### Title

On Measuring and Modeling Physiological Synchrony in Dyads




### Journal

Multivariate Behavioral Research, 53(4)

### ISSN

0027-3171



### Authors

Helm, Jonathan L
Miller, Jonas G
Kahle, Sarah
et al.



### Publication Date

2018-07-04

### DOI

10.1080/00273171.2018.1459292

Peer reviewed

Multivariate Behavioral Research

ISSN: 0027-3171 (Print) 1532-7906 (Online) Journal homepage: http://www.tandfonline.com/loi/hmbr20

# On Measuring and Modeling Physiological Synchrony in Dyads

Jonathan Lee Helm, Jonas G. Miller, Sarah Kahle, Natalie R. Troxel & Paul D. Hastings

**To cite this article:** Jonathan Lee Helm, Jonas G. Miller, Sarah Kahle, Natalie R. Troxel & Paul D. Hastings (2018): On Measuring and Modeling Physiological Synchrony in Dyads, Multivariate Behavioral Research, DOI: 10.1080/00273171.2018.1459292

**To link to this article:** https://doi.org/10.1080/00273171.2018.1459292

Published online: 23 Apr 2018.

MULTIVARIATE BEHAVIORAL RESEARCH
2018, VOL. 0, NO. 0, 1–23
https://doi.org/10.1080/00273171.2018.1459292

Routledge
Taylor & Francis Group



# On Measuring and Modeling Physiological Synchrony in Dyads

Jonathan Lee Helm, Jonas G. Miller, Sarah Kahle, Natalie R. Troxel, and Paul D. Hastings

University of California Davis, Davis, California, USA

**ABSTRACT**

Physiological synchrony within a dyad, or the degree of temporal correspondence between two individuals' physiological systems, has become a focal area of psychological research. Multiple methods have been used for measuring and modeling physiological synchrony. Each method extracts and analyzes different *types* of physiological synchrony, where 'type' refers to a specific manner through which two different physiological signals may correlate. Yet, to our knowledge, there is no documentation of the different methods, how each method corresponds to a specific type of synchrony, and the statistical assumptions embedded within each method. Hence, this article outlines several approaches for measuring and modeling physiological synchrony, connects each type of synchrony to a specific method, and identifies the assumptions that need to be satisfied for each method to appropriately extract each type of synchrony. Furthermore, this article demonstrates how to test for between-dyad differences of synchrony via inclusion of dyad-level (i.e., time-invariant) covariates. Finally, we complement each method with an empirical demonstration, as well as online supplemental material that contains M*plus* code.

**KEYWORDS**

Physiological synchrony; multivariate growth models

Intra-dyad physiological synchrony, or temporal correspondence of two individuals' physiological systems, has become a focal area of psychological research (Timmons, Margolin, & Saxbe, 2015; Palumbo et al. 2017). Broadly speaking, researchers interested in studying physiological synchrony seek to identify how one partner's physiological activity relates to the other's. More specifically, researchers have reported intra-dyad synchrony of different physiological signals (e.g., cortisol, heart rate (HR)) within different contexts (e.g., laboratory tasks, naturalistic settings), and have identified both antecedents (e.g., touching, vocal conversations) and consequences (e.g., development of self-regulation, future relationship dissolution) of physiological synchrony (Feldman, Gordon, & Zagoory-Sharon, 2011; Ferrer & Helm, 2013; Helm, Sbarra, & Ferrer, 2012, 2014; Liu, Rovine, Klein, & Almeida, 2013; Papp, Pendry, & Adam, 2009; Waters, West, & Mendes, 2014).

A variety of methods can measure and model physiological synchrony. These include examining the correspondence between trends (e.g., Liu et al., 2013); correlating partners' physiological responses, and then predicting correlations with dyad-level covariates (e.g., Chatel-Goldman, Congedo, Jutten, & Schwartz, 2014); using multilevel models to predict one partner's raw physiological response from the other's (e.g., Papp et al., 2009); using bivariate lagged models to examine the degree to which one partner's response predicts the other partner's future response (e.g., Helm et al., 2012; 2014); and many others.

Importantly, each method extracts a different *type* of physiological synchrony, or a specific mechanism through which two signals covary (more details in the *Different Types of Synchrony* section below). Yet, research reports rarely denote how the applied method conforms to a specific type of synchrony, what can/cannot be inferred from the type, or how that type differs from other types. Thus, readers may not correctly interpret results, nor be able to fully integrate a set of results across studies. This article helps mitigate these potential misunderstandings by providing conceptual definitions of three types of synchrony commonly investigated in the literature (trend, concurrent, and lagged synchrony), and connecting each type to its respective method.

Furthermore, this article extends each method to predict between-dyad differences in synchrony using time-invariant variables (e.g., biological sex, attachment security). Such extensions enable the investigation of potential causes and consequences of physiological synchrony (e.g., boys show greater synchrony with their mothers than girls). In principle, the methods described here can be extended to include time-varying variables

**CONTACT** Jonathan Lee Helm jhelm@sdsu.edu 153 Life Sciences, 5500 Campanile Drive, San Diego, California 92183, USA.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/r/HMBR.

(e.g., moment-to-moment behavior codes), but we do not discuss these extensions in detail.

Finally, this article provides empirical examples, as well as an online supplemental material with M*plus* code, for each of the methods for detecting physiological synchrony (we assume the reader has a basic understanding of growth modeling; for more information, see Bollen & Curran, 2006; Grimm, Ram, & Estabrook, 2016; Ram & Grimm, 2015).

We limit this article to methods best-suited for shorter physiological series. Shorter series typically contain more dyads (at least 30, hopefully more than 50; see *Discussion* section for issues surrounding power) than physiological measurements per dyad (between 5 and 20). More intensive physiological series – which have many more observations per dyad than total number of dyads (e.g., raw electrodermal activity) – may have different features of theoretical importance, and may therefore require different methods for extracting and modeling physiological synchrony (e.g., dynamical systems models, spectral analysis). This article describes methods for more intensive data in the *Discussion* section; weighing the relative benefits and drawbacks of those methods versus the three focal methods of this article (and directing interested readers to the relevant literature).

We further restrict the methods described in this tutorial to those designed for distinguishable dyads. Distinguishable dyads contain categorically different members, which form two groups across dyads (Kenny, Kashy, & Cook, 2006, pp. 6–7). Examples include mothers and children, or teachers and students. Alternatively, indistinguishable dyads cannot be separated into groups (e.g., identical twins, same-sex romantic partners), and the lack of separation necessitates different analytic methods (Griffin & Gonzalez, 1995; Kennyet al., 2006; Olsen & Kenny, 2006; Woody & Sadler, 2005). The *Discussion* section compares methods for indistinguishable dyads to those described in this manuscript, and directs interested readers toward the pertinent literature.

None of the methods provided within this article are novel (Ferrer & McArdle, 2003; Kenny et al., 2006; MacCallum, Kim, Malarkey, & Kiecolt-Glaser, 1997). However, prior texts do not (1) focus on applications of physiological synchrony, (2) emphasize how each of these methods correspond to different types of synchrony, or (3) caution readers against other methods that aggregate (or confound) different sources of synchrony. Therefore, this article represents a refocusing of modern methods for multivariate longitudinal data analysis, teaching readers how to apply these methods within the context of physiological synchrony.

In the next section, this article describes the empirical example that motivated our examination of these different models. The article follows with recommendations for psychophysiological measurement. Then the article describes, identifies, and illustrates the methods for different types of physiological synchrony. The article concludes with recommendations for researchers interested in modeling physiological synchrony (including issues of power and future directions), and a few cautionary notes on implementation of the methods.

## Motivating example

This article stems from our study of the psychophysiological mechanisms that facilitate the development of prosocial behavior in young children (Hastings, 2009; Miller, Kahle, & Hastings, 2015). Here, we describe the study to aid explanation of the different types of synchrony that can be examined across a set of laboratory tasks.

Parent-child physiological synchrony may serve as a precursor for emotion regulation, such that those children who manifest synchronous physiological responses with their parents tend to show greater emotion regulation across development (Feldman, 2003, 2012; Fogel, 1993; Tronick, 1989). More specifically, children learn to regulate their physiology, behaviors, and emotions via *coregulation* (i.e., external modification) from a parent or caregiver, which manifests (partially) through synchronous physiological responses (Feldman, Singer, & Zagoory, 2010; Moore & Calkins, 2004; Moore et al., 2009). Hence, we launched a study to understand the relation between physiological synchrony and children's emotion regulation.

We devised three laboratory tasks (each 5 min long, and varying in challenge) wherein we measured both parents' and 3 ½ year-old children's autonomic physiology, behavior, and emotional expression: the storybook, puzzle, and origami tasks (see Miller, Kahle, Lopez, & Hastings, 2015). In the storybook task (null to low challenge), children listened to their mothers' improvisation of a story from a book with pictures but no words. In the puzzle task (low to modest challenge), children completed a puzzle intended for older children, and mothers were instructed to provide "as much help as you think your child needs." In the origami task (modest to high challenge), mothers were given an instruction sheet and asked to teach their child how to fold paper into an origami puppy face, but mothers were asked not to touch the paper themselves (numerous other tasks were administered as well; for example, see Kahle, Miller, Lopez, & Hastings, 2016).

Electrocardiography and impedance cardiography were collected throughout the tasks, and used to produce estimates of respiratory sinus arrhythmia (RSA),

pre-ejection period (PEP), and HR across a series of consecutive and non-overlapping epochs. Different signals had different epoch lengths: RSA and PEP were calculated in 30 second epochs (total of 10 epochs), and HR in 2 second epochs (total of 150 epochs), as these are common epoch lengths used in the developmental psychophysiology literature. RSA refers to the component of HR variability that is correspondent with normal breathing rate, which serves as a measure of parasympathetic control over the heart (Berntson, Cacioppo, & Quigley, 1993). PEP represents the amount of time (in ms) between ventricular depolarization and the opening of the aortic valve, and serves as a measure of sympathetic control of the heart (with longer periods indicating less sympathetic activity; Sherwood et al., 1990). HR refers to the number of consecutive heart beats per minute, and serves as a gross indicator of overall autonomic arousal due to parasympathetic, sympathetic and other influences on the heart (Eckberg, 1997). For more information regarding the measurement and interpretation of these different metrics of autonomic chronotropy from a developmental perspective, see Dennis, Buss, and Hastings (2012).

The study included 83 mother-child dyads (age for mothers: $M = 36.52$ years, SD = 5.19; age for children: $M = 3.56$ years, SD = 0.12; including 46 girls, 37 boys). Seventy-two of the dyads were from married two-parent families, three from unmarried two-parent families, and eight were single-mother families. Families were mainly Caucasian (73.5%), and upper-middle class (median income between $75,000 and $90,000; range from less than $15,000 to over $120,000).

## Physiological data collection

Prior to describing the methods for measuring and modeling physiological synchrony, we define and provide suggestions for two aspects of psychophysiological data collection: the sampling rate and the temporal unit of analysis. The sampling rate refers to the length of time needed to produce a single observation in the recorded physiological signal (e.g., raw electrocardiogram or cardiac impedance), whereas the temporal unit of analysis refers to the length of time (i.e., epoch length) required to translate the signal into a meaningful psychophysiological variable (e.g., HR, RSA, PEP).

The rate of sampling must be fast enough to provide a signal that can be translated into a meaningful epoch-by-epoch psychophysiological variable. Most hardware for recording physiological signals offer sampling rates between 500 and 1000 samples per second, which is more than sufficient to obtain a clear and translatable physiological signal. Therefore, we suggest following the hardware's default sampling rate to produce physiological signals.

The ideal temporal unit of analysis (i.e., epoch length) may be identified through (a) the amount of time needed to reliably estimate the psychophysiological variable, and (b) the expected rate of change for the psychophysiological response. For (a), researchers must consult prior literature (e.g., Cacioppo, Tassinary, & Bernston, 2007) to identify the minimal amount of time needed to reliably measure the psychophysiological variable from the raw physiological signal. HR, RSA, and PEP can be measured reliably using 2-, 30-, and 30-second epochs, respectively. For (b), researchers must identify the psychophysiological variable's potential rate of change during the experimental manipulation. For example, changes in physiology due to emotion regulation theoretically occur on a second-by-second basis (see, Field, 1994). The ideal temporal unit of analysis is the epoch length that most closely matches the expected change of the psychophysiological variable (i.e., (b)) that is greater than or equal to the minimum length of time needed to reliably estimate the psychophysiological variable (i.e., (a)).

## Different types of physiological synchrony

This section identifies three different types of synchrony (trend, concurrent, and lagged synchrony), and provides conceptual illustrations for extracting each. Importantly, the conceptual examples are not the suggested methods for estimating synchrony; they merely describe the type of synchrony to be extracted from the data.

### *Trend synchrony*

Trend synchrony refers to a shared physiological response trajectory between dyad members. For example, mothers and children manifest trend synchrony if they show similar HR increases across an experimental task (i.e., a common linear trend). Similarly, mother–child dyads exhibit trend synchrony if the dyad members' RSA responses manifest similar rates of decrease toward a stable set point (i.e., a common quadratic trend). Therefore, trend synchrony measures the degree to which dyad members' physiological responses follow a similar pattern of change (e.g., linear or quadratic) across repeated measures (e.g., the epochs of an experimental task).

Conceptually, one could estimate trend synchrony by calculating intercepts and slopes for each member of a dyad, and then correlating the intercepts and slopes across the members of multiple dyads. The intercepts and slopes summarize each individual's pattern of change across repeated measures, and the correlation across dyads reflects the degree to which partners within dyads

share a similar pattern of change. Within the *Different Models for Physiological Synchrony* section below, we translate this concept into a bivariate growth curve model (GCM) that estimates trend synchrony (i.e., correlations among intercepts and slopes) directly from dyads' physiological responses.

It should be noted that trend synchrony reflects the degree of similarity of intercepts and slopes for a sample of dyads, rather than a single dyad (i.e., we must estimate intercepts and slopes across multiple dyads, and then calculate the correlation among these intercepts and slopes). Thus, an examination of between-dyad differences of trend synchrony requires a comparison across different groups of dyads. For example, we could calculate (and statistically compare) the degree of trend synchrony for mother–son versus mother–daughter dyads. In the *Different Models for Physiological Synchrony* section below, we extend the bivariate latent growth curve to compare trend synchrony across different groups of dyads.

### Concurrent synchrony

Concurrent synchrony refers to a common fluctuation around a trend. For example, a mother and child manifest concurrent synchrony when exhibiting simultaneous increases and decreases in HR. Similarly, a mother and child manifest concurrent synchrony when their RSA responses fluctuate in concert across a conversation. Thus, in contrast to trend synchrony (which summarizes a common pattern across repeated measures), concurrent synchrony examines the moment-to-moment link between partners' physiology.

Conceptually, one could estimate concurrent synchrony by removing a trend from each individual's physiological signal, and then correlating the dyad members' detrended responses. Here, trend removal refers to first fitting a polynomial to an individual's series of physiological responses (i.e., the series of responses as a linear or quadratic function of time), and then collecting the residuals (i.e., detrended responses). Trend removal ensures that the residual correlation reflects the degree of common fluctuation around the trend (see *Why Remove Trends?* sub-section below). In the *Different Models for Physiological Synchrony* section below, we translate this concept into a multilevel path model that quantifies concurrent synchrony directly from the detrended physiological responses.

Concurrent synchrony may be separated into two classes: directional versus non-directional concurrent synchrony. Directional concurrent synchrony assumes that one partner's physiological response causes the other's, whereas non-directional concurrent synchrony does not assume a causal direction. Causation may not be inferred based on the modeling procedure alone; the causal assumption of directional concurrent synchrony is merely the researcher's assumption regarding the generation of the data. Nevertheless, the distinction between directional and non-directional assumptions leads to different modeling approaches, which will be covered in more detail within the *Different Models for Physiological Synchrony*, below.

In contrast to trend synchrony, between-dyad differences in concurrent synchrony may be examined at the level of the dyad. For example, a dyad-level covariate (e.g., mothers' self-reported compassionate love for her child) may be used as a predictor of the between-dyad differences of concurrent synchrony. Therefore, in the *Empirical Examples* section below, we allow the multilevel path model to include dyad-level predictors of concurrent synchrony.

### Lagged synchrony

Lagged synchrony refers to the extent to which each dyad member's current physiological response predicts their partner's future physiological response, after removal of a trend. A mother and child manifest lagged synchrony if the mother's current HR predicts her child's HR a few moments later (and/or vice versa). Similarly, a mother and child show lagged synchrony if the mother's current RSA predicts the child's impending RSA (and/or vice versa). In contrast to concurrent synchrony (which measures the degree of common fluctuation within a single moment), lagged synchrony examines how one dyad member's physiology predicts the other's across moments.

Conceptually, one could estimate lagged synchrony by removing a trend from each individual's physiological responses, and then regressing each member's responses on their own and their partner's previous responses (i.e., accounting for within-person stability), akin to the actor-partner interdependence model (APIM; Cook & Kenny, 2005; Kenny et al., 2006). Analogous to concurrent synchrony, trend removal ensures that the lagged relation reflects the degree to which an individual's moment-to-moment fluctuation around his or her trend predicts their partner's moment-to-moment fluctuation around his or her trend, rather than a common trend across partners (see *Why Remove Trends?* sub-section below). In the *Different Models for Physiological Synchrony* section below, we show how to estimate lagged synchrony for a sample of dyads using a bivariate multilevel path model.

Similar to concurrent synchrony, between-dyad differences in lagged synchrony may be investigated at the level of the dyad. Dyad-level covariates (e.g., child's security

of attachment) may predict between-dyad differences of lagged synchrony. Therefore, in the *Different Models for Physiological Synchrony* section below, we allow the bivariate multilevel path model to include dyad-level predictors of lagged synchrony.

### *Why remove trends?*

Concurrent and lagged synchrony require trend removal because estimates of concurrent and lagged synchrony are contaminated by trend synchrony (Curran & Bauer, 2011; Wang & Maxwell, 2015). Consider Figure 1 Panel A, which has raw physiological responses from two individuals (green and blue) with a common linear trend. The green and blue lines could represent dyad members' HRs while sitting in a cramped room, with increases in HR due to the condition of the room. Figure 1 Panel B contains the detrended responses from Panel A; that is, the individuals' HR after removing a linear trend (Panels A and B have different *y*-axis scales). The dashed red line depicts the mean response across repeated measures, which is the same for both signals in each panel. The raw responses (Panel A) have a significant zero-order correlation ($r = .485$, $p < .01$), whereas the detrended responses do not ($r = .005$, $p = .97$); indicating that the significant relations for the raw responses arise from a common trend (neither a common fluctuation around a trend, nor a lead-lag relation). Given that the correlation quantifies the degree to which two variables co-occur above or below their respective means, data with a common trend (and no other relation) will produce a non-zero correlation and lagged correlation (i.e., in Panel A, both responses tend to be below the red dashed line for the first half of the observations, and above the red line after). Therefore, investigation of concurrent and lagged synchrony requires trend removal to be sure that synchrony does not arise from a common trend which may reflect an extraneous influence (e.g., the condition of the room).

## Different models for physiological synchrony

This section identifies the models that correspond to the different types of physiological synchrony. In principle, the models may be estimated through either a multilevel modeling or structural equation modeling framework (Curran, 2003). Thus, we write the models in their mathematical forms (which can be fitted in either framework), and in the *Empirical Examples* section we use structural equation modeling to show how the models may be implemented within a single framework. We also show how to estimate and interpret effect sizes for each method's estimate of physiological synchrony.

### *Estimating trend synchrony*

Trend synchrony examines the degree of synchrony across intraindividual trends. Hence, the method for estimating trend synchrony simultaneously measures intraindividual trends and summarizes their relation. In this section, we will build the model that performs the simultaneous estimation. We show how to estimate the trend for one individual, then extend it to two individuals in a dyad, and finally describe how to simultaneously estimate trends and their corresponding interrelations across a sample of dyads.

To begin, consider the repeated measures of *just one* individual from *one* dyad. Generically, the individual's repeated responses may be labeled $X_{1t}$, where the 1 indicates that the responses are from the first (rather than the

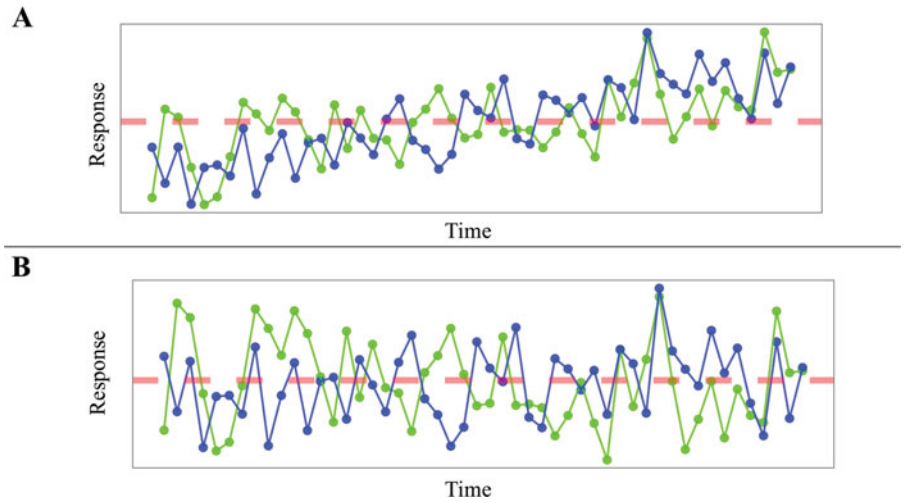


**Figure 1.** Depiction of two signals with a common trend (Panel A), and the same signals detrended (Panel B). The red dashed line refers to the mean of each signal across time. Please note that the *y*-axis is not on the same scale across Panels A and B.

second) member of the dyad, and $t$ refers to the measurement occasion ($t = 1, 2, 3, \ldots, T$; where $T$ is the total number of measurements for the individual). The trend of this individual may be estimated using regression analysis. For example, in

$$X_{1t} = b_0 + \varepsilon_{1t} \tag{1}$$

$b_0$ estimates the individual's average response across repeated measures, and $\varepsilon_{1t}$ refers to the residual at measurement occasion $t$ (i.e., the deviation from the intraindividual average at time $t$). Alternatively, if $time_t$ equals a measurement of time at measurement occasion $t$ (in most cases, $time_t$ equals $t$-1), then for

$$X_{1t} = b_0 + b_1 \text{time}_t + \varepsilon_{1t}, \tag{2}$$

$b_0$ describes the individual's predicted response when $\text{time}_t$ equals zero, and $b_1$ summarizes the individual's linear trend across repeated measures (i.e., the expected change in $X_{1t}$ for a 1-unit change in $time_t$). For both Equations 1 and 2, the model-based parameters (i.e., $b_0$ and $b_1$) describe the individual's trend; Equation 1 summarizes a flat trajectory (using $b_0$), and Equation 2 describes a linear trajectory (using both $b_0$ and $b_1$). Therefore, by predicting an individual's repeated measures (i.e., $X_{1t}$) using functions of time (e.g., '$b_0 + \varepsilon_{1t}$' or '$b_0 + b_1\text{time}_t + \varepsilon_{1t}$') we obtain estimates of the individual's trend. And although we have only shown the intercept-only and linear cases (Equations 1 and 2, respectively), any other function of time may also be used (e.g., quadratic, splines, or exponential) for estimating an individual's trend from their repeated measures (see Ram & Grimm, 2007 for alternative models).

Let us extend the notions embedded in Equations 1–2 to estimating trends for both members of one dyad. In particular, if the second individual in the dyad has responses labeled $X_{2t}$ (where the subscript 2 refers to the second individual in the dyad), then Equation 1 may be extended to a bivariate regression, such that

$$\begin{aligned} X_{1t} &= b_{10} + \varepsilon_{1t} \\ X_{2t} &= b_{20} + \varepsilon_{2t}, \end{aligned} \tag{3}$$

and Equation 2 may be extended to

$$\begin{aligned} X_{1t} &= b_{10} + b_{11}\,\text{time}_t + \varepsilon_{1t} \\ X_{2t} &= b_{20} + b_{21}\,\text{time}_t + \varepsilon_{2t}. \end{aligned} \tag{4}$$

Notice that the subscripts of the model-based coefficients have changed from Equations 1 to 3, and from Equations 2 to 4. The intercepts and slopes now have a leading subscript of 1 or 2 to indicate if they were estimated from partner 1 or 2.

The coefficients from Equations 3 and 4 have similar interpretations to those from Equations 1 and 2, respectively. For Equation 3, $b_{10}$ and $b_{20}$ estimate each partner's average response across repeated measures. And for Equation 4, $b_{10}$ and $b_{20}$ estimate each partner's expected response when $\text{time}_t$ equals 0, whereas $b_{11}$ and $b_{21}$ summarize each partner's linear trend across repeated measures. Thus, by extending the univariate regressions (Equations 1 and 2) to bivariate regressions (Equations 3 and 4), we can obtain estimates of the trend for both members of a dyad (e.g., $b_{10}$, $b_{20}$, $b_{11}$, $b_{21}$).

Next, we will extend the models to simultaneously estimate trends for all members of a sample of dyads, rather than a single dyad. Conceptually, this is similar to estimating bivariate regressions for each dyad separately. In particular, if we have a sample of $D$ dyads, we can write Equations 3 and 4 to have dyad-specific estimates, such that

$$\begin{aligned} X_{1dt} &= b_{10d} + \varepsilon_{1dt} \\ X_{2dt} &= b_{20d} + \varepsilon_{2dt}, \end{aligned} \tag{5}$$

and

$$\begin{aligned} X_{1dt} &= b_{10d} + b_{11d}\,\text{time}_{dt} + \varepsilon_{1dt} \\ X_{2dt} &= b_{20d} + b_{21d}\,\text{time}_{dt} + \varepsilon_{2dt}. \end{aligned} \tag{6}$$

Again, the model-based coefficients include a new subscript; a '$d$' to denote a specific dyad (i.e., the $d$th dyad)[1].

There are two drawbacks to estimating the coefficients of Equations 5 or 6 by separately fitting bivariate regressions to each dyad's repeated measures. First, fitting separate regressions does not produce a parsimonious description of the sample; separate regressions produce a high number of parameter estimates relative to the number of observations (i.e., $D \times 2$ parameter estimates for Equation 5, and $D \times 4$ for Equation 6). Second, the coefficients do not summarize the sample, they characterize each dyad as a separate entity. More specifically, if the goal of the analysis is to describe the sample in a manner that generalizes to other samples, then separate estimates per dyad do not achieve the goal because the dyad-specific coefficients only generalize to their respective dyad. These two drawbacks, along with a few others not described here, represent why separate bivariate regressions are not used in practice. However, the concept underlying separate regressions (i.e., estimation of dyad-specific coefficients) will be carried into the method that solves these drawbacks: GCMs.

[1] As an aside, the growth curve models often contain an estimate of residual covariance (i.e.,$s_{\varepsilon_1,\varepsilon_2}$). This residual covariance can be interpreted as the average level of concurrent synchrony across dyads, but does not vary across dyads (at least not in most structural equation modeling packages).Therefore, it is not the focal method for estimating concurrent synchrony.

GCMs assume that each set of dyad-specific coefficients (e.g., $b_{10d}$ coefficients from $d = 1, 2, 3, \ldots, D$) follows a normal distribution, and that assumption enables the estimation of a parsimonious and generalizable set of results. A normally distributed variable can be fully characterized by its mean and variance. For example, if the $b_{10d}$ coefficients follow a normal distribution (i.e., $b_{10d} \sim N(B_{10},\ s^2_{b_{10}})$), then their mean ($B_{10}$) and variance ($s^2_{b_{10}}$) sufficiently describe $b_{10d}$. Stated differently, if $b_{10d}$ follows a normal distribution, then we only need to estimate $B_{10}$ and $s^2_{b_{10}}$; we do not need estimates of each dyad-specific coefficient. Accordingly, the summary statistics offer a more parsimonious representation of the sample (i.e., two parameter estimates: $B_{10}$ and $s^2_{b_{10}}$) relative to the set of dyad specific coefficients (i.e., $D$ parameter estimates: all $b_{10d}$ coefficients from $d = 1, 2, 3, \ldots, D$). Furthermore, the mean and variance characterize the entire sample in a generalizable manner. For example, for Equation 5, $B_{10}$ describes the expected level for individual 1 across the sample, $s^2_{b_{10}}$ summarizes the spread around that expectation, and those interpretations should generalize to other samples. Thus, GCMs invoke distributional assumptions for the dyad-specific coefficients, and those assumptions lead to a more parsimonious and generalizable set of results.

GCMs also assume that the dyad-specific coefficients jointly follow a multivariate normal distribution, and that assumption provides the estimate of trend synchrony. More specifically, if two (or more) variables follow a multivariate normal distribution, then the relations among those variables may be fully quantified via their covariances. For example, if the dyad-specific coefficients from Equation 5 (i.e., the estimates of $b_{10d}$ and $b_{20d}$ across all dyads) follow a multivariate normal distribution (i.e., $\left[\begin{smallmatrix} b_{10d} \\ b_{20d} \end{smallmatrix}\right] \sim \text{MVN}(\left[\begin{smallmatrix} B_{10} \\ B_{20} \end{smallmatrix}\right],\ \left[\begin{smallmatrix} s^2_{b_{10}} & \\ s_{b_{10},b_{20}} & s^2_{b_{20}} \end{smallmatrix}\right])$), then their relation may be summarized through their covariance ($s_{b_{10},b_{20}}$). Similarly, if the dyad-specific coefficients from Equation 6 follow a multivariate normal distribution, then their relations are summarized through their covariances (i.e., $s_{b_{10},b_{11}}$, $s_{b_{10},b_{20}}$, $s_{b_{10},b_{21}}$, $s_{b_{11},b_{20}}$, $s_{b_{11},b_{21}}$, and $s_{b_{20},b_{21}}$). Thus, GCMs not only summarize dyad-specific coefficients via estimates of means and variances (e.g., $B_{10}$ and $s^2_{b_{10}}$), they also compute the covariances across those estimates (e.g., $s_{b_{10},b_{20}}$).

The covariances across the dyad-specific coefficients measure trend synchrony. For example, the estimate $s_{b_{10},b_{20}}$ that stems from Equation 5 summarizes the extent to which the dyad members' flat trajectories covary (e.g., a positive $s_{b_{10},b_{20}}$ indicates that if one dyad member's flat trajectory is above the mean, the other dyad member's trajectory also tends to be above the mean). Similarly, the estimate $s_{b_{11},b_{21}}$ that stems from Equation 6 summarizes the extent to which dyad members' linear trends covary (e.g., a positive $s_{b_{11},b_{21}}$ indicates that if one dyad member's linear trend is above the mean, the other dyad member's linear trend also tends to be above the mean). It is important to note that not all of the covariances indicate trend synchrony. For example, $s_{b_{10},b_{11}}$ and $s_{b_{20},b_{21}}$ that stem from Equation 6 summarize the extent to which each individual's level when $\text{time}_t$ equals to 0 (e.g., $b_{10d}$) covaries with their own linear slope (e.g., $b_{11d}$). Hence, the covariances that summarize the relation of the dyad-specific coefficients *across* dyad members (e.g., $s_{b_{10},b_{20}}$ and $s_{b_{11},b_{21}}$) may be interpreted as estimates of trend synchrony.

Visualizations of the GCMs aid understanding, and here we present Figures 2A and 2B as illustrations of Equations 5 and 6, respectively. In Figure 2A, the squares

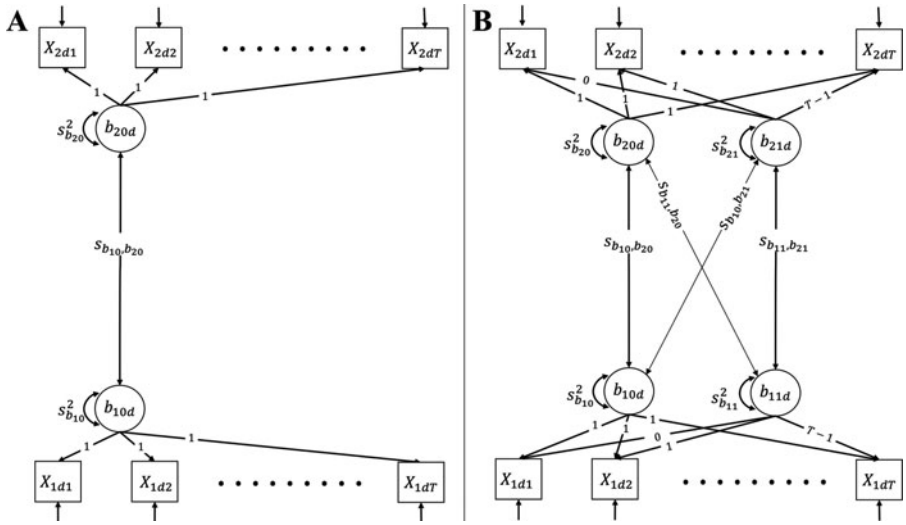


**Figure 2.** Bivariate intercept-only (Panel A) and linear (Panel B) growth curves. Note that, for simplicity, neither latent variable intercepts/means (i.e., $B_{10}$, $B_{11}$, $B_{20}$, $B_{21}$) nor residual covariances (i.e., $s_{\varepsilon_1,\varepsilon_2}$) are depicted.

at the top and bottom depict the physiological responses from individuals in dyad $d$ at measurement occasion $t$. Each of those responses are predicted by the dyad-specific flat trajectory for individual 1 or 2 (i.e., $b_{10d}$ and $b_{20d}$, represented by circles). Importantly, those dyad-specific estimates are not estimated directly (i.e., are latent, not observed), which is why they are represented by circles and not squares. Only their summary statistics are estimated. In particular, their variances ($s^2_{b_{10}}$ and $s^2_{b_{20}}$) and covariance ($s_{b_{10},b_{20}}$) are depicted as double headed arrows. The means of the dyad-specific coefficients (i.e., $B_{10}$ and $B_{20}$) are not depicted in Figure 2A because they are not directly relevant for estimating trend synchrony (but see Ram & Grimm, 2015, for examples of those depictions). Figure 2B extends 2A by introducing a linear trend (i.e., refers to Equation 6 by incorporating $b_{11d}$ and $b_{21d}$), and therefore includes variances for the linear trends (e.g., $s^2_{b_{11}}$) and covariances across all dyad-specific coefficients (e.g., $s_{b_{10},b_{20}}$). These visual representations illustrate how the dyad-specific coefficients relate to each dyad member's physiological responses, and how those coefficients are in turn summarized through sufficient summary statistics.

The magnitude of trend synchrony is difficult to interpret. In particular, the magnitude of a covariance (e.g., $s_{b_{10},b_{20}}$) depends on the magnitude of its underlying variances (e.g., $s^2_{b_{10}}$ and $s^2_{b_{20}}$) and the correlation (e.g., $r_{b_{10},b_{20}}$). Consequently, we recommend estimating trend synchrony in a standardized metric (i.e., correlations rather than covariances), which involves a simple extension of the models in Figures 2A and 2B. In particular, standardized forms of the dyad-specific coefficients (e.g., $z_{10d}$ represents a standardized version of $b_{10d}$) may be computed within the model, and covariances among the standardized coefficients produce correlations (i.e., the covariance between standardized variables equals a correlation). To demonstrate, consider the standardized version of $b_{10d}$, $z_{10d}$, which is commonly calculated as $z_{10d} = \frac{b_{10d} - B_{10}}{s_{b_{10}}}$. That common calculation may be rewritten as $b_{10d} = B_{10} + s_{b_{10}} z_{10d}$, and the rewritten version can be directly incorporated into the GCM. For example, Figure 3A (which extends Figure 2A) contains unobserved variables $z_{10d}$ and $z_{20d}$, which predict $b_{10d}$ and $b_{20d}$ with a weight equal to $s_{b_{10}}$ and $s_{b_{20}}$, respectively (i.e., $b_{10d} = B_{10} + s_{b_{10}} z_{10d}$ and $b_{20d} = B_{20} + s_{b_{20}} z_{20d}$). The variances of $z_{10d}$ and $z_{20d}$ are constrained to equal 1 (i.e., the variances of standardized variables equal 1), and the residual variances of $b_{10d}$ and $b_{20d}$ are now set equal to 0 (i.e., the products $s_{b_{10}} z_{10d}$ and $s_{b_{10}} z_{20d}$ fully account for $b_{10d}$ and $b_{20d}$, respectively). The covariance between $z_{10d}$ and $z_{20d}$ (shown as a double-headed arrow) estimates the correlation between $b_{10d}$ and $b_{20d}$ (labeled $r_{b_{10},\ b_{20}}$). In a similar manner, Figure 3B contains standardized estimates of trend synchrony for the linear case. In particular, by standardizing all dyad-specific coefficients (i.e., $z_{10d}$, $z_{20d}$, $z_{11d}$, and $z_{21d}$), the estimates of trend synchrony are computed as correlations (e.g., $r_{b_{10},b_{20}}$, $r_{b_{11},b_{20}}$, $r_{b_{10},b_{21}}$, and $r_{b_{11},b_{21}}$) rather than covariances. Thus, the respecification in Figure 3A and 3B (relative to Figure 2A and 2B) produces the means, standard deviations, and correlations of the dyad-specific coefficients (as opposed to the means, variances, and covariances); thereby providing a more interpretable estimate of synchrony on a correlation (rather than covariance) metric.

A second benefit of calculating trend synchrony on a standardized metric concerns the reporting of effect sizes. More specifically, the correlation metric may be reported as a measure of effect size (Cohen, 1988, 1992). Ideally,

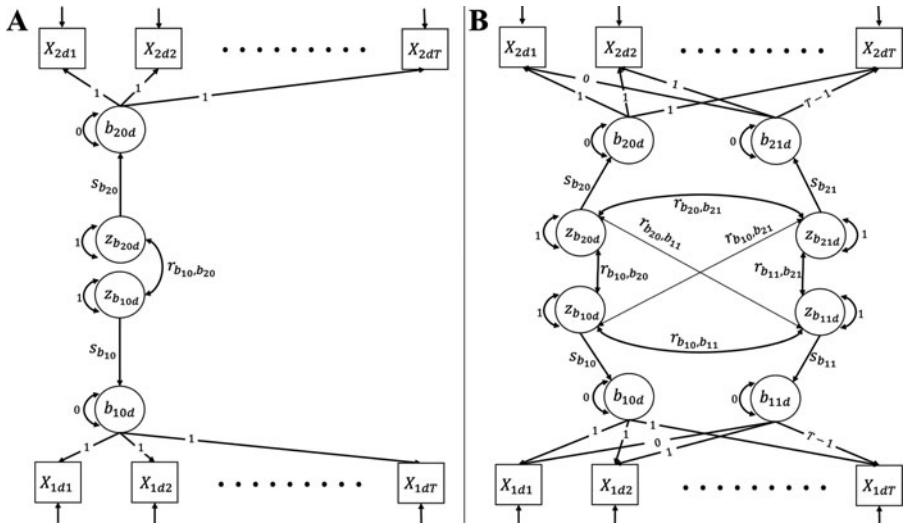


**Figure 3.** Extension of Figure 2 that includes standardization of person-specific coefficients. For simplicity, neither latent variable intercepts/means (i.e., $B_{10}$, $B_{11}$, $B_{20}$, $B_{21}$) nor residual covariances (i.e., $s_{\varepsilon_1,\varepsilon_2}$) are depicted.

the magnitude of an effect size (i.e., small, medium, and large) depends on the specific phenomenon that the effect size describes. However, the literature on physiological synchrony is young, and we find it difficult to confidently define a set (or even a range) of values for small, medium, and large effects. Therefore, we adopt Cohen's (1988, 1992) rules of thumb, and denote correlations of .10, .30, and .50 as small, medium, and large effects, respectively. Importantly, Cohen (1988) noted that a medium effect should be perceptible (when data are plotted) by a knowledgeable observer, small effects are at the lower limit of what might be viewed as theoretically meaningful in the social sciences, and large effects are an equal distance above the medium effect as small effects are below (see Cohen, 1988, pages 77–81 for more detailed explanations of these cut-offs). While recognizing that rules of thumb are never ideal, we hope that these definitions may be useful in interpreting effect sizes until the field moves to a more informed state regarding the expected magnitude of physiological synchrony.

The GCMs described so far produce estimates for a single sample of dyads. Yet, researchers often aim to compare the degree of synchrony across different dyads (i.e., investigate between-dyad differences of physiological synchrony). To make such comparisons, the GCM may be nested within a multiple group framework (Albert, 1994). In particular, if dyads can be separated into two or more different categories (e.g., older versus younger couples, securely versus insecurely attached parents and children), then a multiple group model produces a separate estimate of trend synchrony for each category of dyads which may be statistically compared. We demonstrate this approach within the *Empirical Examples* section below.

The effect size for a difference in trend synchrony across groups may also be computed. Following the same rationale for estimating the effect size of trend synchrony in a sample of dyads, we adopt Cohen's (1988) interpretations of effect size for differences in correlations. Cohen (1988, pp. 109–116) suggests calculating $q$, the difference between correlations on Fisher's z-metric, as the estimated effect size. More specifically, if $r_{\text{group}}$ represents the estimated trend synchrony for a given group, then $z_{\text{group}} = \frac{1}{2}\log_e(\frac{1+r_{\text{group}}}{1-r_{\text{group}}})$, and $q = z_{\text{group 1}} - z_{\text{group 2}}$. Accordingly, Cohen defines $q =$ .10, .30, and .50 as small, medium, and large, respectively. Again, we suggest these general interpretations until the field contains enough information to identify more appropriate (i.e., phenomenon specific) interpretations.

## Estimating concurrent synchrony

Concurrent synchrony measures the degree of common fluctuation across partners' physiological responses, after removing a trend. Therefore, the method begins by removing each individual's trend (e.g., computes residuals around a linear or quadratic trend), and then summarizes the relation between the residuals. As noted in the *Different Types of Physiological Synchrony* section, concurrent synchrony may be modeled as either directional or non-directional. In this section, we build the model for both directional and non-directional concurrent synchrony. We begin with an example of concurrent synchrony for a single dyad, and then describe how to simultaneously estimate concurrent synchrony across a sample of dyads.

Estimation of either directional or non-directional concurrent synchrony begins with trend removal. In particular, for a given member of a dyad, a trend may be removed by fitting a trend (e.g., Equation 1 or 2) directly to the individual's repeated measures, and then saving the residuals (e.g., $\varepsilon_{1dt}$ and $\varepsilon_{2dt}$ from Equation 1 or 2). Given that this tutorial focuses on methods for shorter physiological series, we suggest fitting either linear or quadratic models to each individual's data, and then saving residuals. Then, concurrent synchrony may be estimated by modeling those residuals.

Beginning with a single dyad, directional concurrent synchrony refers to a regression between the partners' responses that conforms to the assumed direction of the effect. For example, if it is assumed that partner 2 affects partner 1, then the data should be modeled as

$$\varepsilon_{1t} = a_1\varepsilon_{2t} + \epsilon_{1t}, \tag{7}$$

whereas if it is assumed that partner 1 affects partner 2, then the data follow

$$\varepsilon_{2t} = a_1\varepsilon_{1t} + \epsilon_{2t}. \tag{8}$$

In Equations 7 and 8, $a_1$ indicates the expected change in the outcome variable for a 1-unit change in the predictor variable; and $\epsilon_{1t}$ and $\epsilon_{2t}$ contain the portion of $\varepsilon_{1t}$ and $\varepsilon_{2t}$ unaccounted for by $\varepsilon_{2t}$ and $\varepsilon_{1t}$, respectively. Thus, for a single dyad, the estimate of $a_1$ summarizes the degree of directional concurrent synchrony within the dyad. Importantly, Equations 7 and 8 do not contain intercepts because the predictor and outcome are residuals after removal of a trend, and therefore the intercept must equal 0.

Extending Equations 7 and 8 to a sample of dyads follows

$$\varepsilon_{1dt} = a_{1d}\varepsilon_{2dt} + \epsilon_{1dt} \tag{9}$$

and

$$\varepsilon_{2dt} = a_{1d}\varepsilon_{1dt} + \epsilon_{2dt}, \tag{10}$$

respectively. Similar to trend synchrony for a sample of dyads, the most intuitive manner to estimate the set of $a_{1d}$ coefficients would be to fit separate regressions to each dyad's physiological responses. However, that approach

would produce a less parsimonious and generalizable set of results (i.e., the same reason why separate regressions are not used for estimating trend synchrony). Therefore, we suggest using a multilevel path analysis (also known as multilevel autoregressive models; Hamaker & Grasman 2015; Jongerling, Laurenceau, & Hamaker, 2015; Liu, 2017; Rovine & Walls, 2006), which introduces a similar assumption as the GCM. In particular, by assuming that the $a_{1d}$ follow a normal distribution (i.e., $a_{1d} \sim N(A_1, s^2_{a_1})$), the dyad-specific estimates of directional concurrent synchrony may be summarized by their mean ($A_1$) and variance ($s^2_{a_1}$). Accordingly, $A_1$ refers to the average degree of concurrent synchrony across all dyads, and $s^2_{a_1}$ indicates the variability of the dyad-specific estimates around the average level of concurrent synchrony; providing both a parsimonious and generalizable set of parameters.

Figure 4 separately depicts both models for directional concurrent synchrony (Equations 9 and 10). Figures 4A1 and 4A2 display the model at the within-dyad level (i.e., the repeated physiological measures in a dyad) for Equations 9 and 10, respectively. In both depictions, one partner's detrended responses are predicted by the other's, and that prediction is governed by a latent variable (i.e., $a_{1d}$). Figures 4B1 and 4B2 describe the between-dyad level, which contains the estimate of concurrent synchrony for each dyad. Note that the dyad-specific estimates of $a_{1d}$ are not estimated directly, rather they are summarized via their sufficient statistics (i.e., $A_1$ and $s^2_{a_1}$). The estimate of the mean corresponds to the regression of a dyad-specific variable on a constant; identical to an intercept in simple regression (i.e., the triangle represents a constant equal to 1 for all dyads, and its prediction of $a_{1d}$ equals $A_1$), and the estimate of the variance is shown as a two-headed arrow.

Now, the directionality assumption embedded in Equations 9 and 10 is not trivial. In particular, results may change depending on which partner is chosen to be the outcome. This may seem counterintuitive because in simple regression (i.e., Equation 7 or 8) selection of the outcome is trivial; the $p$-value of the regression slope (i.e., $a_1$) does not depend on which variable is chosen to be the outcome. However, this property no longer holds when analyzing multiple dyads (i.e., the regression example conforms to analyzing a single dyad, whereas estimates of synchrony examine a sample of dyads). Here, we provide an anecdotal example to show the difference, and later we give an empirical example which also emphasizes the difference. For the anecdotal example, imagine two mother-child dyads, referred to hereafter as dyad $A$ and $B$. Let us say the correlations of the detrended physiological responses (i.e., $r_{\varepsilon_{1dt},\varepsilon_{2dt}}$) for dyads $A$ and $B$ equal 0.60 and 0.30, respectively; the standard deviations for the mother and child in dyad $A$ equal 2 and 1, respectively; and the standard deviations for the mother and child in dyad $B$ equal 1 and 2, respectively. Assuming that mothers are partner 1 and children are partner 2, the estimate of concurrent synchrony in Equation 9 (i.e., $a_{1d} = r_{\varepsilon_{1dt},\varepsilon_{2dt}} \frac{s_{\varepsilon_{1d}}}{s_{\varepsilon_{2d}}}$) for dyads $A$ and $B$ equals 1.2 and .15, respectively; suggesting greater synchrony for dyad $A$ relative to $B$. Yet, concurrent synchrony as estimated by Equation 10 (i.e., $a_{1d} = r_{\varepsilon_{1dt},\varepsilon_{2dt}} \frac{s_{\varepsilon_{2d}}}{s_{\varepsilon_{1d}}}$) equals .30 and .60 for dyads $A$ and $B$, respectively; indicating greater synchrony for dyad $B$ relative to $A$. Therefore, depending on the assumed directionality (and because each individual within the dyad has its own degree of variability), the dyads have differing magnitudes of concurrent synchrony. Thus, to use a directional approach, the research hypothesis must be able to accurately assume the direction of the effect within the dyad. Alternatively, if a specific direction cannot be assumed, then we strongly encourage readers to use non-directional concurrent synchrony.

To estimate non-directional concurrent synchrony (i.e., make the directionality assumption trivial), the

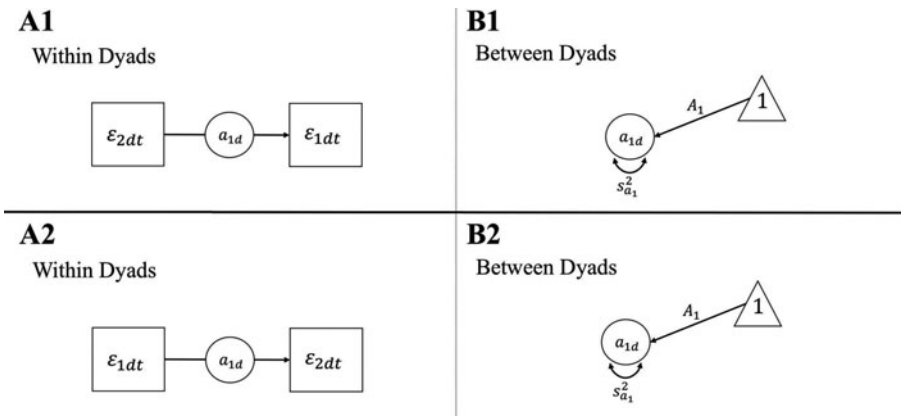


**Figure 4.** Depiction of directional concurrent synchrony. Panels A1 and A2 illustrate within-dyad repeated measures (i.e., repeated physiological responses), and Panels B1 and B2 show the between-dyad variables (i.e., the estimate of concurrent synchrony per dyad).

intraindividual sources of variability must be equated across all individuals. As an example, $\varepsilon_{1dt}$ and $\varepsilon_{2dt}$ may be standardized within individual (i.e., $z_{\varepsilon_{1dt}} = \frac{\varepsilon_{1dt} - \bar{\varepsilon}_{1d}}{s_{\varepsilon_{1d}}}$ and $z_{\varepsilon_{2dt}} = \frac{\varepsilon_{2dt} - \bar{\varepsilon}_{2d}}{s_{\varepsilon_{2d}}}$), and then used within Equations 9 or 10, such that

$$z_{\varepsilon_{1dt}} = a_{1d} z_{\varepsilon_{2dt}} + \epsilon_{z_{1dt}} \tag{11}$$

or

$$z_{\varepsilon_{2dt}} = a_{1d} z_{\varepsilon_{1dt}} + \epsilon_{z_{2dt}}. \tag{12}$$

Standardization of $\varepsilon_{1dt}$ and $\varepsilon_{2dt}$ ensures that the $a_{1d}$ are invariant across Equations 11 and 12 because both approaches estimate the correlation between partners' responses (for Equation 11: $a_{1d} = r_{z_{\varepsilon_{1dt}}, z_{\varepsilon_{2de}}} \frac{s_{z_{\varepsilon_{1dt}}}}{z_{\varepsilon_{2dt}}} = r_{z_{\varepsilon_{1dt}}, z_{\varepsilon_{2dt}}}$; for Equation 12: $a_{1d} = r_{z_{\varepsilon_{1dt}}, z_{\varepsilon_{2dt}}} \frac{s_{z_{\varepsilon_{2dt}}}}{z_{\varepsilon_{1dt}}} = r_{z_{\varepsilon_{1dt}}, z_{\varepsilon_{2dt}}}$). The depiction of Equations 11 and 12 is highly similar to Figure 4A1 and 4B1. Simply replacing $\varepsilon_{1dt}$ and $\varepsilon_{2dt}$ with $z_{\varepsilon_{1dt}}$ and $z_{\varepsilon_{2dt}}$ produces the depiction.

We recognize that within-individual standardization goes against the norms of multilevel structural equation modeling (and multilevel modeling in general; see Moeller, 2015), and we emphasize in the *Discussion* section why breaking these norms is reasonable when calculating non-directional concurrent synchrony. We also note that the method of non-directional concurrent synchrony has conceptual similarities to methods designed for indistinguishable dyads, and we enumerate those similarities (and differences) in the *Discussion* section.

Between-dyad differences of concurrent synchrony (for both directional and non-directional cases) may be examined using dyad-level covariates. Specifically, if we have $m$ dyad-level covariates (generically labeled as $v_{md}$ for the $m^{\text{th}}$ covariate for dyad $d$), then we can structure the dyad-specific covariates as

$$a_{1d} = g_0 + g_1 v_{1d} + g_2 v_{2d} + \cdots + g_m v_{md} + u_{a_{1d}}. \tag{13}$$

In Equation 13, $g_0$ quantifies the expected degree of concurrent synchrony for a dyad that has a value of 0 for all covariates; $g_1$–$g_m$ represent the expected change in concurrent synchrony for a 1-unit change in $v_{1d}$-$v_{md}$, respectively; and $u_{a_{1d}}$ summarizes the portion of $a_{1d}$ unaccounted for by the $m$ dyad-level covariates. Accordingly, a statistically significant coefficient within $g_1$-$g_m$ indicates a non-negligible relation between a specific feature of the dyad (as measured by the $m$th covariate) and concurrent synchrony.

Effect sizes for directional, non-directional, and prediction of concurrent synchrony may be estimated via standardized regression coefficients. For average concurrent or lagged synchrony (i.e., without covariates), standardized regression coefficients may be calculated by first standardizing responses within individual and then re-estimating the model (analogous to estimating a multiple regression with standardized data; see Schuurman, Ferrer, Boer-Sonnenschein, and Hamaker, 2016, for alternative approaches for estimating standardized coefficients). The standardized form for the prediction of concurrent or lagged synchrony may be computed as the square root of the proportion reduction in variance (i.e., $\sqrt{1 - \frac{s^2_{u_{a_1}}}{s^2_{a_1}}}$; analogous to the square root of an $R^2$ from multiple regression). The standardized form of the average concurrent synchrony ($A_1$, for either the directional or non-directional case) reflects the magnitude of the relation between partners' physiological responses. Similarly, the standardized forms of the prediction of concurrent synchrony, $g_1$–$g_m$, reflect the magnitude of the relation between each time-invariant covariate and dyad-specific levels of synchrony. Given that standardized regression coefficients follow the same metric as correlations, we suggest using Cohen's (1988) interpretations of .10, .30, and .50 as small, medium, and large effects. Again, we note that rules of thumb of are not ideal, but may be a useful guide until the field contains enough knowledge to identify more appropriate interpretations of effect size when examining physiological synchrony.

### Lagged synchrony

Lagged synchrony measures the degree to which one partner's current physiological response predicts their partner's response at the next measurement occasion, after removing a trend. Similar to concurrent synchrony, estimation of lagged synchrony begins with removal of each individual's trend, and continues with an analysis of the lagged relations across the detrended responses. Here, we focus on building the model for lagged synchrony following the removal of trends (trend removal is identical for both concurrent and lagged synchrony; readers may refer to the preceding section for details regarding trend removal). We begin with an analysis of a single dyad, and then extend the analysis to accommodate a sample of dyads.

For a single dyad, the model for lagged regression conforms to a bivariate regression, such that

$$\begin{aligned} \varepsilon_{1t} &= c_{11}\varepsilon_{1(t-1)} + c_{12}\varepsilon_{2(t-1)} + \epsilon_{1t} \\ \varepsilon_{2t} &= c_{21}\varepsilon_{2(t-1)} + c_{22}\varepsilon_{1(t-1)} + \epsilon_{2t}, \end{aligned} \tag{14}$$

where $c_{11}$ and $c_{21}$ refer to autoregressive effects for partners 1 and 2, respectively; $c_{12}$ and $c_{22}$ refer to cross-lagged effects for partners 1 and 2, respectively; and $\epsilon_{1t}$ and $\epsilon_{2t}$ contain the portion of $\varepsilon_{1t}$ and $\varepsilon_{2t}$ unaccounted for by $c_{11d}\varepsilon_{1d(t-1)} + c_{12d}\varepsilon_{2d(t-1)}$ and $c_{21d}\varepsilon_{2d(t-1)} + c_{22d}\varepsilon_{1d(t-1)}$, respectively. In Equation 14, $c_{12}$ and $c_{22}$ quantify lagged

synchrony, or the degree to which change in one partner's physiology predicts change in the other's physiology at the next time point (above and beyond each partner's autoregressive effect).

Extending Equation 14 to accommodate a sample of dyad follows

$$\begin{aligned}\varepsilon_{1dt} &= c_{11d}\varepsilon_{1d(t-1)} + c_{12d}\varepsilon_{2d(t-1)} + \epsilon_{1dt}\\ \varepsilon_{2dt} &= c_{21d}\varepsilon_{2d(t-1)} + c_{22d}\varepsilon_{1d(t-1)} + \epsilon_{2dt},\end{aligned} \quad (15)$$

wherein all variables and coefficients match the interpretation from Equation 14, but now refer to a specific dyad *d*. Similar to the estimation of both trend and concurrent synchrony, the dyad-specific coefficients from Equation 15 are assumed to follow a multivariate normal distribution; thereby enabling the same benefits of parsimony and generalizability as compared to fitting separate bivariate regressions for each dyad (for details, see the section above on estimating trend synchrony). More specifically, the multivariate assumption allows the dyad-specific coefficients to be fully characterized through their means (i.e., $C_{11}, C_{12}, C_{21}, C_{22}$), variances (i.e., $s^2_{c_{11}}, s^2_{c_{12}}, s^2_{c_{21}}, s^2_{c_{22}}$), and covariances (e.g., $s_{c_{11},c_{21}}$).

Figures 5A and 5B depict the model for lagged synchrony (i.e., Equation 15). Figure 5A displays the model at the within-dyad level (i.e., the repeated physiological measures in a dyad), wherein each partner's detrended responses are predicted by their own and their partner's detrended responses from the previous measurement occasion. The predictions equal a latent variable (i.e., $c_{11d}, c_{12d}, c_{21d}, c_{22d}$) that differs by dyad. Figure 5B describes the between-dyad level, which contains the estimates of autoregression (i.e., $c_{11d}$and $c_{21d}$) and lagged synchrony (i.e., $c_{21d}$ and $c_{22d}$) for each dyad. Again, the dyad-specific estimates are not estimated directly, instead they are summarized via their sufficient statistics (e.g., $C_{11}$and $s^2_{c_{11}}$). Similar to the depiction of concurrent synchrony, the means (i.e., $C_{11}, C_{12}, C_{21}, C_{22}$) of the dyad-specific coefficients correspond to regressions of dyad-specific variables on a constant (displayed as the prediction from the triangle), and the variances (i.e., $s^2_{c_{11}}, s^2_{c_{12}}, s^2_{c_{21}}, s^2_{c_{22}}$) and covariances (e.g., $s_{c_{11},c_{21}}$) are portrayed as two-headed arrows.

Similar to concurrent synchrony, between-dyad differences of lagged synchrony may be tested via dyad-level covariates. Specifically, a set of *m* dyad-level covariates (generically labeled as $v_{md}$ for the *m*th covariate for dyad *d*) can predict the cross-lagged effects, such that

$$\begin{aligned}c_{12d} &= g_{01} + g_{11}v_{1d} + g_{21}v_{2d} + \cdots + g_{m1}v_{md} + u_{c_{12d}}\\ c_{22d} &= g_{02} + g_{12}v_{1d} + g_{22}v_{2d} + \cdots + g_{m2}v_{md} + u_{c_{22d}}.\end{aligned} \quad (16)$$

The coefficients within Equation 16 have an analogous interpretation to those from Equation 13, and a statistically significant coefficient within $g_{01}$–$g_{m2}$ suggests a meaningful between-dyad difference in lagged synchrony.

Effect sizes for lagged synchrony are virtually identical to those for concurrent synchrony. In particular, effect sizes may be estimated through standardized forms of the average lagged effect (i.e., $C_{12}$ and $C_{22}$), and the standardized prediction of between-dyad differences in lagged synchrony (i.e., $g_{01}$–$g_{m2}$). The same recommendations and caveats as noted in previous sections apply.

## Empirical examples

This section provides empirical examples of trend, concurrent, and lagged synchrony. All models were fitted with *Mplus* 7.11 (Muthén & Muthén, 1998–2012), and syntax files are provided online

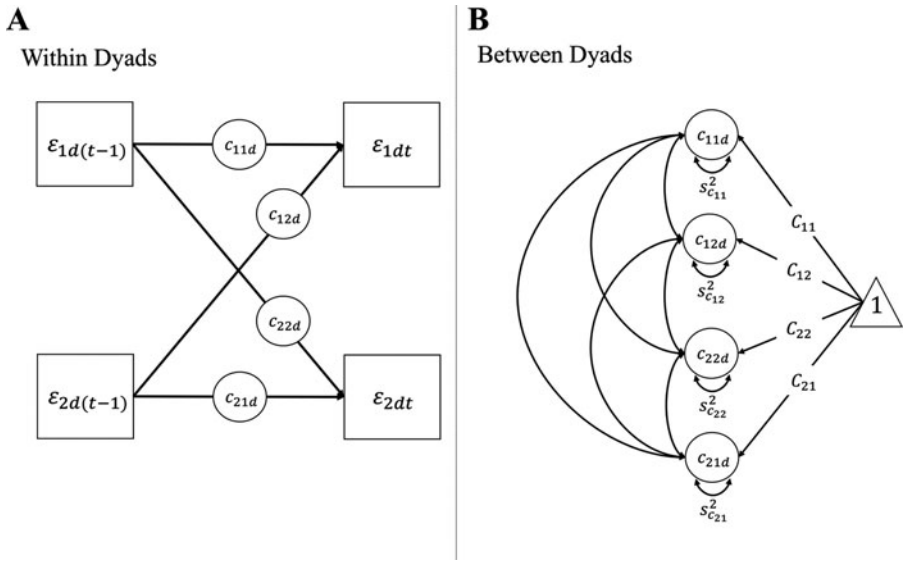


**Figure 5.** Depiction of lagged synchrony. Panel A illustrates the within-dyad repeated measures (i.e., repeated physiological responses), and Panel B shows the between-dyad variables (i.e., estimates of dyad-specific autoregression and lagged synchrony).

(https://jonhelm.org/supplemental-materials/). Parameter estimates are denoted with the hat symbol (e.g., '$A_1$' refers to a specific model parameter, and '$\hat{A}_1$' refers to a sample-based estimate of the model parameter), and '*ES*' refers to a parameter's estimated effect size. Due to limited space, we only report the estimates of physiological synchrony in the text. In practice, we recommend reporting and interpreting all parameter estimates to inform readers of all aspects of the analysis. Finally, the examples represent pedagogical demonstrations of measuring and modeling physiological synchrony, and may not encompass the most meaningful analyses for testing hypothesis from specific theories. We encourage researchers to identify an expected type of synchrony based on theory, and fit the corresponding model.

We use maximum likelihood (ML) estimation for most empirical examples because most multilevel modeling software implements ML estimation, and therefore the parameter estimates from structural equation modeling should be similar to those obtained via multilevel modeling. Furthermore, the $\chi^2$-difference testing performed in the estimation of trend synchrony (see below) is simpler for models fitted via ML estimation (see Satorra & Bentler, 2001). ML estimation assumes multivariate normality for all observed variables, which refers to the observed physiological responses. If normality may not be assumed, then users should implement robust ML (Satorra & Bentler, 1994; Savalei, 2014). As an aside, the results from the empirical examples below produce negligible differences across ML and robust ML (i.e., the parameter estimates are identical across the methods, and standard errors and *p*-values show differences less than .02).

Our only implementation of robust ML occurs for estimation of non-directional concurrent synchrony. Estimation of non-directional concurrent synchrony potentially violates a distributional assumption underlying ML estimation (i.e., all parameter estimates have normal sampling distributions), and robust ML may mitigate that potential violation. We describe that assumption, the scope of its violation, and alternative solutions in the *Discussion*.

## Example of trend synchrony

The example of trend synchrony comes from the PEP data collected within the puzzle task. Responses were initially measured in consecutive 30 s epochs (i.e., one measure of PEP for each epoch), and these measures were averaged within each minute to produce a more reliable physiological index[2]. Bivariate intercept-only, linear, and quadratic models (e.g., Equations 5 and 6 for intercept-only and linear) were fitted and statistically compared. Using the usual $\chi^2$ difference test (see Bollen, 1989), the results indicated that the linear model fitted better than the intercept-only model ($\Delta\chi^2 = 67.95$, $\Delta df = 13$, $p < .01$). Furthermore, fit indices suggested good fit for the linear model (CFI = .98, TLI = .98, RMSEA = .09, 95% C.I. for RMSEA = [.05, .13]). Alternatively, the quadratic model produced correlations outside [−1, 1], indicating its inappropriateness for the data (see Wothke, 1993). Therefore, the linear model (Equation 6) was selected as the best model for the data, and parameter estimates are summarized in Table 1. For those readers interested in quadratic growth models, please refer to De Fraine, Van Damme, & Onghena, 2007, Dietrich, Jokisaari, & Nurmi, 2012, or Mitchell, Beals, & Kaufman, 2006, for clear demonstrations.

**Table 1.** Parameter estimates of trend synchrony for PEP data in the puzzle task.

| *Parameter* | *Estimate* | *s.e.* | *p* |
|---|---|---|---|
| $B_{01}$ | 91.01 | 1.04 | <0.01 |
| $B_{11}$ | − 0.07 | 0.10 | 0.49 |
| $B_{02}$ | 136.31 | 1.73 | <0.01 |
| $B_{12}$ | 1.09 | 0.18 | <0.01 |
| $s_{b_{01}}$ | 8.29 | 0.74 | <0.01 |
| $s_{b_{11}}$ | 0.56 | 0.10 | <0.01 |
| $s_{b_{02}}$ | 14.87 | 1.23 | <0.01 |
| $s_{b_{12}}$ | 0.78 | 0.28 | 0.01 |
| $r_{b_{01},b_{11}}$ | − 0.22 | 0.17 | 0.20 |
| $r_{b_{01},b_{02}}$ | 0.33 | 0.11 | <0.01 |
| $r_{b_{01},b_{12}}$ | 0.10 | 0.26 | 0.68 |
| $r_{b_{11},b_{02}}$ | − 0.17 | 0.18 | 0.35 |
| $r_{b_{11},b_{12}}$ | 0.02 | 0.38 | 0.96 |
| $r_{b_{02},b_{12}}$ | 0.58 | 0.25 | 0.02 |
| $s^2_{\varepsilon_1}$ | 2.57 | 0.28 | <0.01 |
| $s^2_{\varepsilon_2}$ | 17.88 | 1.71 | <0.01 |
| $s_{\varepsilon_1,\varepsilon_2}$ | 0.44 | 0.56 | 0.44 |

*Note*: A second subscript equal to '1' refers to the child, and '2' to the mother. All coefficients refer to Equation 6 and Figure 2B.

For the linear model, parameter estimates $\{\hat{r}_{b_{01},b_{02}}, \hat{r}_{b_{01},b_{12}}, \hat{r}_{b_{11},b_{02}}, \hat{r}_{b_{11},b_{12}}\}$ represent different sources of trend synchrony, or the degree to which one partner's trend relates to the other's. Of the set, only $\hat{r}_{b_{01},b_{02}}$ reached significance ( $\hat{r}_{b_{01},b_{02}} = .33$, $ES = .33$, 95% C.I. = [.11, .55], $p < .01$). Therefore, mothers' and children's average PEP across the puzzle task tended to correspond (i.e., children with above-average PEP tended to have mothers who had above-average PEP), and the estimate of that relation suggests a medium effect size.

[2] Estimation using the original metric of 30 s epochs produced highly unstable parameter estimates and poor model fit for the multiple group model (across different signals and tasks); whereas estimation following aggregation within each minute produced reliable parameter estimates with good fit. Hence, the aggregation performed here was done to enable a demonstration of trend synchrony rather than a suggestion of data aggregation.

Next, we tested for the presence of gender differences for trend synchrony (i.e., examined a between-dyad difference of trend synchrony). Here, we nested the bivariate linear growth curve into a multiple group model that allowed separate estimates for dyads with boys versus dyads with girls. First, we compared models that constrained latent variable means (i.e., $B_{01}$, $B_{11}$, $B_{02}$, $B_{12}$) and then standard deviations (i.e., $s_{b_{01}}$, $s_{b_{11}}$, $s_{b_{02}}$, $s_{b_{12}}$) to a model that freely estimated all parameters across dyads with boys and girls. Neither model showed significantly worse fit to the data (constraining latent variable means: $\Delta\chi^2 = 2.90$, $\Delta df = 4$, $p = .57$; constraining latent variable means and standard deviations: $\Delta\chi^2 = 1.70$, $\Delta df = 4$, $p = .79$), indicating that dyads with boys and girls did not largely differ on their average linear trajectories nor variability around those trajectories. However, constraining across latent variable correlations led to a significant decrease in model fit ($\Delta\chi^2 = 27.30$, $\Delta df = 6$, $p < .01$), potentially suggesting a significant difference in trend synchrony across dyads with boys and girls. Post-hoc examination of differences across each latent variable correlation indicated that only $r_{b_{11},b_{02}}$ differed across dyads with boys and girls (boys: $\hat{r}_{b_{11},b_{02}} = .55$; girls: $\hat{r}_{b_{11},b_{02}} = -.50$; ES = 1.17; 95% C.I. = [.42, 1.68]; $\Delta\chi^2 = 7.73$, $\Delta df = 1$, $p < .01$). These results indicate a large gender difference. Boys with linear increases in PEP (i.e., decreases in sympathetic response) across the puzzle task tended to have mothers with higher than average PEP (i.e., lower than average sympathetic arousal), whereas girls with linear increases in PEP tended to have mothers with lower than average PEP (i.e., higher than average sympathetic arousal).

Overall, the examination of PEP within the puzzle task provided an illustration of both trend synchrony and between-dyad differences in trend synchrony. When averaging across all dyad-specific differences within the sample (i.e., not accounting for gender differences), mothers and children tended to show similar average levels of PEP across the puzzle task, but did not tend to show similar rates of change in PEP. Yet, when focusing in on whether trend synchrony differed by gender, boys' and girls' rates of change related differently to their mothers' average level of PEP. Hence, the general pattern of children's and mother's PEP activity aligned across the puzzle task, and the alignment differed as a function of child's gender.

## Example of concurrent synchrony

Mothers' and children's RSA data within the storybook task demonstrate the importance of detrending before examining concurrent synchrony. With RSA measured in 30 sec epochs, a significant average relation and significant between-dyad variation emerged when analyzing the raw (not detrended) RSA values (for mother predicting child: $\hat{A}_1 = .87$, $ES = .12$, 95% C.I = [.82, .92], $p < .01$; $\hat{s}^2_{a_1} = .04$, 95% C.I. = [.02, .06], $p < .01$; for child predicting mother: $\hat{A}_1 = 1.18$, $ES = .12$, 95% C.I. = [1.12, 1.25], $p < .01$; $\hat{s}^2_{a_1} = .07$, 95% C.I. = [.04, .11], $p < .01$); suggesting that (on average) mothers and children showed noteworthy correspondence between their parasympathetic activity, with some dyads showing greater correspondence relative to others. However, both mothers and children exhibited significant quadratic trajectories (for children: $\hat{B}_{01} = 5.67$, $p < .01$; $\hat{B}_{11} = -.03$, $p = .01$; $\hat{B}_{21} = -.01$, $p < .01$; for mothers: $\hat{B}_{02} = 6.57$, $p < .01$; $\hat{B}_{12} = -.02$, $p = .10$; $\hat{B}_{22} = -.01$, $p < .01$). After removal of each individual's quadratic trend (i.e., fitting a quadratic model to each individual and saving the residuals), the RSA responses no longer produced significant average concurrent synchrony (for mother predicting child: $\hat{A}_1 = -.01$, $ES < .01$, 95% C.I. = [−.09, .08], $p = .85$; for child predicting mother: $\hat{A}_1 = -.01$, $ES < .01$, 95% C.I. = [−.12, .12], $p = .95$), nor between-dyad differences of concurrent synchrony (for mother predicting child: $\hat{s}^2_{A_1} = .02$, 95% C.I. = [−.01, .06], $p = .19$; for child predicting mother: $\hat{s}^2_{A_1} = .07$, 95% C.I. = [−.05, .18], $p = .27$). Thus, the significant concurrent synchrony from the raw RSA responses was attributable to the presence of the common trend across mothers and children, rather than being due to a common fluctuation around the trend; analogous to the anecdotal example within the *Different Models for Physiological Synchrony* section above.

The HR responses from mothers and children in the storybook task can serve to illustrate the distinction between directional and non-directional concurrent synchrony (all parameter estimates are summarized in Table 2). Analyses followed removal of each individual's

**Table 2.** Parameter estimates of concurrent synchrony for heart rate data in the storybook task.

| *Form* | *Direction* | *Parameter* | *Estimate* | *s.e.* | *p* |
|---|---|---|---|---|---|
| *Directional* | | | | | |
| | $M \rightarrow C$ | | | | |
| | | $A_1$ | .05 | .03 | .04 |
| | | $s^2_{a_1}$ | .02 | .01 | .01 |
| | | $s^2_{\epsilon_1}$ | 41.05 | 3.15 | <.01 |
| | $C \rightarrow M$ | | | | |
| | | $A_1$ | .02 | .02 | .15 |
| | | $s^2_{a_1}$ | .006 | <.01 | <.01 |
| | | $s^2_{\epsilon_2}$ | 28.87 | 1.99 | <.01 |
| *Non-directional* | | | | | |
| | *N/A* | | | | |
| | | $A_1$ | .04 | .02 | .04 |
| | | $s^2_{a_1}$ | .01 | <.01 | <.01 |
| | | $s^2_{\epsilon_1}$ | .98 | <.01 | <.01 |

*Note*: '*C*' refers to child, '*M*' to mother. Non-directional concurrent synchrony does not have a direction because parameter estimates are equivalent regardless of the outcome chosen outcome variable.

quadratic trend (i.e., fitting a quadratic model to each individual and saving the residuals). When forming the directional hypothesis that mothers affect children (i.e., mother predicting child), the data supported a significant average effect ($\hat{A}_1$= .05, $ES$ = .03, 95% C.I. = [.002, .10], $p$ = .04) and significant between-dyad variation ($s^2_{a_1}$ = .02, 95% C.I. = [.003, .03], $p$ = .01); suggesting that – on average – increases in a mother's HR co-occurred with increases in her child's HR (i.e., a 1-unit increase in children's HR co-occurred with a .05 increase in mothers' HR), and dyads varied in the strength of this relation. However, the average effect is very small, and its 95% confidence interval is close to zero, indicating that (at best) the magnitude of concurrent synchrony is very small. Yet, when assuming that children's HR affected mothers' (i.e., child predicting mother), the data did not show a significant average effect ($\hat{A}_1$ = .02, $ES$ = .03, 95% C.I. = [−.008, .052], $p$ = .15), but did show significant between-dyad variation ($s^2_{a_1}$ = .006, 95% C.I. = [.002, .01], $p$ < .01); indicating that – on average – increases in a child's HR did not co-occur with changes in his/her mother's HR, but dyads differed in their synchrony (e.g., some dyads may have shown significant positive or negative synchrony). Although these results are not extremely different when examining the effect size and 95% confidence intervals, it would be possible to arrive at different conclusions based on significance. And, more importantly, if the direction of the effect is truly unimportant (i.e., if the researcher cannot assume a causal direction), the results of the models should be identical (i.e., the assumed direction of the effect should have no impact on the results). Therefore, a researcher must exert caution (i.e., provide theoretical and/or empirical justification) when using models that assume a specific direction of the effect from one partner to another.

If a researcher cannot justify a specific direction for concurrent synchrony, then a non-directional approach should be adopted. In that case, intraindividual standardization extinguishes the directionality assumption. Extending the example of concurrent synchrony from above: after completing intraindividual standardization, both modeling approaches (i.e., mother predicting child, or child predicting mother) produced identical parameter estimates and test statistics. In particular, the modeling approaches showed significant average concurrent synchrony ($\hat{A}_1$ = .04, $ES$ = .04, 95% C.I. = [.001, .07], $p$ = .04) and significant between dyad-differences of concurrent synchrony ($s^2_{a_1}$= .01, 95% C.I. = [.002, .01], $p$ < .01), suggesting that mothers and children (on average) manifested a common fluctuation in their HRs, and that the magnitude of correspondence in their fluctuations differed across dyads, with a very small average effect (i.e., average correlation of .04). Nevertheless, intraindividual standardization assures that the modeling approaches produce equivalent summaries of concurrent synchrony, thereby causing the inference to be non-directional, and assuring that the significance and magnitude of the effect are not solely due to intraindividual variances.

The separation of directional and non-directional concurrent synchrony should also be considered when predicting between-dyad differences of concurrent synchrony. Using non-standardized HR data, an examination of HR during the origami task demonstrates this notion. Prediction of concurrent synchrony using a dyad level covariate – security of child's attachment, measured using the observer-rated Attachment Q-Sort (Troxel, 2017) – was not consistent across the models. Child's attachment security significantly moderated concurrent synchrony of HR when mother predicted child ($\hat{g}_1$ = .31, ES = .49, 95% C.I. = [.12, .50], $p$ < .01), but not when child predicted mother ( $\hat{g}_1$= .10, $ES$ = .33, 95% C.I. = [−.02, .05], $p$ = .09). In this case, the effects sizes, confidence intervals, and $p$-values are noticeably different. Therefore, discrepancies among the parameter estimates (and their corresponding tests of significance) may arise across the different directional models when predicting between-dyad differences of concurrent synchrony.

Analogous to estimates of average concurrent synchrony, intraindividual standardization forces prediction of between dyad differences to be consistent across modeling approaches (i.e., mother predicting child, or child predicting mother). Following intraindividual standardization of HR within the origami task, child's attachment security predicted concurrent synchrony in an identical manner across the two modeling approaches (i.e., all parameter estimates and test statistics were equivalent to the fourth decimal place; $\hat{g}_1$ = .17, $ES$ = .41, 95% C.I. = [.04, .31], $p$ = .01). These results suggest that a 1-point attachment security (as quantified via the AQS) correspond to a .17 increase in the correlation between mothers' and children's responses. Thus, intranindividual standardization not only ensures that estimates of concurrent synchrony are non-directional, but that prediction of between-dyad differences also remains invariant across the two modeling approaches.

### *Example of lagged synchrony*

The HR data from the puzzle task also provide an example of lagged synchrony. The model from Equation 15 was fitted to the detrended responses, with child referring to the first partner (i.e., initial subscript = 1) and mother to the second (i.e., initial subscript = 2). A summary of the parameter estimates is given in Table 3. Both mothers and children showed significant average autoregressive

**Table 3.** Parameter estimates of lagged synchrony for heart rate data in the puzzle task.

| *Parameter* | *Estimate* | *s.e.* | *p* |
|---|---|---|---|
| $c_{11}$ | 0.48 | 0.02 | <0.01 |
| $c_{21}$ | 0.39 | 0.03 | <0.01 |
| $c_{12}$ | 0.05 | 0.01 | <0.01 |
| $c_{22}$ | 0.01 | 0.01 | 0.40 |
| $s^2_{c_{11}}$ | 0.01 | 0.00 | <0.01 |
| $s^2_{c_{21}}$ | 0.02 | 0.01 | <0.01 |
| $s^2_{c_{12}}$ | <0.01 | <0.01 | 0.62 |
| $s^2_{c_{22}}$ | <0.01 | <0.01 | 0.31 |
| $s_{c_{11},c_{21}}$ | −0.01 | −1.50 | 0.13 |
| $s_{c_{11},c_{12}}$ | <0.01 | −0.15 | 0.88 |
| $s_{c_{11},c_{22}}$ | <0.01 | −0.79 | 0.43 |
| $s_{c_{21},c_{11}}$ | <0.01 | <0.01 | 0.93 |
| $s_{c_{21},c_{22}}$ | <0.01 | <0.01 | 0.98 |
| $s_{c_{21},c_{22}}$ | <0.01 | <0.01 | 0.24 |
| $s^2_{\epsilon_1}$ | 27.76 | 1.93 | <0.01 |
| $s^2_{\epsilon_2}$ | 22.73 | 2.06 | <0.01 |
| $s_{\epsilon_1,\epsilon_2}$ | −0.42 | 0.38 | 0.26 |

*Note*: An initial subscript '1' refers to the child, and initial subscript '2' to the mother. All coefficients refer to Equation 15.

effects ($\hat{C}_{11} = .48$, $ES = .48$, 95% C.I. = [.44, .52], $p < .01$; $\hat{C}_{21} = .39$, $ES = .39$, 95% C.I. = [.34, .44], $p < .01$), and significant between-dyad variability of those autoregressive effects ($s^2_{c_{11}} = .01$, 95% C.I. = [.004, .02], $p < .01$; $s^2_{c_{21}} = .02$, 95% C.I. = [.01, .04], $p < .01$). A significant average cross-lagged effect appeared for mother's influence on children ($\hat{C}_{12} = .05$, $ES = .04$, 95% C.I. = [.02, .07], $p < .01$), but not vice versa ($\hat{C}_{22} = .01$, $ES = .01$, 95% C.I. = [−.01, .03], $p = .40$). Neither of the cross-lagged partner effects showed significant between-dyad variability ($s^2_{c_{12}} < .01$, 95% C.I. = [−.004, .006], $p = .62$; $s^2_{c_{22}} < .01$, 95% C.I. = [−.001, .004], $p = .31$). These effects suggest that – on average – both mothers and children showed a significant relation between their own HRs from one 2 sec epoch to the next, and that mothers' HR predicted children's subsequent HR above and beyond children's own time-lagged relation (i.e., a 1-unit increase in mothers' current HR predicted a .05 increase in children's HR 2 seconds later). In more detail, the autoregressive effects were fairly large, whereas the estimate of lagged synchrony from mothers to children was very small.

Between-dyad differences in the autoregressive and cross-lagged effects may be examined using dyad-level predictors. Here, we again used child's gender to predict differences in both the autoregressive and cross-lagged effects. Child's gender significantly predicted the cross-lagged effect from mother to child, indicating significant between-dyad differences in the relation between mothers and children. In particular, dyads with boys showed a significant cross-lagged effect (boys: $\hat{C}_{12} = .08$, 95% C.I. = [.04, .12], $p < .01$), whereas those with girls showed a non-significant effect (girls: $\hat{C}_{12} = .02$, 95% C.I. = [−.005, .05], $p = .12$), and there was a significant difference between dyads with boys and girls ($\hat{C}_{12,\text{ girls}} - \hat{C}_{12,\text{ boys}} = -.06$, $ES = .73$, 95% C.I. = [−.11, −.007], $p = .03$). Therefore, the data suggested that the cross-lagged effect from mothers to children was significantly stronger for boys (i.e., a 1-unit increase in mothers' current HR predicted a .08 increase in boys' HR) than for girls (i.e., a 1-unit increase in mothers' current HR predicted a .02 increase in girls' HR), with a reasonably large effect size[3].

## Discussion

This article comes as a response to the varied and disparate methods that have been implemented for measuring and modeling physiological synchrony. In particular, we identified and demonstrated three of the most common forms of physiological covariation for shorter physiological series: trend, concurrent, and lagged synchrony. We showed how each method can be extended to predict between-dyad differences in synchrony via time-invariant variables, enabling researchers to examine between-dyad differences of physiological covariation. Finally, we provide online supplemental material with annotated M*plus* code that corresponds to each method for detecting physiological synchrony (https://jonhelm.org/supplemental-materials/). Accordingly, researchers can use this article as a guide for integrating results across different examinations of physiological synchrony, or for selecting a proper modeling procedure for their own hypotheses.

The different models correspond to different testable hypotheses regarding physiological synchrony, and researchers should choose the model and/or type of synchrony that best aligns with their hypotheses. For example, researchers interested in determining whether mothers and children tend to show a common linear (or quadratic, cubic, etc.) change in HR across an experimental task should examine trend synchrony; researchers focused on identifying the degree of common, moment-to-moment physiology between partners should investigate concurrent synchrony; and researchers geared toward understanding the extent to which one partner's physiology predicts the other's future physiology (controlling for each partner's own physiology at the previous time point) should pursue lagged synchrony. Importantly, one cannot use models for trend synchrony to examine concurrent or lagged synchrony, nor vice

[3] It may seem surprising that a significant prediction occurred for a variable with non-significant variability. However, non-significant between-dyad differences (i.e., $s^2_{d_1}$ with $p > .05$) are not the same as nil between-dyad differences (i.e., $s^2_{d_1} = 0$). Therefore, it is possible, albeit rare, to have significant covariation with a variable that has small variation, and this is what occurred in the investigation of between-dyad differences for lagged synchrony.

versa. Hence, researchers must establish an expectation regarding the generation of physiological synchrony between dyad members, and then select the approach (i.e., type of synchrony and corresponding model) that matches their expectation.

Although the methods described in the article are not new (Kenny et al., 2006; MacCallum et al., 1997; Ferrer & McArdle, 2003), many aspects were novel. The distinction between directional and non-directional concurrent synchrony has not been noted in prior texts, nor the discrepancies that can unfold between those two different modeling approaches. Furthermore, the examples of detrending when examining concurrent or lagged synchrony have not been emphasized within reports of physiological covariation, even though trend synchrony can confound concurrent and lagged synchrony. Accordingly, this article highlights that calculating and examining physiological synchrony via intra-dyad correlations, or the relation between raw physiological responses (i.e., arbitrarily predicting $X_{1dt}$ from $X_{2dt}$, or vice versa), does not determine whether synchrony is due to a common trend, a common fluctuation around the trend, or a combination of the two. Therefore, this article offers a refocusing of modern methods for multivariate longitudinal data analysis by noting the caveats that can arise within the context of intra-dyad physiological synchrony.

### *Standardization in non-directional concurrent synchrony*

The method for non-directional concurrent synchrony prescribes intraindividual standardization (after trend removal); yet intraindividual standardization is often discouraged within longitudinal modeling (Moeller, 2015), as well as statistical modeling in general (Greenland, Maclure, Schlesselman, Poole, & Morgenstern, 1991; King, 1985; Tukey, 1954). This discouragement arises from both interpretational and statistical bases, and we provide counterarguments against these bases when estimating non-directional concurrent synchrony.

The interpretational basis states that parameter estimates provided via unstandardized data are preferred because they describe the functional relation between variables, or how variables actually manifest in the real-world (i.e., a one-unit change in a given predictor co-occurs with an expected level of change in the outcome). Conversely, parameter estimates from standardized data produce coefficients that follow a correlation metric, and do not reflect how the variables are manifested or measured in the real-world (i.e., the units are removed). Yet, as our anecdotal examples (in the *Different Models for Physiological Synchrony* section) and empirical examples illustrate, the estimation of non-directional concurrent synchrony from unstandardized data may violate the scientific property of falsifiability (i.e., the notion that alternative hypotheses can be extinguished via scientific inquiry). More specifically, if a researcher cannot assume a causal direction between partners' physiology, then the researcher's choice of the outcome and predictor variable is ambiguous (i.e., mothers' responses could predict children's, or vice versa). However, a researcher may arrive at different conclusions depending on that choice; thereby creating an unfalsifiable hypothesis. Therefore, we believe that researchers unable to assume a causal direction should use standardized data to be certain that the ambiguity of selecting the outcome variable does not cloud the interpretation of the results, even at the cost of the interpretational benefits provided by unstandardized data (i.e., estimates of the functional relation between variables). Alternatively, those researchers that prefer the benefit of functional relations should provide a justification for the causal direction, and implement directional concurrent synchrony.

The statistical basis indicates that significance testing of standardized coefficients can violate distributional assumptions. In particular, as regression slopes between standardized variables approach the $\pm 1$ boundary, they will fail to have a normal sampling distribution; an assumption that underlies the test of significance (however, as the estimate approaches 0, then the distribution is approximately normal). Given that potential violation, we implemented robust ML estimation for non-directional concurrent synchrony. Nevertheless, this notion has not been fully vetted by prior research, and therefore we are not certain that robust ML provides the best solution for the potential violation to distributional assumptions (i.e., robust ML relaxes distributional assumptions of the data rather than parameter estimates). Hence, future research is needed to determine if robust ML adequately relaxes the potential violation to normality (see below).

Finally, both bootstrapping and likelihood based confidence intervals can relax the distributional assumption underlying the test of significance for non-directional concurrent synchrony (see Goldstein, 2011, and Neale & Miller, 1997, for details). However, neither of these approaches are available in M*plus* for the multilevel path models used to estimate concurrent synchrony, and therefore were not implemented in this tutorial. Interested readers may use the OpenMx package (Neale et al., 2016) in R to either perform bootstrapping or compute likelihood based confidence intervals. Although there is some evidence to suggest that bootstrapping approaches may be preferable (Falk, 2017), future research should compare and contrast robust ML from M*plus*, bootstrapping, and likelihood based confidence intervals to determine if and

when one method outperforms the others in the context of concurrent synchrony.

### Estimating physiological synchrony for longer physiological series

The current tutorial focused on methods for shorter physiological series, or data with more dyads than measurement occasions per dyad. More intensive data sets, such as those containing moment to moment respiration amplitude (measured at 1000 samples per second; see Helm et al., 2012), may be examined using different methods that estimate other types of physiological synchrony. For example, intensive repeated measures are less likely to have stable trends (e.g., linear, quadratic), and are more apt to exhibit fluctuations about a stable set point (see Boker & Nesselroade, 2002). Although concurrent and trend synchrony also examine common fluctuations (and may be applied to longer physiological series), alternative methods for intensive repeated measures quantify synchrony differently, and can therefore offer a distinct summary of physiological synchrony. We provide two examples below (dynamical systems models and spectral analysis), and then compare those examples to the three focal methods (trend, concurrent, and lagged synchrony) of the tutorial.

Dynamical systems models examine the moment-to-moment level (i.e., 0th derivative), speed (i.e., 1st derivative), and acceleration (i.e., 2nd derivative) of one or more signals (Boker & Laurenceau, 2006; Boker, Leibenluft, Deboeck, Virk, & Postolache, 2008; Montpetit, Bergeman, Deboeck, Tiberio, & Boker, 2010). For example, Helm et al. (2012) used a coupled linear oscillator (a specific dynamical systems model; see Boker & Nesselroade, 2002) to examine HR and respiration rate synchrony in opposite-sex romantic partners. Similarly, Ferrer and Helm (2013) used a special adaption of a predator-prey model (another dynamical systems model; see Felmlee & Greenberg, 1999) to quantify HR and respiration synchrony in opposite-sex romantic partners. Both examples had many more within dyad observations than number of dyads, and physiological responses tended to show fluctuations about a stable set point; making a dynamical systems approach more appropriate than the three methods described earlier in the tutorial. Researchers interested in using dynamical systems modeling should consult Boker and Nesselroade (2002), Boker and colleagues (2010), Boker (2012), Deboeck (2010), and Deboeck and colleagues (2013).

Spectral analysis provides a second alternative that summarizes co-occurring oscillations between two signals, and is more commonly implemented for intensive time series. For univariate processes, spectral analysis decomposes a given signal into a sum of sinusoids (i.e., sines and cosines) of different frequencies (i.e., slower and faster rates of oscillation; see Porges et al., 1980). For bivariate processes (or even multivariate processes), spectral analysis estimates coherence, or the degree of common fluctuation across the different sinusoids. As an example, Henning, Boucsein, and Gil (2001) estimated physiological synchrony between teammates' electrodermal activity, HR, and respiration rate using spectral analysis, and reported that teammates with greater synchrony tended to perform better during competition. Similarly, Järvelä, Kivikangas, Kätsyri, and Ravaja (2014) used spectral analysis to find that same-sex university student dyads manifested significant physiological synchrony of HR and electrodermal activity while playing video games. Researchers interested in spectral analysis should consult Gottman (1979, 1981), Porges and coworkers (1980), and Shumway and Stoffer (2006).

The similarities and differences across the methods for shorter (i.e., trend, concurrent, and lagged synchrony) versus longer (dynamical systems analysis and spectral analysis) physiological series concern each method's quantification/operationalization of synchrony. Trend synchrony differs from the rest because it extracts a common pattern of change (e.g., common linear or quadratic slopes) across signals, whereas the remaining methods summarize common fluctuations. Yet, the remaining methods differ in their measurement of common fluctuations across signals. Concurrent and lagged synchrony examine the relations across partners' moment-to-moment level (e.g., a 1-beat increase in a dyad member's HR [i.e., 1-unit increase in level] corresponds to a 2-beat increase in the other dyad member's [i.e., 2-unit increase in level]); dynamical systems models examine relations across variables' derivatives (i.e., levels, velocities, and accelerations; e.g., when one dyad member's HR increases [positive velocity/1st derivative], their partner's HR changes from increasing to decreasing [negative acceleration/2nd derivative]); and spectral analysis summarizes the average degree of coherence across variables' sinusoidal frequencies (e.g., across different sinusoidal frequencies, each member's response accounts for 50% of their partner's). Naturally, different hypotheses will conform to different analyses (i.e., investigating trends, levels, derivatives, or common sinusoidal frequencies), and researchers should choose the methods that most closely match their hypotheses.

### Considerations of power

Statistical power, or the probability of rejecting the null hypothesis when it is false, represents an important consideration for those researchers planning to examine

physiological synchrony. Although – to our knowledge – considerations of power for examining physiological synchrony have not been documented directly, they may be inferred from the larger literature of structural equation modeling and multilevel modeling. In particular, detection of trend synchrony corresponds to the literature of detecting correlations among latent slopes (Hertzog, Lindenberger, Ghisletta, & von Oertzen, 2006), whereas detection of concurrent and lagged synchrony corresponds to a significant level-1 slope in a multilevel model (Maas & Hox, 2005; Snijders & Bosker, 2012).

Power to detect trend, concurrent, or lagged synchrony depends on (1) the residual variability (for trend synchrony: the variability in each partner's raw physiological responses that is not accounted for by the trend; for concurrent or lagged synchrony: the variability in the outcome measure not accounted for by the concurrent or lagged predictors), (2) the number of measurement occasions per dyad, (3) the sample size (i.e., the number of dyads), and (4) the effect size (see the sections on *Different Models for Physiological Synchrony*). These four components must be simultaneously considered to correctly identify the power to detect physiological synchrony (or predict between-dyad differences thereof), and therefore general rules of thumb do not apply across empirical examples. Instead, readers interested in modeling trend synchrony should consult Hertzog and colleagues (2006), which provides projections of power as a function of the four components; whereas readers interested in concurrent or lagged synchrony should consult Scherbaum and Ferreter (2009). Alternatively, researchers may also use M*plus*' simulation facilities to conduct power analyses by specifying different constellations of the four factors affecting power listed at the beginning of this paragraph, simulating data based on the prespecified conditions, and identifying when power reaches a desired threshold (i.e., 80%; Muthén & Muthén, 2002). Examination of these resources should enable researchers to identify the appropriate design (i.e., number of dyads and number of time points per dyad) for examining physiological synchrony for an expected set of conditions (i.e., residual variance and effect size).

## Methods for indistinguishable dyads

As noted in the introduction, the three focal methods in this tutorial are best-suited for distinguishable dyads, or dyads whose members may be separated into two, categorically different, groups (e.g., mothers and children, patients and therapists, or teachers and students).

Data from indistinguishable dyads (e.g., same-sex romantic partners, friends, or twins) require different methods (see Kenny et al., 2006). For indistinguishable dyads, if responses from members of a given dyad were exchanged, the resulting data set would be as plausible as the original. Hence, a range of data sets equally represent a sample of indistinguishable dyads, and one could obtain different results when fitting the same model to each of those data sets (Kenny et al., 2006). Methods for indistinguishable dyads accommodate the indistinguishability by testing if the 'exchangeability' impacts the results. In principle, the methods presented for trend, concurrent, and lagged synchrony may be altered to accommodate indistinguishable dyads (see Kenny et al., 2006; Kashy, Donnellan, Burt, & McGue, 2008), but such alterations are beyond the scope of this tutorial. The one exception occurs for our proposed method for non-directional concurrent synchrony. The method could be applied to indistinguishable dyads because the estimate of synchrony corresponds to the correlation between dyad members' responses (i.e., correlations for a given dyad do not change if the member's responses are exchanged), an unintended benefit of that proposed approach.

## Limitations

An important caveat concerns the different time scales of the physiological responses. Each of the three signals (HR, RSA, and PEP) were produced on different time scales (2 sec epochs for HR, 30sec epochs for RSA and PEP), and were analyzed with different time scales (2 sec epochs for HR, 30sec epochs for RSA, and 1 minute aggregates for PEP). Although the different time scales matched the notions described in the *Physiological Data Collection* section, use of alternative time scales may produce different results. For example, our investigations of concurrent and lagged synchrony of RSA and PEP produced negligible dyadic effects (not reported in this tutorial), whereas HR produced significant effects. These discrepancies may be due to weaker physiological synchrony for RSA and PEP relative to HR, but they may also be due to the differences of time scale and number of epochs (i.e., the relatively small number of responses for RSA and PEP relative to HR). Therefore, future research is needed to identify the minimum sample size required to detect synchrony for the three different methods here, and whether different sample sizes will be needed for different indicators of physiological response systems.

An important distinction between concurrent and lagged synchrony involves the presence of an autoregressive component. The models match prior investigations of physiological synchrony, but the discrepancy raises questions about whether to include an autoregressive component when estimating concurrent synchrony. The inclusion of an autoregressive effect offers the opportunity to estimate the concurrent synchrony that occurs above and

beyond each partner's degree of intraindividual stability, and future research should identify the merits of its inclusion.

## Conclusions

Overall, we believe this article represents the first to summarize different modeling approaches for specific types of physiological synchrony (with short physiological series). We hope researchers will use this as a guide for examining trend, concurrent, and lagged physiological synchrony; that results from previously published studies will be more interpretable; and that others will continue to develop these approaches in order to better understand interpersonal physiology.

## Article information

**Conflict of Interest disclosures:** Each author signed a form for disclosure of potential conflicts of interest. No authors reported any financial or other conflicts of interest in relation to the work described.

**Ethical principles:** The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

**Funding:** This work was supported by Grant 2347.07 from the Fetzer Institute.

**Role of the funders/sponsors:** None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

**Acknowledgments:** The ideas and opinions expressed herein are those of the authors alone, and endorsement by the authors' institutions, or the Fetzer Institute, is not intended and should not be inferred.

## References

Albert, F. D. (1994). Structural equation modeling with longitudinal data: Strategies for examining group differences and reciprocal relationships. *Journal of Consulting and Clinical Psychology*, *62*, 477–487. doi:10.1037/0022-006X.62.3.477

Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1993). Respiratory sinus arrhythmia: Autonomic origins, physiological mechanisms, and psychophysiological implications. *Psychophysiology*, *30*, 183–196. doi:10.1111/j.1469-8986.1993.tb01731.x

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve modeling*. Hoboken, NJ: John Wiley & Sons.

Boker, S. M. (2012). Dynamical systems and differential equation models of change. In H. Cooper, A. Panter, P. Camic, R. Gonzalez, D. Long, & K. Sher (Eds.), *APA handbook of research methods in psychology* (Vol. 3, pp. 323–333). Washington, DC: American Psychological Association.

Boker, S. M., & Laurenceau, J. P. (2006). Dynamical systems modeling: An application to the regulation of intimacy and disclosure in marriage. In T. Walls & J. Schafer (Eds.), *Models for intensive longitudinal data* (pp. 195–218). New York, NY: Oxford University Press.

Boker, S. M., Leibenluft, E., Deboeck, P. R., Virk, G., & Postolache, T. T. (2008). Mood oscillations and coupling between mood and weather in patients with rapid cycling bipolar disorder. *International Journal of Child Health and Human Development*, *1*, 181–202.

Boker, S. M., & Nesselroade, J. R. (2002). A method for modeling the intrinsic dynamics of intraindividual variability: Recovering the parameters of simulated oscillators in multi-wave panel data. *Multivariate Behavioral Research*, *37*, 127–160. doi:10.1207/S15327906MBR3701_06

Boker, S. M., Deboeck, P. R., Edler, C., & Keel, P. K. (2010). Generalized local linear approximation of derivatives from time series. In S. M. Chow, E. Ferrer, & F. Hsieh (Eds.), *Statistical methods for modeling human dynamics: An interdisciplinary dialogue* (pp. 161–178). New York, NY: Routledge/Taylor & Francis Group.

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. (Eds.). (2007). *Handbook of psychophysiology*. United Kingdom: Cambridge University Press.

Chatel-Goldman, J., Congedo, M., Jutten, C., & Schwartz, J. L. (2014). Touch increases autonomic coupling between romantic partners. *Frontiers in Behavioral Neuroscience*, *8*, 95. doi:10.3389/fnbeh.2014.00095

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Earlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159. doi:10.1037/0033-2909.112.1.155

Cook, W. L., & Kenny, D. A. (2005). The actor-partner interdependence model: A model of bidirectional effects in developmental studies. *International Journal of Behavioral Development*, *29*, 101–109. doi:10.1080/01650250444000405

Curran, P. J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research*, *38*, 529–569. doi:10.1207/s15327906mbr3804_5

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

De Fraine, B., Van Damme, J., & Onghena, P. (2007). A longitudinal analysis of gender differences in academic self-concept and language achievement: A multivariate multilevel latent growth approach. *Contemporary Educational Psychology*, *32*, 132–150. doi:10.1016/j.cedpsych.2006.10.005

Deboeck, P. R. (2010). Estimating dynamical systems: Derivative estimation hints from Sir Ronald A. Fisher. *Multivariate Behavioral Research*, *45*, 725–745. doi:10.1080/00273171.2010.498294

Deboeck, P. R., Nicholson, J. S., Bergeman, C. S., & Preacher, K. J. (2013). From modeling long-term growth to short-term

fluctuations: Differential equation modeling is the language of change. In R. E. Millsap, L. A. van der Ark, D. M. Bolt, & C. M. Woods (Eds.), *New Developments in Quantitative Psychology* (pp. 427–447). New York, NY: Springer.

Dennis, T. A., Buss, K. A., & Hastings, P. D. (2012). Physiological measures of emotion from a developmental perspective: State of the science. In *Monographs of the Society for Research in Child Development, 77, Serial No. 303*. New York: Wiley. doi:10.1111/j.1540-5834.2011.00653.x

Dietrich, J., Jokisaari, M., & Nurmi, J. E. (2012). Work-related goal appraisals and stress during the transition from education to work. *Journal of Vocational Behavior*, *80*, 82–92. doi:10.1016/j.jvb.2011.07.004

Eckberg, D. L. (1997). Sympathovagal balance a critical appraisal. *Circulation*, *96*, 3224–3232. doi:10.1161/01.CIR.98.23.2643

Falk, C. F. (2017). Are Robust Standard Errors the Best Approach for Interval Estimation with Nonnormal Data in Structural Equation Modeling? *Structural Equation Modeling*, *25*, 244–266. doi:10.1080/10705511.2017.1367254

Feldman, R. (2003). Infant-mother and infant-father synchrony: The coregulation of positive arousal. *Infant and Mental Health*, *24*, 1–23. doi:10.1002/imhj.10041

Feldman, R. (2012). Bio-behavioral synchrony: A model for integrating biological and microsocial behavioral processes in the study of parenting. *Parenting*, *12*, 154–164. doi:10.1080/15295192.2012.683342

Feldman, R., Gordon, I., & Zagoory-Sharon, O. (2011). Maternal and paternal plasma, salivary, and urinary oxytocin and parent–infant synchrony: Considering stress and affiliation components of human bonding. *Developmental Science*, *14*, 752–761. doi:10.1111/j.1467-7687.2010.01021.x

Feldman, R., Singer, M., & Zagoory, O. (2010). Touch attenuates infants' physiological reactivity to stress. *Developmental Science*, *13*, 271–278. doi:10.1111/j.1467-7687.2009.00890.x

Felmlee, D. H., & Greenberg, D. F. (1999). A dynamic systems model of dyadic interaction. *The Journal of Mathematical Sociology*, *23*, 155–180. doi:10.1080/0022250X.1999.9990218

Ferrer, E., & Helm, J. L. (2013). Dynamical systems modeling of physiological coregulation in dyadic interactions. *International Journal of Psychophysiology*, *88*, 296–308. doi:10.1016/j.ijpsycho.2012.10.013

Ferrer, E., & McArdle, J. (2003). Alternative structural models for multivariate longitudinal data analysis. *Structural Equation Modeling*, *10*, 493–524. doi:10.1207/S15328007SEM1004_1

Field, T. (1994). The effects of mother's physical and emotional unavailability on emotion regulation. *Monographs of the Society for Research in Child Development*, *59*, 208–227. doi:10.1111/j.1540-5834.1994.tb01286.x

Fogel, A. (1993). *Developing through relationships: Origins of communication, self, and culture*. Chicago: University of Chicago Press.

Goldstein, H. (2011). Bootstrapping in multilevel models. In J. J. Hox & J. K. Roberts (Eds.), *Handbook of advanced multilevel analysis* (pp. 163–172). New York, NY: Routledge.

Gottman, J. M. (1979). Detecting cyclicity in social interaction. *Psychological Bulletin*, *86*, 338–348. doi:10.1037/0033-2909.86.2.338

Gottman, J. M. (1981). *Time-series analysis*. Cambridge: CUP.

Greenland, S., Maclure, M., Schlesselman, J. J., Poole, C., & Morgenstern, H. (1991). Standardized regression coefficients: A further critique and review of some alternatives. *Epidemiology*, *2*, 387–392. http://www.jstor.org/stable/20065707. doi:10.1097/00001648-199109000-00015

Griffin, D., & Gonzalez, R. (1995). Correlational analysis of dyad-level data in the exchangeable case. *Psychological Bulletin*, *118*, 430–439. doi:10.1037/0033-2909.118.3.430

Grimm, K. J., Ram, N., & Estabrook, R. (2016). *Growth modeling: Structural equation and multilevel modeling approaches*. New York, NY: Guilford.

Hamaker, E. L., & Grasman, R. P. (2015). To center or not to center?Investigating inertia with a multilevel autoregressive model. *Frontiers in Psychology*, *5*, 1492. doi:10.3389/fpsyg.2014.01492

Hastings, P. D. (2009). *A psychophysiological study of mothers' compassionate love and children's prosocial development. Contract 2347.07*. Kalamazoo, MI: The Fetzer Institute.

Helm, J. L., Sbarra, D., & Ferrer, E. (2012). Assessing cross-partner associations in physiological responses via coupled oscillator models. *Emotion*, *12*, 748–762. doi:10.1037/a0025036

Helm, J. L., Sbarra, D. A., & Ferrer, E. (2014). Coregulation of respiratory sinus arrhythmia in adult romantic partners. *Emotion*, *14*, 522–531. doi:10.1037/a0035960

Henning, R. A., Boucsein, W., & Gil, M. C. (2001). Social–physiological compliance as a determinant of team performance. *International Journal of Psychophysiology*, *40*, 221–232. doi:10.1016/S0167-8760(00)00190-2

Hertzog, C., Lindenberger, U., Ghisletta, P., & von Oertzen, T. (2006). On the power of multivariate latent growth curve models to detect correlated change. *Psychological Methods*, *11*, 244–252. doi:10.1037/1082-989X.11.3.244

Järvelä, S., Kivikangas, J. M., Kätsyri, J., & Ravaja, N. (2014). Physiological linkage of dyadic gaming experience. *Simulation & Gaming*, *45*, 24–40. doi:10.1177/1046878113513080

Jongerling, J., Laurenceau, J. P., & Hamaker, E. L. (2015). A multilevel AR(1) model: Allowing for inter-individual differences in trait-scores, inertia, and innovation variance. *Multivariate Behavioral Research*, *50*, 334–349. doi:10.1080/00273171.2014.1003772

Kahle, S., Miller, J. G., Lopez, M., & Hastings, P. D. (2016). Sympathetic recovery from anger is associated with emotion regulation. *Journal of Experimental Child Psychology*, *142*, 359–371. doi:10.1016/j.jecp.2015.10.004

Kashy, D. A., Donnellan, M. B., Burt, S. A., & McGue, M. (2008). Growth curve models for indistinguishable dyads using multilevel modeling and structural equation modeling: The case of adolescent twins' conflict with their mothers. *Developmental Psychology*, *44*, 316–329. doi:10.1037/0012-1649.44.2.316

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York, NY: Guilford.

King, G. (1985). How not to lie with statistics: Avoiding common mistakes in quantitative political science. *American Journal of Political Science*, *30*, 666–687. doi:10.2307/2111095

Liu, S. (2017). Person-specific versus multilevel autoregressive models: Accuracy in parameter estimates at the population and individual levels. *British Journal of Mathematical and Statistical Psychology*, *70*, 480–498. doi:10.1111/bmsp.12096

Liu, S., Rovine, M. J., Klein, L. C., & Almeida, D. M. (2013). Synchrony of diurnal cortisol patterns in couples. *Journal of Family Psychology*, *27*, 579–588. doi:10.1037/a0033735

Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, *1*, 86–92. doi:10.1027/1614-2241.1.3.86.

MacCallum, R. C., Kim, C., Malarkey, W. B., & Kiecolt-Glaser, J. K. (1997). Studying multivariate change using multilevel models and latent curve models. *Multivariate Behavioral Research*, *32*, 215–253. doi:10.1207/s15327906mbr3203_1

Miller, J. G., Kahle, S., & Hastings, P. D. (2015). Roots and benefits of Costly Giving: Children who are more altruistic have greater autonomic flexibility and less family wealth. *Psychological Science*, *26*, 1038–1045. doi:10.1177/0956797615578476

Miller, J. G., Kahle, S., Lopez, M., & Hastings, P. D. (2015). Compassionate love buffers stress-reactive mothers from fight-or-flight parenting. *Developmental Psychology*, *51*, 36–43. doi:10.1037/a0038236

Mitchell, C. M., Beals, J., & Kaufman, C. E. (2006). Alcohol use, outcome expectancies, and HIV risk status among American Indian youth: A latent growth curve model with parallel processes. *Journal of Youth and Adolescence*, *35*, 726–737. doi:10.1007/s10964-006-9103-0

Moeller, J. (2015). A word on standardization in longitudinal studies: Don't. *Frontiers in Psychology*, *6*, 1389. doi:10.3389/fpsyg.2015.01389

Montpetit, M. A., Bergeman, C. S., Deboeck, P. R., Tiberio, S. S., & Boker, S. M. (2010). Resilience-as-process: Negative affect, stress, and coupled dynamical systems. *Psychology and Aging*, *25*, 631–640. doi:10.1037/a0019268

Moore, G. A., & Calkins, S. D. (2004). Infants' vagal regulation in the still-face paradigm is related to dyadic coordination of mother-infant interaction. *Developmental Psychology*, *40*, 1068–1080. doi:10.1037/0012-1649.40.6.1068

Moore, G. A., Hill-Soderlund, A. L., Propper, C. B., Calkins, S. D., Mills-Koonce, W. R., & Cox, M. J. (2009). Mother-infant vagal regulation in the face-to-face still-face paradigm is moderated by maternal sensitivity. *Child Development*, *80*, 209–223. doi:10.1111/j.1467-8624.2008.01255.x

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén&Muthén.

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, *9*, 599–620. doi:10.1207/S15328007SEM0904_8

Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., … Boker, S. M. (2016). OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, *81*, 535–549. doi:10.1007/s11336-014-9435-8

Neale, M. C., & Miller, M. B. (1997). The use of likelihood-based confidence intervals in genetic models. *Behavior Genetics*, *27*, 113–120. doi:10.1023/A:1025681223921

Olsen, J. A., & Kenny, D. A. (2006). Structural equation modeling with interchangeable dyads. *Psychological Methods*, *11*, 127–141. doi:10.1037/1082-989X.11.2.127

Palumbo, R. V., Marraccini, M. E., Weyandt, L. L., Wilder-Smith, O., McGee, H. A., Liu, S., … Goodwin, M. S. (2017). Interpersonal autonomic physiology: A systematic review of the literature. *Personality and Social Psychology Review*, *21*, 99–141. doi:10.1177/1088868316628405.

Papp, L. M., Pendry, P., & Adam, E. K. (2009). Mother-adolescent physiological synchrony in naturalistic settings: Within-family cortisol associations and moderators. *Journal of Family Psychology*, *23*, 882–894. doi:10.1037/a0017147

Porges, S. W., Bohrer, R. E., Cheung, M. N., Drasgow, F., McCabe, P. M., & Keren, G. (1980). New time-series statistic for detecting rhythmic co-occurrence in the frequency domain: The weighted coherence and its application to psychophysiological research. *Psychological Bulletin*, *88*, 580–587. doi:10.1037/0033-2909.88.3.580

Ram, N., & Grimm, K. (2007). Using simple and complex growth models to articulate developmental change: Matching theory to method. *International Journal of Behavioral Development*, *31*, 303–316. doi:10.1177/0165025407077751

Ram, N., & Grimm, K. J. (2015). Growth curve modeling and longitudinal factor analysis. In W. Damon, R. M. Lerner, W. Overton, & P. C. M. Molenaar (Eds.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (7th ed., pp. 758–788). Hoboken, NJ: Wiley.

Rovine, M. J., & Walls, T. A. (2006). Multilevel autoregressive modeling of interindividual differences in the stability of a process. In T. A. Walls &J. L. Schafer (Eds.), *Models for intensive longitudinal data* (pp. 124–147). New York, NY: Oxford University Press.

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variable analysis: Applications to developmental research* (pp. 399–419). Thousand Oaks, CA: Sage.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514. doi:10.1007/BF02296192

Savalei, V. (2014). Understanding robust corrections in structural equation modeling. *Structural Equation Modeling*, *21*, 149–160. doi:10.1080/10705511.2013.824793

Scherbaum, C. A., & Ferreter, J. M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods*, *12*, 347–367. doi:10.1177/1094428107308906

Schuurman, N. K., Ferrer, E., de Boer-Sonnenschein, M., & Hamaker, E. L. (2016). How to compare cross-lagged associations in a multilevel autoregressive model. *Psychological Methods*, *21*, 206–221. doi:10.1037/met0000062

Sherwood, A., Allen, M. T., Fahrenberg, J., Kelsey, R. M., Lovallo, W. R., & van Doornen, L. J. (1990). Methodological guidelines for impedance cardiography. *Psychophysiology*, *27*, 1–23. doi:10.1111/j.1469-8986.1990.tb02171.x

Shumway, R. H., & Stoffer, D. S. (2006). *Time series analysis and its applications: With R examples*. New York, NY: Springer.

Snijders, T., & Bosker, R. (2012). *Multilevel analysis: An introduction to basic and applied multilevel analysis* (2nd ed.). London: Sage.

Timmons, A. C., Margolin, G., & Saxbe, D. E. (2015). Physiological linkage in couples and its implications for individual and interpersonal functioning: A literature review. *Journal of Family Psychology*, *29*, 720–731. doi:10.1037/fam0000115

Tronick, E. Z. (1989). Emotions and emotional communication in infants. *American Psychologist*, *44*, 112–119. doi:10.1037/0003-066X.44.2.112

Troxel, N. (2017). *Attachment security and childhood adjustment: Mediation by emotional competence* (Unpub-

lished doctoral dissertation). University of California, Davis.

Tukey, J. W. (1954). Causation, regression, and path analysis. In O. Kempthorne, T. A. Bancroft, J. W. Gowen, & J. L.Lush (Eds.), *Statistics and mathematics in biology* (pp. 35–66). Ames: Iowa State College Press.

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

Waters, S. F., West, T. V., & Mendes, W. B. (2014). Stress contagion: Physiological covariation between mothers and infants. *Psychological science*, *25*, 934–942. doi:10.1177/0956797613518352

Woody, E., & Sadler, P. (2005). Structural equation models for interchangeable dyads: Being the same makes a difference. *Psychological Methods*, *10*, 139–158. doi:10.1037/1082-989X.10.2.139

Wothke, W. (1993). Nonpositive definite matrices in structural modeling. *Sage Focus Editions*, *154*, 256–293.