}$	4.41$^{\square}$
	9	80.00	179.94	279.99$^{\square}$	10.29	10.37	10.34	20.77	20.76	20.67	4.24	4.14	4.30
	11	80.03	180.01	279.88$^{\square}$	10.27	10.29	10.32	20.64	20.70	20.64	4.13	4.08	4.18
Time-interval increasing	5	79.88	180.10	274.37$^{\square}$	10.32	9.73	13.04$^{\square}$	20.71	20.39	18.32	4.57$^{\square}$	4.99$^{\square}$	6.20$^{\square}$
	7	80.19	179.82	279.86$^{\square}$	10.42	10.47	10.14	20.66	20.79	19.78	4.29	4.87$^{\square}$	7.03$^{\square}$
	9	79.59	179.06	279.70$^{\square}$	10.07	10.22	10.20	20.33	20.66	20.72	4.17	4.25	4.32
	11	79.89	179.84	279.62$^{\square}$	10.38	10.30	10.47	20.78	20.75	20.68	4.23	4.18	4.13
Time-interval decreasing	5	70.67	179.92	279.63$^{\square}$	15.28$^{\square}$	9.80	10.22	16.63	20.07	20.55	5.48$^{\square}$	5.17$^{\square}$	4.59$^{\square}$
	7	78.23	178.22	279.84$^{\square}$	10.08	10.46	10.39	19.38	20.59	20.69	6.80$^{\square}$	5.09$^{\square}$	4.24
	9	79.95	179.34	278.98$^{\square}$	10.03	10.20	10.05	20.42	20.54	20.28	4.37	4.32	4.19
	11	79.42	179.70	279.52$^{\square}$	10.38	10.13	10.06	20.75	20.45	20.31	4.17	4.16	4.17
Middle-and-extreme spacing	5	71.95	179.61	287.73$^{\square}$	16.78$^{\square}$	10.26	16.74$^{\square}$	15.59	20.61	17.09	6.54$^{\square}$	4.24	8.61$^{\square}$
	7	80.45	180.00	279.15$^{\square}$	13.93$^{\square}$	10.25	13.69$^{\square}$	20.71	20.58	20.61	5.21$^{\square}$	4.16	4.98$^{\square}$
	9	80.28	180.05	279.63$^{\square}$	10.42	10.24	10.24	20.91	20.65	20.85	4.74$^{\square}$	4.26	4.72$^{\square}$
	11	80.19	179.96	279.86$^{\square}$	10.27	10.28	10.15	20.71	20.70	20.71	4.14	4.08	4.16

Table H.1

Parameter Values Estimated for Day- and Likert-Unit Parameters in Experiment 1 (continued)

		θ_{fixed} (Baseline) Pop value = 3.00			θ_{random} (Baseline) Pop value = 0.05			α_{fixed} (Maximal elevation) Pop value = 3.32			α_{random} (Maximal elevation) Pop value = 0.05			ϵ(error) Pop value = 0.03		
Measurement Spacing	Number of Measurements	80	180	280	80	180	280	80	180	280	80	180	280	80	180	280
Equal spacing	5	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval increasing	5	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.33	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval decreasing	5	2.99	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Middle-and-extreme spacing	5	2.99	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.33	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05

Note. Cells shaded in light blue indicate cells where estimation is imprecise (i.e., lower and/or upper whisker lengths exceeding 10% of the parameter's population value. Empty superscript squares ($^{\square}$) indicate biased estimates (i.e., bias exceeding 10% of parameter's population value). Importantly, bias and precision cutoff values for the days-to-halfway elevation parameter (β_{fixed}) are based on a value of 180.00.

H.2 Experiment 2

Table H.2

Parameter Values Estimated in Experiment 2

		β_{fixed} (Days to halfway elevation) Pop value = 180.00						β_{random} (Days to halfway elevation) Pop value = 10.00					
Measurement Spacing	Number of Measurements	30	50	100	200	500	1000	30	50	100	200	500	1000
Equal spacing	5	179.71	179.82	179.53	180.00	179.99	179.64	10.40	10.36	10.04	10.51	10.65	10.74
	7	180.05	179.65	179.53	179.75	179.76	179.99	10.18	10.59	10.49	10.54	10.60	10.58
	9	179.84	180.07	179.94	180.00	180.02	180.03	10.28	10.20	10.30	10.40	10.39	10.36
	11	180.11	180.11	180.01	180.03	179.98	179.98	10.08	10.04	10.28	10.29	10.38	10.29
Time-interval increasing	5	181.81	181.16	181.14	180.27	179.78	179.57	11.24□	10.24	9.93	9.59	9.91	10.22
	7	179.99	179.96	179.73	179.77	179.79	179.83	10.26	10.43	10.50	10.43	10.47	10.47
	9	179.33	179.18	178.99	179.07	179.11	179.13	10.15	10.10	10.17	10.18	10.21	10.29
	11	179.81	179.79	179.86	179.88	179.81	179.82	9.99	10.19	10.32	10.27	10.30	10.30
Time-interval decreasing	5	177.01	178.48	179.13	179.23	179.86	180.37	10.95	11.38□	9.97	9.55	10.36	10.11
	7	178.98	179.68	179.12	179.53	180.07	179.75	10.07	10.31	10.48	10.37	10.46	10.51
	9	179.65	179.01	178.46	179.47	179.64	179.75	10.11	10.16	10.20	10.17	10.28	10.26
	11	179.48	179.68	179.70	179.65	179.64	179.68	9.85	9.98	10.03	10.12	10.13	10.11
Middle-and-extreme spacing	5	177.99	179.65	179.15	179.83	179.61	178.74	10.30	10.24	10.40	10.24	10.28	10.26
	7	179.96	179.82	179.97	179.98	180.02	179.98	10.25	10.20	10.32	10.26	10.29	10.27
	9	179.88	180.07	179.89	179.98	179.98	179.99	10.12	10.16	10.24	10.30	10.24	10.29
	11	180.02	179.96	180.01	179.98	180.01	179.99	10.08	10.35	10.15	10.35	10.30	10.28

Table H.2

Parameter Values Estimated for Day- and Likert-Unit Parameters in Experiment 2 (continued)

		γ_{fixed} (Triquarter-halfway delta) Pop value = 20.00						γ_{random} (Triquarter-halfway delta) Pop value = 4.00					
Measurement Spacing	Number of Measurements	30	50	100	200	500	1000	30	50	100	200	500	1000
Equal spacing	5	18.25	18.11	18.27	19.59	20.27	20.60	17.69□	16.95□	16.41□	15.19□	12.19□	8.51□
	7	20.25	20.53	20.66	20.75	20.81	20.74	9.22□	7.70□	5.77□	4.89□	4.98□	4.34
	9	20.88	20.72	20.73	20.76	20.75	20.73	5.30□	4.99□	4.44□	4.27	4.03	4.00
	11	20.65	20.66	20.73	20.70	20.69	20.71	4.86□	4.49□	4.20	4.10	4.02	4.07
Time-interval increasing	5	18.81	19.11	19.56	20.25	20.80	20.92	6.18□	5.88□	5.25□	4.94□	4.68□	4.42□
	7	20.74	20.74	20.94	20.83	20.83	20.82	7.38□	6.31□	5.45□	5.06□	4.66□	4.45□
	9	20.72	20.65	20.69	20.65	20.63	20.65	5.15□	4.83□	4.44□	4.26	4.16	4.23
	11	20.80	20.69	20.84	20.76	20.78	20.76	4.84□	4.43□	4.25	4.26	4.17	4.14
Time-interval decreasing	5	19.21	18.50	19.21	19.90	20.50	20.79	7.17□	6.01□	5.18□	5.12□	4.91□	4.66□
	7	20.36	20.49	20.57	20.69	21.03	20.76	6.98□	6.18□	5.43□	5.20□	4.67□	4.68□
	9	20.69	20.60	20.55	20.62	20.70	20.63	5.48□	5.12□	4.72□	4.52□	4.72□	4.83□
	11	20.49	20.53	20.38	20.41	20.47	20.41	4.66□	4.57□	4.34	4.20	4.18	4.17
Middle-and-extreme spacing	5	20.80	20.69	20.65	20.67	20.64	20.59	5.21□	4.68□	4.43□	4.18	4.15	4.11
	7	20.76	20.55	20.70	20.63	20.60	20.63	5.07□	4.60□	4.39	4.23	4.19	4.15
	9	20.68	20.71	20.67	20.63	20.58	20.63	4.99□	4.67□	4.49□	4.17	4.13	4.15
	11	20.64	20.74	20.67	20.70	20.66	20.68	4.57□	4.47□	4.22	4.19	4.09	4.07

Table H.2

Parameter Values Estimated for Day- and Likert-Unit Parameters in Experiment 2 (continued)

		θ_{fixed} (Baseline) Pop value = 3.00						θ_{random} (Baseline) Pop value = 0.05					
Measurement Spacing	Number of Measurements	30	50	100	200	500	1000	30	50	100	200	500	1000
Equal spacing	5	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval increasing	5	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval decreasing	5	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
Middle-and-extreme spacing	5	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05

Table H.2

Parameter Values Estimated for Day- and Likert-Unit Parameters in Experiment 2 (continued)

		α_{fixed} (Maximal elevation) Pop value = 3.32						α_{random} (Maximal elevation) Pop value = 0.05					
Measurement Spacing	Number of Measurements	30	50	100	200	500	1000	30	50	100	200	500	1000
Equal spacing	5	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval increasing	5	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval decreasing	5	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Middle-and-extreme spacing	5	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05

Table H.2

Parameter Values Estimated for Day- and Likert-Unit Parameters in Experiment 2 (continued)

		ϵ(error)					
		Pop value = 0.03					
Measurement Spacing	Number of Measurements	30	50	100	200	500	1000
Equal spacing	5	0.05	0.05	0.05	0.05	0.05	0.05
	7	0.05	0.05	0.05	0.05	0.05	0.05
	9	0.05	0.05	0.05	0.05	0.05	0.05
	11	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval increasing	5	0.05	0.05	0.05	0.05	0.05	0.05
	7	0.05	0.05	0.05	0.05	0.05	0.05
	9	0.05	0.05	0.05	0.05	0.05	0.05
	11	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval decreasing	5	0.05	0.05	0.05	0.05	0.05	0.05
	7	0.05	0.05	0.05	0.05	0.05	0.05
	9	0.05	0.05	0.05	0.05	0.05	0.05
	11	0.05	0.05	0.05	0.05	0.05	0.05
Middle-and-extreme spacing	5	0.05	0.05	0.05	0.05	0.05	0.05
	7	0.05	0.05	0.05	0.05	0.05	0.05
	9	0.05	0.05	0.05	0.05	0.05	0.05
	11	0.05	0.05	0.05	0.05	0.05	0.05

Note. Cells shaded in light blue indicate cells where estimation is imprecise (i.e., lower and/or upper whisker lengths exceeding 10% of the parameter's population value. Empty superscript squares ($^{\square}$) indicate biased estimates (i.e., bias exceeding 10% of parameter's population value).

H.3 Experiment 3

Table H.3

Parameter Values Estimated in Experiment 3

		β_{fixed} (Days to halfway elevation) Pop value = 180.00						β_{random} (Days to halfway elevation) Pop value = 10.00					
Time Structuredness	Number of Measurements	30	50	100	200	500	1000	30	50	100	200	500	1000
Time structured	5	179.71	179.67	179.75	179.98	180.00	179.66	10.40	10.27	10.37	10.56	10.73	10.69
	7	180.05	179.59	179.02	179.66	180.03	179.63	10.18	10.42	10.65	10.52	10.76	10.60
	9	179.84	180.01	180.01	179.97	180.01	180.00	10.28	10.28	10.37	10.46	10.42	10.41
	11	180.11	179.91	179.94	180.00	180.00	180.00	10.08	10.32	10.21	10.29	10.36	10.31
Time unstructured (fast response)	5	177.48	177.24	176.74	177.50	177.42	177.06	10.65	10.36	10.38	10.65	10.85	10.96
	7	176.89	177.03	176.37	175.92	177.20	176.95	10.53	10.60	10.88	10.83	10.84	10.84
	9	177.54	177.28	177.27	177.31	177.34	177.33	10.66	10.43	10.44	10.61	10.65	10.59
	11	177.25	177.35	177.27	177.37	177.35	177.30	10.41	10.37	10.37	10.45	10.52	10.51
Time unstructured (slow response)	5	174.13	174.02	173.65	173.85	173.41	173.63	11.23□	10.93	11.22□	11.80□	12.10□	12.07□
	7	173.31	173.63	173.01	173.06	173.55	173.55	11.71□	11.67□	11.88□	11.97□	11.91□	11.94□
	9	173.37	173.37	173.54	173.52	173.50	173.49	11.26□	11.38□	11.42□	11.40□	11.47□	11.46□
	11	173.58	173.56	173.50	173.51	173.49	173.47	10.87	10.98	11.12□	11.18□	11.14□	11.16□
Time unstructured (slow response) with definition variables	5	179.92	179.87	179.97	179.92	179.87	179.88	10.70	10.40	10.35	10.50	10.66	10.61
	7	180.07	179.96	179.96	179.92	179.91	179.94	10.32	10.32	10.33	10.52	10.53	10.50
	9	180.17	179.86	179.88	179.97	179.95	179.98	10.12	10.26	10.43	10.32	10.40	10.38
	11	179.93	180.20	179.94	179.97	179.99	179.99	10.11	10.20	10.34	10.31	10.27	10.32

Table H.3

Parameter Values Estimated for Day- and Likert-Unit Parameters in Experiment 3 (continued)

		γ_{fixed} (Triquarter-halfway delta) Pop value = 20.00						γ_{random} (Triquarter-halfway delta) Pop value = 4.00					
Time Structuredness	Number of Measurements	30	50	100	200	500	1000	30	50	100	200	500	1000
Time structured	5	18.25	18.11	18.46	19.67	20.55	20.65	17.69□	17.05□	16.38□	15.03□	11.63□	9.02□
	7	20.25	20.79	20.67	20.77	20.98	20.93	9.22□	7.32□	6.12□	4.99□	4.45□	4.69□
	9	20.88	20.79	20.84	20.69	20.74	20.71	5.30□	4.95□	4.34	4.13	4.05	3.96
	11	20.65	20.74	20.73	20.69	20.71	20.67	4.86□	4.41□	4.17	4.13	4.09	4.03
Time unstructured (fast response)	5	18.57	18.16	18.59	19.45	20.15	20.58	16.85□	16.21□	14.96□	13.48□	9.94□	7.72□
	7	20.39	20.44	20.67	20.73	20.77	20.77	9.65□	7.07□	6.25□	5.47□	4.61□	4.34
	9	20.54	20.66	20.75	20.71	20.72	20.74	5.27□	4.68□	4.59□	4.08	4.06	4.05
	11	20.77	20.70	20.72	20.70	20.71	20.73	4.85□	4.68□	4.29	4.14	4.16	4.14
Time unstructured (slow response)	5	18.66	17.88	18.34	19.83	20.57	20.67	14.54□	13.26□	11.51□	10.05□	7.89□	6.65□
	7	20.51	20.73	20.75	20.89	20.89	20.86	7.62□	6.65□	5.61□	5.21□	4.83□	4.67□
	9	20.91	20.82	20.82	20.89	20.94	20.89	6.00□	5.32□	4.97□	4.67□	4.74□	4.70□
	11	20.98	20.85	20.90	20.92	20.90	20.90	5.26□	4.92□	4.83□	4.69□	4.75□	4.71□
Time unstructured (slow response) with definition variables	5	20.58	20.64	20.76	20.86	20.90	20.94	11.12□	9.82□	8.51□	6.86□	5.78□	5.17□
	7	20.55	20.68	20.73	20.87	20.81	20.78	6.68□	5.93□	5.14□	4.74□	4.11	4.12
	9	20.69	20.68	20.69	20.74	20.70	20.73	5.22□	4.77□	4.53□	4.24	4.05	4.05
	11	20.66	20.77	20.69	20.69	20.67	20.69	4.79□	4.72□	4.32	4.01	4.14	4.11

Table H.3

Parameter Values Estimated for Day- and Likert-Unit Parameters in Experiment 3 (continued)

		θ_{fixed} (Baseline) Pop value = 3.00						θ_{random} (Baseline) Pop value = 0.05					
Time Structuredness	Number of Measurements	30	50	100	200	500	1000	30	50	100	200	500	1000
Time structured	5	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
Time unstructured (fast response)	5	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
Time unstructured (slow response)	5	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
Time unstructured (slow response) with definition variables	5	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05

Table H.3

Parameter Values Estimated for Day- and Likert-Unit Parameters in Experiment 3 (continued)

		α_{fixed} (Maximal elevation) Pop value = 3.32						α_{random} (Maximal elevation) Pop value = 0.05					
Time Structuredness	Number of Measurements	30	50	100	200	500	1000	30	50	100	200	500	1000
Time structured	5	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Time unstructured (fast response)	5	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Time unstructured (slow response)	5	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Time unstructured (slow response) with definition variables	5	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05

Table H.3

Parameter Values Estimated for Day- and Likert-Unit Parameters in Experiment 3 (continued)

		ϵ(error) Pop value = 0.03					
Time Structuredness	Number of Measurements	30	50	100	200	500	1000
Time structured	5	0.05	0.05	0.05	0.05	0.05	0.05
	7	0.05	0.05	0.05	0.05	0.05	0.05
	9	0.05	0.05	0.05	0.05	0.05	0.05
	11	0.05	0.05	0.05	0.05	0.05	0.05
Time unstructured (fast response)	5	0.05	0.05	0.05	0.05	0.05	0.05
	7	0.05	0.05	0.05	0.05	0.05	0.05
	9	0.05	0.05	0.05	0.05	0.05	0.05
	11	0.05	0.05	0.05	0.05	0.05	0.05
Time unstructured (slow response)	5	0.05	0.05	0.05	0.05	0.05	0.05
	7	0.05	0.05	0.05	0.05	0.05	0.05
	9	0.05	0.05	0.05	0.05	0.05	0.05
	11	0.05	0.05	0.05	0.05	0.05	0.05
Time unstructured (slow response) with definition variables	5	0.05	0.05	0.05	0.05	0.05	0.05
	7	0.05	0.05	0.05	0.05	0.05	0.05
	9	0.05	0.05	0.05	0.05	0.05	0.05
	11	0.05	0.05	0.05	0.05	0.05	0.05

Note. Cells shaded in light blue indicate cells where estimation is imprecise (i.e., lower and/or upper whisker lengths exceeding 10% of the parameter's population value. Empty superscript squares ($^{\square}$) indicate biased estimates (i.e., bias exceeding 10% of parameter's population value).

Appendix I: OpenMx Code for Structured Latent Growth Curve Model With Definition Variables

The code that I used to model logistic pattern of change using definition variables (see definition variables) is shown in Code Block I.1. Note that, the code is largely excerpted from the `run_exp_simulations()` and `create_definition_model()` functions from the `nonlinSims` package, and so readers interested in obtaining more information should consult the source code of this package. One important point to mention is that the model specified in Code Block I.1 can accurately model time-unstructured data because it uses definition variables.

Code Block I.1

OpenMx Code for Structured Latent Growth Curve Model With Definition Variables

```
#Now convert data to wide format (needed for OpenMx)
data_wide <- data[ , c(1:3, 5)] %>%
    pivot_wider(names_from = measurement_day, values_from = c(obs_score,
    actual_measurement_day))

#Definition variable (data. prefix tells OpenMx to use recorded time of observation
for each person's data)
obs_score_days <- paste('data.', extract_obs_score_days(data = data_wide), sep = '')

#Remove . from column names so that OpenMx does not run into error (this occurs
because, with some spacing schedules, measurement days are not integer values.)
names(data_wide) <- str_replace(string = names(data_wide), pattern = '\\.', replacement
= '_')

#Latent variable names (theta = baseline, alpha = maximal elevation, beta =
days-to-halfway elevation, gamma = triquarter-halfway elevation)
latent_vars <- c('theta', 'alpha', 'beta', 'gamma')

def_growth_curve_model <- mxModel(
  model = model_name,
  type = 'RAM', independent = T,
  mxData(observed = data_wide, type = 'raw'),

  manifestVars = manifest_vars,
  latentVars = latent_vars,

  #Residual variances; by using one label, they are assumed to all be equal
  (homogeneity of variance). That is, there is no complex error structure.
  mxPath(from = manifest_vars,
         arrows=2, free=TRUE,  labels='epsilon', values = 1, lbound = 0),

  #Latent variable covariances and variances (note that only the variances are
  estimated. )
  mxPath(from = latent_vars,
         connect='unique.pairs', arrows=2,
         free = c(TRUE,FALSE, FALSE, FALSE,
```

```
                  TRUE, FALSE, FALSE,
                  TRUE, FALSE,
                  TRUE),
         values=c(1, NA, NA, NA,
                  1, NA, NA,
                  1, NA,
                  1),
         labels=c('theta_rand', 'NA(cov_theta_alpha)', 'NA(cov_theta_beta)',
                  'NA(cov_theta_gamma)',
                  'alpha_rand','NA(cov_alpha_beta)', 'NA(cov_alpha_gamma)',
                  'beta_rand', 'NA(cov_beta_gamma)',
                  'gamma_rand'),
         lbound = c(1e-3, NA, NA, NA,
                    1e-3, NA, NA,
                    1, NA,
                    1),
         ubound = c(2, NA, NA, NA,
                    2, NA, NA,
                    90^2, NA,
                    45^2)),

  # Latent variable means (linear parameters). Note that the parameters of beta and
  gamma do not have estimated means because they are nonlinear parameters (i.e., the
  logistic function's first-order partial derivative with respect to each of those two
  parameters contains those two parameters)
  mxPath(from = 'one', to = c('theta', 'alpha'), free = c(TRUE, TRUE), arrows = 1,
         labels = c('theta_fixed', 'alpha_fixed'), lbound = 0, ubound = 7,
         values = c(1, 1)),

  #Functional constraints (needed to estimate mean values of fixed-effect parameters)
  mxMatrix(type = 'Full', nrow = length(manifest_vars), ncol = 1, free = TRUE,
           labels = 'theta_fixed', name = 't', values = 1, lbound = 0,  ubound = 7),
  mxMatrix(type = 'Full', nrow = length(manifest_vars), ncol = 1, free = TRUE,
           labels = 'alpha_fixed', name = 'a', values = 1, lbound = 0,  ubound = 7),
  mxMatrix(type = 'Full', nrow = length(manifest_vars), ncol = 1, free = TRUE,
           labels = 'beta_fixed', name = 'b', values = 1, lbound = 1, ubound = 360),
  mxMatrix(type = 'Full', nrow = length(manifest_vars), ncol = 1, free = TRUE,
           labels = 'gamma_fixed', name = 'g', values = 1, lbound = 1, ubound = 360),

  #Definition variables set for loadings (accounts for time-unstructured data)
  mxMatrix(type = 'Full', nrow = length(obs_score_days), ncol = 1, free = FALSE,
  labels = obs_score_days, name = 'time'),

  #Algebra specifying first-order partial derivatives;
  mxAlgebra(expression = 1 - 1/(1 + exp((b - time)/g)), name="Tl"),
  mxAlgebra(expression = 1/(1 + exp((b - time)/g)), name = 'Al'),
  mxAlgebra(expression = -((a - t) * (exp((b - time)/g) * (1/g))/(1 + exp((b -
  time)/g))^2), name = 'Bl'),
  mxAlgebra(expression =  (a - t) * (exp((b - time)/g) * ((b - time)/g^2))/(1 + exp((b
  -time)/g))^2, name = 'Gl'),

  #Factor loadings; all fixed and, importantly, constrained to change according to
  their partial derivatives (i.e., nonlinear functions)
  mxPath(from = 'theta', to = manifest_vars, arrows=1, free=FALSE,
         labels = sprintf(fmt = 'Tl[%d,1]', 1:length(manifest_vars))),
  mxPath(from = 'alpha', to = manifest_vars, arrows=1, free=FALSE,
         labels = sprintf(fmt = 'Al[%d,1]', 1:length(manifest_vars))),
  mxPath(from='beta', to = manifest_vars, arrows=1,  free=FALSE,
         labels =  sprintf(fmt = 'Bl[%d,1]', 1:length(manifest_vars))),
  mxPath(from='gamma', to = manifest_vars, arrows=1,  free=FALSE,
         labels =  sprintf(fmt = 'Gl[%d,1]', 1:length(manifest_vars))),

  #Fit function used to estimate free parameter values.
  mxFitFunctionML(vector = FALSE)
)
```

```
#Fit model using mxTryHard(). Increases probability of convergence by attempting model convergence by randomly shifting starting values.
model_results <- mxTryHard(def_growth_curve_model)
```

```
library(easypackages)
packages <- c('nonlinSims','tidyverse')
libraries(packages)

theta <- 3
alpha <- 3.32
beta <- 180
gamma <- 20
time <- 0:360

#curve scores
logistic_curve <- theta + (alpha - theta)/(1 + exp((beta-time)/gamma))

logistic_data <- data.frame('score' = logistic_curve,
                            'day' = time)

#day scores
equal_spacing_5 <- data.frame('day' = compute_measurement_schedule(time_period = 360,
num_measurements = 5, smallest_int_length = 30, measurement_spacing =
'equal')$measurement_days)
equal_spacing_7 <- data.frame('day' = compute_measurement_schedule(time_period = 360,
num_measurements = 7, smallest_int_length = 30, measurement_spacing =
'equal')$measurement_days)
equal_spacing_9 <- data.frame('day' = compute_measurement_schedule(time_period = 360,
num_measurements = 9, smallest_int_length = 30, measurement_spacing =
'equal')$measurement_days)
equal_spacing_11 <- data.frame('day' = compute_measurement_schedule(time_period = 360,
num_measurements = 11, smallest_int_length = 30, measurement_spacing =
'equal')$measurement_days)


#provide curve score for each point
equal_spacing_5$score <- logistic_data$score[logistic_data$day %in% equal_spacing_5$day]
equal_spacing_7$score <- logistic_data$score[logistic_data$day %in% equal_spacing_7$day]
equal_spacing_9$score <- logistic_data$score[logistic_data$day %in% equal_spacing_9$day]
equal_spacing_11$score <- logistic_data$score[logistic_data$day %in%
equal_spacing_11$day]

#spacing schedules
equal_spacing_5 <- data.frame('day' = compute_measurement_schedule(time_period = 360,
num_measurements = 5, smallest_int_length = 30, measurement_spacing =
'equal')$measurement_days)
time_inc_5 <- data.frame('day' = compute_measurement_schedule(time_period = 360,
num_measurements = 5, smallest_int_length = 30, measurement_spacing =
'time_inc')$measurement_days)
time_dec_5 <- data.frame('day' = compute_measurement_schedule(time_period = 360,
num_measurements = 5, smallest_int_length = 30, measurement_spacing =
'time_dec')$measurement_days)
mid_ext_5 <- data.frame('day' = compute_measurement_schedule(time_period = 360,
num_measurements = 5, smallest_int_length = 30, measurement_spacing =
'mid_ext')$measurement_days)

time_inc_5$score <- logistic_data$score[logistic_data$day %in% time_inc_5$day]
time_dec_5$score <- logistic_data$score[logistic_data$day %in% time_dec_5$day]
mid_ext_5$score <- logistic_data$score[logistic_data$day %in% mid_ext_5$day]




mid_ext_5 <- ggplot(data = logistic_data, mapping = aes(x = day, y = score)) +
  geom_line(linewidth = 1) +
  geom_point(data = mid_ext_5, size = 5) +
```

```
  scale_x_continuous(name = 'Day', breaks = seq(from = 0, to = 360, by = 60)) +
  scale_y_continuous(name = 'Score', breaks = NULL) +
  theme_classic(base_family = 'Helvetica') +
  theme(axis.text = element_text(color = 'black', size = 16),
        axis.title = element_text(size = 20))

ggsave(filename = 'Figures/mid_ext_5.png', plot = mid_ext_5, width = 8, height = 6)
```