



# “I feel you”: Greater linkage between friends’ physiological responses and emotional experience is associated with greater empathic accuracy

Felicia K. Zerwas<sup>a,\*</sup>, Tabea Springstein<sup>b</sup>, Helena R. Karnilowicz<sup>a</sup>, Phoebe Lam<sup>c</sup>, Emily A. Butler<sup>d</sup>, Oliver P. John<sup>a</sup>, Iris B. Mauss<sup>a</sup>

<sup>a</sup> University of California, Berkeley, United States

<sup>b</sup> Washington University in St. Louis, United States

<sup>c</sup> Northwestern University, United States

<sup>d</sup> University of Arizona, United States

## ARTICLE INFO

### Keywords:

Empathy  
Empathic accuracy  
Physiological linkage  
Friends

## ABSTRACT

How do people come to know others’ feelings? One idea is that affective processes (e.g., physiological responses) play an important role, leading to the prediction that linkage between one’s physiological responses and others’ emotions relates to one’s ability to know how others feel (i.e., empathic accuracy). Participants ( $N = 96$ , 48 female friend pairs) completed a stressful speech task and then provided continuous ratings of their own (as “targets”) and their friend’s (as “perceivers”) emotional experience for the video-taped speeches. We measured physiology-physiology linkage (linkage between perceivers’ and targets’ physiology), physiology-experience linkage (linkage between perceivers’ physiology and targets’ experience), and empathic accuracy (linkage between perceivers’ ratings of targets’ experience and targets’ ratings of their experience). Physiology-experience (but not physiology-physiology) linkage was associated with greater empathic accuracy even when controlling for key potential confounds (random linkage, targets’ and perceivers’ emotional reactivity, and relationship closeness). Results suggest that physiological responses play a role in empathic accuracy.

## 1. Introduction

Humans are inherently social and thus must learn to effectively navigate interactions with others. The ability to accurately know the emotions of others – empathic accuracy – is perhaps one of the most critical skills for smooth social interactions (Ickes, 1993; Singer & Lamm, 2009; Zaki, Bolger, & Ochsner, 2008). Indeed, empathic accuracy is associated with greater use of helpful emotional and instrumental support (Verhofstadt, Buysse, Ickes, Davis, & Devoldre, 2008), more prosocial behavior (Eckland, Huang, & Berenbaum, 2020), and greater relationship satisfaction (for a review, see Sened et al., 2017). Overall, empathic accuracy has meaningful and important consequences for social interactions and relationships, and a growing body of work examines *how* people (“perceivers”) come to know what others (“targets”) are feeling (e.g., Eckland & English, 2019; Ickes et al., 2000; Thomas, Fletcher, & Lange, 1997; Zaki, Bolger, & Ochsner, 2009).

Affective approaches postulate that empathic accuracy is a predominantly emotional process that involves physiological responses when observing targets’ emotion (e.g., Batson, Fultz, & Schoenrade,

1987; Levenson & Ruef, 1997). Physiological responses could be involved in two ways. First, they could serve as an internal signal of targets’ emotional experience, giving rise to empathic accuracy (e.g., Decety & Jackson, 2004). Second, they could be a consequence of empathic accuracy, such that accurately knowing what someone is feeling might give rise to compatible physiological responses (e.g., Dezechache, Jacob, & Grèzes, 2015). No matter whether physiological responses are involved in one or both of these ways, affective theories lead to the prediction that the degree to which perceivers’ physiological responses are linked with targets’ emotional responses should be associated with empathic accuracy.

Broadly, research indeed suggests that perceiving targets’ emotions can result in physiological responses for the perceiver, including increased electrodermal and cardiovascular activity (e.g., Hein, Lamm, Brodbeck, & Singer, 2011; Krebs, 1975; Vaughan & Lanzetta, 1981). However, *overall* physiological responses to a target’s emotions might have nothing to do with the target; instead, it could be simply indicative of the perceiver’s emotional experience. Thus, it is critical to examine whether perceivers’ physiological responses are linked with either

\* Corresponding author.

E-mail address: [fzerwas@berkeley.edu](mailto:fzerwas@berkeley.edu) (F.K. Zerwas).

targets' physiological responses or emotional experience and what this linkage might mean for empathic accuracy. To test the prediction that the degree to which perceivers' physiological responses are linked with targets' emotional responses should be associated with empathic accuracy, one must utilize a within-person approach, measuring continuously both perceivers' and targets' responses and then assessing how correlated they are.

Physiology-physiology linkage<sup>1</sup> (e.g., hearts beating in a correlated way) plays a critical and complex role in predicting romantic relationship satisfaction (for a review, see Timmons, Margolin, & Saxbe, 2015) and thus might also be important for other social outcomes like empathic accuracy. Some work suggests that physiology-physiology linkage relates to trait levels of empathy, subjective reports of perceived empathic accuracy, and accurate judgements of others' likability and confidence (for a review, see Thorson, 2018). This body of work provides important insights on the associations between physiology-physiology linkage and empathic accuracy; however, the approaches utilized in this work have limitations.

Specifically, self-report measures of trait empathy and subjective empathic accuracy do not truly capture the construct of interest because neither measure compares perceivers' ratings with some measure of "the truth" (i.e., targets' ratings). In other words, a more informative measure of empathic accuracy should include a comparison of perceivers' ratings to targets' rating of their own experience so that higher accuracy reflects the ability to predict targets' feelings. Additionally, single ratings of empathic accuracy do not capture the complexity of this process; empathic accuracy requires perceivers to continuously infer the emotions of targets over the course of an interaction. In order to get at the more complex phenomenon of physiology-physiology linkage and empathic accuracy between two people across time, researchers must continuously measure physiological responses and emotional experience from two people to calculate within-person indices of linkage and empathic accuracy. Some studies have examined empathic accuracy in a similar way (e.g., Zaki et al., 2008; Zaki, Bolger et al., 2009; Zaki, Weber, Bolger, & Ochsner, 2009) but did not examine its association with physiological responses.

To our knowledge, only one study has used continuous measures of physiology and emotional experience from two people to examine the association between physiology-physiology linkage and empathic accuracy. In their pioneering work, Levenson and Ruef (1992) used videos from a previous study where participants provided continuous ratings of how they felt during a conversation, as well as continuous physiological measures. Participants ( $N = 31$ ) were each asked to watch two different videos and provide continuous emotion ratings for one of the two target people in the video. They also collected continuous physiology (heart rate, general somatic activity, skin conductance, finger pulse transit time, finger pulse amplitude, and a composite measure of these five measures) from the perceivers while they provided ratings. Overall, greater physiology-physiology linkage (e.g., higher correlations of perceiver physiology and target physiology across time) in skin conductance level, finger pulse transit time, and the composite was positively associated with empathic accuracy for targets' negative emotions. They did not find evidence for the association between

physiology-physiology linkage and empathic accuracy for targets' positive emotions. This innovative work has been influential in theorizing about physiology-physiology linkage and empathic accuracy, but it has not yet been replicated, potentially due to the complex and intensive methodology.

Levenson and Ruef (1992) focused specifically on physiology-physiology linkage in their work; however we could ask about linkage with any aspect of targets' emotional responses (i.e., physiological responses, emotional experience, and expressive behavior; Dolan, 2002; Ekman, 1992; Izard, 1977; Lang, 1988; Lazarus, 1991; Levenson, 1994; Panksepp, 1994). In addition to focusing on physiology-physiology linkage, we also reasoned that empathic accuracy most fundamentally involves understanding the targets' emotional experience rather than the other components of the emotional response process (e.g., physiological responses or behavioral expressions). Thus, the extent to which perceivers' physiology coordinates with targets' experience, rather than with targets' physiology, should be most informative. We are unaware of any work examining physiology-experience linkage and empathic accuracy.

In short, much of the existing work examining the role of physiological responses in empathic accuracy have used limited measures and only one study suggests that physiology-physiology linkage relates to empathic accuracy. Consequently, several gaps remain. At the most fundamental level, we do not have a comprehensive understanding of the extent to which perceivers' physiology links with targets' physiology and emotional experience as well as the extent to which perceivers' ratings of targets' emotional experience link with targets' emotional experience. At a more complex level, we do not have a comprehensive understanding of the extent to which physiology-physiology linkage and physiology-experience linkage relate to empathic accuracy.

### 1.1. The present research

The present study examined the role that physiological responses play in empathic accuracy. To do so, we examined the relationship between physiology-physiology linkage, physiology-experience linkage, and empathic accuracy in female friends. First, all participants completed a validated standardized laboratory stress task, in which they gave a video-recorded speech with little time to prepare (i.e., the "speech task"). Then, in the "empathy task," all participants watched a video recording of their own speech to provide continuous ratings of their own emotional experience (target's emotional experience) and watched a video recording of their friend's speech to provide continuous ratings of their friend's emotional experience (perceiver's ratings of the target's emotional experience) using a cued-recall approach. Physiological responses were continuously collected during the speech task and the empathy task.

We had two preregistered aims (<https://osf.io/yyst48/registrations>): First, we wanted to characterize physiology-physiology linkage, physiology-experience linkage, and empathic accuracy. To do so, we examined average levels of and variation in physiology-physiology linkage (i.e., linkage between perceivers' and targets' physiology), physiology-experience linkage (i.e., linkage between perceivers' physiology and targets' emotional experience), and empathic accuracy (i.e., linkage between perceivers' ratings of targets' emotional experience and targets' ratings of their own emotional experience). Second, we wanted to clarify the associations between physiology-physiology linkage and physiology-experience linkage on the one hand and empathic accuracy on the other hand. To do so, we examined whether physiology-physiology linkage or physiology-experience linkage relate to empathic accuracy and accounted for key confounds (random linkage, targets' and perceivers' emotional reactivity, and relationship closeness).

This approach has several strengths. First, we examined these aims in a community sample of adult female friend pairs instead of pairs of strangers (Levenson & Ruef, 1992). Maintaining close friendships is

<sup>1</sup> We use the term linkage to refer to the level of covariation between two people in their moment-to-moment emotional responses (i.e., physiology or behavior). Other work refers to similar phenomena – covariation of emotional responses either between or within people – as response coherence (Mauss et al., 2005), response system coherence (Ekman, 1992), organization of response components (Frijda, Ortony, Sonnemans, & Clore, 1992; Scherer, 1984; Witherington, Campos, & Hertenstein, 2001), response component syndromes (Averill, 1980; Reisenzein, 2000), concordance (Nesse et al., 1985; Wilhelm & Roth, 2001), organization of response tendencies (Lazarus, 1991; Levenson, 1994), synchrony (Danyluck & Page-Gould, 2018), and coregulation (Butler & Randall, 2013; Timmons et al., 2015).

critical to psychological and physical health (Holt-Lunstad, Robles, & Sbarra, 2017) rendering the examination of these aims within the context of friendships particularly vital. This approach also has high external validity since we often infer the emotions of individuals we know well in our daily experiences.

Second, we used an impromptu speech task that induced strong and variable emotional experiences and physiological responses. This is particularly important given it would be difficult to examine linkage between emotion responses if the elicited emotion response was weak (e.g., flat lines in emotion experience, little physiological activation; cf. Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005). Additionally, this design differs from prior work in that all participants played the role of both target and perceiver instead of having targets complete a task and having a different set of perceivers rate a subset of the targets. To understand the importance of simply sharing the same context (giving a speech), we calculated random versions of our key variables: linkage between perceivers' physiology and random targets' physiology and experience (versus their friend's physiology and experience). These random measures of linkage served as comparisons in Aim 1 and robustness tests in Aim 2.

Third, we examined both physiology-physiology linkage and physiology-experience linkage given how both of these constructs might play a role in empathic accuracy. On the one hand, finding evidence for the relationship between physiology-physiology and empathic accuracy would suggest that perceivers' physiological responses relate to empathic accuracy to the extent that they are linked with targets' physiology. On the other hand, finding evidence for the relationship between physiology-experience linkage and empathic accuracy would suggest that perceivers' physiological responses relate to empathic accuracy to the extent that they are linked with targets' emotional experience.

Fourth, we sampled across multiple measures of autonomic physiological responses. We examined inter-beat interval (IBI) as a key measure of cardiac (mixed sympathetic and parasympathetic) activity, finger pulse amplitude (FPA) as a key measure of sympathetic activation (predominantly adrenergic; Elgendi, 2012), and skin conductance level (SCL) as a key measure of sympathetic activation (predominantly cholinergic; Machado-Moreira et al., 2012). Skin conductance level is also related to emotional arousal (Mendes, 2009). We collected four additional measures to include in a composite measure with IBI, FPA, and SCL to capture autonomic arousal more generally: finger pulse transit time (FPTT), ear pulse amplitude (EPA), ear pulse transit time (EPTT), and skin temperature (SKT). FPTT, EPA, EPTT, and SKT all measure sympathetic activation. Similar composite measures have been utilized in past research on linkage to provide a more comprehensive indicator of overall physiological activation that individual physiological indicators cannot provide (Levenson & Ruef, 1992; Soto & Levenson, 2009).

Fifth, to assess physiology-physiology linkage, physiology-experience linkage, and empathic accuracy, we took a within-person approach by obtaining continuous ratings of targets' emotional experience and physiological responses as well as perceivers' ratings of targets' emotional experience and perceivers' physiological responses. Compared to single measures, the current measures allow us to examine linkage by utilizing within-person indices, minimize measurement error that is due to aggregation across longer periods of time, and reflect the complexity of inferring targets' emotions in unfolding social interactions. Using within-person indices allowed us to account for varying lags between the target and the perceiver by using cross-correlations, since it might take several seconds for the perceivers to react to the targets. Additionally, the current measure of empathic accuracy compares perceivers' ratings of targets' experience to targets' self-reported experience to provide an objective measure of the ability to correctly infer the targets' experience.

In sum, we examined physiology-physiology linkage, physiology-experience linkage, and empathic accuracy using continuous measures

of physiology and experience in a sample of 48 female friend pairs where both participants completed a standardized laboratory stress task and provided ratings, via a cued-recall approach, of their own experience as well as ratings of their friend's experience. The current approach provided a robust examination of our preregistered aims and analytic plan.

## 2. Method

### 2.1. Participants

We preregistered key aspects of the method including the exclusion criteria, selection of measures, and analytic approach. Data were collected during a day of lab procedures as part of a larger study on well-being in 160 participants from the San Francisco Bay Area community. Participants were recruited with a female friend (i.e., 80 female friend pairs). Eligible friends met the following criteria: they were fluent in English and one of the two friends had experienced a stressful life event of at least moderate impact (e.g. change of residence) within the past 6 months. Due to time constraints during the lab session, 32 pairs of participants were not able to complete both of the critical lab tasks described below (i.e., the speech task and the empathy task) and were therefore excluded from the present analyses, as noted in the preregistration. We did not exclude any participants for the second exclusion criterion listed in our preregistration (i.e., for having no variability in their rating dial response) because all the participants who completed both tasks had variation in their rating dial responses. In total, 96 female participants (48 pairs) were included in the present analyses. Their mean age was 42 years ( $SD = 15.4$ ;  $Range = 25-76$ ); 60 % ( $n = 58$ ) identified as European American, 26 % ( $n = 25$ ) as Asian American, and 14 % ( $n = 13$ ) identified as another ethnicity.

### 2.2. Procedures

The Institutional Review Board at the University of California, Berkeley approved the study procedures (protocol ID: 2014-10-6844). The current study is part of a larger project that consisted of an online entrance survey, a lab session, online daily diaries, and an online exit survey; participants received \$230 for completing all parts of the project. The measures for the current study come from the lab session and the online entrance survey completed at least four days earlier, in which participants completed measures of demographics as well as several individual difference measures included in our preregistration of exploratory analyses.<sup>2</sup> The day-long lab session included two tasks designed to measure physiology-physiology linkage, physiology-experience linkage, and empathic accuracy. The first task (referred to as the speech task) took place in the morning at the beginning of the lab session and required participants to deliver a video-recorded speech. The second task (referred to as the empathy task) took place after the speech task and before the joint lunch so that the friends could not discuss their experiences during the stress task with one another. The empathy task required participants to provide ratings of their own and their friend's emotional experience during the speech task. Table 1 provides an overview of the measures we obtained in these tasks and how we used them to assess physiology-physiology linkage, physiology-experience linkage, and empathic accuracy.

#### 2.2.1. Speech task

Participants completed a validated stress induction based on the Trier social stress task (Kirschbaum, Pirke, & Hellhammer, 1993; Mauss, Wilhelm, & Gross, 2003, 2004). Specifically, they were asked to deliver a three-minute speech focused on how their communication skills, both

<sup>2</sup> None of these global self-report measures (empathy, emotional expressivity, and expressive suppression) was related to any of the key variables in this report.

**Table 1**  
Overview of the Measures Obtained from the Speech and Empathy Task.

Construct	Operationalization	Measures used in operationalization of construct	
		Target	Perceiver
Physiology-physiology linkage	Cross-correlation between perceivers' and targets' physiological responses	Physiological responses while giving the speech	Physiological responses while watching the video of the target give the speech
Physiology-experience linkage	Cross-correlation between perceivers' physiological responses and targets' emotional experience	Watching the video of themselves giving the speech and rating their own emotional experience during speech	Physiological responses while watching the video of the target giving the speech
Empathic accuracy	Cross-correlation between perceivers' ratings of targets' emotional experience and targets' ratings of their own emotional experience	Watching the video of themselves giving the speech and rating their own emotional experience during the speech	Watching the video of the target give the speech and rating the target's emotional experience during the speech

*Note.* All participants played the role of both the target and the perceiver; in other words, all participants gave a speech and rated their own speech (i.e., target role) and all participants watched and rated their friend's speech (i.e., perceiver role). The order in which they rated the videos (i.e., self first, friend second) was fixed and all ratings were completed using a cued-recall approach.

verbal and written, qualify them for a fictitious job. They were informed that the speech would be videotaped and that the videotape would be shown to a panel of expert judges, who would rate the quality of their performance, and that their friend would watch the tape later in the day. They were allowed two minutes to prepare for the speech without taking notes. The experimenter then entered the subject room to position the video camera directly in front of the participant.

### 2.2.2. Empathy task

For this task, we used a cued-recall approach (cf., [Gottman & Levenson, 1985](#); [Schulz & Waldinger, 2004](#)) in which all participants watched the video recording of their own speech first and their friend's speech thereafter. While watching the tape, participants provided continuous ratings of their own emotional experience during the speech (i.e., target role); then they watched and provided continuous ratings of their friend's emotional experience during their speech (i.e., perceiver role). We did not counterbalance the order of roles because the available sample size did not suffice to allow an order condition and because we were primarily interested in individual differences (thus making concerns about order that were consistent across participants less of a concern). To gather continuous ratings of emotional experience, we instructed participants in the use of a rating dial (cf. [Mauss et al., 2005](#); [Ruef & Levenson, 2007](#)). Participants completed a brief training period to get familiar with the rating dial and reduce cognitive load during the task itself. They were told to adjust the dial from 0 ("extremely negative") to 5 ("extremely positive") as often as needed so that it would always reflect the amount of emotion they felt; for analyses, we standardized and reverse-scored the ratings such that higher values indicated more negative emotion.

**2.2.2.1. Target ratings.** Participants were instructed to provide continuous ratings of their own emotional experience while they were delivering the speech using a validated cued-recall approach (cf. [Mauss et al., 2005](#)). They received the following instructions:

For the next task, you will be watching a video of the speech task you did earlier. While you watch, please use the rating dial continuously to indicate how you were feeling during the task.

**2.2.2.2. Perceiver ratings.** Next, participants were instructed to provide continuous ratings of their friend's emotional experience while their friend was delivering the speech using a cued-recall approach. They received the following instructions:

Next, you will be watching the video of the speech task that your friend did. Similarly to the previous task, we'd like you to use the rating dial. However, this time we'd like you to use the rating dial to continuously indicate how you think your friend was feeling during the speech task.

### 2.2.3. Physiological responses

During the speech task and the empathy task, physiological channels were sampled at 1000 Hz using laboratory software. Customized analysis software ([Wilhelm, Grossman, & Roth, 1999](#)) was used for physiological data reduction, artifact control, and computation of second-by-second scores for each participant. Second-by-second physiological scores were exported using the coherence module in ANSLAB (anslab.net) after down-sampling to 4 Hz. Once we exported the second-by-second data, we used the programming software R to reverse-score all of the physiological measures except for skin conductance level so that higher scores on all measures indicated more physiological activation. In order to form a physiology composite of the seven physiological measures, we then standardized each of the separate measures and aggregated across them.

We examined inter-beat interval (IBI) as a key measure of cardiac (mixed sympathetic and parasympathetic) activity, finger pulse amplitude (FPA) as a key measure of sympathetic activation (predominantly adrenergic; [Elgendi, 2012](#)), and skin conductance level (SCL) as a key measure of sympathetic activation (predominantly cholinergic; [Machado-Moreira et al., 2012](#)). Skin conductance level is also related to emotional arousal ([Mendes, 2009](#)). We also collected four other measures to include in a composite with IBI, FPA, and SCL: finger pulse transit time (FPTT), ear pulse amplitude (EPA), ear pulse transit time (EPTT), and skin temperature (SKT). FPTT, EPA, EPTT, and SKT all measure sympathetic activation. Similar composites have been utilized in past research on empathic accuracy ([Levenson & Ruef, 1992](#); [Soto & Levenson, 2009](#)).

We measured the electrical signal of the heart with an electrocardiogram (ECG) using a Lead I configuration with three MindWare Disposable ECG 1-1/2" Electrodes: the first sensor was placed in the middle of the participants' right clavicle while the second and third sensors were placed on the bottom of the lowest right and left ribs, respectively. From this channel, we calculated the *inter-beat interval* (IBI) from the interval between successive R-waves in the ECG, which measures the time of one complete heart cycle. Values from ectopic or other kinds of abnormal beats were deleted and replaced by linearly interpolated values. After reverse scoring, higher values indicate faster heart rate.

Additionally, we measured the volumetric variations of blood circulation using a photoplethysmogram (PPG). One sensor was placed on the tip of the non-dominant ring finger to measure volumetric variations in the finger and another sensor was placed on participants' left ear lobe to measure volumetric variations in the ear. We calculated two measures from each of these channels; *finger pulse amplitude* (FPA), *finger pulse transit time* (FPTT), *ear pulse amplitude* (EPA), and *ear pulse transit time* (EPTT). FPA and EPA were measured by obtaining the difference between a maximum and adjacent minimum that represents the amplitude of PPG pulses ([Webster 1997](#)). FPTT and EPTT were calculated by the time, in milliseconds, it takes a pulse wave to travel from the heart to the

finger/ear. After reverse scoring each of these four measures (i.e., FPA, EPA, FPTT, EPTT), higher values indicate more sympathetic nervous system activation.

Skin conductance measurements were acquired using a constant-voltage device that passed 0.5 V between MindWare Disposable GSC 1-1/2" Foam Electrodes on palms of participants' non-dominant hand. From this channel, we calculated *skin conductance levels* (SCL) which measures skin conductance over longer periods of time as opposed to a specific stimulus onset (Mendes, 2009). Higher values indicate more sympathetic nervous system activation.

*Skin temperature* (SKT) was measured using a 19 mm stainless steel disc attached to participants' pinky finger. After reverse scoring, higher values indicate more sympathetic nervous system activation.

### 2.3. Measures

#### 2.3.1. Key study variables

The key variables in this study are physiology-physiology linkage, physiology-experience linkage, and empathic accuracy. As indicated in Table 1, we operationalized (a) physiology-physiology linkage as the cross-correlation (with the target leading the perceiver) between perceivers' and targets' physiological responses, (b) physiology-experience linkage as the cross-correlation (with the target leading the perceiver) between perceivers' physiological responses and targets' emotional experience, and (c) empathic accuracy as the cross-correlation (with the target leading the perceiver) between perceivers' ratings of targets' emotional experience and targets' ratings of their own emotional experience. Each of these three measures captures the linkage between two separate time series variables (e.g., empathic accuracy is the linkage between perceivers' ratings of the targets' emotional response and targets' emotional response). Thus, we calculated two within-dyad indices of each measure (i.e., one for each participant in the dyad acting as perceiver, since each participant was both a perceiver and a target).

Fig. 1 summarizes the main steps in the procedure, which are similar to those used in Butler, Gross, & Barnard (2014) and Mauss et al. (2005, 2011).

We used cross-correlations since time series data are expected to be non-stationary with mean and variance varying over time as people become more or less reactive. Cross-correlations are an appropriate indicator of linkage between two sets of time-series variables because they take into account between-variable associations that occur due to shared mean, slope, or variance changes, shared oscillations or shared momentary fluctuations (Butler et al., 2014).

The first step when using cross-correlations is to calculate lagged correlations for pairs of responses (Dan-Glauser & Gross, 2013; Mauss et al., 2005; 2011; Sze, Gyurak, Yuan, & Levenson, 2010) as illustrated in Step 1 of Fig. 1. One decision that needed to be made was the time-window to consider for the lags between responses. We preregistered a window of -15 to 0 s for two reasons. First, we did not allow for positive lags so that the target responses always led the perceiver responses, since we were interested in the perceivers' responses to targets. As an example, determining whether a perceiver is empathically accurate relies on understanding if the perceiver is reacting to the cues from the target and then accurately rating the target's emotions. A positive lag would suggest that the target is reacting to the perceiver, which is not possible in this design. Second, we chose the length of the window to be 15 s primarily because physiological responding to an event can take up to 15 s (e.g., Gratton, 2000; Kettunen, Ravaja, Näätänen, Keskivaara, & Keltikangas-Järvinen, 1998; Levenson, 1988). We used the same lag window for empathic accuracy to be consistent with the linkage measures that include physiology and because it is possible that it takes up to 15 s for a perceiver to react to the emotional cues provided by the targets.

To obtain the most accurate index of association among responses for each measure, we chose (and preregistered) the highest value of the resulting 16 cross-correlations for each participant (Step 2 of Fig. 1). In

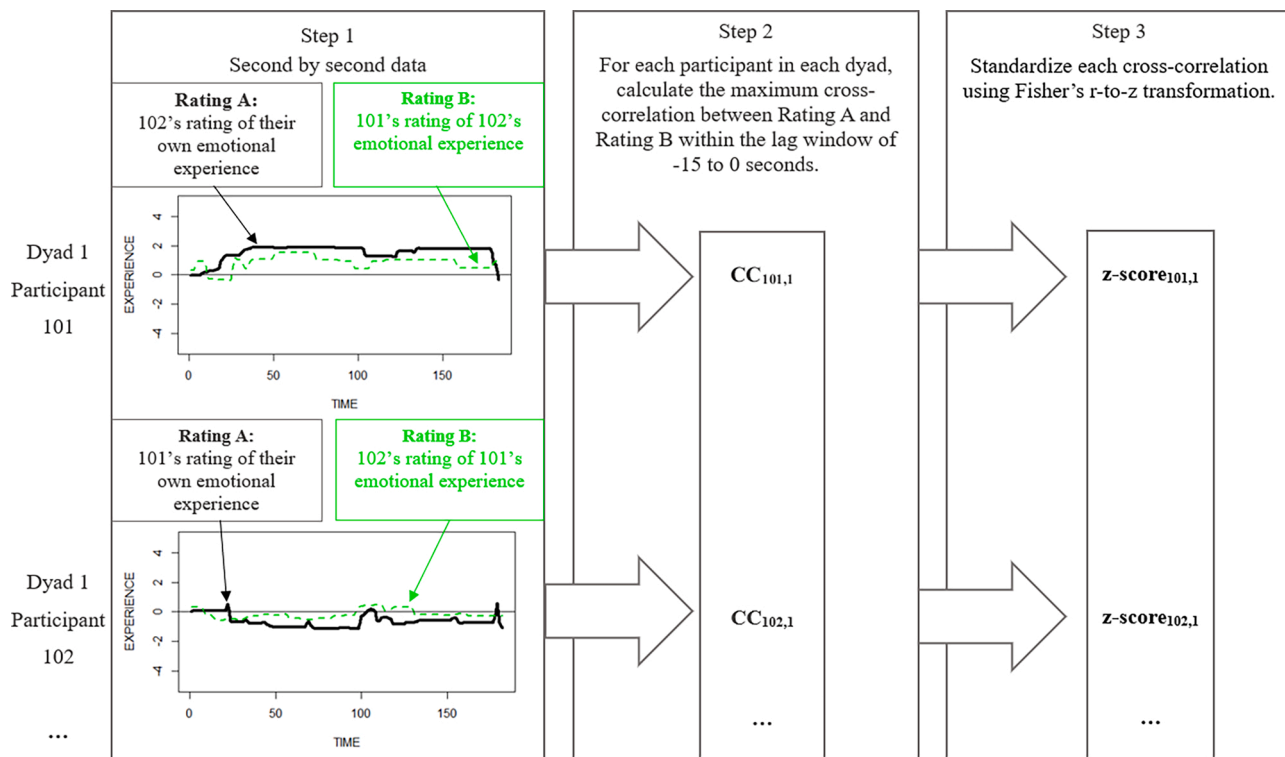


Fig. 1. An outline of the data processing steps to obtain indices of empathic accuracy. The solid line indicates Rating A and the dashed line indicates Rating B. This figure only focuses on empathic accuracy as an example; however, we used the same procedures for physiology-physiology linkage, physiology-experience linkage, random physiology-physiology linkage, and random physiology-experience linkage. The ellipses indicate that the same procedures were conducted for each dyad and participant.

other words, we chose the cross-correlation at the lag time within the window of  $-15$  to  $0$  s where the two responses were most strongly aligned with one another. The resulting cross-correlation for each individual indicates the extent to which their time series covaried with their friend's time series, while taking into account potential lags between measures. To render these individual-level indices linearly scaled and usable in individual-difference analyses, we standardized them using Fisher's  $r$ -to- $z$  transformation (Step 3 of Fig. 1). The cross-correlations for each measure (i.e., physiology-physiology linkage, physiology-experience linkage, empathic accuracy) can theoretically range from  $1.0$  (perfect linkage between responses) to  $0$  (no linkage between responses) to  $-1.0$  (perfect inverse linkage between responses). See Fig. 2 for examples of time series used to compute the cross-correlations for empathic accuracy.

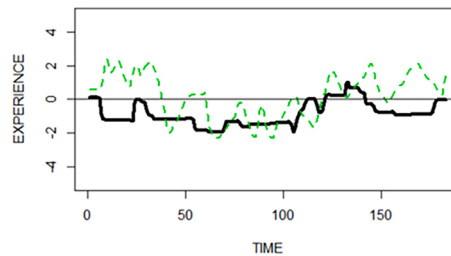
### 2.3.2. Measures for robustness tests

To ensure that the associations between empathic accuracy and the

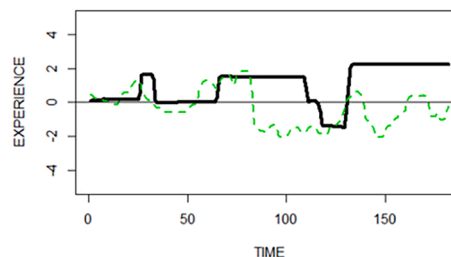
two predictor variables (i.e., physiology-physiology linkage and physiology-experience linkage) were specific to linkage within dyads versus more general linkage-like effects (e.g., those attributable to being in a shared context), we examined random physiology-physiology linkage and random physiology-experience linkage as comparisons (for Aim 1) and as control variables (for Aim 2). Random linkage (of either physiology or experience) was operationalized as the cross-correlation between a perceiver's responses and a random target's (i.e., *not* their friend) responses. We randomly paired participants with another participant by assigning random dyad IDs using R's sample function. Once we had randomly paired participants, we used the same procedures regarding cross-correlations outlined above to calculate the two variables (i.e., random physiology-physiology linkage and random physiology-experience linkage).

Next, to ensure that significant associations were not due to perceivers' or targets' emotional reactivity (versus linkage), we calculated perceivers' peak physiological activation during the empathy task and

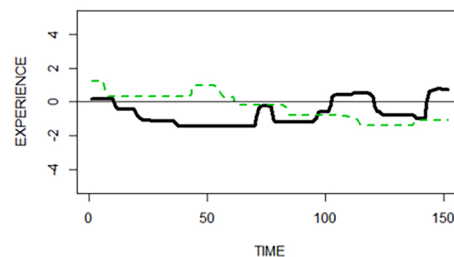
- a.) Positive cross-correlation ( $CC_{117,9} = .56$ ) between the target's rating of their own emotional experience (solid line) and the perceiver's rating of the target's emotional experience (dashed line). This is an example of relatively high empathic accuracy.



- b.) Null cross-correlation ( $CC_{181,41} = .005$ ) between the target's rating of their own emotional experience (solid line) and the perceiver's rating of the target's emotional experience (dashed line). This is an example of relatively low empathic accuracy.



- c.) Negative cross-correlation ( $CC_{204,52} = -.31$ ) between the target's rating of their own emotional experience (solid line) and the perceiver's rating of the target's emotional experience (dashed line). This is an example of relatively low empathic accuracy.



**Fig. 2.** Examples of positive, null, and negative cross-correlations for empathic accuracy. Higher experience values indicate more negative emotional experience, whereas lower experience values indicate more positive emotional experience.

targets' peak emotional experience during the speech task.

Finally, because a perceiver's closeness to the target might influence how well the perceiver can read the target's emotions, we statistically controlled for the perceiver's rating of relationship closeness.

#### 2.4. Data-analytic plan

For Aim 1 (i.e., to test the magnitude and variability of linkage and empathic accuracy), we conducted multilevel intercept-only models using the nlme package (version 3.1.148) in R (version 4.0.2) with participants nested within dyads to examine the average magnitude and variability of the various linkage measures (physiology-physiology, random physiology-physiology, physiology-experience, random physiology-experience, and empathic accuracy). Equation 1 shows an example equation to examine the magnitude and variability of IBI-IBI linkage.

$$\begin{aligned} (\text{IBI-IBI linkage})_{ij} &= \beta_0j + \zeta_j + \varepsilon_{ij}; \\ (\zeta_j) &\sim N(0, \psi); \\ (\varepsilon_{ij}) &\sim N(0, \theta) \end{aligned}$$

IBI-IBI linkage for a particular individual as perceiver (i) in a specific dyad (j) is modeled by:  $\beta_0j$  is the overall mean of IBI-IBI linkage,  $\zeta_j$  is the deviation of the mean linkage for dyad j from the overall mean and  $\varepsilon_{ij}$  is the residual for the IBI-IBI linkage score of individual ij. The variance of  $\zeta_j$  from dyad to dyad is  $\psi$ . This represents the between-dyad variance in IBI-IBI linkage. The variance of  $\varepsilon_{ij}$  is  $\theta$  which is the residual variability in IBI-IBI linkage across individuals. This represents the within-dyad variance.

For Aim 2 (i.e., to test the associations between linkage and empathic accuracy), we conducted multilevel random intercept models with participants nested within dyads and either the z-transformed physiology-physiology linkage cross-correlations or the z-transformed physiology-experience linkage cross-correlations predicting the z-transformed empathic accuracy scores. We did not center the linkage variables because they are z-transformed cross-correlations which means that the zero point of the z-scored variables is the average level of linkage for the raw cross-correlations. Thus, the intercept in these models refers to the predicted value of empathic accuracy at the average value of the raw cross-correlations. Equation 2 shows an example equation to examine the association between empathic accuracy as the outcome variable and IBI-IBI linkage as the predictor variable.

$$\begin{aligned} (\text{Empathic Accuracy})_{ij} &= \beta_0j + \beta_1jX(\text{IBI-IBI linkage})_{ij} + \zeta_j + \varepsilon_{ij}; \\ (\zeta_j | X(\text{IBI-IBI linkage})_{ij}) &\sim N(0, \psi); \\ (\varepsilon_{ij} | X(\text{IBI-IBI linkage})_{ij}) &\sim N(0, \theta) \end{aligned}$$

Empathic accuracy for a particular individual as perceiver (i) in a specific dyad (j) is modeled by:  $\beta_0j$  is the overall mean of empathic accuracy,  $\beta_1j$  represents the coefficient for the IBI-IBI linkage score,  $\zeta_j$  is the deviation of the mean empathic accuracy for dyad j from the overall mean, and  $\varepsilon_{ij}$  is the residual for the empathic accuracy score of individual ij. The variance of  $\zeta_j$  from dyad to dyad is  $\psi$ . This represents the between-dyad variance in empathic accuracy. The variance of  $\varepsilon_{ij}$  is  $\theta$  which is the residual variability in empathic accuracy across individuals. This represents the within-dyad variance.

Overall the missingness per variable was 1.0 % for empathic accuracy, 4.2 % for IBI-IBI linkage, 2.1 % for IBI-experience linkage, 8.3 % for SCL-SCL linkage, 5.2 % for SCL-experience linkage, 4.2 % for FPA-FPA linkage, 3.1 % for FPA-experience linkage and 0% for both physiology composite-physiology composite linkage and physiology composite-experience linkage. Due to the relatively low levels of missingness, we used listwise deletion for all analyses.

### 3. Results

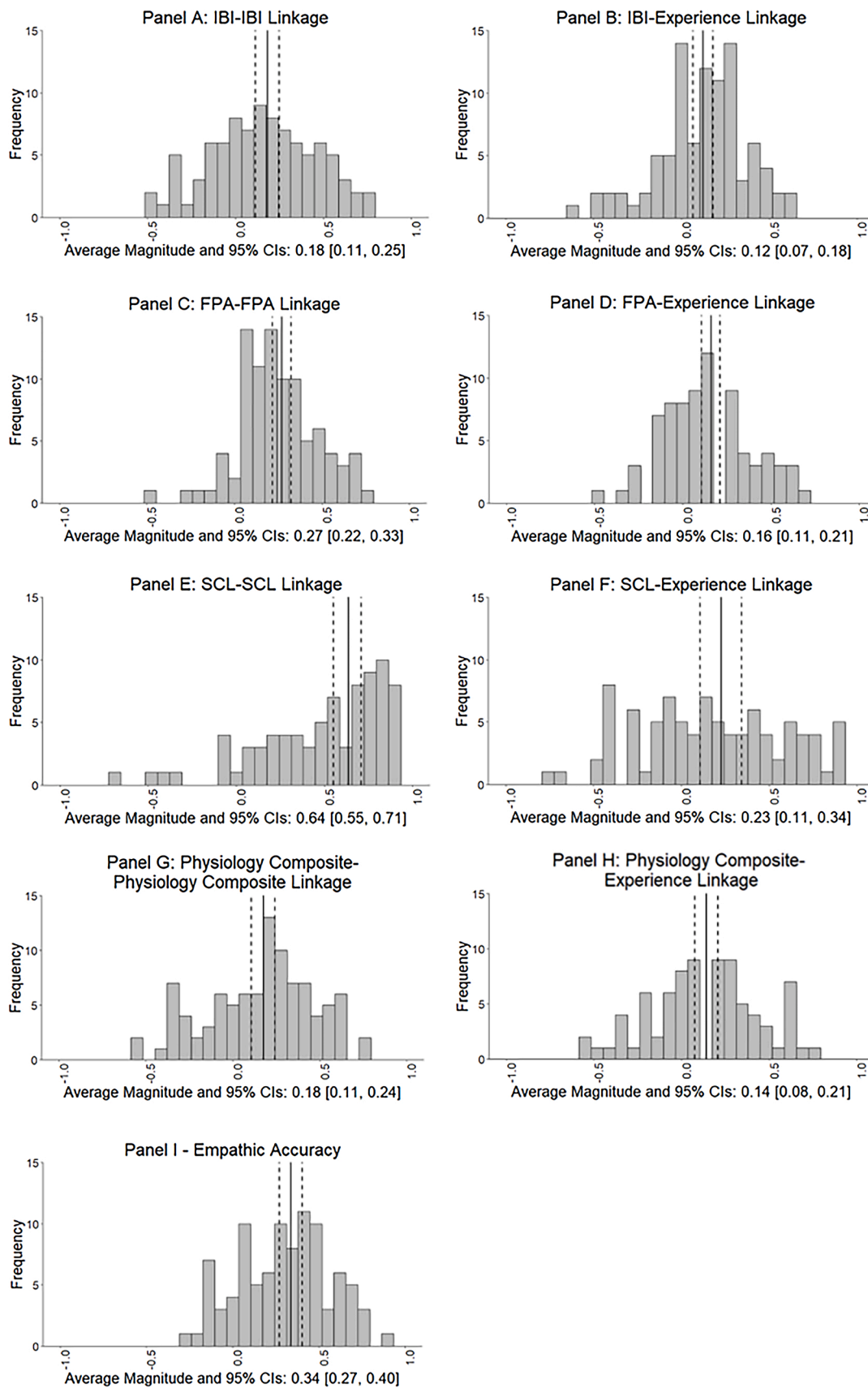
#### 3.1. Aim 1: magnitude and variability of linkage and empathic accuracy

Fig. 3 shows the distribution of the cross-correlations for each linkage variable and empathic accuracy. In order to provide the means and 95 % confidence intervals for those distributions, we transformed the intercept estimates from the random-intercept models from z-scores back into cross-correlations. When interpreting these results, it is important to remember that we sampled across multiple measures of autonomic physiological responding. We examined inter-beat interval (IBI) as a key measure of cardiac (mixed sympathetic and parasympathetic) activity, finger pulse amplitude (FPA) as a key measure of sympathetic activation (predominantly adrenergic; [Elgendi, 2012](#)), skin conductance level (SCL) as a key measure of sympathetic activation (predominantly cholinergic; [Machado-Moreira et al., 2012](#)), and a composite of physiological measures that reflect multiple physiological systems and responses (i.e., inter-beat interval, finger pulse amplitude, skin conductance level, finger pulse transit time, ear pulse amplitude, ear pulse transit time, and skin temperature). Thus, the magnitudes of the linkage variables are not directly comparable; however, the magnitudes are still informative for each individual measure.

Overall, as shown in panels A, C, E, and G of Fig. 2, we found significant average levels of physiology-physiology linkage for IBI-IBI linkage ( $M$  [95 % CI] = 0.18 [0.11, 0.25]), FPA-FPA linkage ( $M$  [95 % CI] = 0.27 [0.22, 0.33]), SCL-SCL linkage ( $M$  [95 % CI] = 0.64 [0.55, 0.71]), and physiology composite-physiology composite linkage ( $M$  [95 % CI] = 0.18 [0.11, 0.24]). Each of the average physiology-physiology linkage values were positive, suggesting that greater activation in the perceiver was linked with greater activation in the target. The intraclass correlation coefficients for the physiology-physiology linkage variables were all below .08 suggesting that there was very little dyad-level variance and very high individual-level variance. In other words, most variability was on the level of the individual who was acting as the perceiver and not at the level of the dyad.

Additionally, we found positive and significant levels of random physiology-physiology linkage for random IBI-IBI linkage ( $M$  [95 % CI] = 0.15 [0.10, 0.21]), random FPA-FPA linkage ( $M$  [95 % CI] = 0.26 [0.21, 0.32]), random SCL-SCL linkage ( $M$  [95 % CI] = 0.58 [0.49, 0.66]), and random physiology composite-physiology composite linkage ( $M$  [95 % CI] = 0.15 [0.09, 0.21]). The levels of random physiology-physiology linkage were not significantly different from the levels of nonrandom physiology-physiology linkage given the overlapping confidence intervals across all measures of linkage. Observing significant levels of physiology-physiology linkage between random dyads suggests that the shared context (e.g., completing the same speech task) might have played an important role in the emergence of physiology-physiology linkage.

Overall, as shown in panels B, D, F, and H of Fig. 2, we found significant average levels of physiology-experience linkage for IBI-experience linkage ( $M$  [95 % CI] = 0.12 [0.07, 0.18]), FPA-experience linkage ( $M$  [95 % CI] = 0.16 [0.11, 0.21]), SCL-experience linkage ( $M$  [95 % CI] = 0.23 [0.11, 0.34]), and physiology composite-experience linkage ( $M$  [95 % CI] = 0.14 [0.08, 0.21]). Each of the average physiology-experience linkage values were positive indicating that higher physiological activation in the perceiver was linked with higher negative emotional experience in the target. The intraclass correlation coefficients for IBI-experience linkage, FPA-experience linkage, and physiology composite-experience linkage were below .08, again suggesting that there was very little dyad-level variance and very high individual-level variance. In other words, most variability was on the level of the individual who was acting as the perceiver and not at the level of the dyad. In contrast, the intraclass correlation coefficient for SCL-experience linkage was .29; this value suggests that more of the variance in the cross-correlations was due to dyad-level differences than for the other measures of linkage; however, the majority of the variance



**Fig. 3.** Distributions of the cross-correlations with the average estimate (solid line) for each variable and the 95 % confidence intervals (dashed lines) for the average estimate. IBI = inter-beat interval; FPA = finger pulse amplitude; SCL = skin conductance level.



was still due to individual-level differences in who was acting as the perceiver.

Additionally, we found significant levels of random physiology-experience linkage for random IBI-experience linkage ( $M$  [95 % CI] = 0.17 [0.10, 0.23]), random FPA-experience linkage ( $M$  [95 % CI] = 0.11 [0.06, 0.16]), and random physiology composite-experience linkage ( $M$  [95 % CI] = 0.12 [0.06, 0.19]); however, we did not find significant levels of random SCL-experience linkage ( $M$  [95 % CI] = 0.10 [-0.01, 0.21]). The levels of random physiology-experience linkage were not significantly different from the levels of nonrandom physiology-experience linkage given the overlapping confidence intervals across all measures of linkage. Similar to above, observing significant levels of physiology-experience linkage between random dyads for three of the four measures of linkage suggests that the shared context (e.g., completing the same speech task) might have played an important role in the emergence of physiology-experience linkage.

Next, as shown in Panel I of Fig. 2, we found significant average levels of empathic accuracy ( $M$  [95 % CI] = 0.34 [.27, .40]) and that 85 % of the empathic accuracy scores were positive. Overall, this suggests that perceivers were relatively accurate when rating targets' emotional experience. The intraclass correlation coefficient for empathic accuracy was .04, again suggesting that there was very little dyad-level variance and very high individual-level variance. In other words, most variability was on the level of the individual who was acting as the perceiver and not at the level of the dyad.

In sum, we observed positive and significant levels of each measure such that perceivers' physiology was linked with targets' physiology and emotional experience; additionally, perceivers' ratings of targets' emotional experience was linked with targets' ratings of their own emotional experience. Overall, across all measures of linkage, we observed little dyad-level variance and very high individual-level variance in the cross-correlations. In other words, most variability in linkage was due to differences in terms of which person was acting as the perceiver.

### 3.2. Aim 2: associations between key variables

Table 2 shows the estimates, t-values, and p-values for all models.

None of the physiology-physiology linkage measures were significantly associated with empathic accuracy ( $ps > .37$ ). Thus, the extent to which one's physiology was linked with one's friend's physiology did not significantly relate to empathic accuracy.

In contrast, all of the physiology-experience linkage measures were significantly associated with empathic accuracy. Broadly speaking, as shown in Table 2, the more strongly perceivers' physiological activation was linked with targets' experience of negative emotions, the greater

**Table 2**  
Multi-Level Models of Physiology-Physiology Linkage and Physiology-Experience Linkage Computed with Different Physiological Indicators Predicting Empathic Accuracy.

	<i>b</i>	<i>t</i> (df)	<i>p</i>
<i>Physiology-physiology linkage using...</i>			
Inter-beat interval (IBI)	-0.04	-0.35 (44)	.73
Finger pulse amplitude (FPA)	-0.08	-0.57 (42)	.57
Skin conductance level (SCL)	-0.05	-0.90 (38)	.37
Physiology composite	-0.04	-0.36 (46)	.72
<i>Physiology-experience linkage using...</i>			
Inter-beat interval (IBI)	0.28	2.13 (44)	.04
Finger pulse amplitude (FPA)	0.34	2.42 (43)	.02
Skin conductance level (SCL)	0.21	3.14 (41)	<.01
Physiology composite	0.29	2.65 (46)	.01

Note. Each predictor was entered separately. The physiology composite includes inter-beat interval, finger pulse amplitude, skin conductance level, finger pulse transit time, ear pulse amplitude, ear pulse transit time, and skin temperature. *b* is the unstandardized slope estimate.

was perceivers' empathic accuracy.

### 3.3. Robustness tests

To ensure that the associations between physiology-experience linkage and empathic accuracy were specific to linkage between friends versus more general linkage-like effects (e.g., those attributable to being in a shared context), we conducted another set of multilevel random intercept models accounting for random physiology-experience linkage as a predictor variable. As indicated in Table 3, results were comparable to the original analyses when accounting for random physiology-experience linkage; however, the associations between two of the physiological indicators (IBI-experience linkage, FPA-experience linkage) and empathic accuracy became marginal. None of the random physiology-experience linkage measures significantly predicted empathic accuracy ( $ps > .48$ ).

Next, to ensure that the associations between physiology-experience linkage and empathic accuracy were not due to perceivers' emotional reactivity (versus linkage), we conducted an additional set of the multilevel models accounting for perceivers' peak emotional reactivity (i.e., maximum physiological activation during the empathy task). As indicated in Table 3, results were comparable to the original analyses when accounting for perceivers' peak reactivity. None of the peak reactivity measures significantly predicted empathic accuracy ( $ps > .27$ ).

Next, to ensure that the associations between physiology-experience linkage and empathic accuracy were not due to targets' emotional reactivity during the speech task, we conducted another set of multilevel models accounting for targets' peak emotional reactivity (i.e., targets' peak negative emotional experience) during the speech task. As indicated in Table 3, results were comparable to the original analyses when accounting for targets' peak emotional reactivity. Targets' peak emotional reactivity did not significantly predict empathic accuracy ( $ps$

**Table 3**  
Multi-Level Models Testing the Robustness of Physiology-experience Linkage Predicting Empathic Accuracy.

	<i>b</i>	<i>t</i> (df)	<i>p</i>
<i>Physiology-experience linkage predicting empathic accuracy controlling for random physiology-experience linkage</i>			
Inter-beat interval (IBI)	0.27	2.01 (43)	.05
Finger pulse amplitude (FPA)	0.30	2.02 (42)	.05
Skin conductance level (SCL)	0.23	3.31 (40)	<.01
Physiology composite	0.29	2.68 (45)	.01
<i>Physiology-experience linkage predicting empathic accuracy controlling for perceivers' peak emotional reactivity during the empathy task for each physiological measure</i>			
Inter-beat interval (IBI)	0.28	2.11 (43)	.04
Finger pulse amplitude (FPA)	0.36	2.50 (41)	.02
Skin conductance level (SCL)	0.17	2.37 (37)	.02
Physiology composite	0.31	2.79 (45)	.01
<i>Physiology-experience linkage predicting empathic accuracy controlling for targets' peak emotional reactivity during the speech task</i>			
Inter-beat interval (IBI)	0.27	2.08 (43)	.04
Finger pulse amplitude (FPA)	0.33	2.35 (42)	.02
Skin conductance level (SCL)	0.21	3.14 (40)	<.01
Physiology composite	0.29	2.62 (45)	.01
<i>Physiology-experience linkage predicting empathic accuracy controlling for perceivers' ratings of relationship closeness</i>			
Inter-beat interval (IBI)	0.28	2.12 (43)	.04
Finger pulse amplitude (FPA)	0.33	2.33 (42)	.02
Skin conductance level (SCL)	0.20	2.93 (40)	.01
Physiology composite	0.29	2.70 (45)	.01

Note. The physiology composite includes inter-beat interval, finger pulse amplitude, skin conductance level, finger pulse transit time, ear pulse amplitude, ear pulse transit time, and skin temperature. *b* is the unstandardized slope estimate.

> .34).

Finally, we conducted another set of multilevel models accounting for perceivers' ratings of relationship closeness. As indicated in Table 3, results were comparable to the original analyses when accounting for perceivers' relationship closeness. In addition, perceivers' relationship closeness did not significantly predict empathic accuracy ( $ps > .21$ ).

In sum, we found that physiology-physiology linkage was not associated with empathic accuracy. However, physiology-experience linkage was consistently associated with empathic accuracy, which suggests that perceivers' physiological responses related to empathic accuracy to the extent that they were linked with targets' negative emotional experience. SCL-experience linkage and physiology composite-experience linkage were the most robust predictors of empathic accuracy when accounting for potential confounds.

#### 4. Discussion

The present study examined the role that physiological responses play in empathic accuracy. First, we characterized physiology-physiology linkage, physiology-experience linkage, and empathic accuracy. To do so, we examined average levels and variability of physiology-physiology linkage (i.e., linkage between perceivers' and targets' physiology), physiology-experience linkage (i.e., linkage between perceivers' physiology and targets' emotional experience), and empathic accuracy (i.e., linkage between perceivers' ratings of targets' emotional experience and targets' ratings of their own emotional experience). We observed positive and significant levels of each measure such that on average perceivers' physiology was linked with targets' physiology and emotional experience; additionally, perceivers' ratings of targets' emotional experience were linked with targets' ratings of their own emotional experience. Furthermore, across all measures of linkage, we observed little dyad level variance and high individual level variance in the cross-correlations. In other words, most variability in the measures of linkage was due to differences in terms of which person was acting as the perceiver.

Second, we examined the associations between physiology-physiology linkage and physiology-experience linkage on the one hand and empathic accuracy on the other hand. We found that physiology-physiology linkage was not a significant predictor of empathic accuracy. However, physiology-experience linkage consistently predicted empathic accuracy, which suggests that perceivers whose physiological responses were more linked with the target's emotional experience were more accurate in predicting how the target felt. Skin conductance level-experience linkage and physiology composite-experience linkage were the most robust predictors of empathic accuracy when accounting for linkage between randomly paired partners and when controlling for targets' and perceivers' emotional reactivity, suggesting that these results are robust to influences of shared context.

In contrast to previous work (Levenson & Ruef, 1992), physiology-physiology linkage consistently did not relate to empathic accuracy. Our study design differed from this previous work in critical ways that might explain the seemingly inconsistent results. First, our sample was comprised of friend pairs rather than strangers. We chose this sample because empathic accuracy is particularly important for maintaining close relationships (e.g., Sened et al., 2017). Second, all participants in our sample played the role of both perceiver and target and thus had some knowledge about the task. Third, we did not calculate empathic accuracy for negative and positive emotion separately because the speech task was meant to induce mostly negative emotion. Given the many differences in approach, it is possible that both physiology-physiology linkage and physiology-experience linkage play a role in empathic accuracy; however, our results provide strong evidence that physiology-experience linkage is more relevant for empathic accuracy in this particular context.

#### 4.1. Implications

The results for Aim 1 provide basic insight into the nature of physiology-physiology linkage, physiology-experience linkage, and empathic accuracy in a sample of female friend pairs. Physiology-physiology linkage has a relatively rich literature consistent with the present findings (for a review, see Palumbo et al., 2017), compared to physiology-experience linkage and empathic accuracy. Specifically, other studies have reported significant levels of IBI-IBI linkage and SCL-SCL linkage (e.g., Järvelä, Kivikangas, Kätsyri, & Ravaja, 2014) as well as physiology composite-physiology composite linkage (e.g., Soto & Levenson, 2009). Consistent with these findings, the current results suggest that perceivers' physiology can become linked with targets' physiology while perceivers observe targets.

To our knowledge, this is the first study to examine physiology-experience linkage between two people. We observed significant and positive levels of physiology-experience linkage with large individual-level variation suggesting that perceivers' physiological responses do coordinate with targets' emotional experience and the extent to which this linkage occurs varies by who is the perceiver. These findings suggest that perceivers' physiology can also become linked with targets' emotional experience, beyond targets' physiology, while perceivers observe targets.

Importantly, we also found significant and positive levels of random physiology-physiology linkage (linkage of perceivers' physiology with random targets' physiology) and random physiology-experience linkage (linkage of perceivers' physiology with random targets' emotional experience) for all variables except SCL-experience linkage. These findings suggest that linkage might be driven, in part, by the participants' shared environment. Ultimately, these findings highlight the importance of considering the impact that a shared environment might have on contributing to physiological linkage; in other words, observing physiological linkage does not necessarily reflect a shared experience unique to two people and instead could reflect environmental demands from the task at hand (c.f., Danyluck & Page-Gould, 2019).

Using cross-correlations that can range from  $-1$  to  $1$ , we observed significant and positive levels of empathic accuracy ( $M$  [95% CI] =  $0.34$  [.27, .40]). A few studies (Zaki et al., 2008; Zaki, Bolger et al., 2009; Zaki, Weber et al., 2009) have used the same methodological approach to calculate empathic accuracy without measuring physiological responding and found similar levels of empathic accuracy (.46-.47). These findings suggest that perceivers are relatively accurate when rating targets' emotional experience.

The results for Aim 2 provide insight into the role that physiological responses play in empathic accuracy. Physiology-physiology linkage consistently, across four measures of physiological responding, did not relate to empathic accuracy, while physiology-experience did relate to empathic accuracy. In other words, perceivers' physiological responses relate to empathic accuracy to the extent that they are linked with targets' negative emotional experience. Why might this be? Physiology-experience linkage might be a more proximal component of empathic accuracy such that positive linkage between perceivers' physiology and targets' experience would suggest that perceivers have an accurate physiological signal of the other person's emotional experience and that might in turn translate to empathic accuracy.

Our study design differed from previous work by having all participants fulfill the role of both perceiver and target. Thus, all participants had some knowledge about how they felt during the task and this shared experience might contribute to linkage between participants. In order to account for this, we calculated and accounted for the effect of random physiology-physiology linkage (linkage of perceivers' physiology with random targets' physiology) and random physiology-experience linkage (linkage of perceivers' physiology with random targets' emotional experience). Overall, we observed significant levels of random linkage suggesting that general linkage-like effects existed in our sample and that we needed to account for these effects in our analyses. Although the

levels of random linkage were generally similar to the levels of non-random linkage, the random linkage measures never significantly predicted empathic accuracy. Taken together, these findings suggest that the association between physiology-experience linkage was not simply because participants shared the same context but plausibly because they were linked with their friend's emotional experience.

The results for Aim 2 are consistent with two possible ways that physiological responses might be involved in empathic accuracy. First, perceivers' physiology could serve as an internal signal of targets' emotional experience, giving rise to empathic accuracy (e.g., Decety & Jackson, 2004). Second, accurately knowing what someone is feeling might give rise to a compatible physiological response in the perceiver (e.g., Dezechache et al., 2015). No matter whether physiological responses are involved in one or both of these ways, our results suggest that perceivers' physiology plays an important role in empathic accuracy.

Given the ubiquity and significance of social interactions in our daily lives, it is important to consider the functional impact of our findings. For example, imagine a discussion between two individuals where one individual (i.e., the target) is experiencing several negative emotions. The way in which the other individual (i.e., the perceiver) responds to the target will depend, in part, on the perceiver's ability to accurately discern the target's emotional state. If the perceiver thinks the target is not experiencing negative emotions (i.e., low empathic accuracy), they might provide less helpful responses than if the perceiver thinks the target is experiencing several negative emotions (i.e., high empathic accuracy). Our research suggests that physiological responses play a role in this process and could potentially contribute to more helpful responses and greater overall relationship satisfaction (Levenson & Gottman, 1983; Sened et al., 2017; Verhofstadt et al., 2008).

#### 4.2. Limitations and future directions

The current study provided a strong examination of our two pre-registered aims, but several limitations and directions for future research are noteworthy. First, due to the complex and labor-intensive study design, our sample size did not provide optimal sensitivity to examine how the results might depend on the ethnicity of the perceivers and targets. A study by Soto and Levenson (2009) found that levels of empathic accuracy were equivalent when perceivers viewed targets of the same ethnicity and targets of a different ethnicity; however, levels of physiology-physiology linkage were dependent on target ethnicity such that perceivers evidenced greater physiology-physiology linkage when targets were of the same ethnicity (versus a different ethnicity). The influence of ethnicity might be minimized in close friends that have known each other for a long time; however, future work should examine how ethnicity match might influence the associations between physiology-physiology linkage, physiology-experience linkage, and empathic accuracy.

Second, the current sample consisted of only female participants, which increased our statistical power by reducing sample heterogeneity, but also limits the generalizability of the current results. Existing work suggests that our results would hold for men given the lack of gender differences in objective empathy-related processes (e.g., Baez et al., 2017; Eisenberg & Lennon, 1983); however, future work might replicate the current findings in male friend pairs and also mixed-sex friend pairs (Ickes et al., 1990).

Third, the standardized laboratory tasks had several strengths, but they also had some limitations. The speech task was limited in that it primarily evoked negative emotion. Previous research found that physiology-physiology linkage predicted empathic accuracy for negative emotions but not positive emotions in a mixed-emotion task (Levenson & Ruef, 1992); however, they did not examine physiology-experience linkage. Future research might examine the associations between physiology-experience linkage and empathic accuracy in a predominantly positive emotion context to examine the

generalizability of these associations.

The empathy task was limited in that it did not include an actual interaction between the friend pairs. Thus, we cannot say whether our findings would generalize to in-person interactions. The associations between physiology-experience and empathic accuracy could potentially be stronger during in-person interactions because perceivers would ostensibly have access to more of the targets' emotional cues (not just cues captured on the video frame). On the other hand, the associations could be weaker or null because the perceiver would not be passively perceiving the target's cues and competing demands (e.g., regulating their response to the target) could hinder linkage.

Fourth, the rating dial measure captured only emotional valence (i.e., positive-negative ratings) and not emotional arousal. We expect that the results would replicate for emotional arousal because the empathy task would still ultimately involve accurately reading an aspect of a person's emotional experience. Thus, the extent to which perceivers' physiology links up with targets' emotional arousal should be associated with greater empathic accuracy for targets' emotional arousal. Future work might try to capture ratings of both emotional valence and arousal to examine whether the effects do replicate (c.f., Rattel, Mauss, Liedlgruber, & Wilhelm, 2020).

Finally, we examined several physiological channels (inter-beat interval, finger pulse amplitude, and skin conductance level) as well as a composite of seven physiological channels (inter-beat interval, finger pulse amplitude, skin conductance level, finger pulse transit time, ear pulse amplitude, ear pulse transit time, and skin temperature). This approach allowed us to broadly capture physiological responses but most of the measures of linkage (with the exception of inter-beat interval) index sympathetic activation. Linkage in parasympathetic activation (e.g., increased respiratory sinus arrhythmia) might also play an important role in empathic accuracy, with recent work suggesting that both sympathetic activation and parasympathetic activation are involved in empathic responding (Stellar, Anderson, & Gatchpazian, 2020). Future work might consider examining both types of physiological activation.

#### 4.3. Conclusion

The current results provide evidence for the notion that physiological responses play a role in empathic accuracy. Specifically, these results are consistent with the ideas that one's physiological responses might serve as a signal of others' emotional experiences and that when one reads others' emotions accurately one's physiological responses become aligned with their emotions.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgements

This research was supported by National Institutes of Health grant R01AG043592 and National Science Foundation grant BCS-1941868 awarded to I.B.M.

#### References

- Averill, J. R. (1980). A constructivist view of emotion. In R. Plutchik, & H. Kellerman (Eds.), *Emotion: Theory, research, and experience*, 1 pp. 305–339. New York: Academic Press.
- Baez, S., Flichtentrei, D., Prats, M., Mastandueno, R., García, A. M., Cetkovich, M., et al. (2017). Men, women... who cares? A population-based study on sex differences and gender roles in empathy and moral cognition. *PLoS One*, 12(6), Article e0179336.
- Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and empathy: Two qualitatively distinct vicarious emotions with different motivational consequences. *Journal of Personality*, 55, 19–39. <https://doi.org/10.1111/j.1467-6494.1987.tb00426.x>

- Butler, E. A., & Randall, A. K. (2013). Emotional coregulation in close relationships. *Emotion Review*, 5(2), 202–210. <https://doi.org/10.1177/1754073912451630>
- Butler, E. A., Gross, J. J., & Barnard, K. (2014). Testing the effects of suppression and reappraisal on emotional concordance using a multivariate multilevel model. *Biological Psychology*, 98, 6–18. <https://doi.org/10.1016/j.biopsycho.2013.09.003>
- Dan-Glauser, E. S., & Gross, J. J. (2013). Emotion regulation and emotion coherence: Evidence for strategy-specific effects. *Emotion*, 13(5), 832–842. <https://doi.org/10.1037/a0032672>
- Danyluck, C., & Page-Gould, E. (2018). Intergroup dissimilarity predicts physiological synchrony and affiliation in intergroup interaction. *Journal of Experimental Social Psychology*, 74, 111–120. <https://doi.org/10.1016/j.jesp.2017.08.001>
- Danyluck, C., & Page-Gould, E. (2019). Social and physiological context can affect the meaning of physiological synchrony. *Scientific Reports*, 9(1), 1–13. <https://doi.org/10.1038/s41598-019-44667-5>
- Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, 3(2), 71–100. <https://doi.org/10.1177/1534582304267187>
- Dezecache, G., Jacob, P., & Grèzes, J. (2015). Emotional contagion: Its scope and limits. *Trends in Cognitive Sciences*, 19(6), 297–299. <https://doi.org/10.1016/j.tics.2015.03.011>
- Dolan, R. J. (2002). Emotion, cognition, and behavior. *Science*, 298(5596), 1191–1194. <https://doi.org/10.1126/science.1076358>
- Eckland, N. S., & English, T. (2019). Trait-level emotion regulation and emotional awareness predictors of empathic accuracy. *Motivation and Emotion*, 43(3), 461–470. <https://doi.org/10.1007/s11031-018-9741-z>
- Eckland, N. S., Huang, A. B., & Berenbaum, H. (2020). Empathic accuracy: Associations with prosocial behavior and self-insecurity. *Emotion*, 20(7), 1306–1310.
- Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin*, 94(1), 100–131.
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6, 169–200. <https://doi.org/10.1080/02699939208411068>
- Elgendi, M. (2012). On the analysis of fingertip photoplethysmogram signals. *Current Cardiology Reviews*, 8(1), 14–25. <https://doi.org/10.2174/157340312801215782>
- Frijda, N. H., Ortony, A., Sonnemans, J., & Clore, G. L. (1992). The complexity of intensity: Issues concerning the structure of emotion intensity. In M. S. Clark (Ed.), *Emotion* (pp. 60–89). Newbury Park: Sage Publications.
- Gottman, J. M., & Levenson, R. W. (1985). A valid procedure for obtaining self-report of affect in marital interaction. *Journal of Consulting and Clinical Psychology*, 53(2), 151–160. <https://doi.org/10.1037/0022-006X.53.2.151>
- Gratton, G. (2000). Biosignal processing. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 900–923). Cambridge, England: Cambridge University Press.
- Hein, G., Lamm, C., Brodbeck, C., & Singer, T. (2011). Skin conductance response to the pain of others predicts later costly helping. *PLoS One*, 6(8), Article e22759.
- Holt-Lunstad, J., Robles, T. F., & Sbarra, D. A. (2017). Advancing social connection as a public health priority in the United States. *American Psychologist*, 72(6), 517–530. <https://doi.org/10.1037/amp0000103>
- Ickes, W. (1993). Empathic accuracy. *Journal of Personality*, 61(4), 587–610. <https://doi.org/10.1111/j.1467-6494.1993.tb00783.x>
- Ickes, W., Buysse, A. N. N., Pham, H. A. O., Rivers, K., Erickson, J. R., Hancock, M., et al. (2000). On the difficulty of distinguishing “good” and “poor” perceivers: A social relations analysis of empathic accuracy data. *Personal Relationships*, 7(2), 219–234.
- Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology*, 59(4), 730–742.
- Izard, C. E. (1977). *Human emotions*. New York: Plenum Press.
- Järvelä, S., Kivikangas, J. M., Kätsyri, J., & Ravaja, N. (2014). Physiological linkage of dyadic gaming experience. *Simulation & Gaming*, 45(1), 24–40. <https://doi.org/10.1177/1046878113513080>
- Kettunen, J., Ravaja, N., Näätänen, P., Keskivaara, P., & Keltikangas-Järvinen, L. (1998). The synchronization of electrodermal activity and heart rate and its relationship to energetic arousal: A time series approach. *Biological Psychology*, 48(3), 209–225. [https://doi.org/10.1016/s0301-0511\(98\)00017-9](https://doi.org/10.1016/s0301-0511(98)00017-9)
- 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. *Neuropsychobiology*, 28(1–2), 76–81. <https://doi.org/10.1159/000119004>
- Krebs, D. (1975). Empathy and altruism. *Journal of Personality and Social Psychology*, 32, 1134–1146.
- Lang, P. J. (1988). What are the data of emotion? In V. Hamilton, G. H. Bower, & N. H. Frijda (Eds.), *Cognitive perspectives on emotion and motivation* (pp. 173–191). New York: Kluwer Academic/Plenum Press.
- Lazarus, R. S. (1991). *Emotion and adaptation*. London: Oxford University Press.
- Levenson, R. W. (1994). Human emotions: A functional view. In P. Ekman, & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 123–126). New York: Oxford University Press.
- Levenson, R. W. (1988). Emotion and the autonomic nervous system: A prospectus for research on autonomic specificity. In H. L. Wagner (Ed.), *Social psychophysiology and emotion: Theory and clinical implications* (pp. 17–42). New York: Wiley.
- Levenson, R. W., & Gottman, J. M. (1983). Marital interaction: Physiological linkage and affective exchange. *Journal of Personality and Social Psychology*, 45(3), 587–597.
- Levenson, R. W., & Ruef, A. M. (1992). Empathy: A physiological substrate. *Journal of Personality and Social Psychology*, 63(2), 234–246.
- Levenson, R. W., & Ruef, A. M. (1997). Physiological aspects of emotional knowledge and rapport. In W. J. Ickes (Ed.), *Empathic accuracy* (pp. 44–72). The Guilford Press.
- Machado-Moreira, C. A., McLennan, P. L., Lillioja, S., van Dijk, W., Caldwell, J. N., & Taylor, N. A. S. (2012). The cholinergic blockade of both thermally and non-thermally induced human eccrine sweating. *Experimental Physiology*, 97, 930–942. <https://doi.org/10.1113/expphysiol.2012.065037>
- Mauss, I. B., Wilhelm, F. H., & Gross, J. J. (2003). Autonomic recovery and habituation in social anxiety. *Psychophysiology*, 40, 648–653.
- Mauss, I. B., Wilhelm, F. H., & Gross, J. J. (2004). Is there less to social anxiety than meets the eye? Emotion experience, expression, and bodily responding. *Cognition and Emotion*, 18, 631–662.
- Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion*, 5(2), 175–190.
- Mauss, I. B., Shallcross, A. J., Troy, A. S., John, O. P., Wilhelm, F. H., & Gross, J. J. (2011). Don't hide your happiness! Positive emotion dissociation, social connectedness, and psychological functioning. *Journal of Personality and Social Psychology*, 100(4), 738–748. <https://doi.org/10.1037/a0022410>
- Mendes, W. B. (2009). Assessing autonomic nervous system activity. In E. Harmon-Jones, & J. S. Beer (Eds.), *Methods in social neuroscience* (pp. 118–147). Guilford Press.
- Nesse, R. M., Curtis, G. C., Thyer, B. A., McCann, D. S., Huber-Smith, M. J., & Knopf, R. F. (1985). Endocrine and cardiovascular responses during phobic anxiety. *Psychosomatic Medicine*, 47(4), 320–332. <https://doi.org/10.1097/00006842-198507000-00002>
- Palumbo, R. V., Marraccini, M. E., Weyandt, L. L., Wilder-Smith, O., McGee, H. A., Liu, S., et al. (2017). Interpersonal autonomic physiology: A systematic review of the literature. *Personality and Social Psychology Review*, 21(2), 99–141. <https://doi.org/10.1177/1088868316628405>
- Panksepp, J. (1994). The basics of basic emotion. In P. Ekman, & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 237–242). New York: Oxford University Press.
- Rattel, J. A., Mauss, I. B., Liedlgruber, M., & Wilhelm, F. H. (2020). Sex differences in emotional concordance. *Biological Psychology*, 151, 1–12. <https://doi.org/10.1016/j.biopsycho.2020.107845>
- Reisenzein, R. (2000). Exploring the strength of association between the components of emotion syndromes: The case of surprise. *Cognition and Emotion*, 14(1), 1–38. <https://doi.org/10.1080/026999300378978>
- Ruef, A. M., & Levenson, R. W. (2007). Continuous measurement of emotion: The affect rating dial. In J. A. Coan, & J. J. B. Allen (Eds.), *Series in affective science. Handbook of emotion elicitation and assessment* (pp. 286–297). Oxford University Press.
- Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In K. R. Scherer, & P. Ekman (Eds.), *Approaches to emotion* (pp. 293–317). Hillsdale, NJ: Erlbaum.
- Schulz, M. S., & Waldinger, R. J. (2004). Looking in the mirror: Participants as observers of their own and their partners' emotions in marital interactions. In P. Kerig, & D. Baucom (Eds.), *Couple observational coding systems* (pp. 259–272). Hillsdale, NJ: Erlbaum.
- Sened, H., Lavidor, M., Lazarus, G., Bar-Kalifa, E., Rafaeli, E., & Ickes, W. (2017). Empathic accuracy and relationship satisfaction: A meta-analytic review. *Journal of Family Psychology*, 31(6), 742–752.
- Singer, T., & Lamm, C. (2009). The social neuroscience of empathy. *Annals of the New York Academy of Sciences*, 1156(1), 81–96. <https://doi.org/10.1111/j.1749-6632.2009.04418.x>
- Soto, J. A., & Levenson, R. W. (2009). Emotion recognition across cultures: The influence of ethnicity on empathic accuracy and physiological linkage. *Emotion*, 9(6), 874–884.
- Stellar, J. E., Anderson, C. L., & Gatchpazian, A. (2020). Profiles in empathy: Different empathic responses to emotional and physical suffering. *Journal of Experimental Psychology: General*, 149(7), 1398–1416.
- Sze, J., Gyurak, A., Yuan, J. W., & Levenson, R. W. (2010). Coherence between emotional experience and physiology: Does body awareness training have an impact? *Emotion*, 10(6), 803–814.
- Thomas, G., Fletcher, G. J. O., & Lange, C. (1997). On-line empathic accuracy in marital interaction. *Journal of Personality and Social Psychology*, 72(4), 839–850.
- Thorson, K. R. (2018). *Understanding the relationship between empathic accuracy and physiological synchrony: A theoretical and analytic guide with evidence from three contexts*. New York University. Available from ProQuest Dissertations Publishing (UMI No. 10637769).
- 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.
- Vaughan, K. B., & Lanzetta, J. T. (1981). The effect of modification of expressive displays on vicarious emotional arousal. *Journal of Experimental Social Psychology*, 17, 16–30.
- Verhofstad, L. L., Buysse, A., Ickes, W., Davis, M., & Devoldre, I. (2008). Support provision in marriage: The role of emotional similarity and empathic accuracy. *Emotion*, 8(6), 792–802. <https://doi.org/10.1037/a0013976>
- Wilhelm, F. H., Grossman, P., & Roth, W. T. (1999). Analysis of cardiovascular regulation. *Biomedical Sciences Instrumentation*, 35, 135–140.
- Wilhelm, F. H., & Roth, W. T. (2001). The somatic symptom paradox in DSM-IV anxiety disorders: Suggestions for a clinical focus in psychophysiology. *Biological Psychology*, 57(1–3), 105–140. [https://doi.org/10.1016/s0301-0511\(01\)00091-6](https://doi.org/10.1016/s0301-0511(01)00091-6)
- Witherington, D. C., Campos, J. J., & Hertenstein, M. J. (2001). Principles of emotion and its development in infancy. In G. Bremner, & A. Fogel (Eds.), *Blackwell handbook of infant development* (pp. 427–464). Malden, MA: Blackwell Publishers.

- Zaki, J., Bolger, N., & Ochsner, K. (2008). It takes two: The interpersonal nature of empathic accuracy. *Psychological Science*, 19(4), 399–404. <https://doi.org/10.1111/j.1467-9280.2008.02099.x>
- Zaki, J., Bolger, N., & Ochsner, K. (2009). Unpacking the informational bases of empathic accuracy. *Emotion*, 9(4), 478–487.

- Zaki, J., Weber, J., Bolger, N., & Ochsner, K. (2009). The neural bases of empathic accuracy. *Proceedings of the National Academy of Sciences*, 106(27), 11382–11387. <https://doi.org/10.1073/pnas.0902666106>