Measuring physiological influence in dyads: A guide to designing, implementing, and analyzing dyadic physiological studies

Katherine R. Thorson[a,1]

Tessa V. West[a]

Wendy Berry Mendes[b]

In Press at *Psychological Methods*

[a]Department of Psychology, New York University, New York, NY, USA

[b]Department of Psychiatry, University of California San Francisco, San Francisco, CA, USA

[1]Correspondence concerning this article should be addressed to Katherine Thorson, 6 Washington Place, New York, NY 10003. Email: katherine.thorson@nyu.edu.

Author's note: The authors thank William Brady and Diego Reinero for their feedback on an earlier version of this manuscript. This manuscript was not shared on any social media platform, website, or listserv prior to submission.

Abstract

Scholars across domains in psychology, physiology, and neuroscience have long been interested in the study of shared physiological experiences between people. Recent technological and analytic advances allow researchers to examine new questions about how shared physiological experiences come about. Yet, comprehensive guides that address the theoretical, methodological, and analytic components of studying these processes are lacking. The goal of this paper is to provide such a guide. We begin by addressing basic theoretical issues in the study of shared physiological states by presenting five guiding theoretical principles for making psychological inferences from *physiological influence*—the extent to which one dyad member's physiology predicts the other dyad member's physiology at a future time point. Second, keeping theoretical and conceptual concerns at the forefront, we outline considerations and recommendations for designing, implementing, and analyzing dyadic psychophysiological studies. In so doing, we discuss the different types of physiological measures one could use to address different theoretical questions. Third, we provide three illustrative examples in which we estimate physiological influence, using the stability and influence model. We conclude by providing detail about power analyses for the model and by comparing the strengths and limitations of this model to pre-existing models.

Keywords: Physiological Influence, Interpersonal Physiology, Dyadic Interaction, Multilevel Modeling

Scholars have utilized physiological approaches to capture psychological experiences of individuals—including emotions, motivations, and attention—since the early 20th Century (e.g., Cohen & Patterson, 1937; Darrow, 1929; Jacobson, 1930; Mittleman & Wolff, 1929). For example, early work by Carl Jung examined electrodermal activity as a measure of attention to different stimuli in healthy and clinical samples (Ricksher & Jung, 1908). Beginning in the 1950s, social scientists started to collect data from two or more people in interpersonal interactions to measure interdependence between their physiological states. Early work focused on how similarity between patients' and therapists' heart rates mapped onto behavioral processes such as rapport and antagonism (Coleman, Greenblatt, & Soloman, 1956; DiMascio, Boyd, & Greenblatt, 1957). Since that time, physiological influence has been used to study romantic couples, parent-child dyads, and newly-acquainted dyads and teams, and influence has been associated with relationship quality, individual differences like attachment, and the development of self-regulation and trust (Hill-Soderlund et al., 2008; Levenson & Gottman, 1983; Mitkidis, McGraw, Roepstorff, & Wallot, 2015; Suveg, Shaffer, & Davis, 2016; for reviews see Timmons, Margolin, & Saxbe, 2015; Palumbo et al., 2016).

A primary strength of studying physiological influence in interpersonal encounters is that it allows scholars to test theoretical questions that are not testable using traditional measures of self-report or behavioral recordings alone. For example, physiological measures can provide continuous information about participants' emotional states—including those that are outside of awareness and may not be readily observable (Blascovich & Mendes, 2010); they are also not subject to the same demand effects that can bias self-reported data. In addition, because they are recorded unobtrusively, physiological data allow researchers to measure psychological processes without disrupting the natural dynamics of an interaction, which is critical for collecting

ecologically valid interpersonal data (e.g., how upset people feel during a conflict conversation or how engaged they are during a negotiation task). Lastly, concomitant psychophysiological measurements of dyads and groups can lead to novel insights and theoretical advancements regarding interpersonal dynamics, such as how quickly emotions "spread" through social groups and how shared emotional experiences in classrooms improve learning for students.

Given the number of benefits of collecting physiological data in dyadic interactions, interest in studying shared physiological experiences in interactions has increased, and so too has the number of methodological papers that discuss how dyadic physiological data can be analyzed (e.g., Butner, Amazeen, & Mulvey, 2005; Gates & Liu, 2016; Gottman, 1990; Helm, Sbarra, & Ferrer, 2012; Liu, Zhou, Palumbo, & Wang, 2016; McAssey, Helm, Fushing, Sbarra, & Ferrer, 2013). There have also been many theoretical and empirical papers that emphasize what psychological processes one can assess by examining similarities in partners' physiological states (e.g., Butler, 2011; Levenson & Ruef, 1992; Sbarra & Hazan, 2008; Waters, West, & Mendes, 2014; see Palumbo et al., 2016 and Timmons et al., 2015 for reviews). However, to our knowledge there is no single paper that provides a comprehensive framework for conducting a dyadic physiological program of research that addresses the critical steps of study design, data analysis, and interpretation of data, keeping theoretical considerations at the forefront of methodological decisions.

The goal of this paper is to provide a guide for researchers who plan to study the interplay between two (or more) individuals' physiological states. We focus on one type of physiological overlap in particular—*physiological influence*—the extent to which one dyad member's physiology predicts the other dyad member's physiology at a future time point, yet, many of the basic principles we discuss can be applied to other conceptualizations of

physiological interdependence. In using the term physiological influence, we do not mean to imply that one dyad member's physiological response causes the other's physiological response without any psychosocial process occurring between them. Rather, the term influence implies a specific temporal pattern—that one dyad member's physiology precedes another's in time—and, as we elaborate in the guiding principles of this paper, can be associated with psychosocial processes that occur between the dyad members (e.g., one person's heart rate influencing a partner's heart rate via a verbal outburst of anger).

The present approach of examining physiological influence differs from many models of physiological interdependence, which assess the degree of similarity between dyad members at the same time point (e.g., Chatel-Goldman, Congedo, Jutten, & Schwartz, 2014; Waters et al., 2014). We focus on this form of interdependence because it allows researchers to examine which member of a dyad influences the other and how. For example, researchers could examine how mothers soothe infants (Bernard, Kashy, Levendosky, Bogat, & Lonstein, 2017), how spouses regulate each other's emotions (Reed, Barnard, & Butler, 2015), and how higher-status people influence others (Kraus & Mendes, 2014).

This approach is particularly well-suited for experimental methods where researchers manipulate a process in one dyad member and then examine the impact that process has on the other dyad member via physiological influence. For example, Waters et al. (Waters, West, Karnilowicz, & Mendes, 2017) tested whether infants become more physiologically influenced by their stressed mothers when those mothers were able to touch their infants—a behavior that is theorized to communicate stress and therefore potentiates "stress contagion" via physiological influence. Manipulating stress and touch allowed the researchers to gain insight into what factors can cause the infant to become physiologically influenced by the mother and, critically, to test

the hypothesis that the mother's physiological state at one moment preceded the infant's physiological state.

In addition, most dyadic research involves distinguishable dyads (i.e., dyads where the two members can be distinguished on a meaningful theoretical factor, such as a parent and a child or a teacher and a student; Kenny, Kashy, & Cook, 2006), and distinguishing factors may have theoretical implications for physiological influence. For example, a parent may be more likely to influence a child than vice versa under certain conditions (e.g. when the parent is soothing a child), but the child may be more likely to influence the parent under other conditions (e.g., when the child is throwing a tantrum). A model of physiological influence can uncover directionality and can also incorporate features of the interaction that may change over time (e.g., which minutes involve soothing and which involve the child having tantrums) to test more nuanced hypotheses about when one dyad member (e.g., the child) influences the other. Approaches that only assess similarity at the same time point can restrict the range of possible theoretical questions that one can test regarding who is influencing whom and when.

We start by providing five guiding theoretical principles for making psychological inferences from physiological influence. Second, we address methodological issues in study design, including selecting the most appropriate physiological measure for certain theoretical questions, "scoring" one's data, and choosing behavioral measures that can provide insight into the process through which physiological influence is potentiated. Third, we present an analytic approach to measuring the correspondence between two partners' physiological states based on the well-established Actor-Partner Interdependence Model (APIM; Kashy & Kenny, 2000). We use a version of the APIM, known as the stability and influence model, which considers both how a person's physiology at one time point is predicted by his or her own physiology at the

prior time point (the stability effect) and by his or her partner's physiology at the prior time point (the influence effect). In this way, the model examines both autoregressive (stability) and cross-lagged (influence) effects. We illustrate this stability and influence approach with three examples and include annotated syntax. We then provide a guide of how to conduct power analyses for the model. Finally, we compare the strengths and limitations of this model to pre-existing models, with an emphasis on the different types of theoretical questions they address. By bridging the gap between theoretical and analytic issues in the study of physiological interdependence between dyad members, we hope to provide a guide for researchers who vary in their levels of expertise of physiological measures and dyadic data—from novices to well-seasoned scholars.

State of the Field: Past Approaches to Measuring Dyadic Correspondence of Physiological States

Before presenting guiding principles for studying physiological influence specifically, we provide a brief overview of the different ways in which the interdependence between two people's physiological responses has been conceptualized and measured more generally. To begin, there are a variety of terms for measures of interdependence in partners' physiology, including attunement, concordance, contagion, coregulation, coupling, covariation, entrainment, influence, linkage, and synchrony (Bachrach, Fontbonne, Joufflineau, & Ulloa, 2015; Chatel-Goldman et al., 2014; Helm, Sbarra, & Ferrer, 2014; Hill-Soderlund et al., 2008; Levenson & Gottman, 1983; Papp, Pendry, & Adam, 2009; Stratford, Lal, & Meara, 2012; Wass, Clackson, Cook, & de Barbaro, 2015; Waters et al., 2014). As noted by others (e.g., Bernard et al., 2017; Butler, 2011; Palumbo et al., 2016; Timmons, et al., 2015), occasionally these terms imply conceptual differences in what is being measured or are used to refer to specific analytic techniques or theoretical approaches but this is inconsistent. For example, the terms *compliance,*

coupling, and *synchrony* have all been used to refer to processes analyzed by the same approach: correlating two partners' physiology at the same time point (Chatel-Goldman et al., 2014; Henning, Boucsein, & Gil, 2011; Suveg et al., 2016). In contrast, the term *coregulation* has been used to refer to processes analyzed by different statistical methods: the degree that one partner's physiology predicts another's at a following time point (Helm et al., 2014) versus the same time point (Lunkenheimer et al., 2015).

Just as a variety of terms have been used to refer to the relationships between two or more people's physiological responses, there has also been diversity in the analytic options used to assess these relationships. We outline four differences here. First, some techniques consider the relationship between two people's physiology at the same time point (e.g., Papp, Pendry, Simon, & Adam, 2013; Saxbe & Repetti, 2010; Waters et al., 2014), which is important for examining shared experiences. Other techniques use a time-lagged design to examine whether one partner's physiology at one time point predicts the other partner's physiology at a later time point (e.g., Helm et al., 2014; Kraus & Mendes, 2014; Liu, Rovine, Klein, & Almeida, 2013), which is important for examining whether one partner's physiological state is predicted by the other partner's state.

Second, analytic approaches differ in how they handle autocorrelation—how stable or similar individuals' physiological responses are from one moment to the next. Some approaches model autocorrelation as a fixed effect when estimating the degree of shared physiology between two or more people (e.g., Feldman, Magori-Cohen, Galili, Singer, & Louzoun, 2011; Helm et al., 2014; Suveg et al., 2016), while others do not (e.g., Chatel-Goldman et al., 2014; Reed, Randall, Post, & Butler, 2012). The way in which autocorrelation is handled in analytic models of influence can impact other effects in the model (e.g., the significance and direction of influence

between partners), and so even if it is not of interest theoretically, it should be of interest empirically.

Third, approaches differ in how they handle the repeated nature of physiological data. In some models, repeated measurements are first averaged over time, usually by calculating a correlation for each dyad that represents the degree to which physiological states are similar, where X values are one dyad member's physiology at each time point and Y values are the other dyad member's physiology at each time point (e.g., Ebisch et al., 2012). These approaches typically provide one value (such as a correlation) for each dyad across all of the time points (e.g., Henning et al., 2011). These dyad-level estimates can then be analyzed in a subsequent model, perhaps to see whether dyad-level estimates are associated with other outcomes, like rapport, for example. Such a method is useful for answering questions about influence that do not involve hypotheses about changes over time, but rather, involve correlating influence with other processes of interest (e.g., rapport). In contrast, with other methods (such as multilevel modeling; e.g., West, Koslov, Page-Gould, Major, & Mendes, 2017), each time point of data from each participant is used in an analysis (rather than having been averaged ahead of the analysis for each dyad) and an average estimate of correspondence for the whole sample is obtained. Such a method is useful for answering questions about influence that involve hypotheses about changes over time or association with other time-variant processes.

A fourth difference is whether dyadic behaviors can be incorporated into one's model, which are important because they can provide insight into how and when two people share physiological states (e.g., Ham & Tronick, 2009; Reed et al., 2013). For example, Feldman and colleagues (Feldman et al., 2011) incorporated behaviors between mothers and infants to examine how these behaviors potentiated physiological influence. They found that interbeat

intervals (IBIs) between mothers and infants were most related to one another when mothers and infants were matched on positive affect and emitted positive vocalizations at the same time.

In summary, the study of shared physiological states has been around for decades, resulting in considerable diversity in how dyadic physiological data are analyzed and in the terminology used to refer to shared physiological states (for a similar conclusion, see also Gates & Liu, 2016, Palumbo et al., 2016, and Timmons et al., 2015). In this paper, we outline many of these differences and discuss the numerous methodological and analytic decisions that researchers must make when studying physiological influence. We emphasize thorough reporting of these decisions, with the goal of improving the synthesis of findings across the field.

Part 1: Guiding Principles

We next provide five guiding principles for the study of physiological influence in dyads to help researchers as they confront the challenge of making psychological inferences from physiological influence. Throughout, we emphasize the importance of context, as well as the collection and analysis of multiple streams of data, acknowledging that no one measure can perfectly capture a person's psychological experience.

Principle 1: The same physiological response may be associated with different psychological processes for each member of a dyad.

Many psychological states are associated with similar changes in physiological responses (Kreibig, 2010; Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2008; Mendes, 2016). For example, sympathetic nervous system (SNS) activity, which generally represents intensity of affective states (Mendes, 2016), can be interpreted as stress when individuals undergo a social-evaluative task, like the Trier Social Stress Task (Kirschbaum, Pirke, & Hellhammer, 1993), as positive emotion if someone is viewing exciting images (Shiota, Neufeld, Yeung, Moser, &

Perea, 2011) or as anger if someone is being harassed (Mauss, Cook, Cheng, & Gross, 2007). Two partners may even experience the same context very differently, with their SNS arousal representing different psychological states that they are experiencing. For example, after a social-evaluative task, two colleagues may both show high SNS arousal, but the high-status colleague may subjectively experience this as excitement, while the lower-status colleague feels stressed. As such, researchers should not assume that the same physiological response in two partners indicates that they are both experiencing the same psychological state.

In addition, emotion theory has a long history on the temporal relationship between subjective experiences and physiological changes dating back to at least James-Lange and Cannon-Bard theories (see Gross & Barrett, 2011; Manstead, 2012 for contemporary perspectives)—for example, does the high-status colleague experience a feeling of excitement which manifests as SNS arousal or does the colleague experience SNS arousal and subsequently label that arousal as excitement? While a full review of these theories and their evidence is beyond the scope of this paper, we note that the relationships among physiology, subjective experiences, and behavior are complex and may not always follow a specific temporal pattern (e.g., subjective experience precedes physiological response) and that these processes may also unfold differently in two members of a dyad. Thus, in this paper, we assume that the researchers have a theoretical understanding of the temporal relationship between the specific physiological and psychological states of interest as they begin study design.

Principle 2: The psychological process(es) that are associated with a particular physiological response in one member of a dyad might be different than the psychological processes that are associated with physiological influence from one dyad member to the other.

Given Principle 1, influence on the same physiological response does not necessarily imply that two dyad members are sharing a similar psychological or affective state. Using the example of two colleagues who experience SNS arousal in response to a social-evaluative task, influence during a dyadic interaction may imply that excitement is being passed from the high-status colleague to the low-status colleague, but it is also possible that influence on SNS activity reflects psychological attunement between people (i.e., the process of attending to one's partner's emotional or psychological states) rather than a shared experience of excitement. Indeed, several studies have focused on the relation between physiological influence of SNS activity and psychological processes related to attunement such as empathy and social sensitivity (Järvelä, Kivikangas, Kätsyri, & Ravaja, 2013; Guastello, Pincus, & Gunderson, 2006; Levenson & Ruef, 1992; Marci, Ham, Moran, & Orr, 2007). In this case, a psychological process that may be associated with SNS activity within an individual (e.g., stress) may be associated with interest and engagement in the partner, suggesting that influence can occur not because stress, specifically, is transferred, but because partners are attuned more generally with the fluxes and flows of each other's affective states.

We suggest that in order to fully understand what dyadic process physiological influence is capturing, and to understand the psychological states of each partner, researchers must go beyond identifying influence in responses and attempt to identify the individual experiences that participants are having. Though no one type of variable (e.g., physiological, behavioral, or subjective reports) can perfectly capture a person's experience, the consideration of multiple streams of information together can help researchers triangulate on a more precise understanding of psychological experiences. We discuss how the consideration of context, as well as measuring other variables, can be used to do this in Part 2 of the paper.

Principle 3: For physiological influence to provide information about a social process between two individuals during the time studied, the physiological state or process in one person (called the "sender") needs to result in information that can be "picked up" by the partner (called the "receiver").

Physiological influence between two dyad members is often related to social processes, such as the spread of emotions between people or the sensitivity of one dyad member to another (see Palumbo et al., 2016). For ease of clarity, we refer to the dyad member being influenced as the receiver and the person who is influencing as the sender. We posit that for physiological influence to be related to a social process between two people, physiology in the sender must be associated with signals that the receiver detects, consciously or non-consciously, that potentiate influence. Specifically, these signals can either 1) be effortlessly or automatically detected without awareness (i.e., via a low-level route), such as sensorial signals like odor, touch, visual displays, or voice frequency (see Liu et al., 2016), or 2) require effortful and motivated detection (i.e., via a high-level route) such as understanding whether crying is due to excitement or anxiety, that might necessitate attention or experience in decoding affective responses.

Principle 3 is predicated on two psychological processes that are thought to give rise to reading another person's psychological states during interpersonal encounters: one, the *perceptivity* of the receiver— the ability to be attuned to the sender's psychological states—and two, the *expressivity* of the sender— the quality and quantity of the cues that the sender gives off that reflect his or her psychological states (similar to Funder's RAM model; 1995, which refers to "good judges" and "good targets" in the study of accuracy, and also Gosling, Ko, Mannarelli, & Morris, 2002 in the study of personality detection). Because perceptivity and expressivity may be enhanced among close relationship partners (i.e., close relationship partners may be both more

perceptive of each other's psychological states and more expressive of their psychological states when with each other), physiological influence may be stronger between close relationship partners and people with existing relationships than between unacquainted people (Sbarra & Hazan, 2008).

Principle 4: Physiological influence may reflect exposure to shared stimuli or similar physiological milieu and not a social process between two people.

Physiological influence can occur in the absence of shared information about a partner's physiological state (e.g., when the two partners are in a separate room and are not communicating) but this process may reflect exposure to shared stimuli or similar physiological milieu and not a social process between the two individuals. Dyad researchers have long considered how a shared environment or stimulus can affect partners in simultaneous ways, which leads to interdependence in responses (e.g., "common fate model," also known as the latent group model; Gonzalez & Griffin, 2002; Kenny et al., 2006; Kenny & La Voie, 1985). For example, the correspondence between the physiology of two people may be high if those two individuals are watching the same movie, exposed to the same stressor, or are engaging in the same task, even if those two individuals do not have any contact with one another.

In addition, physiological influence in some contexts may be the result of a social process, but the observable cues that have created the influence may have occurred at a time point prior to the study context. Models of physiological and emotional convergence posit that one way close relationship partners become physiologically and emotionally linked is via shared appraisals (Anderson, Keltner, & John, 2003; Parkinson, 2011). For instance, a mother who fears dogs may teach her child to fear dogs. When both the mother and her child encounter a dog, they may both experience a fear response to the dog. Their physiological influence may be the result

of a social process (i.e., the mother teaching her child to fear dogs), but it may not explain how influence happens *during* the encounter (indeed, mother and child can be exposed to two different dogs in two different cities and show similar physiological responses). The social cues that create influence are found outside of the current encounter. In summary, to interpret influence as the result of a social process, researchers may need to rule out the influence of a shared environment first (see also Liu et al., 2016). If they want to interpret effects as being due to shared psychological process during the study, they also need to rule out shared processes that might have occurred prior to the study (especially when dyad members are close others).

Principle 5: To understand the process of influence, it is critical to incorporate psychological and/or behavioral variables from the receiver (the person doing the "catching") and the sender (the person doing the "transmitting") into the study design.

Elaborating on Principle 3, we argue in Principle 5 that to understand what potentiates influence, scholars need to show *when* influence occurs, and that optimal experimental methods of testing for factors that potentiate influence are those that directly target variables related to expressivity (e.g., access to cues) and perceptivity (e.g., motivation to and ability to attend to the partner). We suggest measuring or manipulating variables that are related to cues to partners' physiological states (such as their behaviors during the interaction), as well as variables related to perceiving those cues (e.g., motivation to attend to the partner, visibility of the partner, actual attention such as looking time) to understand the processes underlying influence.

We focus on how four different patterns of psychological and/or behavioral variables from the receiver, sender, or both can provide insight into the underlying psychological process of influence (see Table 1). Note that these variables can be either time-invariant (e.g., student's motivation to attend to teacher) or time-variant (e.g., mutual gaze between students and

teachers); both can be incorporated into one's analysis. These patterns are certainly not exhaustive but are meant to give researchers guidance of where to start. We use the example of a researcher conducting a study that examines physiological influence between students (the receivers) and teachers (the senders). The patterns describe variables related either to the receiver, the sender, or both that could potentiate physiological influence.

Table 1. Four patterns of variables that can be used to understand the processes underlying physiological influence.

Variable that potentiates influence	Factors captured	Example[a]
1. Receiver-only variable	Perceptivity of the receiver in detecting cues	Students' motivation to attend to teacher
2. Sender-only variable	Expressivity of the sender in providing cues	Teacher disciplining a student
3. Interaction between receiver and sender variables	Interaction between perceptivity of receiver and expressivity of sender	Students' motivation to attend to the teacher and teacher disciplining a student
4. Dyad-level variable	Transmission and perception of cue	Mutual gaze between students and teachers

Note. [a]Students are the receivers and teachers are the senders; students are physiologically influenced by their teachers (i.e., teachers' physiology predicts students').

First, receiver-only variables are related to perceptivity and could include those that involve empathic ability or motivation to attend to the sender. For example, a researcher might hypothesize that students who are more motivated to attend to their teachers are more influenced by them. Second, sender-only variables are related to the expressivity of the sender and map onto the theoretical construct of cue *validity*—that is, they capture the quality and quantity of behavioral indicators of the sender's physiological state. For example, a researcher might hypothesize that when a teacher engages in clear behaviors that are relevant to students (e.g., telling students to stop talking), her students will be more influenced by her.

Third, interactions between receiver and sender variables can be used to test whether the combination of perceptivity and expressivity give rise to influence; for example, students and teachers may influence each other when students are motivated to pay attention to teachers *and* when those teachers engage in behaviors relevant to students. An interaction between these variables builds on the prior two patterns by including the main effects of receiver and sender variables. A significant interaction implies that the combination of receiver and sender variables catalyzes influence.

Fourth, dyad-level variables that capture the expression and perception of a cue may explain physiological influence. For instance, students may become attuned to teachers through eye contact, leading to greater influence when students and teachers are looking at each other. In this case, mutual gaze both conveys the psychological state of the sender and potentiates attunement of that affective state by the receiver. Other dyad-level variables that capture the ability of the dyad members to perceive each other's cues, as well as their tendency to be expressive with each other—such as relationship length—may also be associated with influence.

In summary, we have reviewed four potential patterns that could underlie physiological influence, all of which are empirically testable within the framework that we outline in Part 3. Note that these variables can be either time-invariant (e.g., student's motivation to attend to teacher) or time-variant (e.g., mutual gaze between students and teachers); both can be incorporated into one's analysis. Having reviewed five guiding principles that can help researchers make psychological inferences from physiological influence, we now outline critical concerns for the design phase of research.

Part 2: Study Design

Physiological Measures

One of the first decisions for researchers to make is which physiological responses to measure for the two dyad members. We assume that most researchers will be interested in measures of reactivity from a baseline or resting phase where dyad members are separated or do not interact to a phase where both dyad members are interacting. As noted by others (Helm et al., 2014; Timmons et al., 2015), the inferences that can be drawn from physiological influence are largely determined by the response measured. The extent to which a given physiological response (or pattern of responses) reveals the presence or strength of a psychological process varies widely, and researchers should do their best to understand how a response and its context contribute to the psychological inferences one can make.

Dimensions of psychophysiological relationships. Cacioppo and colleagues have provided three dimensions along which psychophysiological relationships can be assessed: generality, specificity, and sensitivity (Cacioppo, Tassinary, and Berntson, 2007b). All of these dimensions help researchers consider the extent to which a physiological response relates to a psychological process. While a full review of these dimensions is beyond the scope of this paper, we review them briefly here. Generality refers to the extent to which a relationship is context-dependent—for example, respiratory sinus arrhythmia (RSA) decreases may indicate psychological distress during a stressor (Fisher & Newman, 2013) but cognitive effort and attention during executive functioning tasks (Hansen, Johnson, & Thayer, 2003). Specificity refers to the extent that a particular psychological process (and only that process) is associated with a physiological response. At the highest level of specificity, when a psychological process is activated, so too is the physiological response (or pattern of responses) and vice versa—the

process and response have a one-to-one relationship. Lastly, sensitivity refers to the extent to which a physiological response varies in degree with the psychological process. Highly sensitive relationships exist when small increases in the psychological process are reflected in changes in the physiological response. For instance, skin conductance is considered a sensitive measure because it changes with low levels of psychological and affective states and may vary with the intensity of the state. Cortisol reactivity, on the other hand, is a less sensitive measure, and small increases in acute stress may not result in detectable increases in cortisol.

Selecting a measure. We recommend two approaches that can be used when choosing which responses to measure and how to interpret them. As outlined in the guiding principles, additional measures, such as behaviors or self-reports, as well as consideration of the context should be used when taking either of these approaches. With the first approach, researchers use measures of physiological activity as an indicator of attunement between individuals. For example, because SNS activity (measured, for example, with pre-ejection period or skin conductance) can be interpreted broadly as a measure of affective intensity (Mendes, 2016), it is particularly useful for this purpose, especially in less-specific contexts where SNS activity may indicate any one of multiple psychological states. Rather than physiological influence indicating that individuals are becoming synchronized on a particular psychological state, physiological influence may reflect more general attunement to the fluxes and flows in a partner's affective intensity, indicating that both partners are experiencing similar intensity of affect (or one partner's intensity is following another partner's intensity) without experiencing the same affective state.

Using a previous example, a high-status colleague may feel excitement and show SNS arousal, and her lower-status colleague may attune to this arousal but interpret it as anxiety. As a

result, the lower-status colleague may feel anxiety and show SNS arousal. Although the two individuals did not experience the same psychological state, their SNS responses may nevertheless influence each other. Variability in the context and how it is interpreted may result in different affective experiences for two members of a dyad (Barrett, 2013; Quigley & Barrett, 2014) that could manifest in influence on the same physiological response. Though we have focused on influence in SNS measures as an indicator of attunement, future research may reveal other exciting measures that can be used for this purpose, such as electrical activity at certain frequencies from the brain (Dikker et al., 2016).

The second approach involves measuring the response(s) that is most likely to reflect the psychological state or process on which individuals are becoming synchronized. With this approach, influence reflects psychological attunement between partners (as in the first approach), but influence also goes beyond attunement to indicate a more specific state that individuals are sharing. Thus, the measure chosen should be carefully matched to the psychological process researchers are interested in capturing. For example, if mothers and children are becoming synchronized in the escalation of stress over a period of minutes, influence of SNS responses like pre-ejection period or skin conductance levels, which are related to arousal and sensitive to quick changes in affect (Mendes, 2016), might best capture that process (e.g., Manini et al., 2013; Waters et al., 2014). Perhaps a researcher is interested in another process altogether: understanding how parents soothe their children and socially bond with them. In this case, influence of parasympathetic nervous system (PNS) reactivity, which can be related to positive emotions and relaxation (Bazhenova, Plonskaia, & Porges, 2001; Houtveen, Rietveld, & De Geus, 2002), may be best (e.g., Lunkenheimer et al., 2015; Waters et al., 2017). We encourage researchers to consider the dimensions provided by Cacioppo and colleagues (Cacioppo et al.,

2007b) when selecting a measure for this approach. For instance, because the relationship between RSA and psychological distress is context-dependent (i.e., it is low on generality), the use of RSA influence as an indicator of shared psychological distress likely only applies in certain contexts.

When choosing a measure, researchers should keep in mind that variability in participants' responding must be present in order to detect physiological influence. Variability in responding may be limited by the measure (i.e., less sensitive measures provide less variability), the situation (e.g., some study contexts elicit stable responses throughout) or individual characteristics of participants (e.g., older and overweight individuals have less physiological flexibility and may have sluggish or dampened responses to shifts in context or affective states; Mendes, 2009).

Here, we have discussed the use of one physiological response in influence analyses (e.g., either pre-ejection period, cortisol, RSA, heart rate, electrodermal activity, etc.). Some research has relied on an array of physiological responses and determined the percentage of responses for which influence is observed (e.g., Levenson & Ruef, 1992). If using this approach, we recommend that researchers justify why each measure was collected or why the measures have been grouped together. There can be very large correlations between different measures of SNS activity, for example, and the extent to which the measures overlap might yield exaggerated evidence of influence. Ideally, in the interest of parsimony, we recommend that researchers attempt to capture the purest measure of the type of activity in which they are most interested theoretically (e.g., pre-ejection period for SNS activity or RSA for PNS activity) and analyze influence on that measure only and then expand to other measures that are closely aligned with a different biological system.

In sum, the choice of one's measure has to be informed by the psychological process one is studying. We encourage the interested reader to see Blascovich, Vanman, Mendes, & Dickerson, 2011; Cacioppo, Tassinary, & Berntson, 2007a; Mendes, 2016; and Siegel et al., 2016, for more extensive information on the relationship between various physiological measures and psychological states and processes. If researchers are unsure of what physiological measure best reflects the psychological process they want to study, we encourage pilot testing one's measures in the context and with the participants one intends to study, with careful attention paid to other variables—such as self-reports and behaviors—to make as accurate an inference as possible.

Cues Potentiating Influence

As outlined in Principle 3, physiological influence that is the result of a social process occurs when people exhibit cues that can be observed by their partners. Specific behavioral cues can be thought of as representing a latent construct that researchers aim to capture. For example, behavioral cues of hand fidgeting, nail biting, and hair twirling might all be considered cues that tap into the latent construct of anxiety. Prior research and pilot testing may be useful during this phase of research to figure out 1) the cues that are a result of one partner's psychological and physiological experience, 2) the time frame when those cues occur (e.g., immediately, 30 seconds later, days later, etc.), 3) whether those cues are either unconsciously or consciously picked up on by interaction partners, and 4) whether those cues are measurable.

There are several challenges with measuring behaviors and, just as with physiology and self-reports, no one behavioral measure provides a perfect window into someone's psychological experience. As noted above, some behaviors may be quite subtle, making them difficult to observe and to measure. Furthermore, different people may perceive the same behavioral cues in

different ways, making it important for researchers to establish inter-rater reliability prior to coding data. With dyadic psychophysiology studies, researchers may have to balance the desire to obtain an accurate physiological signal that is not distorted by movement with the desire to let participants freely engage in behaviors that may potentiate physiological influence. Some physiological signals are more robust to movement artifacts than others. For example, blood pressure responses and skin conductance are highly sensitive to movement, but, in general, electrocardiography and impedance cardiography are more robust to physical movement. Our recommendation is that researchers choose an ecologically valid study context in which natural social behaviors would not dramatically compromise the physiological signal being measured, and within this context, allow participants to move as freely as possible. For more information on behavioral coding and reliability statistics, we suggest Heyman, Lorber, Eddy, and West, 2013. Sometimes the cues that potentiate influence are too difficult to measure directly, and so one must manipulate variables necessary for the transmission of these cues. For example, pheromones communicated via smell may be one pathway through which anxiety contagion occurs. One could manipulate the ability to smell (e.g., using a mentholated topical cream under participants' noses) and then hypothesize that people who cannot smell will be less influenced by partners who have been stressed.

In this paper, we outline how cues and the guiding principles from Part 1 can be used within a stability and influence framework, but these suggestions and principles are also relevant for other conceptualizations of physiological interdependence. For example, researchers using within-time point correlations (i.e., estimating the degree of similarity between receivers and senders at the same time point) can examine receiver and sender variables as moderators of influence. Researchers who compute one correlation per dyad that represents similarity in

physiology can manipulate variables that capture the processes underlying influence (e.g., mutual gaze) between dyads and use that manipulation to predict the strength of correlations across dyads.

In sum, there are several ways researchers can use measured cues from a dyadic interaction to uncover the psychological meaning behind influence. Mapping behavioral cues onto physiological influence is relatively uncharted territory, and thus, these recommendations are not hard-and-fast rules for determining what influence means, but rather guidelines for using one's data to make the most accurate inferences possible.

Interval Length

When measuring physiology, researchers must consider the amount of time that one observation represents—interval length. Some measurements are obtained at one point in time (e.g., cortisol measurements from saliva samples and spot blood pressure recordings). However, other measurements are obtained continuously and commonly get averaged across a segment of time. Averaging across a specified time interval accomplishes the goal of obtaining a best estimate of the physiological reactivity for that time point, and, in some cases, the averaging process leads to more reliable estimates. We recommend three points to consider when choosing the length of time across which to average measurements.

First, we recommend prioritizing an accurate signal over more frequent measurements. For example, when estimating pre-ejection period (PEP) there may be an advantage in accurate estimation with averaged waves (ensembled) over single wave (beat-to-beat) estimates (Sherwood, et. al., 1990). Second, we suggest using the smallest interval in which a response could change without compromising the quality of the signal. The more slowly the response changes, the longer the interval can be. For example, measures from facial electromyography

(EMG) and electroencephalography (EEG) are detected within milliseconds after a stimulus, whereas SNS responses like PEP or skin conductance respond within a few seconds, and end products of hypothalamic-pituitary-adrenal (HPA) activation, like cortisol, take 15 to 20 minutes to detect psychological changes in the biological response. Third, we advise researchers to consider when a particular psychological state begins and ends and to use a corresponding interval length. For instance, in a study where participants solve a new math problem of varying difficulty every 30 seconds, we chose an interval length of 30 seconds for PEP to match the interval's beginning and end to the beginning and end of a psychological experience for the participant (Thorson, Forbes, Magerman, & West, 2017).

Longer intervals may mask meaningful changes in responding and have the potential to reverse the direction of the influence estimate (i.e., the effect of one dyad member's physiology on the other's at a following time point) obtained. In Figure 1, we show potential patterns in responding for two dyad members (termed "receiver" and "sender") who are both fluctuating around the grand mean but are doing so in opposition at the same time point. All else being equal, shorter intervals (as in Panel A) would likely yield a negative influence estimate: higher values from the sender are associated with negative values from the receiver at the following time point. However, longer intervals (as in Panel B) would likely yield a positive influence estimate: higher values from the sender are associated with positive values from the receiver at the following time point. We recommend graphing physiological values for individual dyad members over time to aid in the decision regarding interval length. If responding is relatively stable, shorter intervals may not be necessary. Syntax for graphing individual estimates is in the Supplemental Material (see Figure S1).

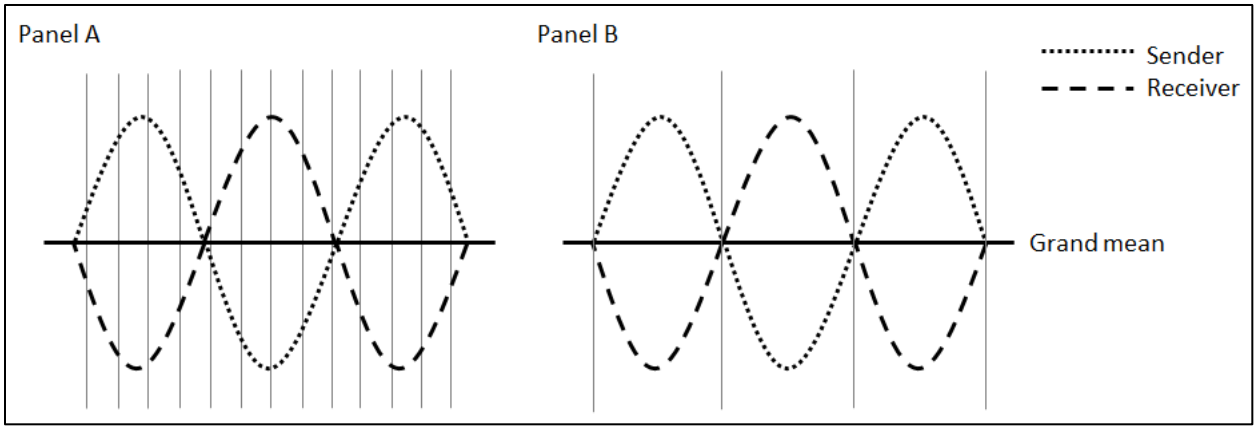

Figure 1. Example pattern of responding for two dyad members with different interval lengths. The interval length may change the direction of the influence estimate—the effect of one dyad member's physiology on the other's at a following time point. All else being equal, the shorter intervals in Panel A would produce a negative influence estimate, while the longer intervals in Panel B would produce a positive influence estimate.

Part 3: Analytic Model

In Parts 1 and 2, we presented guiding principles and considerations for study design that are relevant for anyone conducting a dyadic physiological program of research. Though a wide range of techniques can be used to estimate physiological influence, we next outline one statistical model that can be used in tandem with the guiding principles we presented. As described in Part 1, a major focus of our theoretical approach to studying influence is the integration of dyadic cues (for example, behaviors) to gain insight into how and when partners become physiologically linked. Thus, a primary strength of the model that we present is the flexibility to incorporate these cues, making the model particularly ideal for researchers intending to empirically test the processes underlying influence. This is especially important because relatively little work has examined these processes (see Liu et al., 2016). In this section, we now discuss how physiology of both partners, in combination with measured cues, can be incorporated into a single analytical approach. In Example 1, we demonstrate a basic model in which the stability and influence paths are estimated. In Example 2, we elaborate this model to

consider a behavioral variable as a moderator of influence; this model allows one to directly test the different ways in which influence is potentiated discussed in principles three and five. Lastly, in Example 3, we illustrate how time can be incorporated into the stability and influence model to examine whether stability and influence increase or decrease over time.

Stability and Influence Model

The stability and influence model (a special case of the APIM; Kenny et al., 2006) can be utilized when physiological measurements are collected from two members of a dyad repeatedly over time. We recommend that the approach is used when researchers have three or more repeated measures of autonomic physiology, as well as neuroendocrine measures like cortisol and testosterone and neurological measures like EEG. In this model, a participant's physiology score at one time point is treated as a function of his/her own physiology score at a prior time point (the stability path, which is an autoregressive effect) and his/her partner's physiology score at a prior time point (the influence path, which is a cross-lagged effect; see Figure 2). In this paper, we demonstrate how to estimate the stability and influence model using multilevel modeling (MLM). One can also use multilevel structural equation modeling (MSEM) to estimate the model. We note that MSEM requires many more data points than MLM (Hox, Maas, and Brinkhuis, 2010; see also Hox, 2013) and so an MSEM approach is more appropriate for large sample sizes.

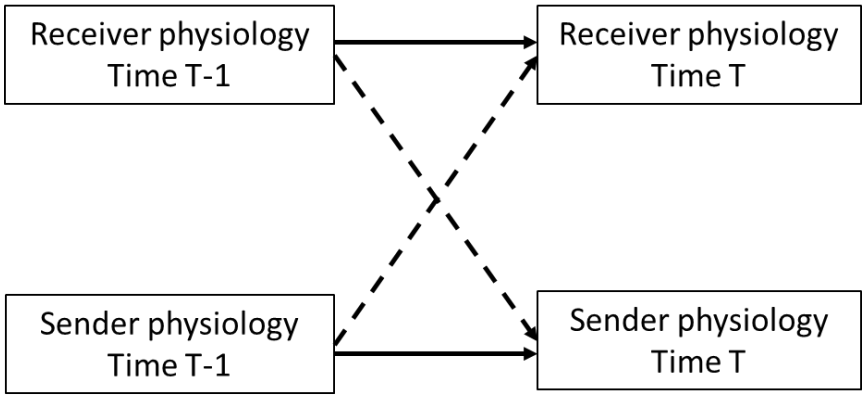

Figure 2. Stability and influence model. The solid lines represent the stability or autoregressive paths, where a dyad member's physiology at one time point predicts their own physiology at a later time point. The dashed lines represent the influence or cross-lagged paths, where a dyad member's physiology at one time point predicts the other dyad member's physiology at a later time point.

We next provide overviews of three topics that are important for building an analysis model: the structure of repeated measures dyadic data, centering of variables, and whether to use raw values or change scores. We then describe the analytic model, followed by three examples of utilizing the model for the distinguishable case, with an example of the indistinguishable case in the Supplemental Materials. We assume the reader has a basic knowledge of multilevel modeling, dyadic data analysis, and moderation, and we recommend the following resources if this is not the case: Aiken & West, 1991; Bauer & Curran, 2005; Bolger & Laurenceau, 2013; Fitzmaurice, Laird, & Ware, 2011; Kenny et al., 2006.

Structure of Repeated Measures Dyadic Data

When researchers collect data that contain physiological measures at multiple time points for both partners, there are three factors to consider: person, dyad, and repeated measure. If dyad members provide physiological data at the same time points, then the level of repeated measure is the same for both members of the dyad, and repeated measure and person are crossed (not nested). The stability and influence model assumes a two-level crossed design, which allows one

to test the correlation of errors within time point for the two dyad members (e.g., if one partner is particularly reactive at time T is the other partner also particularly reactive at time T). If dyad members provide data at *different* time points, then a three-level model can be estimated (see Chapter 13 of Kenny et al., 2006, for more details). To structure one's data, we refer to West (2013), who discusses the analysis of repeated measures dyadic data. As shown in Table 2 below, the data need to be structured as a *person period pairwise* file, with each person having a line of data for each time point. Structuring the data in this manner allows researchers to estimate stability and influence for both receivers and senders simultaneously. Lines should be inserted for missing data (e.g., data missing for a dyad member or for a particular time point), and these should be marked as missing in a program-specific manner (e.g., left blank in SAS or listed as "NA" in R).

Table 2. Example subset of data in a person period pairwise file format.

Dyad	Partner	Time	Receiver Condition	Sender Condition	Receiver PEP Reactivity	Sender PEP Reactivity	Receiver PEP Reactivity Lagged	Sender PEP Reactivity Lagged
1	1	1	-1	1	-7	1		
1	1	2	-1	1	-6	0	-7	1
1	1	3	-1	1	-3	-1	-6	0
1	1	4	-1	1	-2	2	-3	-1
1	2	1	1	-1	1	-7		
1	2	2	1	-1	0	-6	1	-7
1	2	3	1	-1	-1	-3	0	-6
1	2	4	1	-1	2	-2	-1	-3

Centering of Variables

Predictor terms in the stability and influence model can be centered in several different ways that necessitate different interpretations of the coefficients. We highlight three possibilities, none of which we recommend over the other, but each of which can be used to answer different theoretical questions of interest.

First, one can grand mean-center both the stability and influence terms (i.e., receiver physiological data at time T-1 and sender physiological data at time T-1). A positive influence coefficient indicates that higher values (i.e., values above the grand mean) of sender physiology are associated with higher values of receiver physiology at the following time point. A negative influence coefficient indicates that higher values of sender physiology are associated with lower values of receiver physiology at the following time point.

Second, a person-centering approach can be done by calculating means for receiver physiology and sender physiology across all time points for each person and subtracting these means from receiver and sender predictor terms. A positive influence coefficient indicates that when the sender is higher than he or she is on average, the receiver is higher at a following time point. A negative influence coefficient indicates that when the sender is higher than he or she is on average, the receiver is lower at a following time point.

Third, grand mean-centered and person-centered stability and influence terms can be included in one model if researchers are interested in disentangling between-person variability in physiological influence from within-person variability. The grand mean-centered influence terms provide information about between-person variability (i.e., how a sender having a higher or lower physiological value than the average person affects the receiver's physiology). Person-centered influence terms provide information about within-person variability (i.e., how a sender

having a higher or lower physiological value that his or her own average affects the receiver's physiology). Including both grand mean-centered and person-centered terms in the model simultaneously allows researchers to model between-person and within-person variability, while accounting for the influence of the other (for an example of this approach, see Inauen, Shrout, Bolger, Stadler, & Scholz, 2016).

Change Score versus Raw Value

Researchers' theoretical questions should dictate whether they use change scores (e.g., reactivity scores from a baseline interval to the dyadic interaction) or raw values in their analysis. If researchers want to understand influence on change scores (i.e., the physiological variable of interest is a change from one time point, like a baseline interval, to another, like the dyadic interaction), then a change score should be used as the dependent variable, as well as for the lagged receiver and sender predictor variables. Likewise, if researchers want to understand influence on raw values, then the raw value should be used as the dependent variable and the lagged receiver and sender variables.

If using change scores, researchers must decide what the initial measurement should be. A common approach in psychophysiological research is to subtract a resting/baseline value from each value during the period of interest so that observations represent changes from a resting state (Blascovich et al., 2011). However, researchers may be interested in changes from another point during the study: for example, from a stressful task prior to a dyadic interaction. Conceptually, physiological influence on these two types of change scores would be different. The first approach represents influence in reactivity, whereas the second approach represents influence in recovery responses. Researchers may not necessarily be interested in reactivity to a stimulus (or from a baseline) but instead may be interested in seeing how physiological

responses unfold over time in response to many different stimuli. In this case, a change score may not be appropriate for the conceptual question researchers are seeking to address and raw values can be used instead.

Lag Length

Thus far, we have mentioned lag lengths of one interval, such that responses at time T predict responses at time T+1. This decision makes sense when researchers predict that responding at one time point influences responding at the following time point. However, it is possible that influence does not happen at this speed and, instead, occurs over a longer distance in time. For example, lags of greater than one time point might be appropriate when the behaviors that facilitate influence are removed in time from the physiological response associated with them. For example, if a participant experiences elevated PEP reactivity during time T, but the psychological state associated with this does not appear behaviorally until time T+2, a lag of two or three intervals might be more appropriate. Longer lags may also make sense when it takes longer for people to pick up on the experiences of their partners or if participants exhibit a delay in physiological responding—potentially because the physiological response itself is slow-moving (e.g., electrogastrography) or because individuals show less flexibility in physiological responding (e.g., the elderly).

Finally, lags greater than one may be used in studies where the context is not the same during each interval and alternates in a similar manner across time. For example, an analysis of cortisol measured in parents and children three times a day (in the morning, after school, and at night) might be better-suited with a lag of three, such that morning responses predict morning responses etc., than a lag of one. The choice of lag length should be closely related to decisions regarding interval length, and both the study context and response measured should influence

how these choices are made. Although we focus on the study context and the response measured as considerations for lag length, we note that lag length is completely data-driven in some analytic approaches to modeling physiological interdependence (e.g., Reed et al., 2012; Scarpa et al., 2017). We include SAS syntax for lagging variables in the Supplemental Materials.

Stability and Influence Model to Estimate Physiological Influence

We now present the stability and influence model which can be used to yield fixed effects estimates of physiological influence, as well as autoregressive effects (see also Butler, 2011; Levenson & Ruef, 1992 on the importance of considering autocorrelation). The level 1 equation for the basic model with distinguishable dyads—one male and one female—is Equation 1. The outcome is the receiver's data at one time point. "R" represents the receiver's own data. "S" represents the sender's physiological data. The terms in the model are described in Table 3. The model presented in Equations 1 through 7 and Example 1 is a "two-intercept model," where intercepts for both males and females are estimated (Kenny et al., 2006). One advantage of this model is that fixed effects for both males and females can be obtained directly from the output. The disadvantage is that one cannot test whether these fixed effects differ as a function of gender. After describing the "two-intercept" approach, we show a different approach where one can test whether the fixed effects differ as a function of gender (or any other distinguishing factor).

$$Y_{ijt} = (b_{0mj})M + (b_{0fj})F + (b_{1mj})M(R_{mj(t-k)}) + (b_{1fj})F(R_{fj(t-k)}) + (b_{2mj})M(S_{mj(t-k)}) + (b_{2fj})F(S_{fj(t-k)}) + M(e_{mjt}) + F(e_{fjt}) \quad (1)$$

Table 3. Terms in Equation 1.

Term	Description
Y_{ijt}	Outcome for person *i* in dyad *j* at time *t*
b_{0mj}	Intercept for the male in dyad *j*
M	Dummy code; males are coded as 1 and females as 0
b_{0fj}	Intercept for the female in dyad *j*
F	Dummy code; females are coded as 1 and males as 0
b_{1mj}	Slope for the male in dyad *j* for receiver lag
$R_{mj(t-k)}$	Receiver value for the male in dyad *j* at time *t – k*, where k is the lag length
b_{1fj}	Slope for the female in dyad *j* for receiver lag
$R_{fj(t-k)}$	Receiver value for the female in dyad *j* at time *t – k*, where k is the lag length
b_{2mj}	Slope for the male in dyad *j* for sender lag
$S_{mj(t-k)}$	Sender value for the male in dyad *j* at time *t – k*, where k is the lag length
b_{2fj}	Slope for the female in dyad *j* for sender lag
$S_{fj(t-k)}$	Sender value for the female in dyad *j* at time *t – k*, where k is the lag length
e_{mjt}	Residual error for the male in dyad *j* at time *t*
e_{fjt}	Residual error for the female in dyad *j* at time *t*

The level 1 equation can be broken into six level-two equations (equations 2 through 7). Each equation has a fixed effects component plus a random effects component, which are outlined in Table 4. The random effects components are also outlined in Table S1 of the Supplemental Materials.

$$b_{0mj} = a_{0m} + u_{0mj} \quad (2)$$

$$b_{0fj} = a_{0f} + u_{0fj} \quad (3)$$

$$b_{1mj} = c_{0m} + u_{1mj} \quad (4)$$

$$b_{1fj} = c_{0f} + u_{1fj} \quad (5)$$

$$b_{2mj} = d_{0m} + u_{2mj} \quad (6)$$

$$b_{2fj} = d_{0f} + u_{2fj} \quad (7)$$

Table 4. Terms in Equations 2 through 7.

Term	Description	Estimated as fixed or random?
a_{0m}	Intercept for males	Fixed
u_{0mj}	Deviation in the intercept for males at the dyad level	Random
a_{0f}	Intercept for females	Fixed
u_{0fj}	Deviation in the intercept for females at the dyad level	Random
c_{0m}	Slope for males for the receiver term	Fixed
u_{1mj}	Deviation in the slope for males for the receiver term at the dyad level	Random
c_{0f}	Slope for females for the receiver term	Fixed
u_{1fj}	Deviation in the slope for females for the receiver term at the dyad level	Random
d_{0m}	Slope for males for the sender term	Fixed
u_{2mj}	Deviation in the slope for males for the sender term at the dyad level	Random
d_{0f}	Slope for females for the sender term	Fixed
u_{2fj}	Deviation in the slope for females for the sender term at the dyad level	Random

The random effects can be correlated, such that the random intercept can be correlated with the random effects for receiver and sender lag, answering whether participants' initial levels of physiology (assuming time is centered at zero) are associated with the degree to which they experience stability or influence, respectively. The random effect for receiver lag can also be correlated with the random effect for sender lag, indicating whether the degree to which participants are stable in their physiological levels is associated with the extent to which they experience physiological influence to their partner. We discuss which random effects to estimate in more detail below. Additional predictors, such as time or an experimental manipulation, may be added to the model to examine potential moderators of stability and influence, yielding insight into both processes and outcomes associated with physiological influence (see Example 2).

Which Random Effects to Include?

As noted above, three types of random effects can be estimated in the stability and influence model to examine within-person and within-dyad processes: variances, within-person

covariances, and between-person covariances. Ideally, the random effects would be fully saturated, resulting in 24 random effects for distinguishable dyads or 14 random effects for indistinguishable dyads (see Supplemental Materials for the full list of effects). However, researchers are likely to find that such a model will not converge, especially with samples of few participants (e.g., in Liu et al., 2016). If this is the case, we recommend first trimming out covariances. In our experience, trimming covariances that involve the influence slope is often most helpful as there tends to be less variance in the influence slope to begin with. Trimming variances should only be done as a last resort. We note that fixed effects estimates can change dramatically as a result of the random effects estimated. Thus, we recommend making an *a priori* decision to model all random effects and only trim effects that make it difficult for the model to converge—potentially because they do not account for a lot of variance. Failures of convergence may also be due to multicollinearity, and so we encourage researchers struggling with model convergence to inspect their data for collinearity issues and also to make sure their data are structured properly.

Distinguishability in Dyadic Data

With dyadic data, dyad members may be conceptually distinguishable from one another on a meaningful dichotomous variable (Kenny et al., 2006). For example, in a study of heterosexual couples where all dyads contain one man and one woman, gender is a distinguishing variable. In a study where one partner receives an experimental manipulation and the other does not, experimental condition is a distinguishing variable. We describe how to conduct a formal test of "distinguishability" in the Supplemental Material.

Example 1: Distinguishable Dyads

In the following example, we analyze data from a study of 70 male-female dyads (college students) who solved math problems together (Thorson et al., 2017), while we measured PEP reactivity from baseline as a measure of SNS activity (Schachinger, Weinbacher, Kiss, Ritz, & Langewitz, 2001). We measured PEP in 30-second intervals while participants were given 30 seconds to solve each problem, resulting in a total of 54 time points of PEP reactivity from baseline. For this example, gender may be a distinguishing factor and so we include dummy codes for males and females (see Equation 1 and Tables 3 to 6) to obtain separate fixed and random effects for males and females. We describe the variables used in the model in Table 5 and the syntax using PROC MIXED in SAS in Table 6. For ease of presentation, statements have been written on multiple lines. A semicolon indicates the end of a statement.

Table 5. Variables used in Example 1.

Variable name	Description
Dyad	A unique identification number for each dyad, which is the same for each member of a dyad.
Obs_id	A unique identification number for each pair of observations that occur at the same time point for the same dyad. It is calculated as "time + nt(dyad-1)" where "time" represents the time point of the observation, "nt" is the number of time points, and "dyad" is the unique identification number for each dyad.
Gender_class	Coded as -1 for females and 1 for males.
Pepreact_R	The dependent variable: receiver PEP reactivity. Reactivity scores were created by subtracting PEP during the last 30 seconds of baseline from each of the subsequent 54 time points.
Male	Coded as 0 for females and 1 for males.
Female	Coded as 0 for males and 1 for females.
Pep_lag_RC	Receiver PEP reactivity at the prior time point centered on the grand mean; also called "receiver lag."
Pep_lag_SC	Sender PEP reactivity at the prior time point centered on the grand mean; also called "sender lag."

Table 6. Annotated syntax for a two-level crossed model with distinguishable dyads.

PROC MIXED COVTEST;	COVTEST requests that standard errors and Wald tests for covariance parameters be displayed in the output.
CLASS dyad obs_id gender_class;	The CLASS statement indicates which variables are categorical.
MODEL pepreact_R = male female	The MODEL statement specifies that PEP reactivity of the receiver is the outcome variable and is predicted by the male dummy code and female dummy code. The corresponding terms in Equations 2 and 3 are a_{0m} and $a_{0f.}$ These terms will produce the fixed intercepts for males and for females.
male*pep_lag_RC female*pep_lag_RC male*pep_lag_SC female*pep_lag_SC	Interaction terms are included to see whether receiver lag and sender lag are significant for males and for females. The corresponding terms in Equations 4 through 7 are c_{0m}, c_{0f}, d_{0m}, and d_{0f}. These terms will produce the fixed "stability" and "influence" slopes, respectively, for males and for females.
/NOINT CL S DDFM=satterth;	NOINT requests no intercept in the model (this is done because our male and female dummy codes will provide intercepts for males and females, respectively). CL requests 95% confidence intervals on the fixed effects estimates. S requests that SAS output the estimates for the fixed effects. DDFM=satterth requests that degrees of freedom be estimated using the Satterthwaite method (Satterthwaite, 1946).
RANDOM	The RANDOM statement specifies the random effects in the model and their patterns of covariation (i.e., the G matrix).
male female	Random intercepts for males and females are specified. The corresponding terms in Equations 2 and 3 are u_{0mj} and u_{0fj}, and the random intercepts are the variances of these terms.
male*pep_lag_RC female*pep_lag_RC	Random slopes for receiver lag for men and women are specified. The corresponding terms in Equations 4 and 5 are u_{1mj} and u_{1fj}, and the random slopes are the variances of these terms.
/SUB=dyad TYPE=un;	SUB=dyad indicates that there is independence in random effects from dyad to

	dyad. TYPE=un specifies an unstructured variance/covariance matrix in which covariances between all of the random effects listed in that statement are estimated.
RANDOM male*pep_lag_SC female*pep_lag_SC	A second RANDOM statement specifies additional random effects and their patterns of covariation. We use a second statement because the patterns of covariation for these random effects are different than what is specified in the first. Random slopes for sender lag for men and women are specified. The corresponding terms in Equations 6 and 7 are u_{2mj} and u_{2fj}, and the random slopes are the variances of these terms.
/SUB=dyad TYPE=vc;	TYPE=vc specifies that variances should be estimated for the variables listed in the statement.
REPEATED gender_class	The REPEATED statement specifies the Level 1 residuals and their pattern of covariation (i.e., the R matrix). On this line, we indicate the variable that distinguishes between members of each dyad so that the errors between two dyad members at the same time point can be correlated.
/ TYPE=csh SUB=dyad*obs_id;	TYPE=csh specifies compound symmetry heterogeneous, which allows the degree of unexplained variance for dyad members to be different. SUB=dyad*obs_id correlates the errors across dyad members at the same time point.
RUN;	

Results for the fixed effects of the model are shown in Table 7. The results are all unstandardized. The intercepts ("male" and "female") represent the average PEP reactivity for male and female receivers, respectively, when their own PEP reactivity at the prior time point (receiver lag) is centered on the grand mean and their partner's PEP reactivity at the prior time point (sender lag) is centered on the grand mean. The estimates for receiver lag (the stability slopes; "male*pep_lag_RC" and "female*pep_lag_RC") indicate that, for both males and females, higher values of receiver PEP reactivity at one time point (i.e., those above the grand

mean of PEP reactivity) are associated with higher values of receiver PEP reactivity at the following time point. The estimate for sender lag for males (the influence slope; "male*pep_lag_SC") indicates that higher values of sender PEP reactivity at one time point (i.e., those above the grand mean of PEP reactivity) are associated with higher values of receiver PEP reactivity at the following time point.

Table 7. Fixed effects estimates for Example 1.

Effect	Estimate	*SE*	df	*t*	*p*	Lower CI	Upper CI
Male	-2.52	0.50	40.3	-5.05	< .001	-3.52	-1.51
Female	-3.57	0.41	40.1	-8.70	< .001	-4.40	-2.74
Male*pep_lag_RC	0.31	0.03	58.7	10.14	< .001	0.25	0.37
Female*pep_lag_RC	0.46	0.03	53.8	13.68	< .001	0.40	0.53
Male*pep_lag_SC	0.07	0.03	52.5	2.57	.01	0.01	0.12
Female*pep_lag_SC	-0.02	0.02	54.0	-0.86	.39	-0.05	0.02

The random effects specified in this model are outlined in Table 8 and the results are in Table 9. Note that the output from SAS will list the random effects as "UN(1,1)", "UN(1,2)", and so on. These numbers correspond to the variables listed in the first RANDOM statement, such that UN(1,2), for example, refers to the covariance between the first variable listed (the intercept for male) and the second variable listed (the intercept for female).

Table 8. Random effects specified for Example 1.

Effect	Interpretation
Variances, for males and females	
• Intercept	Do people vary in their levels of reactivity?
• Receiver effect	Do people vary in how stable they are?
• Sender effect	Do people vary in how much they are influenced by their partners?
Between-person covariances	
• The intercept for males with the intercept for females	Do the two partners have similar reactivity levels?
• The intercept for females with receiver effect for males	If females have higher reactivity scores, do they have male partners who are more/less stable?
• The intercept for males with receiver effect for females	If males have higher reactivity scores, do they have female partners who are more/less stable?
• Receiver lag for males with receiver effect for females	If one dyad member is stable, is the other dyad member stable?
Within-person covariances	
• The intercept for males with receiver effect for men	If a male has a higher reactivity score, is he more/less stable?
• The intercept for females with receiver effect for females	If a female has a higher reactivity score, is she more/less stable?
Common covariance	Are the two partner's reactivity scores similar within a given time point? Similar to an intra-class correlation.

Table 9. Random effects estimates for Example 1.

Random effects ([co-]variances)	Estimate	*SE*	*Z*	*P*
Variance of intercept for males	12.40	2.96	4.18	< .001
Between-person covariance of intercept for males and intercept for females	-5.14	1.84	-2.80	.01
Variance of intercept for females	8.38	2.02	4.15	< .001
Within-person covariance of intercept for males and receiver effect for males	0.09	0.10	0.89	.38
Between-person covariance between intercept for females and receiver effect for males	-0.17	0.09	-1.82	.07
Variance of receiver effect for males	0.03	0.01	3.01	.001
Between-person covariance of intercept for males and receiver effect for females	-0.21	0.13	-1.63	.10
Within-person covariance of intercept for females and receiver effect for females	0.07	0.11	0.65	.51
Between-person covariance of receiver lag for males and receiver effect for females	-0.01	0.01	-1.21	.23
Variance of receiver effect for females	0.04	0.01	3.72	< .001
Variance of sender effect for males	0.02	0.01	2.29	.01
Variance of sender effect for females	0.003	0.003	1.07	.14
Residual variance for females	10.04	0.31	32.38	< .001
Residual variance for males	12.78	0.40	31.83	< .001
Common covariance	0.06	0.02	2.90	.004

As we noted before, the model in Equations 1 through 7 and in Tables 3 through 6 is a "two-intercept model," where one disadvantage is that one cannot test whether the fixed stability and influence effects differ as a function of gender. To do this, one would need to include the main effect of a gender variable (either a dummy-coded gender variable or an effect-coded gender variable, e.g., -1 for females and 1 for males), and interact the receiver and sender physiological variables with this variable. This model tests whether the intercepts for males and females are different and whether the stability and influence slopes for males and females are different. The main effects of stability and influence refer to people "on average" (across men

and women) if effect-coding is used; they refer to whoever is coded as zero if dummy coding is used. SAS syntax for this type of model is presented in Figure 3. Regardless of the approach used, asymmetric influence estimates can be generated for the different dyad members—a particular benefit of using the stability and influence model in general.

```
PROC MIXED COVTEST;
CLASS  obs_id dyad gender_class;
MODEL pepreact_R = pep_lag_RC pep_lag_SC
gender gender*pep_lag_RC gender*pep_lag_SC
/CL S DDFM=SATTERTH ;
RANDOM male female male*pep_lag_RC female*pep_lag_RC /SUB=dyad TYPE=un;
RANDOM male*pep_lag_SC female*pep_lag_SC /SUB=dyad TYPE=vc;
REPEATED gender_class /SUB=dyad*obs_id TYPE=csh;
RUN;
```

Figure 3. Alternative syntax for Example 1, which tests whether the fixed effects statistically differ as a function of gender.

Example 2: Behavior as a moderator

For researchers who are interested in understanding the processes that contribute to physiological influence or the conditions under which physiological influence occurs, one useful approach to addressing these questions may be to include verbal or nonverbal behaviors as moderators of influence. Though mediation is typically thought of as the statistical procedure used to gain insights into process or mechanism, moderation can also be used for this purpose because it allows scholars to turn a process "on" and "off" and measure when the process is stronger or weaker. For example, in a study where women taking the bar exam are theorized to pass anxiety to their husbands through behavioral cues of anxiety, wives' behavioral cues may moderate the physiological influence path. If influence from the wife to the husband (i.e., the wife's physiology predicts the husband's physiology) is stronger when the wife engages in anxiety-related behaviors, this evidence suggests that influence occurs through anxiety cues.

To demonstrate this approach analytically, we use data from Example 1, where trained coders counted the number of questions that participants asked each other about the math task

they were completing together. We note that this example involves a time-invarying moderator, but it is possible to also use time-varying moderators (e.g., we could code for number of questions for each math question, resulting in 27 data points of questions). Syntax for this analysis is provided in Figure 4; we use the approach where the distinguishing factor (gender) is a moderator and is an effect code (similar to the syntax in Figure 3). Following the recommendations of Ledermann and colleagues (Ledermann, Macho, & Kenny, 2011), this model is fully saturated in that it contains both receiver and sender behaviors as moderators of both stability and influence. Receiver and sender variables often have empirical overlap, and so showing that these variables distinctly moderate the influence path (adjusting for the moderating role of the stability path) can strengthen the argument that they uniquely explain when influence occurs.

```
PROC MIXED COVTEST; (1)
CLASS  obs_id dyad gender_class; (2)
MODEL pepreact_R = gender pep_lag_RC pep_lag_SC (3)
qasked_RC qasked_SC  (4)
gender*pep_lag_RC gender*pep_lag_SC (5)
gender*qasked_RC gender*qasked_SC  (6)
qasked_RC*pep_lag_RC qasked_SC*pep_lag_RC (7)
qasked_RC*pep_lag_SC qasked_SC*pep_lag_SC  (8)
gender*qasked_RC*pep_lag_RC gender*qasked_SC*pep_lag_RC (9)
gender*qasked_RC*pep_lag_SC gender*qasked_SC*pep_lag_SC (10)
/CL S DDFM=SATTERTH ;  (11)
RANDOM male female male*pep_lag_RC female*pep_lag_RC /SUB=dyad TYPE=un; (12)
RANDOM male*pep_lag_SC female*pep_lag_SC /SUB=dyad TYPE=vc;  (13)
REPEATED gender_class /SUB=dyad*obs_id TYPE=csh;  (14)
RUN;
```

Figure 4. Syntax for a two-level crossed model with distinguishable dyads with receiver and sender behaviors as moderators.

Results of this model reveal that sender questions asked moderates physiological influence and that this varies by gender (see the interaction term "gender*qasked_SC*pep_lag_SC" in Table 10). Follow-up analyses revealed that males were physiologically influenced by their female partners and that this influence was weaker the more

their female partners asked questions of them. In contrast, females were not physiologically influenced by their male partners, regardless of how many questions they asked.[1] We describe how one would graph influence in this example in the Supplemental Materials.

Table 10. Fixed effects estimates for Example 2.

Effect	Estimate	*SE*	Df	*t*	*P*	Lower CI	Upper CI
Intercept	-3.14	0.25	42.8	-12.33	< .001	-3.66	-2.63
Gender	0.51	0.41	39	1.24	0.22	-0.32	1.34
pep_lag_RC	0.37	0.02	47.3	16.09	< .001	0.32	0.41
pep_lag_SC	0.02	0.02	97.4	0.95	0.34	-0.02	0.05
qasked_RC	-0.06	0.05	3913	-1.26	0.21	-0.15	0.03
qasked_SC	-0.11	0.05	3851	-2.34	0.02	-0.20	-0.02
gender*pep_lagRC	-0.08	0.03	44.1	-3.00	0.004	-0.13	-0.03
gender*pep_lagSC	0.04	0.02	94.9	2.54	0.01	0.01	0.07
gender*qasked_RC	0.02	0.05	3868	0.46	0.64	-0.07	0.11
gender*qasked_SC	-0.05	0.05	3811	-1.04	0.29	-0.14	0.04
qasked_1C*pep_lag_RC	0.001	0.01	2794	0.12	0.91	-0.01	0.01
qasked_2C*pep_lag_RC	0.002	0.01	2723	0.31	0.76	-0.01	0.02
qasked_1C*pep_lag_SC	0.01	0.01	2123	1.17	0.24	-0.01	0.02
qasked_2C*pep_lag_SC	-0.01	0.01	1867	-1.58	0.11	-0.02	0.003
gender*qasked_1C*pep_lag_RC	-0.01	0.01	2772	-1.62	0.10	-0.03	0.002
gender*qasked_2C*pep_lag_RC	-0.01	0.01	2737	-1.29	0.20	-0.02	0.005
gender*qasked_1C*pep_lag_SC	0.002	0.01	2133	0.32	0.75	-0.01	0.01
gender*qasked_2C*pep_lag_SC	-0.01	0.01	1903	-2.06	0.04	-0.03	-0.001

The random effects estimates for this example are very similar to those observed in Example 1 without behaviors. Given that both examples included the same random effects and only the addition of other fixed effects to the model, this is not surprising.

Example 3: Time as a moderator

One question researchers may be interested in examining is whether physiological influence changes over time or during different portions of a study. The role of time in one's model largely depends on the study design and the theoretical question of interest. If researchers suspect that influence increases in a linear fashion—perhaps as initially-unacquainted dyad members get to know one another—then a linear term for time can be included as a fixed effect in the model and in interaction terms with receiver physiology and sender physiology. Higher-order effects for time can also be included if researchers suspect those patterns—for example, a quadratic term may be useful if researchers suspect that influence eventually levels off. A different technique for incorporating time can be used if researchers have participants engage in several different tasks during one study. If the tasks are counterbalanced, it may be more suitable to include a term representing the different tasks in one's model and interact that term with receiver and sender physiology to see whether influence is stronger during different time periods.

In this example, we analyze data from a study of 29 mother-child dyads who discussed a topic of conflict together, while PEP reactivity was measured from baseline as a measure of SNS reactivity (Waters et al., 2017). We measured PEP in 30-second intervals while participants were talking with one another, resulting in ten time points of PEP reactivity from baseline. We include receiver role (mother or child; mothers are coded as -1 and children as 1) as a distinguishing variable in the following analysis. This variable would be similar to the "receiver condition" variable in Table 2. Syntax for this analysis is provided in Figure 5. The main effect of time

(centered at the study mid-point) and its two-way interactions with role (mother or child), receiver lag, and sender lag are included on line 5. Line 6 includes three-way interactions among time, role, receiver lag, and sender lag.

```
PROC MIXED COVTEST SCORING=15; (1)
CLASS obs_id dyad role_class;   (2)
MODEL pepreact_R = pep_lag_RC pep_lag_SC role (3)
pep_lag_RC*role pep_lag_SC*role  (4)
timeCmid timeCmid*role pep_lag_RC*timeCmid pep_lag_SC*timeCmid  (5)
pep_lag_RC*role*timeCmid pep_lag_SC*role*timeCmid  (6)
/CL S DDFM=SATTERTH;  (7)
RANDOM parent child  (8)
/SUB=dyad TYPE=un; (9)
RANDOM parent*pep_lag_RC child*pep_lag_RC (10)
parent*pep_lag_SC child*pep_lag_SC  (11)
/SUB=dyad TYPE=vc;  (12)
REPEATED role_class /TYPE=csh sub=dyad*obs_id; (13)
RUN;
```

Figure 5. Syntax for a two-level crossed model with time as a moderator.

Results of this model reveal that influence is moderated by both time and role (see the interaction term "pep_lag_SC*timeCmid*role" in Table 11). Follow-up analyses testing the effect of physiological influence over time for both children and mothers revealed that physiological influence did not significantly change for mothers over time, but strengthened for children over time (see Figure 6). As the study went on, children became more positively linked to their mothers' physiology, such that higher values of mothers' physiology at one time point predicted higher values of children's physiology at the following time point. Note that the outcome in Figure 6 is physiological influence and not the outcome of the model (receiver PEP reactivity); we describe how to graph influence in the Supplemental Materials.

Table 11. Fixed effects estimates for Example 3.

Effect	Estimate	*SE*	df	*T*	*p*	Lower CI	Upper CI
Intercept	-2.48	0.77	16.5	-3.24	0.01	-4.09	-0.86
pep_lag_RC	0.19	0.05	57.7	3.78	<.001	0.09	0.29
pep_lag_SC	-0.06	0.04	58.3	-1.36	0.18	-0.15	0.03
Role	-0.45	0.71	13.9	-0.64	0.54	-1.97	1.07
pep_lag_RC*role	0.03	0.05	57.3	0.53	0.60	-0.07	0.13
pep_lag_SC*role	0.04	0.04	57.4	0.90	0.37	-0.05	0.13
timeCmid	0.03	0.07	202	0.47	0.64	-0.11	0.17
role*timeCmid	-0.10	0.06	197	-1.65	0.10	-0.22	0.02
pep_lag_RC*timeCmid	-0.03	0.01	280	-3.46	0.001	-0.04	-0.01
pep_lag_SC*timeCmid	0.004	0.01	353	0.50	0.62	-0.01	0.02
pep_lag_RC*timeCmid*role	0.005	0.01	282	0.59	0.56	-0.01	0.02
pep_lag_SC*timeCmid*role	-0.02	0.01	364	2.14	0.03	0.001	0.03

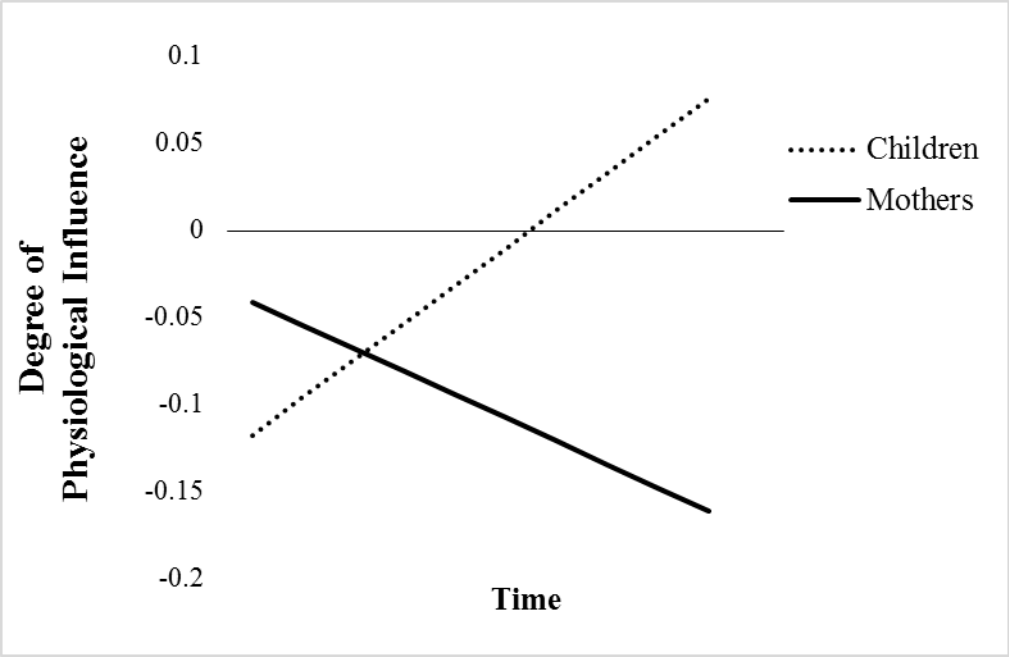

Figure 6. Example 3: Physiological influence as a function of time and participant role.

Presentation of Results

In Table 12, we provide a checklist of information researchers should present in a paper—either in the main text or in supplemental materials—using the stability and influence model to analyze their data. We have outlined a number of decisions researchers must make during the design, implementation, and analysis of physiological influence studies, and we encourage readers to outline how they have made these choices in order to facilitate the comparison and synthesis of results. We strongly recommend that researchers present the full specification of their model and random effects estimates when publishing their data because this allows for a more complete understanding of the circumstances under which the fixed effects estimates were obtained.

Table 12. Information to be presented in a paper using the stability and influence model.

1. Lag length used.
2. Interval length used.
3. Centering method used.
4. Whether change scores or raw values were used.
5. Information regarding individual physiological responding over time.
6. Rate of missing data for all variables included in the model, with reasons why data are missing (e.g., due to shared behaviors that obstruct quality of signal).
7. All fixed effects estimated in the model, including main effects and interactions.
8. All random effects estimated in the model.
9. Fixed effects estimates.
10. Random effects estimates.

Final Concerns

Sample size. When deciding on sample size, researchers must consider 1) how many dyads and 2) how many time points are needed to have sufficient statistical power to detect a hypothesized effect. Increasing upper-level units (i.e., the number of dyads) typically boosts power more than increasing the number of lower-level units (i.e., time points; Bolger, Stadler, & Laurenceau, 2011). Importantly, power does not exist for a study as a whole but rather for individual effects, and, if conducting *a priori* power analyses to determine sample size, researchers should make sure they have enough power to detect their hypothesized effect of interest. Typically, higher-order effects (e.g., a three-way interaction) require a greater sample size than lower-order effects (e.g., a main effect). In addition, researchers may wish to take into account the closeness of the pairs being studied, as well as the study context. For instance, romantic couples may have a strong motivation to attend to their partners or to be sensitive to their psychological states, creating a larger influence effect. This may not be the case with strangers, unless they are placed in a context where they are particularly motivated to attend to each other (e.g., due to shared performance goals).

When deciding on sample size, researchers should keep in mind that missing data have a multiplicative impact in the stability and influence model. If a measurement is missing at time T for the receiver, then three observations are missing: 1) time T for the receiver (because the dependent variable is missing), 2) time T+1 for the receiver (because the stability predictor is missing), and 3) time T+1 for the sender (because the influence predictor is missing). The number of missing physiological measurements may not result in three times that many missing observations in the analysis, however, if the missing data "overlap" (e.g., if both the receiver and the sender are missing measurements at time T, this will result in four missing observations in

the analysis, not six). Lastly, there is no upper limit to the number of time points researchers can have with this model, although researchers might be limited by their computer's processing capabilities and increasing the number of lower-level units can only go so far in increasing statistical power (Bolger et al., 2011).

Power analysis. One approach to conducting power analyses when using the stability and influence model is a simulation method (see Bolger et al., 2012, and Lane & Hennes, 2016). This technique has the advantage of being flexible and able to accommodate many types of models. Because a substantial amount of information is needed to conduct such analyses, we strongly recommend pilot testing before conducting these analyses to gain a sense of values of the stability and influence paths and the random effects in the model.

Here, we outline four steps for a researcher who has collected pilot data and wants to conduct a power analysis using a simulation method to plan the final sample size. We present annotated SAS syntax for each of these steps in our Supplemental Materials. The first step would be to run the model one intends to use on a full sample of dyads on a pilot sample of dyads. The pilot data can then be analyzed to get the fixed and random effects estimates to simulate data for 1000 hypothetical studies with the same number of dyads and time points as the pilot study. When the data have been simulated, the third step is to analyze each of those 1000 samples individually, using the model that was run on the pilot data. Initially, we suggest using the same number of dyads and time points in one's pilot study to check the estimates, standard errors, and degrees of freedom obtained in the power analysis against the results obtained from the pilot study. This can identify mistakes in one's syntax at this stage of the power analysis. Next, one can then document the number of times a hypothesized effect is significant, using the percentage of significant effects across all 1000 studies as an estimate of power. One can then go back to the

second step of simulating data and change the number of dyads and time points per sample and repeat the 3rd and 4th steps, discovering the final sample size needed for a hypothesized effect to have sufficient (typically 80%) power. We fully outline how to implement each of these four steps, as well as tips and potential pitfalls, for the stability and influence model in our Supplemental Materials.

Discussion

We have presented a guide for researchers planning to study the correspondence between two individuals' physiological states. We have outlined five principles for making psychological inferences from physiological influence and have provided guidance for researchers as they design, implement, and analyze dyadic psychophysiological data. We highly recommend that scholars consider these principles in the design and implementation stage of the study, as these choices can affect parameters in the analytical model. For example, in Principles 3 and 5, we discuss how scholars need to measure signals that potentiate physiological influence. If a researcher has a good understanding of what these signals are and does a sufficient job measuring or manipulating them, then they can be used as moderators of the linkage paths in the analytical model to gain insight into how and when influence occurs. Importantly, because these principles can be used to guide researchers in the study design phase, they can certainly be applied to other analytical approaches other than the one we describe here. For example, other approaches also allow for the inclusion of time-varying predictors (see Chow, Ferrer, and Hsieh, 2010), and so one can use Principles 3 and 5 to inform how different cues are incorporated into the model, using other approaches. In Part 2, we present important considerations for study design, and lastly, we provide an analytic approach that allows researchers to answer questions

regarding interpersonal influence in physiological states, demonstrating the model through three examples. We conclude by highlighting strengths and limitations of the model.

Strengths of the Model

We highlight seven strengths of the stability and influence model approach. One, with the stability and influence model, researchers can empirically test the processes that underlie influence. By including observable cues that are measured over time as moderators of influence, researchers can address critical questions regarding how and when influence occurs. Two, researchers can model asymmetric influence estimates for each member of a dyad. Therefore, researchers can measure or manipulate processes in one dyad member at a time, examining how these processes may affect each dyad member in separate ways. Three, for researchers interested in interpersonal influence, the stability and influence model allows researchers the opportunity to test which dyad member influences the other. Four, the stability and influence model provides both group-level information—for the fixed effects—and dyad-level information—for the random effects, providing information in one model (i.e., about both average trends and how much variability there is across dyads) that other approaches which take only a group-level or dyad-level approach cannot.

Five, the stability and influence model estimates physiological influence, while also estimating and accounting for stability (or autoregressive effects). Autoregressive effects are often quite strong and can impact both the significance and the direction of influence between partners. Therefore, even if these effects are not of interest theoretically, they should be of interest empirically. In addition, by estimating both stability and influence in the same model, researchers can examine important questions regarding within-person and within-dyad relationships between stability and influence. For example, are people who are more influenced

by their partners also less stable in their own physiological responses? Positive within-person relationships between stability and influence might occur in some contexts (e.g., when a receiver's emotions are being regulated around a stable set point by a sender's), whereas negative relationships might occur in others (e.g., when a receiver is influenced by an emotionally labile sender, resulting in instability for the sender). Although questions regarding the relationship between stability and influence have received little empirical interest so far, they have exciting potential as future directions in the study of physiological interdependence broadly.

Six, the stability and influence model allows researchers to examine changes over time in both stability and influence. For example, romantic partners might strengthen in their influence of each other during an argument in which they are closely attending to each other, but then decline in influence as the argument comes to a resolution. This type of influence would be considered "morphogenic" in that the strength of the influence changes over time (vs. "morphostatic" influence that does not change over time; Butler, 2011). In a study with various tasks or time periods (e.g., a study where mothers discuss a topic of conflict with children and then play with their children), a variable representing the task period could be included to examine whether influence differs during specific tasks. Researchers can also use piecewise linear models (Fitzmaurice et al., 2011) to examine changes in influence over time within certain periods. Pairwise correlations that provide one correlation for the dyad during the whole study or during a particular time frame, cannot examine these effects.

Finally, the stability and influence model technically requires only three time points of data to be collected, unlike other methods used for estimating physiological interdependence, such as time series analysis (Levenson & Gottman, 1983), which require many more. This may

be particularly helpful for researchers collecting costly neuroendocrine measures, such as cortisol (e.g., Bernard et al., 2017). There are two important caveats to this point, however. First, fewer observations necessitate simpler models. For example, fewer random effects can be specified when researchers have obtained fewer observations per person (potentially variances only and no within-person or between-person covariances). Similarly, with fewer time points, researchers may not be able to model non-linear changes in stability and influence over time. Second, fewer observations compromise the accuracy of the estimates, especially for random effects. Thus, one can estimate this model with fewer time points relative to other models of physiological interdependence, but researchers should be aware that having fewer time points may come with significant disadvantages.

Limitations of the Model

One limitation of the approach presented here is that researchers must specify the functional form of the model. Other models that take a nonparametric approach may be more useful in an exploratory phase of research, where although researchers have collected intensive data, they are aiming to uncover whether similarity in responding exists (e.g., dynamical correlation, Liu et al., 2016). Approaches like the one here can be used when researchers are ready to specify more precise models. Relatedly, researchers must also specify the length of each time interval with the stability and influence model. The interval that most closely captures changes in psychological responding can be difficult to identify and can also affect estimates of influence. Researchers who are unsure of an appropriate interval length to use (based on the task, equipment, and participant demographics, for example), may prefer methods that can model continuous responses without specifying a particular interval length (e.g., empirical mode decomposition, McAssey et al., 2013).

Because the unit of analysis in the stability and influence model is individual time points, data that do not adequately capture physiological responding—due to missing data or measurements that are too infrequent—may result in inaccurate estimates of influence. Functional data analytic methods where the unit of analysis is the underlying curve (e.g., dynamic correlation, Liu et al., 2016) may be better when the sampling interval is less frequent or there is a lot of missing data, as these models assume that the data come from an underlying smooth function of time. Such approaches may therefore be particularly useful during exploratory phases of research when researchers are unsure whether an adequate number of time points have been measured and simply want to document whether similarity between physiological responses exists (e.g., Liu et al., 2016).

Although one of the strengths of the stability and influence model is the ability to estimate a fixed stability effect (i.e., an autoregressive effect), this strength may also be viewed as a limitation because it provides conservative estimates of influence. If a large portion of the variance in participants' physiology is already accounted for by their own physiology at a prior time point, there is less variance that can be accounted for by a partner's physiology. As mentioned previously, this is one reason why researchers studying physiological influence need to design studies in which participants can experience variability in their physiological responses and have sufficient opportunity to influence and be influenced by another dyad member (e.g., interact with another person face-to-face).

An additional limitation of this model, in contrast to others such as the coupled linear oscillator model (Butner et al., 2005; Reed, Barnard, & Butler, 2015), is an inability to provide information about the level of dyad members' responses. In other words, the stability and influence model can indicate whether two participants are physiologically influencing each other,

but it cannot provide information on the level or range of the physiological responses—for example, whether two dyad members who are influencing each other are both at very high levels of SNS reactivity or showing very little reactivity at all. Thus, it cannot provide information about two kinds of influence—termed coregulation (when dyad members keep each other regulated around a set-point) and codysregulation (when dyad members cause each other's physiology to continually escalate or fall from a set-point). Such models may be particularly useful for researchers interested in understanding the physical and psychological health implications of social relationships.

Finally, the stability and influence model is less well-suited for understanding the outcomes of influence that occur beyond those measured during the study. With idiographic approaches, researchers can correlate the estimates of similarity for each dyad with outcome variables, such as performance on a dyadic task (e.g., Henning et al., 2011). In the stability and influence model, the best way to address questions regarding the outcomes of influence is to use those outcomes as moderators of influence in one's analysis. For example, researchers could moderate influence by performance, showing that positive influence occurs during periods of high performance, inferring that performance is improved by positive influence between dyad members. It would be difficult to make causal claims here—for instance, about whether influence causes performance or performance causes influence.

Conclusion

The stability and influence model provides a flexible approach for researchers seeking to understand physiological dynamics in dyads and the psychological processes that underlie them. Our goal is that this paper can be used as a guide for anyone who is interested in studying interpersonal processes by examining shared physiology. We walked researchers through critical

decision points in their research and provided recommendations for the best ways to make psychological inferences from physiological influence. Our goal is that this comprehensive guide can help scholars answer questions regarding emotion contagion, psychological attunement, and interpersonal influence (to name a few) in their programs of dyadic psychophysiological research.

References

Aiken, L. S. & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Newbury Park, CA: Sage Publications.

Anderson, C., Keltner, D., & John, O. P. (2003). Emotional convergence between people over time. *Journal of Personality and Social Psychology, 84*(5), 1054-1068. doi:10.1037/0022-3514.84.5.1054

Bachrach, A., Fontbonne, Y., Joufflineau, C., & Ulloa, J. L. (2015). Audience entrainment during live contemporary dance performance: Physiological and cognitive measures. *Frontiers in Human Neuroscience, 9*(179), 1-13. doi:10.3389/fnhum.2015.00179

Barrett, L. F. (2013). Psychological construction: A Darwinian approach to the science of emotion. *Emotion Review, 5,* 379-389. doi: 10.1177/1754073913489753

Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research, 40*(3), 373-400. doi: 10.1207/s15327906mbr4003_5.

Bazhenova, O. V., Plonskaia, O., & Porges, S. W. (2001). Vagal reactivity and affective adjustment in infants during interaction challenges. *Child Development, 72*(5), 1314-1326. doi:10.1111/1467-8624.00350

Bernard, N. K., Kashy, D. A., Levendosky, A. A., Bogat, G. A., & Lonstein, J. S. (2017). Do different data analytic approaches generate discrepant findings when measuring mother-infant HPA axis attunement? *Developmental Psychobiology, 59,* 174-184. doi: 10.1002/dev.21474

Blascovich, J., & Mendes, W. B. (2010). Social psychophysiology and embodiment. In: S. T. Fiske, D. T. Gilbert, G. Lindzey, (Eds.). *The handbook of social psychology, 5th Edition* (pp. 194-227). New York, NY: John Wiley & Sons Inc.

Blascovich, J., Vanman, E. J., Mendes, W. B., & Dickerson, S. (2011). *Social psychophysiology for social and personality psychology*. London: SAGE Publications.

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

Bolger, N., Stadler, G., & Laurenceau, J. P., (2011). Power analysis for intensive longitudinal studies. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 285-301). New York, NY: Guilford.

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., Borsboom, D., & Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLOS One, 8*(4). doi: 10.1371/journal.pone.0060188

Butler, E. A., (2011). Temporal interpersonal emotion systems: The "TIES" that form relationships. *Personality and Social Psychology Review, 15*(4), 367-393. doi:10.1177/1088868311411164

Butner, J., Amazeen, P. G., & Mulvey, G. M., (2005). Mutlilevel modeling of two cyclical processes: Extending differential structural equation modeling to nonlinear coupled systems. *Psychological Methods, 10*(2), 159-177. doi:10.1037/1082-989X.10.2.159

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M. &, Ito, T. A. (2000). The psychophysiology of emotion. In M. Lewis & J. M. Haviland-Jones (Eds.). *Handbook of Emotions, 2nd edition* (pp 173-191). New York: The Guilford Press.

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (2007a). *Handbook of psychophysiology, 3rd Edition.* New York, NY: Cambridge University Press.

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (2007b). Psychophysiological science: Interdisciplinary approaches to classic questions about the mind. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology, 3rd Edition,* (pp. 1-16). New York, NY: 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,* 1-12. doi:10.3389/fnbeh.2014.00095

Chow, S.M., Ferrer, E., & Hsieh, F. (2010). *Statistical methods for modeling human dynamics: An interdisciplinary dialogue*. New York, NY: Taylor and Francis.

Cohen, L. H., & Patterson, M. (1937). Effect of pain on the heart rate of normal and schizophrenic individuals. *The Journal of General Psychology*, *17*(2), 273-289. doi: 10.1080/00221309.1937.9918000

Coleman, R., Greenblatt, M., & Solomon, H. C. (1956). Physiological evidence of rapport during psychotherapeutic interviews. *Diseases of the Nervous System, 17*(3), 71-77.

Darrow, C. W. (1929). Differences in the physiological reactions to sensory and ideational stimuli. *Psychological Bulletin, 26*(4), 192-201. doi:10.1037/h0074053

Dikker, S. Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., Van Bavel, J. J., Ding, M, & Poeppel, D. (in press). Brain-to-brain synchrony during real-world dynamic group interactions: A classroom EEG study. *Current Biology.*

DiMascio, A., Boyd, R., & Greenblatt, M. (1957). Physiological correlates of tension and antagonism during psychotherapy: A study of interpersonal physiology. *Psychosomatic Medicine, 19*(2), 99-104. doi:10.1097/00006842-195703000-00002

Ebisch, S. J., Aureli, T., Bafunno, D., Cardone, D., Romani, G. L., & Merla, A. (2012). Mother and child in synchrony: Thermal facial imprints of autonomic contagion. *Biological Psychology, 89*(1), 123-129. doi:10.1016/j.biopsycho.2011.09.018

Feldman, R., Magori-Cohen, R., Galili, G., Singer, M., & Louzoun, Y. (2011). Mother and infant coordinate heart rhythms through episodes of interaction synchrony. *Infant Behavior & Development, 34*(4), 569-577. doi:10.1016/j.infbeh.2011.06.008

Fisher, A. J., & Newman, M. G. (2013). Heart rate and autonomic response to stress after experimental induction of worry versus relaxation in healthy, high-worry, and generalized anxiety disorder individuals. *Biological Psychology, 93*(1), 65-74. doi:10.1016/j.biopsycho.2013.01.012

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2011). *Applied longitudinal analysis*, 2nd Edition. New Jersey: John Wiley & Sons, Inc.

Funder, D. C., (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102*(4), 652-670. doi:10.1037/0033-295X.102.4.652

Gates, K. M., & Liu, S. (2016). Methods for quantifying patterns of dynamic interactions in dyads. *Assessment, 23*(4): 459-471. doi: 10.1177/1073191116641508.

Gonzalez, R., & Griffin, D., (2002). Modeling the personality of dyads and groups. *Journal of Personality, 70*(6), 901-924. doi: 10.1111/1467-6494.05027

Gosling, S. M., Ko, S. J., Mannarelli, T. & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology, 82*(3), 379-398. doi:10.1037//0022-3514.82.3.379

Gottman, J. M., (1990). How marriages change. In G. R. Patterson (Ed.), *Depression and aggression in family interaction* (pp. 75-101). Hillsdale, NJ: Lawrence Erlbaum.

Gross, J. J. & L. F. Barrett. (2011). Emotion generation and emotion regulation: One or two depends on your point of view. *Emotion Review, 3*(1), 8-16. doi:10.1177/1754073910380974

Guastello, S. J., Pincus, D., & Gunderson, P. R. (2006). Electrodermal arousal between participants in a conversation: Nonlinear dynamics and linkage effects. *Nonlinear Dynamics, Psychology, and Life Sciences, 10*(3), 365-399. doi.org/10.1037/e404812008-001

Ham, J., & Tronick, E., (2009). Relational psychophysiology: Lessons from mother–infant physiology research on dyadically expanded states of consciousness. *Psychotherapy Research, 19*(6), 619-632. doi:10.1080/10503300802609672

Hansen, A. L., Johnson, B. H., & Thayer, J. F., (2003). Vagal influence on working memory and attention. *International Journal of Psychophysiology, 48*(3), 263-274. doi:10.1016/S0167-8760(03)00073-4

Helm, J. L., Sbarra, D. A., & Ferrer, E., (2012). Assessing cross-partner associations in physiological responses via coupled oscillator models. *Emotion, 12*(4), 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*(3), 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*(3), 221-232. doi:10.1016/S0167-8760(00)00190-2

Heyman, R. E., Lorber, M. F., Eddy, J. M., & West, T. V., (2014). Behavioral observation and coding. In H. T. Reis & C. M. Judd (Eds). *Handbook of research methods in social and personality psychology, second edition* (pp. 345-372). New York: NY Cambridge University Press.

Hill-Soderlund, A. L., Mills-Koonce, W. R., Propper, C., Calkins, S. D., Granger, D. A., Moore, G. A., & Cox, M. J. (2008). Parasympathetic and sympathetic responses to the strange situation in infants and mothers from avoidant and securely attached dyads. *Developmental Psychobiology, 50*(4), 361-376. doi:10.1002/dev.20302

Houtveen, J. H., Rietveld, S., & De Geus, E. J. (2002). Contribution of tonic vagal modulation of heart rate, central respiratory drive, respiratory depth, and respiratory frequency to respiratory sinus arrhythmia during mental stress and physical exercise. *Psychophysiology, 39*(4), 427-436. doi:10.1017.S0048577202394022

Hox, J. J. (2013). Multilevel regression and multilevel structural equation modeling. In T. D. Little (Ed.), *The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2: Statistical Analysis* (pp. 281-294). New York, NY: Oxford University Press.

Hox, J. J., Maas, C. J. M., & Brinkhuis, M. J. S. (2010). The effect of estimation method and sample size in multilevel structural equation modeling. *Statistica Neerlandica, 64,* 157-170.

Inauen, J., Shrout, P. E., Bolger, N., Stadler, G., & Scholz, U. (2016). Mind the gap? An intensive longitudinal study of between-person and within-person intention-behavior relations. *Annals of Behavioral Medicine, 50,* 516-522. doi:10.1007/s12160-016-9776-x

Jacobson, E. (1930). Electrical measurements of neuromuscular states during mental activities. *American Journal of Physiology, 95*(3), 694-702.

Järvelä, S., Kivikangas, J. M., Kätsyri, J., & Ravaja, N. (2013). Physiological linkage of dyadic gaming experience. *Simulation & Gaming, 45*(1), 24-40. doi:10.1177/1046878113513080

Kashy, D. A., & Kenny, D. A. (2000). The analysis of data from dyads and groups. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social psychology*. New York, NY: Cambridge University Press.

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

Kenny, D. A., & La Voie, L. J. (1985). Separating individual and group effects. *Journal of Personality and Social Psychology, 48*(2), 339-348. doi:10.1037/0022-3514.48.2.339

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H., (1993). The 'Trier Social Stress Test': A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychology, 29*(1-2), 76-81. doi.org/10.1159/000119004

Kraus, M. W., & Mendes, W. B., (2014). Sartorial symbols of social class elicit class-consistent behavioral and physiological responses: A dyadic approach. *Journal of Experimental Psychology: General, 143*(6), 2330-2340. doi:10.1037/xge0000023

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology, 84,* 394-421. doi: 10.1016/j.biopsycho.2010.03.010.

Lane, S. P., & Hennes, E. P. (in press). Power struggles: Estimating sample size for multilevel relationships research. *Journal of Social and Personal Relationships.*

Ledermann, T., Macho, S., & Kenny, D. A. (2011). Assessing mediation in dyadic data using the actor-partner interdependence model. *Structural Equation Modeling, 18,* 595-612.

Levenson, R. W., & Gottman, J. M. (1983). Marital interaction: Physiological linkage and affective exchange. *Journal of Personality and Social Psychology, 45*(3), 587-597. doi:10.1037/0022-3514.45.3.587

Levenson, R. W., & Ruef, A. M. (1992). Empathy: A physiological substrate. *Journal of Personality and Social Psychology, 63*(2), 234-246. doi:10.1037/0022-3514.63.2.234

Liu, S., Zhou, Y., Palumbo, R., & Wang, J. L. (2016). Dynamical correlation: A new method for quantifying synchrony with multivariate intensive longitudinal data. *Psychological Methods, 21*(3), 291-308. doi:10.1037/met0000071

Liu, S., Rovine, M. J., Cousino Klein, L., & Almeida, D. M. (2013). Synchrony of diurnal cortisol pattern in couples. *Journal of Family Psychology, 27*(4), 579 –588. doi:10.1037/a0033735

Lunkenheimer, E., Tiberio, S. S., Buss, K. A., Lucas-Thompson, R. G., Boker, S. M., & Timpe, Z. C. (2015). Coregulation of respiratory sinus arrhythmia between parents and preschoolers: Differences by children's externalizing problems. *Developmental Psychobiology, 57*(8), 994-1003. doi:10.1002/dev.21323

Manini, B., Cardone, D., Ebisch, S. J., Bafunno, D., Aureli, T., & Merla, A. (2013). Mom feels what her child feels: Thermal signatures of vicarious autonomic response while watching children in a stressful situation. *Frontiers in Human Neuroscience, 7* (299), 1-10. doi: doi.org/10.3389/fnhum.2013.00299

Manstead, A. S. R. (2012). A history of affect and emotion research in social psychology. In A. W. Kruglanski & W. Stroebe (Eds.), *Handbook of the history of social psychology* (pp. 177-198). New York: Psychology Press.

Marci, C. D., Ham, J., Moran, E., & Orr, S. P. (2007). Physiologic correlates of perceived therapist empathy and social-emotional process during psychotherapy. *Journal of Nervous and Mental Disease, 195*(2), 103-111. doi:10.1097/01.nmd.0000253731.71025.fc

Mauss, I. B., Cook, C. L., Cheng, J. Y., & Gross, J. J. (2007). Individual differences in cognitive reappraisal: Experiential and physiological responses to an anger provocation. *International Journal of Psychophysiology*, *66*(2), 116-124. doi: doi.org/10.1016/j.ijpsycho.2007.03.017

McAssey, M. P., Helm, J., Fushing, H., Sbarra, D. A., & Ferrer, E. (2013). Methodological advances for detecting physiological synchrony during dyadic interactions. *Methodology*, *9*(2), 41-53. doi: 10.1027/1614-2241/a000053

Mendes, W. B. (2009). Assessing the autonomic nervous system. In E. Harmon-Jones & J. S. Beer (Eds.) *Methods in social neuroscience* (pp. 118-147). New York, NY: Guilford Press.

Mendes, W. B. (2016). Emotion and the autonomic nervous system. In L. E. Barrett, M. Lewis, & J. Haviland-Jones (Eds.), *Handbook of emotions, Fourth edition* (pp. 166-181). New York, NY: Guilford Press.

Mitkidis, P., McGraw, J. J., Roepstorff, A., & Wallot, S., (2015). Building trust: Heart rate synchrony and arousal during joint action increased by public goods game. *Physiology & Behavior, 149,* 101-106. doi:10.1016/j.physbeh.2015.05.033

Mittleman, B., & Wolff, H. G. (1939). Affective states and skin temperature: Experimental study of subjects with “cold hands” and Raynaud’s syndrome. *Psychosomatic Medicine,* 1, 271-292. doi:10.1097/00006842-193904000-00003

Palumbo, R. V., Marraccini, M. E., Weyandt, L. L., Wilder-Smith, O., McGee, H. A., Liu, S., & Goodwin, M. S. (2016). Interpersonal autonomic physiology: A systematic review of the literature. *Personality and Social Psychology Review,* 1-43. 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*(6), 882–894. doi: 10.1037/a0017147

Papp, L. M., Pendry, P., Simon, C. D., & Adam, E. K. (2013). Spouses’ cortisol associations and moderators: Testing physiological synchrony and connectedness in everyday life. *Family Process, 52*(2), 284 –298. doi:10.1111/j.1545-5300.2012.01413.x

Parkinson, B. (2011). Interpersonal emotion transfer: Contagion and social appraisal. *Social and Personality Psychology Compass, 5*(7), 428-439. doi:10.1111/j.1751-9004.2011.00365.x

Quer, G., Daftari, J., & Rao, R. R. (2016). Heart rate wavelet coherence analysis to investigate group entrainment. *Pervasive and Mobile Computing, 28*(C), 21-34. doi:10.1016/j.pmcj.2015.09.008

Quigley, K. S., & Barrett, L. F. (2014). Is there consistency and specificity of autonomic changes during emotional episodes? Guidance from the Conceptual Act Theory and psychophysiology. *Biological Psychology, 98,* 82-94. doi:10.1016/j.biopsycho.2013.12.013

Reed, R. G., Barnard, K., & Butler, E. A. (2015). Distinguishing emotional coregulation from codysregulation: An investigation of emotional dynamics and body weight in romantic couples. *Emotion, 15*(1): 45-60. doi: 10.1037/a0038561

Reed, R. G., Randall, A. K., Post, J. H., & Butler, E. A., (2013). Partner influence and in-phase versus anti-phase physiological linkage in romantic couples. *International Journal of Psychophysiology, 88*(3), 309-316. doi:10.1016/j.ijpsycho.2012.08.009

Ricksher, C., & Jung, C. G. (1908). Further investigations on the galvanic phenomenon and respiration in normal and insane individuals. *The Journal of Abnormal Psychology, 2*(5), 189-217. doi:10.1037/h0073786

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin, 2,* 110-114. doi: 10.2307/3002019

Saxbe, D., & Repetti, R. L. (2010). For better or worse? Coregulation of couples' cortisol levels and mood states. *Journal of Personality and Social Psychology, 98*(1), 92–103. doi:10.1037/a0016959

Sbarra, D. A., & Hazan, C. (2008). Coregulation, dysregulation, self-regulation: An integrative analysis and empirical agenda for understanding adult attachment, separation, loss, and recovery. *Personality and Social Psychology Review*, *12*(2), 141-167. doi:10.1177/1088868308315702

Scarpa, A., Ashley, R. A., Waldron, J. C., Zhou, Y., Swain, D. M., Dunsmore, J. C., & Bell, M. A. (2017). Side by side: Modeling dyadic physiological linkage in strangers. *Emotion.* Advance online publication. doi: 10.1037/emo0000340.

Schachinger, H., Weinbacher, M., Kiss, A., Ritz, R., & Langewitz, W., (2001). Cardiovascular indices of peripheral and central sympathetic activation. *Psychomatic Medicine, 63*(5), 788-796. doi:10.1097/00006842-200109000-00012

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

Shiota, M. N., Neufeld, S. L., Yeung, W. H., Moser, S. E., & Perea, E. F. (2011). Feeling good: autonomic nervous system responding in five positive emotions. *Emotion*, *11*(6), 1368-1378. doi: 10.1037/a0024278

Siegel, E. H., Sands, M. K., Condon, P., Chang, Y., Dy, J., Quigley, K. S., & Barrett, L. F. (2016). *Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic features of emotion categories*. Manuscript submitted for publication.

Silver, R., & Parente, R., (2004). The psychological and physiological dynamics of a simple conversation. *Social Behavior and Personality, 32*(5), 413-418. doi:10.2224/sbp.2004.32.5.413

Stratford, T., Lal, S., & Meara, A. (2012). Neuroanalysis of therapeutic alliance in the symptomatically anxious: The physiological connection revealed between therapist and client. *American Journal of Psychotherapy*, *66*(1), 1-21.

Suveg, C., Shaffer, A., & Davis, M. (2016). Family stress moderates relations between physiological and behavioral synchrony and child self-regulation in mother–preschooler dyads. *Developmental Psychobiology*, *58*(1), 83-97. doi:10.1002/dev.21358

Thorson, K. R., Forbes, C. E., Magerman, A., & West, T. V. (2017). *Under threat but engaged: Stereotype threat leads women to engage with female but not male partners in STEM.* Manuscript submitted for publication.

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*(5), 720-731. doi:10.1037/fam0000115

Wass, S. V., Clackson, K., Cook, C., & de Barbaro, K. (2015). Tonic and phase co-variation of peripheral arousal indices in infants. *Biological Psychology, 111*, 26-39. doi: 10.1016/j.biopsycho.2015.08.006

Waters, S. F., West, T. V., Karnilowicz, H., & Mendes, W.B. (2017). *Affect contagion between mothers and babies: Exploring the role of touch and valence.* Manuscript submitted for publication.

Waters, S. F., West, T. V., & Mendes, W. B. (2014). Stress contagion physiological covariation between mothers and infants. *Psychological Science, 25*(4), 934-942. doi:10.1177/0956797613518352

West, T. V. (2013). Repeated measures with dyads. In J. A. Simpson & L. Campbell (Eds.), *The Oxford handbook of close relationships* (pp. 731-749). New York: Oxford University Press.

West, T. V., Koslov, K., Page-Gould, E., Major, B., & Mendes, W. B. (in press). Contagious anxiety: Anxious European Americans can transmit their physiologic reactivity to African Americans. *Psychological Science.*

Footnotes

[1]These results were found with only a subset of participants from this study and are only used for demonstration purposes. They should not be cited for their empirical merit.