AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychological Methods

ISSN: 1082-989X

https://doi.org/10.1037/met0000749

Timescale Mismatch in Intensive Longitudinal Data: Current Issues and Possible Solutions Based on Dynamic Structural Equation Models

Xiaohui Luo, Yueqin Hu, and Hongyun Liu
Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education (Beijing Normal University), Faculty of Psychology, Beijing Normal University

Abstract

Intensive longitudinal data have been increasingly used to examine dynamic bidirectional relations between variables. However, the problem of timescale mismatch between variables faced by applied researchers remains understudied. Under the dynamic structural equation modeling framework, previous studies used the partial-path model and the average-score model, respectively, to explore the dynamic interaction processes and overall reciprocal effects between variables with mismatched timescales. The present study aimed to evaluate the performance of the existing modeling approaches and the effectiveness of the improved approaches (i.e., the full-path model, the factor model, and the adjusted factor model). Study 1 showed that the full-path model, which considered the cross-lagged effects of all time points of variables with denser timescales, better reflected dynamic interaction processes and time-specific effects between variables than the partial-path model. Study 2-1 found that the estimates of autoregressive and cross-lagged effects between timescale mismatched variables were biased in the average-score model, but accurate in the factor model. Study 2-2 further suggested that when there were regression effects between different time points of variables with denser timescales, the adjusted factor model obtained less bias than the factor model, yet the difference is negligible when the regression effects are small. Study 3 used empirical data with timescale mismatched variables to illustrate the differences of all modeling approaches. This study identified the important problem of timescale mismatch in intensive longitudinal data and its possible solutions, providing methodological guidance and valuable insights for data collection and analysis of variables with mismatched timescales.

Translational Abstract

In intensive longitudinal studies, some variables are measured more frequently, for example, four times a day, while others are measured less often, such as once a day. How can we investigate the dynamic relations between two variables with different measurement frequencies? This study aims to provide an answer. Several existing modeling approaches (namely, the partial-path model and the average-score model) were evaluated, and improved solutions (the full-path model, the factor model, and the adjusted factor model) were also proposed. The full-path model, which fully considers the cross-lagged effects of all time points of variables with denser timescales, thereby offering a more precise depiction of dynamic interactions and effects at specific time points, significantly outperforms partial path models. Additionally, the average-score model exhibits greater bias in estimating autoregressive and cross-lagged effects compared to the factor model, whereas the adjusted factor model can further reduce bias. Finally, these models were demonstrated using empirical data on the relation between stress feelings at different times of the day and daily sleep quality. The results indicated that the improved solutions are better suited to identifying the most influential timings of stress feelings within a day. Methodological guidance for data collection and modeling strategies, as well as Mplus code, were also provided.

Emilio Ferrer served as action editor.

Yueqin Hu https://orcid.org/0000-0002-5533-9277

Hongyun Liu https://orcid.org/0000-0002-3472-9102

This research was supported by the National Natural Science Foundation of China (32471145 and 32171089). This study was not preregistered. All data, code, and materials have been uploaded to the Open Science Framework repository (https://osf.io/fdpsj).

Xiaohui Luo served as lead for data curation, formal analysis, visualization, and writing–original draft. Xiaohui Luo, Yueqin Hu, and Hongyun Liu contributed equally to conceptualization and methodology. Xiaohui Luo and Hongyun Liu contributed equally to software. Yueqin Hu and Hongyun Liu contributed equally to funding acquisition, supervision, and writing–review and editing.

Correspondence concerning this article should be addressed to Yueqin Hu, Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education (Beijing Normal University), Faculty of Psychology, Beijing Normal University, 19, Xin Jie Kou Wai Street, Hai Dian District, Beijing 100875, People's Republic of China, or Hongyun Liu, Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education (Beijing Normal University), Faculty of Psychology, Beijing Normal University, 19, Xin Jie Kou Wai Street, Hai Dian District, Beijing 100875, People's Republic of China. Email: yueqinhu@bnu.edu.cn or hyliu@bnu.edu.cn

Keywords: timescale mismatch, intensive longitudinal data, dynamic structural equation model, dynamic bidirectional relation, reciprocal effects

Supplemental materials: https://doi.org/10.1037/met0000749.supp

In recent years, intensive longitudinal data (ILD) have been increasingly used in social and behavioral sciences such as psychology (Hamaker & Wichers, 2017; Luo et al., 2024; Zhou et al., 2021). It helps researchers understand dynamic processes and interactions between individuals' states in natural settings. In contrast to retrospective surveys and laboratory studies, ILD offers advantages like reduced recall bias and increased ecological validity (Bolger et al., 2003; Shiffman et al., 2008; Trull & Ebner-Priemer, 2013). For data collection, researchers commonly employ daily diaries, ecological momentary assessment (EMA), and experience sampling method (Bolger & Laurenceau, 2013; Bolger et al., 2003; Shiffman et al., 2008). Besides using self-reported questionnaires as subjective measures of individuals' states, the enhanced accessibility of advanced techniques allows for more objective measures using digital devices (e.g., sensors and mobile phones; Schick et al., 2023). Regarding data analysis, a novel modeling approach, dynamic structural equation modeling (DSEM; Asparouhov et al., 2018), has been proposed and has gained widespread use in recent years for analyzing ILD. This modeling framework integrates the strengths of multilevel modeling, structural equation modeling, and time series modeling to more effectively investigate the within-person processes between variables as well as the between-person differences of these dynamic processes.

Despite the effectiveness of recently developed statistical methods for modeling ILD, researchers have identified several methodological issues in ILD analyses (Hamaker & Wichers, 2017). Among these issues, some researchers have concentrated on problems related to the timing of repeated measurements in ILD, such as the problem of time-interval dependency (Kuiper & Ryan, 2018; McNeish & Hamaker, 2020) and temporal misalignment (Luo & Hu, 2024). Time-interval dependency refers to the issue where estimated parameters depend on the specific time intervals chosen for the study. If the time intervals are equal, the estimated parameters can only be interpreted within the context of that specific interval. If the time intervals are unequal, treating them as equal can bias the parameter estimates. Temporal misalignment refers to situations where two variables have the same density of timescales (e.g., both mood and sleep measured at 1-day intervals) but are measured at different times (e.g., mood measured during the day and sleep at night). Both problems (i.e., time-interval dependency and temporal misalignment) consider variables with the same density of timescales. However, another practical problem faced by empirical researchers on ILD—the problem of timescale mismatch (which is described in detail in the next section)—remains understudied.

Timescale Mismatch in ILD

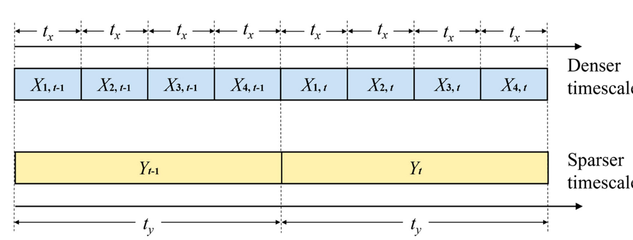
In ILD, if the time interval between consecutive measurements of one variable differs from that of another variable, there is a timescale mismatch between the two variables (see Figure 1). For example, a researcher may be interested in the dynamic bidirectional relation between individuals' affect (X) and exercise (Y). Considering that affect fluctuates more over time and that its measurement is usually based on the current moment, the researcher measures affect four times a day, with 4 hr between each measurement. In contrast, given that exercise takes time and its measurement is usually based on a longer period of time, the researcher measures exercise once a day (e.g., participants are asked to report the total number of minutes of exercise each day before sleep). In this case, there is a 4-hr interval between consecutive measurements of affect and a 1-day interval between consecutive measurements of exercise, suggesting a mismatch in timescales between the two variables.

In general, variables with mismatched timescales have the following characteristics. Variables with denser timescales fluctuate more over time and are usually assessed on shorter time intervals (e.g., hours or minutes). For example, self-reported positive and negative affect may change within a few hours and are usually measured with a time reference "at the moment." In addition, physical and physiological states, such as physical activity and sedentary behavior, can also change substantially in a short period of time, and their fluctuations are often better captured with real-time sensors (e.g., accelerometers; Liao et al., 2017; Maes et al., 2023). In contrast, variables with sparser timescales show weaker fluctuations over time and are usually assessed on longer time intervals (e.g., 1 day). First, some variables can only be measured on sparser timescales due to objective constraints. For example, due to natural rhythms, a person's sleep duration can only be measured at 1-day intervals. Second, some variables occur less frequently or need to be accumulated over a long period of time, and are therefore more suitable for assessment on sparser timescales. For example, measuring a person's exercise on longer time intervals (e.g., 1 day) may better reflect fluctuations in that person's exercise. Third, some variables occur at specific times of the day. For example, researchers may be interested in the dynamic associations between an individual's state at a specific time point (e.g., a child's school refusal behavior each morning) and his or her other states at other times of the day (e.g., the child's affective well-being at school). In this case, researchers are likely to use different timescales to assess different state variables.

The importance of choosing appropriate timescales has been emphasized in dynamic studies. Researchers have examined the impacts of sampling frequency on parameter estimation accuracy

Figure 1

Illustration of Timescale Mismatch



Note. X has a denser timescale, while Y has a sparser timescale. See the online article for the color version of this figure.

(Batra et al., 2023) and reliability (i.e., the standard errors of the parameter estimates (Adolf et al., 2021). Batra et al. (2023) showed that the estimation of autoregressive effects can be highly biased if the sampling timescales are sparser than the true timescales (data-generating timescales) of given processes. Furthermore, Adolf et al. (2021) demonstrated how to determine the optimal sampling frequency that leads to minimal standard errors. Regarding digital-phenotyping data, Langener et al. (2024) demonstrate the impacts of time-related decisions (e.g., the choice of temporal resolution to aggregate data of variables with denser timescales) on the estimated associations between variables of interest. By summarizing existing literature and expert opinions, Velozo et al. (2024) provided practical guidelines for designing studies that combine experience sampling data (typically with sparser timescales) with wearables and passive sensing data (typically with denser timescales). These help researchers make informed decisions about the timescales on which to sample the variables of interest. However, researchers face methodological challenges when attempting to investigate the dynamic relations between these variables.

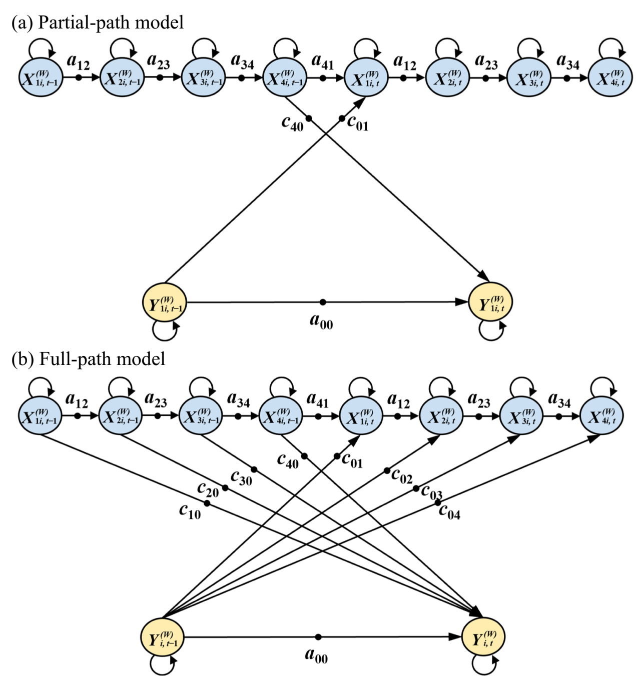
Two modeling frameworks are primarily considered for analyzing (intensive) longitudinal data: discrete-time models and continuous-time models. Discrete-time models (e.g., DSEM) assume equal time intervals between consecutive measures. However, this assumption is often violated in ILD, where time intervals may vary between repeated measures within an individual and across individuals (Voelkle et al., 2012). Although some practical solutions have been proposed to address this issue (e.g., by inserting phantom variables at missing occasions (Voelkle & Oud, 2015), and rescaling variables to equal time intervals and then filling in missing data (Asparouhov et al., 2018)), parameter estimates of discrete-time models still suffer from time-interval dependency (i.e., parameter interpretation depends on the time interval between consecutive measures; Kuiper & Ryan, 2018). Continuous-time models (e.g., stochastic differential equation models; Oravecz et al., 2011; Voelkle et al., 2012) can circumvent these issues. By assuming an underlying continuous function for the dynamic process, continuous-time models treat time as a continuum of real values, rather than a sequence of discrete integer values. This allows for unequal time intervals between consecutive measures and enables flexible parameter interpretation and comparison by rescaling them to different time intervals (Voelkle et al., 2012). Currently, several software implementations support fitting continuous-time models. For example, R packages such as dynr (Ou et al., 2019), ctsem (Driver et al., 2017), and OpenMx (Neale et al., 2016) can be used for continuous-time modeling, while Mplus recently provides a continuous-time version of DSEM (Asparouhov & Muthén, 2024). Despite the conceptual advantages and growing implementation of continuous-time models, they are notably more complex to understand compared to other discrete-time models (e.g., DSEMs) commonly used for ILD analyses by psychologists. This complexity is further exacerbated when examining the dynamic interplay between variables with mismatched timescales. Therefore, existing studies have mainly used discrete-time modeling approaches to address two types of research questions regarding timescale mismatched variables, one focusing on denser timescales and the other on sparser timescales, which are further elaborated in the following two paragraphs.

Focusing on the denser timescales, some researchers were interested in the detailed processes of the dynamic interplay between variables with mismatched timescales. For example, a study explored the dynamic relation between accelerometer-assessed physical activity and individuals' self-reported affective, intentional, and physical states (Maes et al., 2023). Participants completed a questionnaire 6 times a day to evaluate their states, and their physical activity was captured by the accelerometer 15, 30, 60, and 120 min after each questionnaire. For each of the four time lags, the researchers constructed separate multilevel models to explore the lagged effects of individuals' self-reported states on their accelerometer-assessed physical activity. In another study examining the dynamic bidirectional relation between self-rated health and sleep (Lücke et al., 2023), researchers measured self-rated health six times per day and sleep once per day. In their additional analyses, researchers assumed that individuals' last self-rated health before sleep predicted that day's sleep, and the night sleep predicted the first self-rated health after waking. Therefore, they used DSEM to estimate only the cross-lagged effects between the temporally closest time points of self-rated health and sleep (in other words, the cross-lagged effects corresponding to the shortest time interval) to examine their dynamic process (see Figure 2a for a demonstrative model).

Focusing on the sparser timescales, other researchers aimed to examine the overall reciprocal effects between variables with mismatched timescales. For example, in a study exploring the dynamic reciprocal effects between affective well-being and sleep (Neubauer

Figure 2

Dynamic Structural Equation Models for Studying Dynamic Interplay Processes Between Variables with Mismatched Timescales



Note. Current practice (a) and its improvement (b) for studying dynamic interplay processes between variables with mismatched timescales. For simplicity, only the within-person models are presented here. The between-person models include all random effects (corresponding to the solid dots in the within-person models) and estimate the correlations among X_1, X_2, X_3, X_4, and Y. See the online article for the color version of this figure.

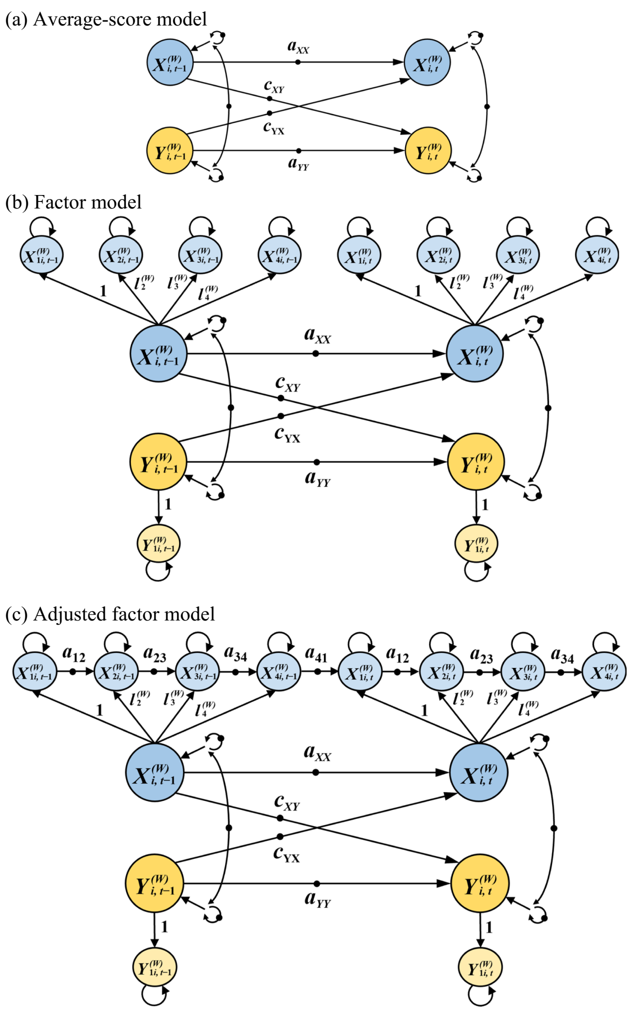
et al., 2021), researchers measured positive and negative affect four times per day and sleep duration and quality once per day. To estimate the overall effects of the dynamic interaction between affect and sleep, the researchers aggregated the four daily measures of affect (morning, late morning, afternoon, and evening) into one average score that reflects an individual's overall affect state each day. After adjusting affect to the same 1-day timescale as sleep, the researchers estimated the cross-lagged effects between affect and sleep in a DSEM (see Figure 3a for a demonstrative model).

Current Issues and Possible Solutions

Although studies have proposed some approaches to address the problem of timescale mismatch between variables, each approach has its limitations. On the one hand, in studies focusing on detailed processes of dynamic interactions between variables, analyses based on separate multilevel models (corresponding to different time lags between variables) did not take into account the time-series processes (or, autoregressive processes) of variables with denser timescales, and were unable to estimate the effects of one time point after controlling for the effects of other time points. In addition, under the framework of DSEM, the partial-path model (see Figure 2a) that considers only the cross-lagged effects between the temporally closest time points between variables relied on strong assumptions about the dynamic interaction process between variables. These cross-lagged effects may not adequately reflect the dynamic interplay processes between variables. Moreover, the reciprocal effects between variables did not necessarily decrease as their corresponding time intervals increased, and each time point (not only the closest time point) of the variables with denser timescales may have a cross-lagged effect with the variables with sparser timescales. This suggested that the partial-path model may neglect time-specific reciprocal effects between timescale mismatched variables. In contrast, the full-path model (see Figure 2b) takes into account all possible reciprocal effects between variables and is more flexible to apply in different situations. Therefore, in order to better understand the process of dynamic interactions between variables with mismatched timescales, we believe it is more appropriate to examine the full-path model (Figure 2b) than the partial-path model (Figure 2a).

Figure 3

Dynamic Structural Equation Models for Studying Overall Reciprocal Effects Between Variables with Mismatched Timescales



Note. Current practice (a) and its improvement (b) and (c) for studying overall reciprocal effects between variables with mismatched timescales. For simplicity, only the within-person models are presented here. The between-person models are detailed in the Simulation Conditions section in Study 2-1 and Study 2-2. See the online article for the color version of this figure.

On the other hand, regarding the overall reciprocal effects between timescale mismatched variables, existing studies primarily adopt the average-score model (see Figure 3a) to match the timescales of the two variables. However, this approach has several limitations. First, the average-score model fails to fully utilize the information of variables with denser timescales. Second, calculating the average score for the variable with a denser timescale does not take into account the different contributions of the variable's different time points (e.g., morning and evening) to its overall level over a longer period (e.g., a day; McNeish & Wolf, 2020). Third, modeling based on average scores cannot reflect the temporal relations between different time points of the variable with a denser timescale (Hamaker & Wichers, 2017; Zhou et al., 2021). For the first two limitations, a common solution within the context of structural equation modeling is to construct latent factors (Fan, 2003; Oh & Jahng, 2023). Therefore, we propose a factor model (see Figure 3b), in which the differences in the contributions of different time points to the overall level can be reflected by the factor loadings of different time points on the latent state factor. The latent factors of the two variables are matched in their timescales, which enables the examination on the overall cross-lagged effects between the two variables.

Furthermore, concerning the third limitation of the average-score model, we believe that a more reasonable and flexible assumption (i.e., a weaker assumption) is that after accounting for the common factor of different time points of the variable with a denser timescale, there is still temporal dependence between these time points. Therefore, we constructed an adjusted factor model (see Figure 3c), incorporating regression effects[1] between different time points of the variable with denser timescales into the factor model.

[1] In this study, we use the regression effect (i.e., a_{12}, a_{23}, a_{34}, and a_{41} in Figure 3c) to refer to the temporal dependency between different time points of the variable with denser timescales in the adjusted factor model.

The inclusion of these regression effects suggests that different time points of the variable with denser timescales may not be completely independent indicators of a latent factor, as posited by traditional measurement models, but are temporally interdependent, highlighting a dynamic temporal process.

Taken together, from the perspective of constructing latent factors, we examined the factor model (see Figure 3b) and the adjusted factor model (see Figure 3c). The only difference between these two models is whether they take into account the within-person regression effects between different time points of variables with denser timescales and the individual differences in these effects. By comparing the three modeling approaches (i.e., the average-score model, the factor model, and the adjusted factor model), we aim to offer effective and practical solutions for empirical studies examining the overall reciprocal effects between variables with mismatched timescales.

The Present Study

The main objective of this study is to explore possible solutions for timescale mismatch in ILD based on DSEMs. Previous research has primarily focused on two types of relevant research questions (i.e., the detailed processes of dynamic interactions between variables, and the overall reciprocal effects of bidirectional relations between variables) and proposed preliminary approaches (i.e., the partial-path model, and the average-score model) to model the dynamic relations between variables with mismatched timescales. The study assessed the limitations of existing approaches and proposed possible improvements (i.e., the full-path model, and the (adjusted) factor model).

Specifically, three simulation studies were conducted to demonstrate the issues in existing approaches and to evaluate the effectiveness of the proposed approaches. In Study 1, we focused on the detailed processes of dynamic interactions between variables, comparing the partial-path model and the full-path model. We expected that the full-path model would outperform the partial-path model. In Study 2, we focused on the overall reciprocal effects of bidirectional relations between variables. Study 2-1 compared the average-score model with the factor model, and Study 2-2 compared the factor model with the adjusted factor model. We expected the factor model to outperform the average-score model, while we did not have clear anticipation regarding the choice between the factor model and the adjusted factor model. In Study 3, we applied empirical data with timescale mismatched variables to illustrate the differences in the modeling approaches for the two types of research questions (i.e., the partial-path model, and the full-path model; the average-score model, the factor model and the adjusted factor model).

Study 1: Dynamic Interaction Processes of Timescale Mismatched Variables

Method

Transparency and Openness

Data were analyzed using Mplus Version 8.10 (Muthén & Muthén, 1998–2017), and R, Version 4.2.2 (R Core Team, 2020), the package ggplot2, Version 3.4.0 (Wickham, 2011), and the package MplusAutomation, Version 1.1.0 (Hallquist & Wiley, 2018). All data, code, and materials have been uploaded to the Open Science Framework repository (https://osf.io/fdpsj; Luo, 2024). This study was not preregistered.

Procedure

To examine how to better describe dynamic processes between variables with mismatched timescales, we first simulated the data of two variables. Specifically, we adjusted the number of significant cross-lagged effects under different simulation conditions so that the true models were the partial-path model (see Figure 2a) and the full-path model (see Figure 2b), respectively. The timescale of one variable (i.e., X) in the data was several times denser than that of the other variable (i.e., Y). The data were generated in Mplus. The simulation was replicated for 100 times.

Then, we fit the simulated data to the partial-path model and the full-path model, respectively. The model parameters were estimated in Mplus using Bayesian estimation. Two Gibbs-sampler chains were used, each with a minimum number of iterations of 3,000 and a maximum number of iterations of 10,000 (given that the parameter estimation was time consuming[2]). Model convergence was determined according to the potential scale reduction (PSR) of the parameters. In addition, we examined the parameter estimates for both models to exclude outliers (i.e., a variance or covariance of any variable or parameter greater than 2 was considered an outlier as the true values of (co)variances were (much) less than 1). Finally, the parameter estimation results from converged models and without outliers were used to evaluate the performance of the partial-path model and the full-path model. The R package MplusAutomation (Hallquist & Wiley, 2018) was used to replicate the parameter estimation for 100 times.

We evaluated the performance of the partial-path model and the full-path model based on their parameter estimation, power, and model convergence. First, we calculated the mean ($\overline{\varphi}$) of 100 parameter estimates for the autoregressive and cross-lagged parameters. To evaluate the estimation accuracy, we calculated the bias ($\overline{\varphi} - \Phi$) and relative bias $\left(\frac{\overline{\varphi}-\Phi}{\Phi}\right)$ of each parameter, where Φ denotes the true values of the parameters. Then, we calculated the root-mean-squared error to evaluate the estimation efficiency. In addition, for parameters whose true values are not 0, we calculated their power as the percentage of significant cases (i.e., 95% credible intervals excluding zero) in the 100 replications. For parameters with a true value of 0, their type I errors (i.e., the percentage of significant cases over the 100 replications) were calculated. Regarding model convergence, models with PSR less than 1.1 for all parameters were considered convergent. We calculated the percentage of the number of models that converged in 100 parameter estimations.

Simulation Conditions

Study 1 considered 18 simulation conditions (see Table 1). We first set a reference condition (Condition 1) in which the autoregressive effects of the variables with the denser (i.e., X) and the sparser

[2] Study 1 and Study 2-1 used a server with a 16-core central processing unit and 90GB dynamic random-access memory, and the simulation studies took approximately 70 hr and 150 hr, respectively. Study 2-2 used a server with a 64-core central processing unit and 128GB dynamic random-access memory, and the simulation took approximately 900 hr.

Table 1
Simulation Conditions in Study 1

Condition	a_{12}	a_{23}	a_{34}	a_{41}	a_{00}	c_{10}	c_{20}	c_{30}	c_{40}	c_{01}	c_{02}	c_{03}	c_{04}	N	T	Ratio	Missing (%)
Reference																	
1	0.3	0.3	0.3	0.3	0.1	0.15	0.2	0.25	0.3	0.3	0.25	0.2	0.15	200	15	4:1	0
Test CR																	
2	0.3	0.3	0.3	0.3	0.1	**0**	**0**	**0**	0.3	0.3	**0**	**0**	**0**	200	15	4:1	0
3	0.3	0.3	0.3	0.3	0.1	**0**	**0**	**0**	0.3	0.3	0.25	0.2	0.15	200	15	4:1	0
4	0.3	0.3	0.3	0.3	0.1	0.15	0.2	0.25	0.3	0.3	**0**	**0**	**0**	200	15	4:1	0
Test AR																	
5	**0.1**	**0.1**	**0.1**	**0.1**	0.1	0.15	0.2	0.25	0.3	0.3	0.25	0.2	0.15	200	15	4:1	0
6	**0.5**	**0.5**	**0.5**	**0.5**	0.1	0.15	0.2	0.25	0.3	0.3	0.25	0.2	0.15	200	15	4:1	0
7	0.3	0.3	0.3	0.3	**0.3**	0.15	0.2	0.25	0.3	0.3	0.25	0.2	0.15	200	15	4:1	0
8	0.3	0.3	0.3	0.3	**0.5**	0.15	0.2	0.25	0.3	0.3	0.25	0.2	0.15	200	15	4:1	0
Test *N*																	
9	0.3	0.3	0.3	0.3	0.1	0.15	0.2	0.25	0.3	0.3	0.25	0.2	0.15	**100**	15	4:1	0
10	0.3	0.3	0.3	0.3	0.1	0.15	0.2	0.25	0.3	0.3	0.25	0.2	0.15	**300**	15	4:1	0
Test *T*																	
11	0.3	0.3	0.3	0.3	0.1	0.15	0.2	0.25	0.3	0.3	0.25	0.2	0.15	200	**7**	4:1	0
12	0.3	0.3	0.3	0.3	0.1	0.15	0.2	0.25	0.3	0.3	0.25	0.2	0.15	200	**30**	4:1	0
Test *N*&*T*																	
13	0.3	0.3	0.3	0.3	0.1	0.15	0.2	0.25	0.3	0.3	0.25	0.2	0.15	**100**	**30**	4:1	0
14	0.3	0.3	0.3	0.3	0.1	0.15	0.2	0.25	0.3	0.3	0.25	0.2	0.15	**300**	**7**	4:1	0
Test ratio																	
15	0.3	0.3	0.3	0.3	0.1	0.15	0.225	0.3		0.3	0.225	0.15		200	15	**3:1**	0
16	0.3	0.3	0.3	0.3	0.1	0.15	0.1875	0.225	0.2625[a]	0.3	0.2625	0.225	0.1875[a]	200	15	**5:1**	0
Test missing																	
17	0.3	0.3	0.3	0.3	0.1	0.15	0.2	0.25	0.3	0.3	0.25	0.2	0.15	200	15	4:1	**20**
18	0.3	0.3	0.3	0.3	0.1	0.15	0.2	0.25	0.3	0.3	0.25	0.2	0.15	200	15	4:1	**40**

Note. Bold values indicate the parameters varied in each simulation condition. a_{12}, a_{23}, a_{34}, and a_{41} = autoregressive effects of *X*; a_{00} = autoregressive effects of *Y*; c_{10}, c_{20}, c_{30}, and c_{40} = cross-lagged effects of *X* on *Y*; c_{01}, c_{02}, c_{03}, and c_{04} = cross-lagged effects of *Y* on *X*; N = sample size; T = the number of time points per subject for the variable with the sparser timescale (i.e., *Y*); ratio = timescale mismatch ratio between *X* and *Y*; missing = proportion of missing data for *X* and *Y*; CR = cross-lagged effect; AR = autoregressive effect.
[a] When the timescale mismatch ratio was equal to 5, the cross-lagged effects between *X* and *Y* were set as 0.3 (c_{50} and c_{01}), 0.2625 (c_{40} and c_{02}), 0.225 (c_{30} and c_{03}), 0.1875 (c_{20} and c_{04}), and 0.15 (c_{10} and c_{05}).

(i.e., *Y*) timescale were set to 0.3 and 0.1, respectively, the cross-lagged effects between *X* and *Y* were set as 0.3, 0.25, 0.2, and 0.15 (assuming that the cross-lagged effects diminished as the corresponding time interval increases), the sample size was set to 200, the number of time points per subject for *Y* was set to 15, the timescale mismatch ratio was set to 4:1, and the proportion of missing data for *X* and *Y* was set to 0%.

Then, we varied the autoregressive effects of *X* and *Y*, the cross-lagged effects between *X* and *Y*, the sample size, the number of time points per subject for *Y*, the timescale mismatch ratio, and the proportion of missing data to compare the performance of the partial-path model and the full-path model. Specifically, for the autoregressive effects of *X* and *Y*, we set them to 0.1, 0.3, and 0.5 to reflect small, medium, and large autoregressive effects. For the cross-lagged effects between *X* and *Y*, we set Condition 2 to reflect the situation where the cross-lagged effect is significant only between *X* and *Y* that are closest in time (i.e., the true model is the partial-path model), and set Conditions 3 and 4 to represent common situations in empirical studies where one variable has a greater lagged effect on the other. Regarding sample size, we used 100, 200, and 300 to represent the small, medium, and large samples in intensive longitudinal studies that examine dynamic bidirectional relations between variables (Luo et al., 2024). Regarding the number of time points per subject for *Y*, we used 7, 15, and 30 to represent the small (e.g., Lücke et al., 2023), medium, and large (e.g., Neubauer et al., 2021) number of time points in intensive longitudinal studies on variables with mismatched timescales. Additionally, we set Conditions 13 and 14 to examine the interplay of small/large sample size and large/small number of time points. In terms of the timescale mismatch ratio, we set 3:1, 4:1, and 5:1 to reflect different degrees of mismatch between the timescale of variables (Luo et al., 2024). To investigate the impact of missing data, we set the proportion of missing data for *X* and *Y* to 20% and 40% in Conditions 17 and 18, respectively, following the missing completely at random missingness mechanism.

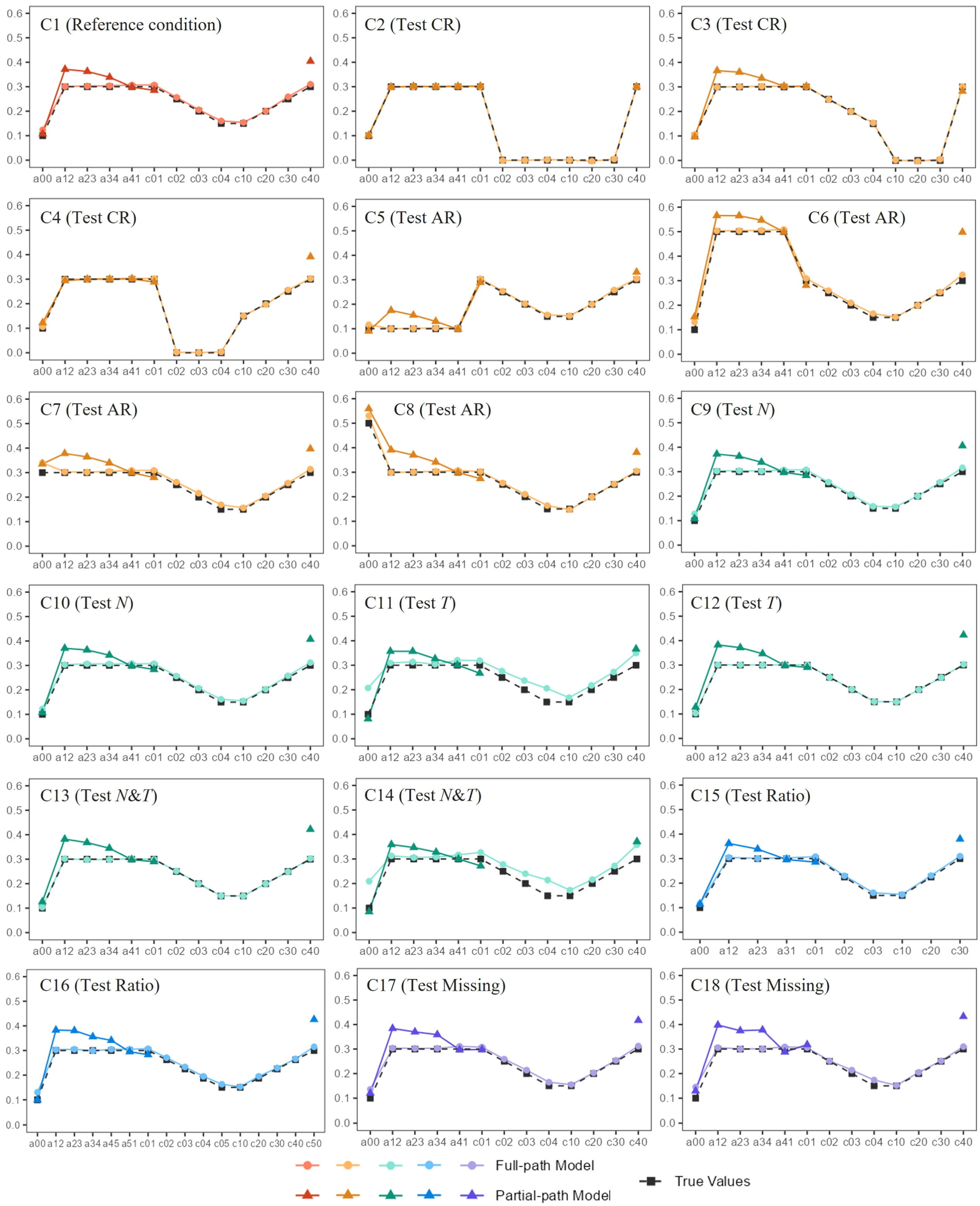
In addition, other model parameters were fixed in all conditions. At the within-person level, the residual variances of *X* and *Y* were fixed to 0.5. All autoregressive and cross-lagged effects were allowed to vary across individuals, and their random variances were equal to their fixed effects multiplied by 0.01. At the between-person level, the correlations between *X* at different time points (i.e., X_1–X_4) were fixed to .4, and the correlations of *X* at different time points with *Y* were fixed to .2. The residual variances of *X* and *Y* were fixed to .5.

Results

Figure 4 shows the true and estimated values of the autoregressive and cross-lagged parameters in the partial-path model and the full-path model. In the reference condition (Condition 1), all parameters in the partial-path model deviated from the true values to varying degrees, while the parameters in the full-path model accurately recovered the true values. As the number of significant cross-lagged

Figure 4

Parameter Estimation Results in Study 1



Note. The vertical axis is the true or estimated value of each parameter. Reference condition: $a_{12} = a_{23} = a_{34} = a_{41} = 0.3$; $a_{00} = 0.1$; $c_{10} = 0.15$; $c_{20} = 0.2$; $c_{30} = 0.25$; $c_{40} = 0.3$; $c_{01} = 0.3$; $c_{02} = 0.25$; $c_{03} = 0.2$; $c_{04} = 0.15$; $N = 200$; $T = 15$; ratio = 4:1; missing = 0%; CR denotes cross-lagged effect; AR denotes autoregressive effect; N denotes sample size; T denotes the number of time points per subject for the variable with the sparser timescale (i.e., Y). See the online article for the color version of this figure.

effects in the true model decreased (Condition 1 → Condition 3 and 4 → Condition 2), the accuracy of parameter estimation in the partial-path model increased. In Condition 2, the partial-path model accurately estimated all parameters and the full-path model also estimated the parameters with a true value of 0 very well. However, the parameter estimates of the partial-path model still showed large biases as the autoregressive effects of X and Y varied, and the bias of the cross-lagged effect of X on Y (i.e., c_{40}) increased as the autoregressive effects of X increase (Condition 5 → Condition 1 → Condition 6). The sample size and the number of time points per subject of X did not substantially affect the accuracy of parameter estimates for either model, except that the estimation accuracy for both models was lower when the number of time points per subject was smaller, Condition 11 ($N = 200$, $T = 7$) and Condition 14 ($N = 300$, $T = 7$). In addition, parameter estimates in the partial-path model showed greater biases as the timescale mismatch ratio (Condition 15 → Condition 1 → Condition 16) or the proportion of missing data (Condition 1 → Condition 17 → Condition 18) increased.

Regarding the power, all parameters in both models had power equal to or larger than 94% in all conditions, except for the autoregressive effect of Y (i.e., a_{00}) in the partial-path model, which had a power of 57% in Condition 11 ($N = 200$, $T = 7$) and a power of 75% in Condition 14 ($N = 300$, $T = 7$). The type I error of the parameters with true values of 0 was equal to or less than 6%. In terms of model convergence, the full-path model in Conditions 8 (80%), 11(91%), and 14 (93%) had relatively lower convergence, while the models in other conditions converged well (equal to or higher than 97%). The full results for Study 1 are presented in the online supplemental materials (S1).

The results suggested that the parameter estimation accuracy of the partial-path model was not as good as that of the full-path model under a variety of conditions. Their parameter estimation accuracies were similar only when the cross-lagged effect between X and Y exists solely at temporally adjacent time points (Condition 2). Both models performed similarly well in terms of parameter power and model convergence. In addition, it was worth noting that when the number of time points per subject became smaller ($T = 7$), the performance of both models in terms of parameter estimation accuracy, power, and model convergence decreased, suggesting that this number of time points may be too small for analyzing dynamic processes between variables that are mismatched in timescales.

Study 2: Dynamic Reciprocal Effects Between Timescale Mismatched Variables

Study 2-1: Comparison Between the Average-Score Model and the Factor Model

Method

Procedure. We first simulated the data of two variables with mismatched timescales based on the factor model (see Figure 3b). The timescale of one variable (i.e., X) was several times denser than that of the other variable (i.e., Y). The data were generated in Mplus with the same settings as in Study 1. The data were then fit to the average-score model and the factor model, respectively. The model parameters were estimated in Mplus with the same settings as in Study 1. The parameter estimation results from converged models and without outliers (i.e., valid results) were used to evaluate the performance of the average-score model and the factor model. Generated data sets with improper solutions (i.e., data sets yielding results with outliers when fitted to the data-generating models) were discarded, and the data generation and parameter estimation process was replicated until 100 valid results were obtained for both models. The number of data sets generated to obtain 100 valid results is presented in the online supplemental materials (S2).

To compare the performance of the average-score model and the factor model, we examined the parameter estimate of the autoregressive and cross-lagged effects of the latent state factors of X and Y, as well as the within-person factor loadings (in the factor model only). The estimation accuracy and power of these parameters, as well as the model convergence, were calculated in the same way as in Study 1.

Simulation Conditions. Table 2 presents 18 simulation conditions we considered in Study 2. We first set a reference condition (Condition 1) in which within-person factor loadings of X at different time points (i.e., $X_1{\sim}X_4$) were fixed to 1 and set as 0.7, 0.7, and 0.7, respectively, the autoregressive and cross-lagged effects of X and Y were set as 0.3, the sample size was set to 200, and the number of time points per subject for Y was set to 15, the timescale mismatch ratio was set to 4:1, and the proportion of missing data for X and Y was set to 0%.

Then, we varied the within-person factor loadings of X, the autoregressive and cross-lagged effects of X and Y, the sample size, and the number of time points per subject for Y, the timescale mismatch ratio, and the proportion of missing data to compare the performance of average-score model and factor model. Specifically, for the within-person factor loadings of X, we set Conditions 2 and 3 to reflect the situation where there is some difference between the within-person factor loadings of X and where they are exactly equal, respectively. We expected Condition 3 to be the condition least affected by the differences in within-person factor loadings of X and most likely to support the average-score model. For the autoregressive effects of X and Y, we set Conditions 4 and 5 to reflect common situations in empirical studies where the autoregressive effect of one variable is larger than the other. For the cross-lagged effects between X and Y, we set Condition 6 to reflect the situation where both cross-lagged effects are relatively small. In addition, we also set Conditions 7 and 8 to represent common situations in empirical studies where one variable has a greater lagged effect on the other. Regarding the sample size, the number of time points per subject, the timescale mismatch ratio, and the proportion of missing data, we set Conditions 9–18 based on the same rationale as in Study 1.

In addition, other model parameters were fixed in all conditions. At the within-person level, the autoregressive and cross-lagged effects were allowed to vary between persons, and their random variances were equal to their fixed effects multiplied by 0.01. The residual variances of $X_1{\sim}X_4$ and Y were fixed to 0.1. The fixed effects and random variances of the residual variances of the within-person latent factors (i.e., latent state factors) of X and Y were fixed to −0.4 and 0.01, respectively (log-transformed values set in Mplus, corresponding to true fixed effects of 0.674 and true random variances of 0.006; Schuurman & Hamaker, 2019). The fixed effect and random variance of the residual covariance between the two factors were fixed to −0.7 and 0.01, respectively (log-transformed values set in Mplus, corresponding to true fixed effects of 0.500 and

Table 2
Simulation Conditions in Study 2-1

Condition	$l_1^{(W)}$	$l_2^{(W)}$	$l_3^{(W)}$	$l_4^{(W)}$	a_{XX}	a_{YY}	c_{XY}	c_{YX}	N	T	Ratio	Missing (%)
Reference												
1	1	0.7	0.7	0.7	0.3	0.3	0.3	0.3	200	15	4:1	0
Test loadings												
2	1	**0.9**	**0.8**	**0.7**	0.3	0.3	0.3	0.3	200	15	4:1	0
3	1	**1**	**1**	**1**	0.3	0.3	0.3	0.3	200	15	4:1	0
Test AR												
4	1	0.7	0.7	0.7	**0.2**	**0.4**	0.3	0.3	200	15	4:1	0
5	1	0.7	0.7	0.7	**0.4**	**0.2**	0.3	0.3	200	15	4:1	0
Test CR												
6	1	0.7	0.7	0.7	0.3	0.3	**0.1**	**0.1**	200	15	4:1	0
7	1	0.7	0.7	0.7	0.3	0.3	0.3	**0.1**	200	15	4:1	0
8	1	0.7	0.7	0.7	0.3	0.3	**0.1**	0.3	200	15	4:1	0
Test *N*												
9	1	0.7	0.7	0.7	0.3	0.3	0.3	0.3	**100**	15	4:1	0
10	1	0.7	0.7	0.7	0.3	0.3	0.3	0.3	**300**	15	4:1	0
Test *T*												
11	1	0.7	0.7	0.7	0.3	0.3	0.3	0.3	200	**7**	4:1	0
12	1	0.7	0.7	0.7	0.3	0.3	0.3	0.3	200	**30**	4:1	0
Test *N&T*												
13	1	0.7	0.7	0.7	0.3	0.3	0.3	0.3	**100**	**30**	4:1	0
14	1	0.7	0.7	0.7	0.3	0.3	0.3	0.3	**300**	**7**	4:1	0
Test ratio												
15	1	0.7	0.7	0.7	0.3	0.3	0.3	0.3	200	15	**3:1**	0
16	1	0.7	0.7	0.7	0.3	0.3	0.3	0.3	200	15	**5:1**	0
Test missing												
17	1	0.7	0.7	0.7	0.3	0.3	0.3	0.3	200	15	4:1	**20**
18	1	0.7	0.7	0.7	0.3	0.3	0.3	0.3	200	15	4:1	**40**

Note. Bold values indicate the parameters varied in each simulation condition. $l_1^{(W)}$, $l_2^{(W)}$, $l_3^{(W)}$, and $l_4^{(W)}$ = within-person factor loadings of X at different time points (i.e., $X_1 \sim X_4$); a_{XX} and a_{YY} = autoregressive effects of X and Y; c_{XY} and c_{YX} = cross-lagged effects between X and Y; N = sample size; T = number of time points per subject for Y; ratio = timescale mismatch ratio between X and Y; missing = proportion of missing data for X and Y; AR = autoregressive effect; CR = cross-lagged effect.

true random variance of 0.003; Schuurman & Hamaker, 2019). At the between-person level, the between-person factor loadings of X at different time points (i.e., $X_1{\sim}X_4$) and of Y were all fixed to 1. The residual variances of $X_1{\sim}X_4$ and Y were fixed to 0.1. The residual variances of the between-person latent factors (i.e., latent trait factors) of X and Y were fixed to 0.5, and their covariance was fixed to 0.3.

Results

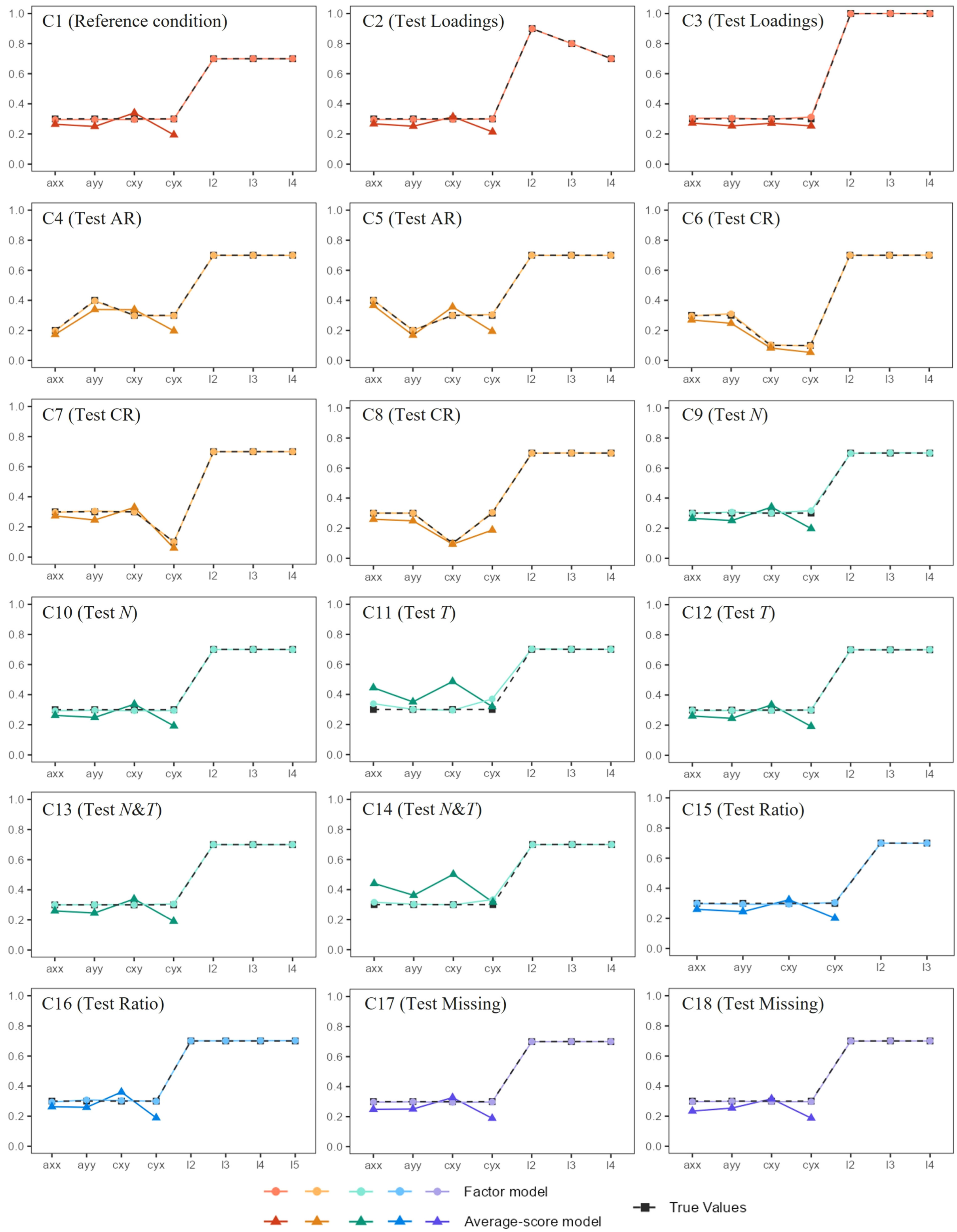
The true and estimated values of the autoregressive and cross-lagged effects and the within-person factor loadings in the average-score model and the factor model are presented in Figure 5. In the reference condition (Condition 1), the average-score model shows bias in the estimation of autoregressive and cross-lagged effects. Moreover, it overestimates one cross-lagged effect and underestimates the other, which may affect the comparison of the reciprocal effects between X and Y in empirical studies, and thus distort the conclusion of their causal dominance. In contrast, the factor model accurately estimates the autoregressive and cross-lagged effects and the within-person factor loadings. Although the estimation bias of the average-score model decreases as the difference between within-person factor loadings in the true model decreases, the average-score model still underestimates the autoregressive and cross-lagged effects even when the within-person factor loadings are equal (Condition 3). In fact, under all conditions that vary the differences between within-person factor loadings (Conditions 2 and 3), autoregressive effects (Conditions 4 and 5), cross-lagged effects (Conditions 6~8), sample size and the number of time points (Conditions 9~14), timescale mismatch ratio (Conditions 15 and 16), and the proportion of missing data (Conditions 17 and 18), the average-score model has varying degrees of parameter estimation bias, whereas the factor model always recovers the true values of the parameters better. It is worth noting that in Conditions 11 ($N = 200$, $T = 7$) and 14 ($N = 300$, $T = 7$), the estimates of autoregressive and cross-lagged effects between X and Y are biased in both models, but the parameter estimates are particularly biased in the average-score model.

In terms of the power of parameters, all parameters with true values equal to or greater than 0.3 had a power of 100%. The power of parameters with lower truth values (i.e., 0.1) was higher than 89% in the factor model and 78% in the average-score model. In terms of model convergence, both models had convergent rates higher than 72% in all conditions. In Conditions 11 ($N = 200$, $T = 7$) and 14 ($N = 300$, $T = 7$), both models had relatively low rates of convergence (the average-score model = 92.3% and 82.4%, the factor model = 91.6% and 72.1%). The full results for Study 2 are presented in the online supplemental materials (S2).

Overall, the bias in the parameter estimates of the average-score model was consistently larger than that of the factor model, even when the factor loadings of X at different time points (i.e., $X_1{\sim}X_4$) were equal (Condition 2). Moreover, the power of the parameters with smaller true values (i.e., 0.1) was slightly lower in the average-score model than in the factor model. The model convergence of both models was acceptable. It should be noted that, since the factor

Figure 5

Parameter Estimation Results in Study 2-1



Note. The vertical axis is the true or estimated value of each parameter. Reference condition: $l_1^{(W)} = 1$; $l_2^{(W)} = 0.7$; $l_3^{(W)} = 0.7$; $l_4^{(W)} = 0.7$; $a_{XX} = a_{YY} = c_{XY} = c_{YX} = 0.3$; $N = 200$; $T = 15$; ratio = 4:1; missing = 0%; AR denotes autoregressive effect; CR denotes cross-lagged effect; N denotes sample size; T denotes the number of time points per subject for the variable with the sparser timescale (i.e., Y). See the online article for the color version of this figure.

model was the true model, it may have an inherent advantage over the average-score model in this comparison, making the comparison not entirely fair. This limitation should be taken into account when interpreting the simulation results. In addition, similar to Study 1, when the number of time points per subject for X was only seven (Conditions 11 and 14), the parameter estimation accuracy, power, and model convergence decreased for both models, suggesting that this number of time points may not be sufficient to explore the reciprocal effects between variables with mismatched timescales.

Study 2-2: Comparison Between the Factor Model and the Adjusted Factor Model

Method

Procedure. After confirming that the factor model outperformed the average-score model, we further investigated whether the factor model or the adjusted factor model was more suitable for analyzing variables with mismatched timescales. The data were first simulated with the timescale of one variable (i.e., X) was several times denser than that of the other variable (i.e., Y). Specifically, we manipulated whether the regression effects between X at different time points were zero, making the true model a factor model (Figure 3b) or an adjusted factor model (Figure 3c). The data were then fitted to the factor model and the adjusted factor model, respectively. The data generation and parameter estimation were performed in Mplus with the same settings as in Study 1. Data generation and parameter estimation were replicated until 100 valid results were obtained for both models. The number of data sets generated to obtain 100 valid results for each simulation condition is shown in the online supplemental materials (S3).

Finally, we examined the estimation accuracy and power of the autoregressive and cross-lagged effects of the latent state factors of X and Y, the within-person factor loadings, the autoregressive effects between X at different time points (i.e., $X_1 \sim X_4$; in the adjusted factor model only), as well as the model convergence. The evaluation indexes were calculated in the same way as in Study 1.

Simulation Conditions. We compared the performance of the factor model and the adjusted factor model under 42 simulation conditions (see Table 3). We first considered three main types of conditions, with no, small, or large regression effects (i.e., *aw*; a_{12}, a_{23}, a_{34}, and a_{41}) between X at different time points (i.e., $X_1 \sim X_4$). In each type of condition, we set a reference condition (Condition 1) in which the autoregressive and cross-lagged effects of X and Y were set as 0.3, the sample size was set to 200, and the number of time points per subject for Y was set to 15, the timescale mismatch ratio was set to 4:1, and the proportion of missing data for X and Y was set to 0%.

Then, we varied the autoregressive and cross-lagged effects of X and Y (Conditions 2∼6), the sample size and the number of time points (Conditions 7∼10), the timescale mismatch ratio (Conditions 11 and 12), and the proportion of missing data (Conditions 13 and 14) in each type of conditions based on the same rationale as in Study 2-1. We only considered a larger number of time points per subject for Y (Conditions 9 and 10; $T = 30$) for two main reasons. First, both Study 1 and Study 2-1 suggested that a smaller number of time points (i.e., $T = 7$) may not be appropriate for examining the dynamic interplay between variables that were mismatched in timescales. Second, the number of data sets generated (i.e., 430 and 581 in Conditions 11 and 14, respectively) to obtain 100 valid results was quite large for the condition with a smaller number of time points in the factor model in Study 2-1 (see Tables S2.11 and S2.14 in the online supplemental materials), which required considerable computational resources. In the adjusted factor model, it could be expected that more data sets may needed. Therefore, we did not test the condition with a smaller number of time points (i.e., $T = 7$) in Study 2-2.

Table 3
Simulation Conditions in Study 2-2

Condition	a_{XX}	a_{YY}	c_{XY}	c_{YX}	N	T	Ratio	Missing (%)
Reference								
1	0.3	0.3	0.3	0.3	200	15	4:1	0
Test AR								
2	**0.2**	**0.4**	0.3	0.3	200	15	4:1	0
3	**0.4**	**0.2**	0.3	0.3	200	15	4:1	0
Test CR								
4	0.3	0.3	**0.1**	**0.1**	200	15	4:1	0
5	0.3	0.3	0.3	**0.1**	200	15	4:1	0
6	0.3	0.3	**0.1**	0.3	200	15	4:1	0
Test *N*								
7	0.3	0.3	0.3	0.3	**100**	15	4:1	0
8	0.3	0.3	0.3	0.3	**300**	15	4:1	0
Test *T*								
9	0.3	0.3	0.3	0.3	200	**30**	4:1	0
Test *N*&*T*								
10	0.3	0.3	0.3	0.3	**100**	**30**	4:1	0
Test ratio								
11	0.3	0.3	0.3	0.3	200	15	**3:1**	0
12	0.3	0.3	0.3	0.3	200	15	**5:1**	0
Test missing								
13	0.3	0.3	0.3	0.3	200	15	4:1	**20**
14	0.3	0.3	0.3	0.3	200	15	4:1	**40**

Note. Bold values indicate the parameters varied in each simulation condition. These conditions are nested within *aw* (with three possible values 0, 0.1, and 0.3). a_{XX} and a_{YY} = autoregressive effects of X and Y; c_{XY} and c_{YX} = cross-lagged effects between X and Y; N = sample size; T = the number of time points per subject for Y; ratio = timescale mismatch ratio between X and Y; missing = proportion of missing data for X and Y; AR = autoregressive effect; CR = cross-lagged effect.

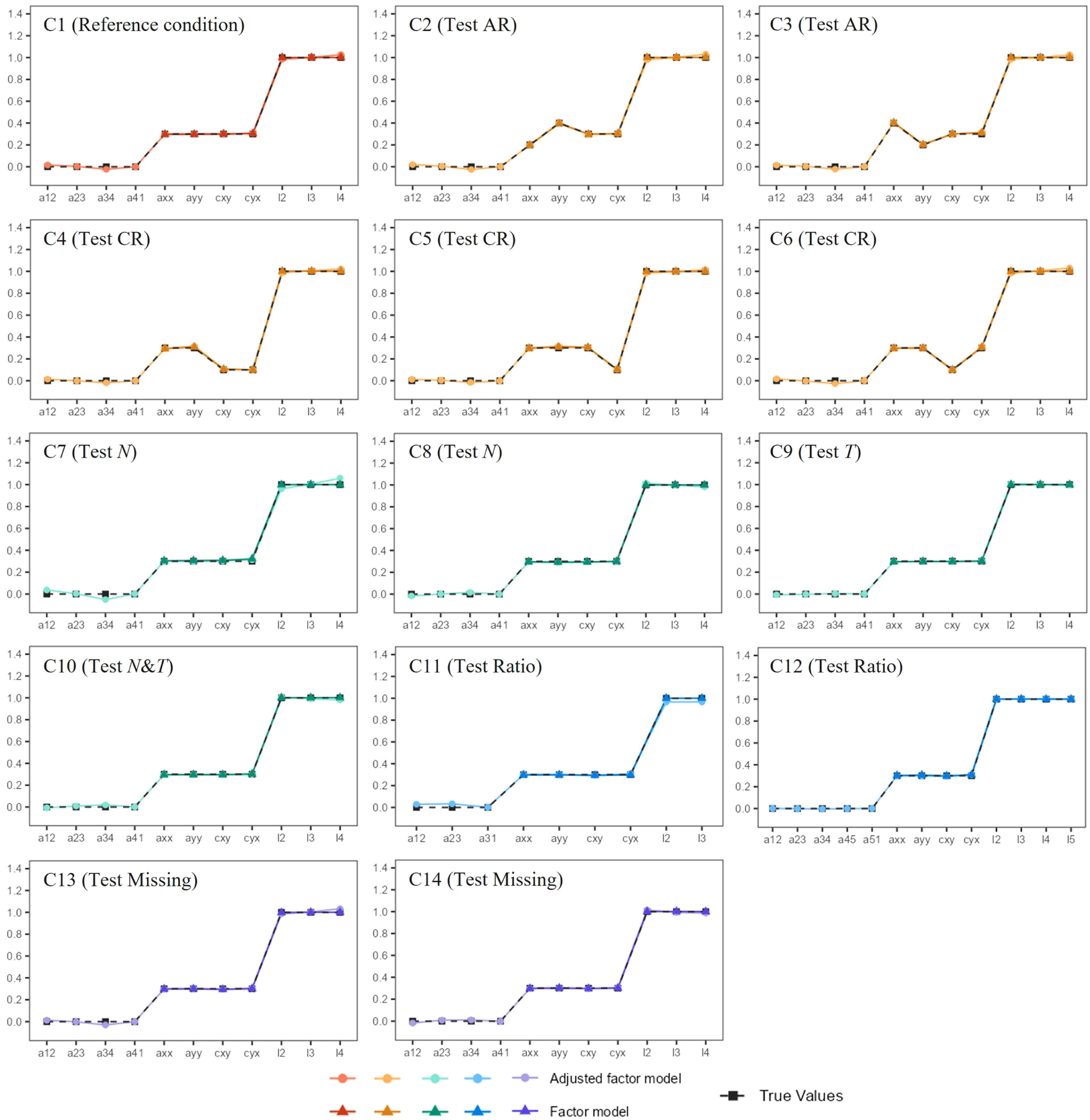
In addition, other model parameters were fixed in all conditions. At the within-person level, the autoregressive and cross-lagged effects were allowed to vary from person to person, and their random variances were equal to their fixed effects multiplied by 0.01. The residual variances of $X_1 \sim X_4$ and Y, the fixed effects and random variances of the residual variances and covariance of the latent state factors of X and Y were set to the same values as in Study 2-1. At the between-person level, the between-person factor loadings, and the residual variances of X at different time points (i.e., $X_1 \sim X_4$) and of Y, as well as the residual variances and covariance of the latent trait factors of X and Y, were fixed to the same values as in Study 2-1.

Results

Figures 6 ($aw = 0$), 7 ($aw = 0.1$), and 8 ($aw = 0.3$) show the true and estimated values of the autoregressive and cross-lagged effects, the within-person factor loadings, and the regression effects between X at different time points (in the adjusted factor model only). At first glance, we found that as the regression effects (i.e., *aw*) between X at different time points increased (from 0, 0.1, to 0.3), the bias of the parameter estimates in the factor model became larger, especially for the within-person factor loadings, whereas the adjusted factor

Figure 6

Parameter Estimation Results in Study 2-2 (aw = 0)

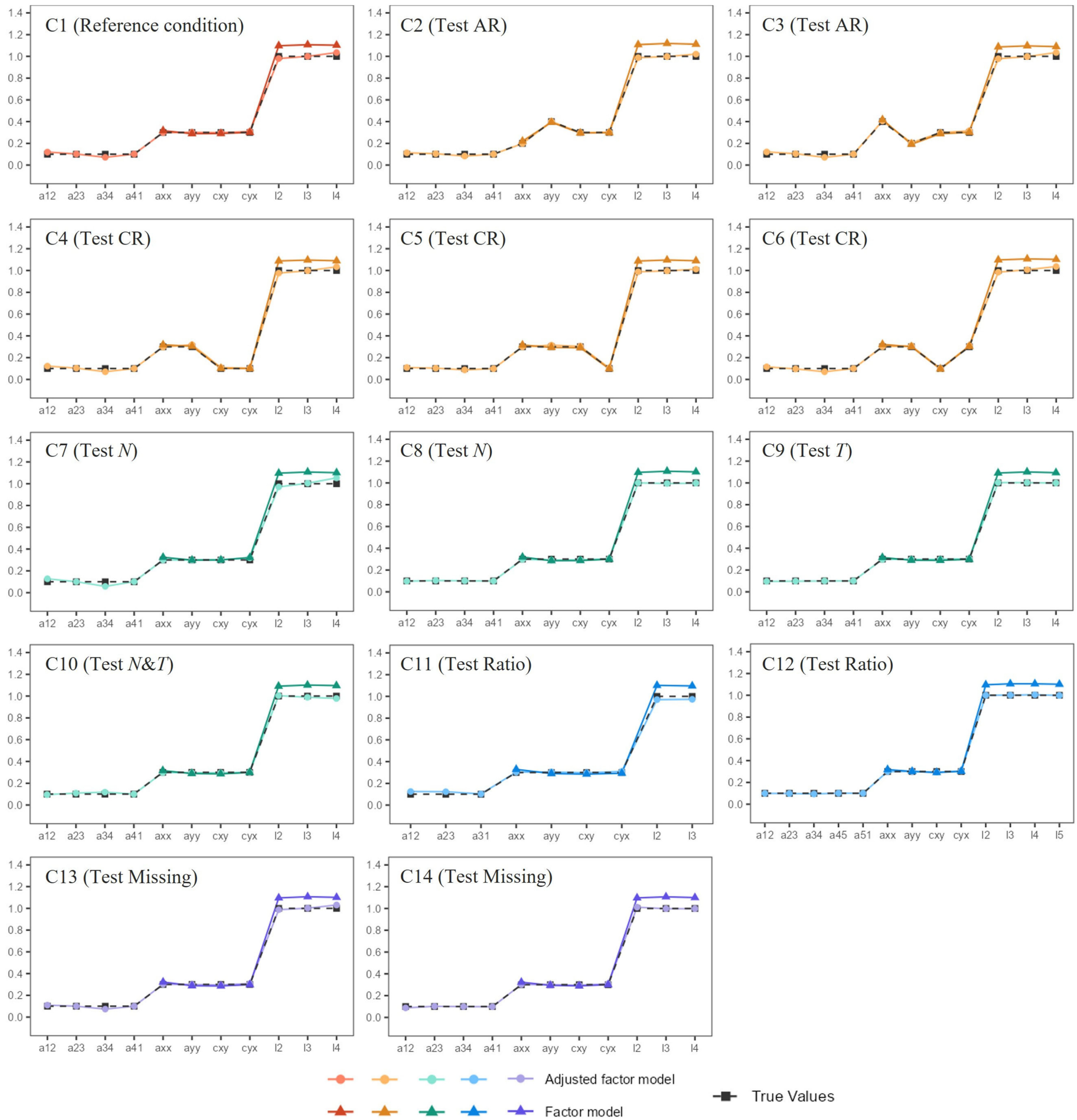


Note. The vertical axis is the true or estimated value of each parameter. Reference condition: $a_{XX} = a_{YY} = c_{XY} = c_{YX} = 0.3$; $N = 200$; $T = 15$; ratio = 4:1; missing = 0%; AR denotes autoregressive effect; CR denotes cross-lagged effect; N denotes sample size; T denotes the number of time points per subject for the variable with the sparser timescale (i.e., Y). See the online article for the color version of this figure.

model always recovered the parameter estimates accurately. This result held across all conditions with different values of the true autoregressive and cross-lagged effects of X and Y, sample size, number of time points per subject, timescale mismatch ratio, and proportion of missing data. Specifically, in the adjusted factor model, the deviations of the estimates from the true values were less than 0.024 for the autoregressive and cross-lagged effects (with true values from 0.1 to 0.4), and 0.095 for the within-person factor loadings (with a true value of 1). In contrast, in the factor model, the deviations of the autoregressive and cross-lagged effects from the true values were as large as 0.068, and the overestimation of the within-person factor loadings from the true values were as large as 0.393.

Figure 7

Parameter Estimation Results in Study 2-2 ($aw = 0.1$)



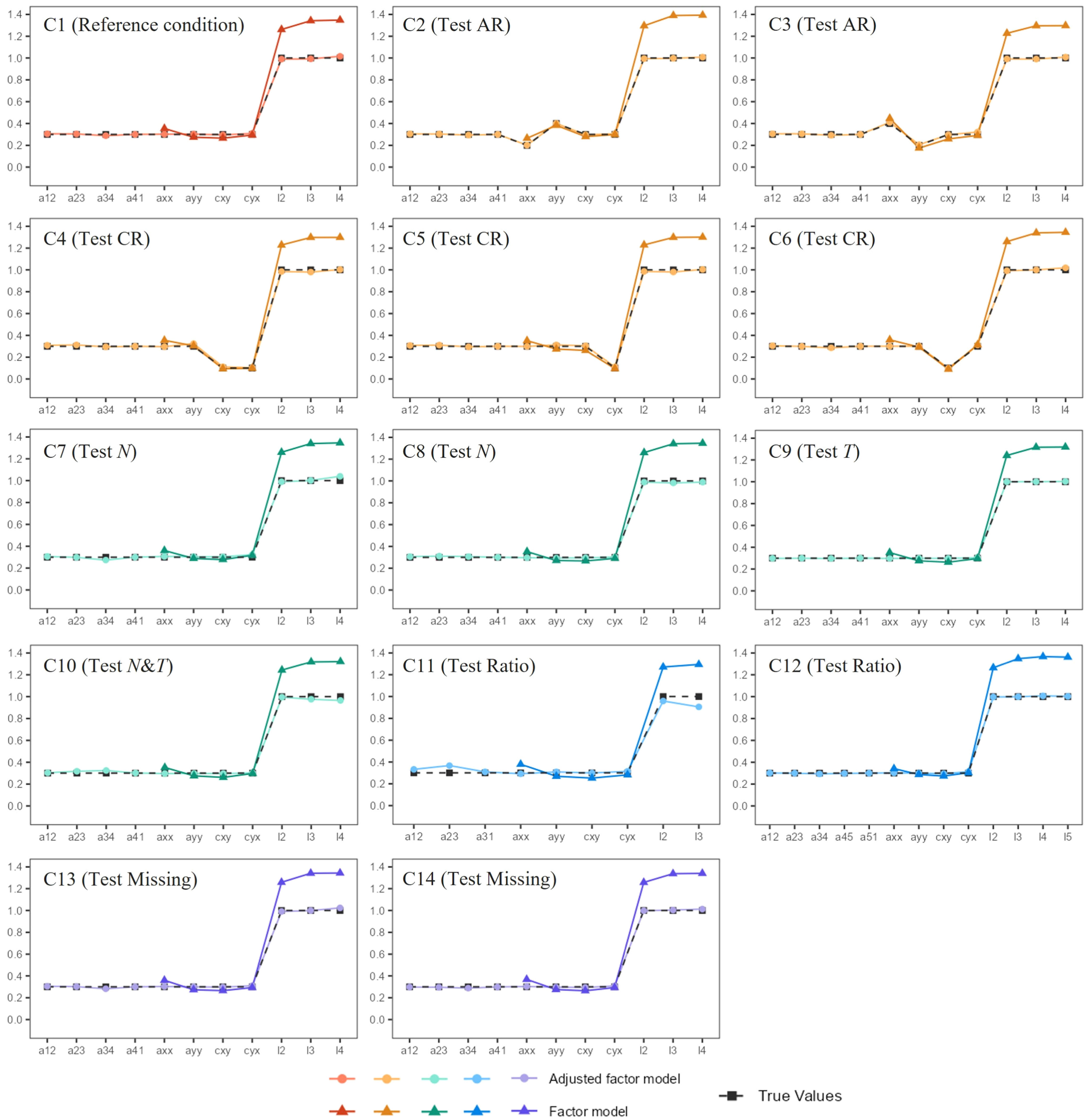
Note. The vertical axis is the true or estimated value of each parameter. Reference condition: $a_{XX} = a_{YY} = c_{XY} = c_{YX} = 0.3$; $N = 200$; $T = 15$; ratio = 4:1; missing = 0%; AR denotes autoregressive effect; CR denotes cross-lagged effect; N denotes sample size; T denotes the number of time points per subject for the variable with the sparser timescale (i.e., Y). See the online article for the color version of this figure.

However, a further look revealed that when there was no or a small regression effect between X at different time points (i.e., $aw = 0$ or 0.1), both models had small estimation biases (less than 0.019 for the adjusted factor model, and 0.029 for the factor model) for the autoregressive and cross-lagged effects between X and Y, which were of most interest in empirical studies. Nonetheless, in these cases, the within-person factor loadings in the factor model were still somewhat overestimated (with a bias of about 0.1).

In terms of the power and type I error, the autoregressive and cross-lagged effects as well as the within-person factor loadings

Figure 8

Parameter Estimation Results in Study 2-2 (aw = 0.3)



Note. The vertical axis is the true or estimated value of each parameter. Reference condition: $a_{XX} = a_{YY} = c_{XY} = c_{YX} = 0.3$; $N = 200$; $T = 15$; ratio = 4:1; missing = 0%; AR denotes autoregressive effect; CR denotes cross-lagged effect; N denotes sample size; T denotes the number of time points per subject for the variable with the sparser timescale (i.e., Y). See the online article for the color version of this figure.

had a power equal to or larger than 89%. For the regression effects between X at different time points, the type I error of the regression effects with a true value of 0 was lower than 21%. The power of the regression effects with a true value of 0.1 was equal to or greater than 64% in most cases, except for the regression effect between the third and fourth time points of X (i.e., a_{34}) and the regression effects in Conditions 7, 10 ($N = 100$), 11 (mismatch ratio = 3:1), 13, and 14 (proportions of missing data of 20% and 40%, respectively). The power of the regression effects with a true value of 0.3 was equal to or greater than 92%, except for some regression effect of X in Conditions 7 (51%) and 14 (73%). In terms of model convergence, the convergence of the factor model was slightly better than that of the adjusted factor model, but both models had acceptable rates of convergence (90%–100% for the factor model; 88%–100% for the adjusted factor model).

Overall, as the regression effects (i.e., *aw*) between *X* at different time points increased (from 0, 0.1, to 0.3), the estimation bias of the parameters in the factor model became larger, whereas the adjusted factor model accurately estimated the parameters. This was particularly evident for the within-person factor loadings, indicating that the within-person factor loadings estimated with the factor model did not accurately represent the contribution of *X* at different time points to its overall latent level.

However, when we focused on the autoregressive and cross-lagged effects between *X* and *Y*, which were key parameters of great interest in empirical studies, we found that both models showed low estimation bias, especially when there was no or a small regression effect between *X* at different time points (i.e., $aw = 0$ or 0.1). In addition, both models had performed similarly well in terms of the power of the autoregressive and cross-lagged effects. The convergence rate of the factor model was slightly higher than that of the adjusted factor model. These results suggest that the factor model may be useful for estimating the overall reciprocal effect between *X* and *Y*, especially when the regression effect (i.e., *aw*) was small between *X* at different time points. Thus, the factor model may provide researchers with a practical method for estimating dynamic bidirectional effects between variables with mismatched timescales when the adjusted factor model does not converge successfully.

Study 3: Application to Timescale Mismatch Data

Empirical Data

The EMA data from the UT1000 project (Wu et al., 2021), a multimodal data collection study conducted at the University of Texas at Austin, were used to show the differences in parameter estimation results between different modeling methods. We first selected data from October 10, 2018, to October 31, 2018, based on daily subject participation rates. Then, we removed individuals with low compliance rates to avoid a high percentage of missingness. Finally, data from 381 individuals for 22 consecutive days were used for the analyses. In this study, we focused on the dynamic bidirectional relation between stress feelings and sleep quality, two variables with mismatched timescales. Participants reported their feelings of stress at 9 a.m., 12 p.m., 15 p.m., and 18 p.m. each day (i.e., "I am feeling stress"), from 1 = *not at all* to 4 = *very much*, and reported their sleep quality the previous night each morning ("How restful was your sleep?") from 1 = *not at all restful* to 4 = *very restful*.

We demonstrated the two types of research questions that may be of interest in empirical studies on timescale mismatch variables. On the one hand, to investigate the dynamic interaction processes between stress feelings and sleep quality, we fitted the data to the partial-path model and the full-path model. On the other hand, to examine the overall reciprocal effects between stress feelings and sleep quality, we used the average-score model, the factor model, and the adjusted factor model. We used Bayesian estimation and the Markov chain Monte Carlo algorithm in Mplus 8.10. Two Gibbs-sampler chains were used, each with a minimum number of iterations of 5,000. Models with the PSR less than 1.1 for all parameters were considered converged. Data and Mplus syntax are available at https://osf.io/fdpsj.

Results

The results of the unstandardized parameter estimation for the two types of research questions and the corresponding models are presented in Tables 4 and 5, respectively. As shown in Table 4, stress feelings and sleep quality had significant positive autoregressive effects at different time intervals. Both models showed that better sleep quality predicted less stress feelings in the morning, and more stress feelings in the evening predicted lower sleep quality the next day. However, the cross-lagged effects estimated in the partial-path model ($c_{01} = -0.147[-0.177, 0.118]$; $c_{40} = -0.099[-0.131, -0.066]$) and the full-path model ($c_{01} = -0.148[-0.179, -0.118]$; $c_{40} = -0.070[-0.110, -0.030]$) were different. The partial-path model exaggerated the cross-lagged effect of stress feelings in the evening on sleep quality. Moreover, the full-path model indicated that there were two additional significant cross-lagged effects of sleep quality on stress feelings, and one significant cross-lagged effect of stress feelings on sleep quality that day. Specifically, there were negative cross-lagged effects between sleep quality and the third measure of stress feelings each day (i.e., at 15 p.m.), suggesting a dynamic bidirectional relation between sleep and stress feelings in the afternoon. Additionally, sleep quality had a weaker lagged effect on the fourth measures of stress feelings each day (i.e., at 18 p.m.), suggesting that sleep quality might even mitigate stress levels later in the next day. These results illustrate the importance of using the full-path model to better understand the dynamic processes between timescale mismatched variables, as it not only helps to accurately estimate the dynamic interaction between sleep quality and stress feelings at the closest time points, but also reveals time-specific cross-lagged effects between them.

In Table 5, we found significant positive autoregressive effects of stress feelings and sleep quality and significant negative cross-lagged effects between stress feelings and sleep quality across days. However, the autoregressive effect of stress feelings was

Table 4

Unstandardized Parameter Estimates of the Bidirectional Relation Between Stress Feelings and Sleep Quality Based on Partial-Path Model and Full-Path Model

	Partial-path model		Full-path model	
Parameter	*B*	95% CI	*B*	95% CI
a_{00}	0.078	[0.049, 0.109]	0.070	[0.039, 0.100]
a_{12}	0.538	[0.506, 0.569]	0.537	[0.506, 0.569]
a_{23}	0.565	[0.533, 0.597]	0.559	[0.527, 0.592]
a_{34}	0.554	[0.522, 0.586]	0.551	[0.518, 0.584]
a_{41}	0.391	[0.358, 0.424]	0.391	[0.357, 0.424]
c_{01}	−0.147	[−0.177, 0.118]	−0.148	[−0.179, −0.118]
c_{02}			−0.003	[−0.027, 0.022]
c_{03}			−0.049	[−0.072, −0.025]
c_{04}			−0.026	[−0.051, −0.001]
c_{10}			−0.007	[−0.049, 0.033]
c_{20}			−0.009	[−0.052, 0.036]
c_{30}			−0.045	[−0.086, −0.002]
c_{40}	−0.099	[−0.131, −0.066]	−0.070	[−0.110, −0.030]

Note. CI = confidence interval; a_{00} = autoregressive effect of sleep quality at a 24-hr time interval; a_{12}, a_{23}, a_{34}, and a_{41} = autoregressive effects of stress feelings at a 3-hr time interval; c_{01}, c_{02}, c_{03}, and c_{04} = cross-lagged effects of sleep quality on stress feelings; c_{10}, c_{20}, c_{30}, and c_{40} = cross-lagged effects of stress feelings on sleep quality.

Table 5

Unstandardized Parameter Estimates of the Bidirectional Relation Between Stress Feelings and Sleep Quality Based on Average-Score Model, Factor Model, and Adjusted Factor Model

	Average-score model		Factor model		Adjusted factor model	
Parameter	*B*	95% CI	*B*	95%CI	*B*	95% CI
a_{12}					0.197	[0.148, 0.246]
a_{23}					0.078	[0.010, 0.140]
a_{34}					0.113	[0.036, 0.183]
a_{41}					0.178	[0.142, 0.214]
a_{XX}	0.435	[0.402, 0.466]	0.575	[0.540, 0.609]	0.551	[0.510, 0.590]
a_{YY}	0.073	[0.043, 0.103]	0.073	[0.040, 0.106]	0.071	[0.039, 0.108]
c_{XY}	−0.111	[−0.148, −0.076]	−0.143	[−0.193, −0.094]	−0.191	[−0.259, −0.128]
c_{YX}	−0.080	[−0.099, −0.060]	−0.070	[−0.093, −0.050]	−0.052	[−0.073, −0.034]
l_2			1.076	[1.025, 1.129]	1.026	[0.928, 1.129]
l_3			1.078	[1.078, 1.138]	1.260	[1.098, 1.423]
l_4			0.987	[0.933, 1.047]	1.090	[0.913, 1.299]

Note. CI = confidence interval; a_{12}, a_{23}, a_{34}, and a_{41} = within-day regression effects of stress feelings; a_{XX} and a_{YY} = across-day autoregressive effects of stress feelings and sleep quality, respectively; c_{XY} and c_{YX} = cross-lagged effects between stress feelings and sleep quality; l_2, l_3, and l_4 = loadings of stress feelings at different time points within a day on the latent factor of stress feelings of that day.

stronger in the factor and adjusted factor models compared to the average-score model. Furthermore, the negative cross-lagged effect of stress feelings on sleep quality became stronger, and the negative cross-lagged effect of sleep quality on stress feelings became weaker from the average model, the factor model, to the adjusted factor model. Since the adjusted factor model showed small to moderate regression effects between stress feelings at different time points each day, the results from the adjusted factor model were more reliable according to the simulation results in Study 2-2. This selection was further supported by examining the estimates of within-person factor loadings of stress feelings at different time points. The factor model showed similar loadings across time points, whereas the adjusted factor model indicated that the third measure of stress feelings each day (i.e., at 15 p.m.) had the strongest loading on the latent state factor of stress feelings. The latter was in accord with the results from the full-path model, which suggested the relatively large contribution of the third time point of stress feelings on the dynamic reciprocal relation between sleep quality and stress feelings. Although this model selection did not alter the main conclusions on the dynamic interplay between sleep and stress (i.e., the significance of autoregressive and cross-lagged effects was not affected by the model choice), the adjusted factor model promotes a deeper understanding of the contribution and temporal dependency between different time points of stress feelings within a day.

Discussion

Despite the rapid development of methods for analyzing ILD, some issues related to ILD analyses in real-world data have not received much attention. Applied researchers have been confronted with the problem of timescale mismatch between variables, but have lacked appropriate methodological guidance to address this issue. In this study, we reviewed the current practices for modeling variables with mismatched timescales (i.e., the partial-path model, and the average-score model), examined their limitations, and proposed possible solutions (i.e., the full-path model, the factor model, and the adjusted factor model) to timescale mismatch based on DSEMs. Through three simulation studies, we showed the problems in the existing approaches and tested the effectiveness of the proposed approaches under various conditions. Our preliminary explorations into the problem of timescale mismatch encourage more attention and examination of this practical problem in ILD and provide feasible solutions for applied researchers.

Dynamic Interaction Processes of Timescale Mismatched Variables

One of the main goals in studies collecting data with timescale mismatched variables is to understand the detailed processes of the dynamic interaction between variables. To this end, previous studies have used separate multilevel models for different time lags between variables with mismatch timescales. However, this modeling approach did not consider the temporal relations between the different time points of the variable with a denser timescale and may not accurately describe the dynamic interactions between variables. Based on a more integrated modeling framework (i.e., DSEM), other researchers have constructed the partial-path model, which considered the autoregressive effects between different time points of the variable with a denser timescale, but estimated only the cross-lagged effect between the closest time points between timescale mismatched variables (Lücke et al., 2023). For this partial-path model, our simulation results suggested that it underestimated the cross-lagged effect from the variable with a sparser timescale to the variable with a denser timescale, and overestimated the cross-lagged effect on the other direction. In contrast, the full-path model proposed in this study considered the cross-lagged effects of all time points of the variable with a denser timescale with the variable with a sparser timescale, which better reflected the processes of their dynamic interactions. More importantly, the full-path model examined the unique effect of different time points, which helped to reveal time-specific effects in the dynamic interplay between timescale mismatch variables. Therefore, we suggest that applied researchers adopt the full-path model rather than the partial-path model to better examine the dynamic interaction processes of timescale mismatched variables.

Dynamic Reciprocal Effects Between Timescale Mismatched Variables

The other research question of great interest in studies on timescale mismatched variables is the dynamic reciprocal effects between variables. Since variables with mismatched timescales may not be easily fitted to models commonly used in ILD analyses, an intuitive approach used in previous studies is aggregating variables with denser timescales to transform the variables to sparser timescales (Langener et al., 2024; Seizer et al., 2024), and then fit an average-score model (Neubauer et al., 2021). However, our simulation study found that the average-score model underestimated the autoregressive effects between variables. Even worse, it overestimated cross-lagged effects in one direction and underestimated them in the other. Since researchers may infer the casual relations (i.e., Granger causal relations; Granger, 1969) between variables based on the comparison of their dynamic reciprocal effects, the biased estimation of cross-lagged effects in the average-score model may alter the results of the comparison of cross-lagged effects and mislead researchers of the causal dominance between variables.

In this study, we propose two other modeling approaches (i.e., the factor model, and the adjusted factor model) to match the timescales of variables by constructing a latent state factor for the variable with a denser timescale. The simulation results suggested that the factor model had higher parameter estimation accuracy than the average-score model. In addition, both the factor model and the adjusted factor model accurately estimated the autoregressive and cross-lagged effects between variables with mismatched timescales when there were no or small regression effects between different time points of variables with denser timescales. When these regression effects became larger (e.g., 0.3), the factor model estimated the autoregressive and cross-lagged effects less accurately, while the adjusted factor model still performed well. Therefore, it is recommended to construct an adjusted factor model to examine the dynamic reciprocal effects between timescale mismatched variables.

However, it should be noted that in our simulation study, the convergence rate of the adjusted factor model was relatively lower than that of the factor model. Moreover, it can be expected that the adjusted factor model is more difficult to converge than the factor model in empirical studies. Considering the relatively small bias of the factor model in estimating autoregressive and cross-lagged effects, we believe that the factor model is an acceptable and practical approach for investigating the bidirectional relations between variables with mismatched timescales, especially when the adjusted factor model fails to converge. Still, it should be cautioned that the within-person factor loadings estimated in the factor model should not be interpreted as contributions from different time points of variables with denser timescales, as our simulation results showed that the bias in their estimation increased as the regression effects between different time points of variables with denser timescales increased.

In addition, the impacts of sample size, number of time points, and their interplay on model performance should be noted. Most conditions (except $N = 200$, $T = 7$; or $N = 300$, $T = 7$) produced results similar to the reference condition ($N = 200$ and $T = 15$). This suggests that our main findings are robust across varying combinations of sample sizes and number of time points. However, conditions with fewer time points ($T = 7$) showed greater biases in parameter estimation, as well as lower power and model convergence compared to the reference condition, regardless of sample size ($N = 200$ or 300). This indicates that this number of time points may be insufficient to explore the reciprocal effects between variables with mismatched timescales and that a larger sample size may not be able to compensate for such a small number of time points.

Contributions and Implications

Researchers have emphasized the importance of selecting appropriate measurement timescales to effectively capture dynamic processes of various variables in intensive longitudinal studies (Adolf et al., 2021; Batra et al., 2023; Langener et al., 2024). However, most previous studies measured all variables using the same timescale (Luo et al., 2024), which may be due to the lack of methodological guidance for analyzing the dynamic relations between variables with mismatched timescales. This approach can lead to several problems. On the one hand, using a sparser timescale to assess all variables may not accurately capture fluctuations in some frequently changing variables (e.g., affective states). Moreover, employing measurement intervals that are sparser than the true intervals at which the processes operate can result in highly biased autoregressive estimates (Batra et al., 2023). On the other hand, assessing all variables with a denser timescale may be unnecessary or even impractical, as it may impose an excessive burden on participants, which reduces their compliance (Vachon et al., 2019) and data quality. Therefore, it is recommended to choose appropriate timescales for different variables based on the characteristics of their dynamic processes.

Nevertheless, it should be recognized that researchers do not always have a precise understanding of the true underlying process, and that the measurement timescales used to collect data do not necessarily reflect the theoretical timescales of the processes of interest. To address this issue, recent studies have offered some valuable suggestions. First, researchers can select measurement timescales guided by relevant theoretical and conceptual considerations (Seizer et al., 2024; Velozo et al., 2024), as well as previous studies or reviews that provide evidence and recommendations on the sampling timescales of the same variable (e.g., heart rate variability; Shaffer & Ginsberg, 2017). Second, in cases where the choice of measurement timescales is unclear and multiple timescales are plausible, researchers can consider conducting a multiverse analysis (Steegen et al., 2016) to explore the results from different timescales and assess the robustness of their findings (Langener et al., 2024). Third, when theoretical knowledge is very limited, researchers can use the method proposed by Adolf et al. (2021) to estimate an optimal sampling interval that leads to highest estimation reliability (i.e., minimal standard errors) based on previous estimates of the autoregressive effects. By integrating theoretical and empirical knowledge and utilizing relevant statistical methods, researchers can determine appropriate measurement timescales that better reflect theoretical timescales and gain more confidence in further analyses of the dynamic interplay between variables with mismatched timescale.

In addition, recent technology advances have opened up new opportunities to more accurately and objectively measure the dynamics of everyday experiences. Fine-grained and multimodal ILD collected by wearable trackers, sensors, and smartphones are able to capture fluctuations in state variables on denser timescales, which may help us deepen our understanding of the dynamic processes and interplay mechanisms between variables. However, previous studies have typically examined these data at the aggregate level and have not fully exploited the potential of these valuable data (Kim et al., 2018; Schick et al., 2023). In addition, studies that used appropriate statistical methods to synthesize data from subjective reports and objective measures and explore dynamic bidirectional relations between variables are still limited. Technology developments call for effective methods to better integrate and analyze data on variables with mismatched timescales.

Focusing on two types of research questions explored by previous studies with timescale mismatched variables (i.e., dynamic interaction processes and overall reciprocal effects between variables), we demonstrated the limitations of current practices and tested the effectiveness of the proposed approaches through three simulation studies. Based on the findings, we provide feasible methods and practical suggestions for applied researchers who collect data of variables with mismatched timescales. This could not only facilitate the use of appropriate timescales for collecting and analyzing data according to the characteristics of different variables but also the wider application of objective measures in intensive longitudinal studies and their combination with subjective reports.

Limitations and Future Directions

The study has limitations that need to be noted. First, this study examined a relatively small degree of timescale mismatch between the two variables (from 3:1 to 5:1). The reason for setting these ratios of timescale mismatch between variables is because it is common in EMA studies to take measurements three to six times per day (Luo et al., 2024). However, for variables with large differences in timescale densities (e.g., one variable is 10 times denser than the other), other suitable solutions may be needed, especially if one of the variables is based on objective measurements using digital devices. One possible solution is to construct a three-level model. For example, Kanning and Schoebi (2016) examined the dynamic associations between affective states and physical activity with activity by constructing a three-level multilevel model, with physical activity assessed at 5-min time intervals (Level 1) being nested within momentary affects assessed at 45-min time intervals (Level 2), which were nested within persons (Level 3). However, it remains to be explored how to construct a three-level model under the DSEM framework, where Level 1 (e.g., within-day level) includes the autoregressive process of the variable with a denser timescale, Level 2 (e.g., between-day level) focuses on the dynamic bidirectional relation between the overall level of the denser variable and the variable with a sparser timescale, and Level 3 (e.g., between-person level) considers the between-person association of the two variables.

In addition, the generalizability of our findings may be limited to the simulation conditions considered in this study. In the three simulation studies, we mainly varied the values of critical parameters, the sample size, the number of time points per subject, the timescale mismatch ratio, and the proportion of missing data to explore whether our findings held in different situations. For example, regarding the proportion of missing data, while we have confirmed the robustness of our findings with relatively small proportions of missing data (e.g., 20% and 40%), it remains unclear whether our conclusions still hold under conditions with higher proportions of missing data (e.g., 60%) and different missing mechanisms (i.e., missing at random, and missing not at random). Given the considerable time and computational resources required for these simulations, examining the impacts of all potentially relevant factors may not be feasible. Nevertheless, real-life studies may be more diverse than the conditions we set, and other potentially influencing factors may need to be further considered in future studies on timescale mismatch.

References

Adolf, J. K., Loossens, T., Tuerlinckx, F., & Ceulemans, E. (2021). Optimal sampling rates for reliable continuous-time first-order autoregressive and vector autoregressive modeling. *Psychological Methods*, *26*(6), 701–718. https://doi.org/10.1037/met0000398

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 359–388. https://doi.org/10.1080/10705511.2017.1406803

Asparouhov, T., & Muthén, B. (2024). Continuous time dynamic structural equation models. https://statmodel.com/download/CTRDSEM.pdf

Batra, R., Johal, S. K., Chen, M., & Ferrer, E. (2023). Consequences of sampling frequency on the estimated dynamics of AR processes using continuous-time models. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000595

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, *54*(1), 579–616. https://doi.org/10.1146/annurev.psych.54.101601.145030

Bolger, N., & Laurenceau, J. P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. Guilford Press.

Driver, C. C., Oud, J. H., & Voelkle, M. C. (2017). Continuous time structural equation modeling with R package *ctsem*. *Journal of Statistical Software*, *77*(5), 1–35. https://doi.org/10.18637/jss.v077.i05

Fan, X. (2003). Two approaches for correcting correlation attenuation caused by measurement error: Implications for research practice. *Educational and Psychological Measurement*, *63*(6), 915–930. https://doi.org/10.1177/0013164403251319

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, *37*(3), 424–438. https://doi.org/10.2307/1912791

Hallquist, M. N., & Wiley, J. F. (2018). Mplusautomation: An R package for facilitating large-scale latent variable analyses in M*plus*. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(4), 621–638. https://doi.org/10.1080/10705511.2017.1402334

Hamaker, E. L., & Wichers, M. (2017). No time like the present: Discovering the hidden dynamics in intensive longitudinal data. *Current Directions in Psychological Science*, *26*(1), 10–15. https://doi.org/10.1177/0963721416666518

Kanning, M. K., & Schoebi, D. (2016). Momentary affective states are associated with momentary volume, prospective trends, and fluctuation of daily physical activity. *Frontiers in Psychology*, *7*, Article 744. https://doi.org/10.3389/fpsyg.2016.00744

Kim, J., Marcusson-Clavertz, D., Togo, F., & Park, H. (2018). A practical guide to analyzing time-varying associations between physical activity and affect using multilevel modeling. *Computational and Mathematical Methods in Medicine*, *2018*, Article 8652034. https://doi.org/10.1155/2018/8652034

Kuiper, R. M., & Ryan, O. (2018). Drawing conclusions from cross-lagged relations: Re-considering the role of the time-interval. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(5), 809–823. https://doi.org/10.1080/10705511.2018.1431046

Langener, A. M., Stulp, G., Jacobson, N. C., Costanzo, A., Jagesar, R. R., Kas, M. J., & Bringmann, L. F. (2024). It's all about timing: Exploring different temporal resolutions for analyzing digital-phenotyping data. *Advances in Methods and Practices in Psychological Science*, *7*(1), Article 25152459231202677. https://doi.org/10.1177/25152459231202677

Liao, Y., Chou, C. P., Huh, J., Leventhal, A., & Dunton, G. (2017). Examining acute bi-directional relationships between affect, physical feeling states, and physical activity in free-living situations using electronic ecological momentary assessment. *Journal of Behavioral Medicine*, *40*(3), 445–457. https://doi.org/10.1007/s10865-016-9808-9

Lücke, A. J., Wrzus, C., Gerstorf, D., Kunzmann, U., Katzorreck, M., Hoppmann, C., & Schilling, O. K. (2023). Bidirectional links of daily sleep quality and duration with pain and self-rated health in older adults' daily lives. *The Journals of Gerontology: Series A*, *78*(10), 1887–1896. https://doi.org/10.1093/gerona/glac192

Luo, X. (2024, August 26). *Timescale Mismatch in Intensive Longitudinal Data*. https://osf.io/fdpsj

Luo, X., & Hu, Y. (2024). Temporal misalignment in intensive longitudinal data: Consequences and solutions based on dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *31*(1), 118–131. https://doi.org/10.1080/10705511.2023.2207749

Luo, X., Liu, H., & Hu, Y. (2024). From cross-lagged effects to feedback effects: Further insights into the estimation and interpretation of bidirectional relations. *Behavior Research Methods*, *56*(4), 3685–3705. https://doi.org/10.3758/s13428-023-02304-0

Maes, I., Mertens, L., Poppe, L., Vetrovsky, T., Crombez, G., De Backere, F., & Brondeel, R. (2023). Within-person associations of accelerometer-assessed physical activity with time-varying determinants in older adults: Time-based ecological momentary assessment study. *JMIR Aging*, *6*, Article e44425. https://doi.org/10.2196/44425

McNeish, D., & Hamaker, E. L. (2020). A primer on two-level dynamic structural equation models for intensive longitudinal data in M*plus*. *Psychological Methods*, *25*(5), 610–635. https://doi.org/10.1037/met0000250

McNeish, D., & Wolf, M. G. (2020). Thinking twice about sum scores. *Behavior Research Methods*, *52*(6), 2287–2305. https://doi.org/10.3758/s13428-020-01398-0

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.).

Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., Estabrook, R., Bates, T. C., Maes, H. H., & Boker, S. M. (2016). Openmx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, *81*(2), 535–549. https://doi.org/10.1007/s11336-014-9435-8

Neubauer, A. B., Kramer, A. C., Schmidt, A., Könen, T., Dirk, J., & Schmiedek, F. (2021). Reciprocal relations of subjective sleep quality and affective well-being in late childhood. *Developmental Psychology*, *57*(8), 1372–1386. https://doi.org/10.1037/dev0001209

Oh, H., & Jahng, S. (2023). Incorporating measurement error in the dynamic structural equation modeling using a single indicator or multiple indicators. *Structural Equation Modeling: A Multidisciplinary Journal*, *30*(3), 501–514. https://doi.org/10.1080/10705511.2022.2103703

Oravecz, Z., Tuerlinckx, F., & Vandekerckhove, J. (2011). A hierarchical latent stochastic differential equation model for affective dynamics. *Psychological Methods*, *16*(4), 468–490. https://doi.org/10.1037/a0024375

Ou, L., Hunter, M. D., & Chow, S. M. (2019). What's for *dynr*: A package for linear and nonlinear dynamic modeling in R. *The R Journal*, *11*(1), 91–111. https://doi.org/10.32614/RJ-2019-012

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org

Schick, A., Rauschenberg, C., Ader, L., Daemen, M., Wieland, L. M., Paetzold, I., Postma, M. R., Schulte-Strathaus, J. C. C., & Reininghaus, U. (2023). Novel digital methods for gathering intensive time series data in mental health research: Scoping review of a rapidly evolving field. *Psychological Medicine*, *53*(1), 55–65. https://doi.org/10.1017/S0033291722003336

Schuurman, N. K., & Hamaker, E. L. (2019). Measurement error and person-specific reliability in multilevel autoregressive modeling. *Psychological Methods*, *24*(1), 70–91. https://doi.org/10.1037/met0000188

Seizer, L., Schiepek, G., Cornelissen, G., & Löchner, J. (2024). A primer on sampling rates of ambulatory assessments. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000656

Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in Public Health*, *5*(1), Article 258. https://doi.org/10.3389/fpubh.2017.00258

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, *4*(1), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, *11*(5), 702–712. https://doi.org/10.1177/1745691616658637

Trull, T. J., & Ebner-Priemer, U. (2013). Ambulatory assessment. *Annual Review of Clinical Psychology*, *9*(1), 151–176. https://doi.org/10.1146/annurev-clinpsy-050212-185510

Vachon, H., Viechtbauer, W., Rintala, A., & Myin-Germeys, I. (2019). Compliance and retention with the experience sampling method over the continuum of severe mental disorders: Meta-analysis and recommendations. *Journal of Medical Internet Research*, *21*(12), Article e14475. https://doi.org/10.2196/14475

Velozo, J. D. C., Habets, J., George, S. V., Niemeijer, K., Minaeva, O., Hagemann, N., Herff, C., Kuppens, P., Rintala, A., Vaessen, T., Riese, H., & Delespaul, P. (2024). Designing daily-life research combining experience sampling method with parallel data. *Psychological Medicine*, *54*(1), 98–107. https://doi.org/10.1017/S0033291722002367

Voelkle, M. C., & Oud, J. H. (2015). Relating latent change score and continuous time models. *Structural Equation Modeling: A Multidisciplinary Journal*, *22*(3), 366–381. https://doi.org/10.1080/10705511.2014.935918

Voelkle, M. C., Oud, J. H., Davidov, E., & Schmidt, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. *Psychological Methods*, *17*(2), 176–192. https://doi.org/10.1037/a0027543

Wickham, H. (2011). ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*, *3*(2), 180–185. https://doi.org/10.1002/wics.147

Wu, C., Fritz, H., Bastami, S., Maestre, J. P., Thomaz, E., Julien, C., Castelli, D. M., de Barbaro, K., Bearman, S. K., Harari, G. M., Cameron Craddock, R., Kinney, K. A., Gosling, S. D., Schnyer, D. M., & Nagy, Z. (2021). Multi-modal data collection for measuring health, behavior, and living environment of large-scale participant cohorts. *GigaScience*, *10*(6), Article giab044. https://doi.org/10.1093/gigascience/giab044

Zhou, L., Wang, M., & Zhang, Z. (2021). Intensive longitudinal data analyses with dynamic structural equation modeling. *Organizational Research Methods*, *24*(2), 219–250. https://doi.org/10.1177/1094428119833164

Received April 15, 2024
Revision received January 2, 2025
Accepted January 29, 2025 ■