Physiological Synchrony Predicts Observational Threat Learning in Humans

Philip Pärnamets*[1,2], Lisa Espinosa[1], and Andreas Olsson[1]

[1]Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, 171 77, Stockholm, Sweden
[2]Department of Psychology, New York University, 6 Washington Place New York, NY 10003, USA

Abstract

Understanding how information about threats in the environment is shared and transmitted between individuals is crucial for explaining adaptive, survival-related behavior in humans and other animals, and for developing treatments for phobias and other anxiety disorders. Research across species has shown that observing a conspecific's, a "demonstrator's", threat responses causes strong and persistent threat memories in the "observer". Such observational threat learning is hypothesized to be supported by empathic processes. Here, we examined if physiological synchrony between demonstrator and observer can serve as a biomarker of the hypothesized empathic transfer. We measured synchrony between demonstrators' and observers' phasic electrodermal signals during learning, which directly reflects autonomic nervous system activity. Synchrony predicted the strength of the observer's later skin conductance responses to threat predicting stimuli. Dynamic coupling between an observer's and a demonstrator's autonomic nervous system activity may reflect empathic experience sharing processes facilitating the formation of observational threat associations.

In social species, like humans, knowledge about threats and dangers is often acquired through various forms of social transmission, for example through observation. Research across species has shown that observing a conspecific's - a "Demonstrator's" - threat responses to a previously neutral stimulus can cause strong and persistent threat memories in the "Observer" [1–12]. Such memories are expressed by heightened autonomic nervous system activity in the observer when later facing that stimulus alone. Observational threat learning is efficient, and minimizes risks to the individual

*philip.parnamets@ki.se

arising from directly interacting with potential dangers [13]. Understanding how threat information in the environment is observationally acquired is central to explaining adaptive, as well as maladaptive, survival related behavior in humans and other animals.

For observational learning to occur, the demonstrator's reactions to the threatening stimulus must function as unconditioned stimulus for the observer. It has been hypothesized that empathic processes, enabling the sharing or understanding of affective states between demonstrator and observer, facilitate observational learning [4, 10, 14–17]. Previous work has implicated trait measures of cognitive empathy in observational learning [10, 18], emphasizing the role individual differences in the observer's capacity to take the perspective of the demonstrator plays in learning. This overlooks another component of empathy known as experience sharing, which is characterized by an observer having similar emotional experiences as a target, even if the observer is not in the same situation as the target (e.g., when an audience member feels vicarious anxiety about their friend's performance on stage) [19–21]. It is possible that direct sharing of affective states between demonstrator and observer promotes learning and that momentary fluctuations in shared experience matter and state empathy more than empathic traits. However, there is no direct evidence for the role of experience sharing in human threat learning. Here we attempted to remedy this and provide evidence that synchronous patterns of arousal influences observational threat learning.

Synchrony is a pervasive natural phenomenon and occurs when two systems become coupled so that their trajectories develop temporal interdependence [22, 23]. It is also a fundamental feature of interpersonal coordination and social cognition [21, 24–30]. For example, in humans, synchrony has been observed over multiple levels of analysis including intrapersonal limb coordination[23], interpersonal eye-movements during communication [31], shared attention [32], and postural sway in dyadic coordination tasks [33]. Importantly, synchrony has been related to interpersonal cohesion and cooperative outcomes [26, 34, 35]. Related individuals will show greater synchrony in their heart rates compared to non-related individuals during hazardous social rituals [34] and dyads who show higher degrees of synchrony will show greater team cohesion [35]. Recent advances in the cognitive neurosciences have led to the discovery of coupling in BOLD fMRI signals in several contexts [21, 36–38]. For example, neural coupling between individuals has been found to predict successful verbal communication, such that the more a listener's brain activity correlated with the speaker's, the better comprehension the listener would report [37]. Despite considerable evidence for the role of synchrony in diverse intra- and interpersonal processes, its role, if any, in social learning is not understood.

Past experiments have all employed artificial situations involving confederate demonstrators, either in live settings [4–6, 16, 17], or more recently,

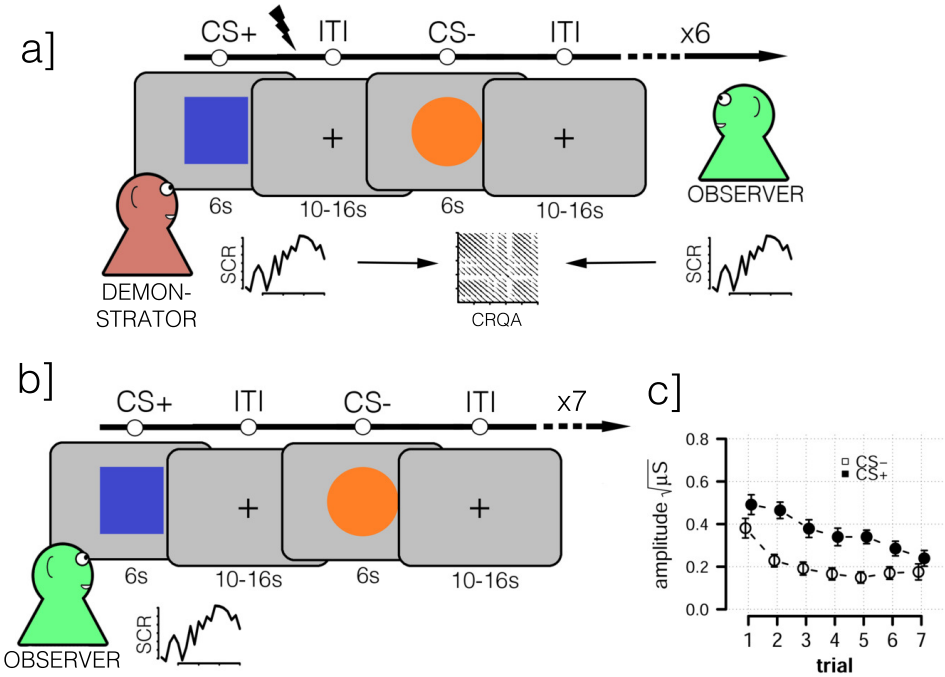

Figure 1: Overview of experiment. **a]** Learning phase. Demonstrator watched two images, one which terminated with an uncomfortable electrical shock during 4 of 6 presentations (CS+) and another that never terminated with a shock (CS-). Each CS presentation lasted 6s and inter-trial interval (ITI) varied between 10 and 16s. Valence (CS+/-) of the first image presented was randomly varied. While the Demonstrator watched the CS presentations and received shocks, the Observer watched both the Demonstrator and the CS's. The Observer was instructed to learn the shock contingency. The Observer received no shocks during the learning phase. Electrodermal activity was continuously recorded from both the Demonstrator and Observer. **b]** Testing phase. Immediately following the learning phase, the Observer was repeatedly shown both CS's again, instructed that they would receive shocks to the same image as the Demonstrator had received shocks to. Importantly, only the 7th, final presentation of the CS+ terminated with shock, to not interfere with the measurement of the vicariously acquired threat response. Greater SCR to the CS+ compared to the CS- in the Observer, in this phase, indicates successful threat learning. **c]** Average trial-by-trial data from testing phase showing average skin conductance responses to the CS+ (dark circles) and to the CS- (light circles) for the Observer. Error bars are 95% confidence intervals.

displayed via video recording [8, 10, 12, 39], and this may be on One reason for the lack of evidence concerning synchrony and observational threat learning. Hence, to investigate synchrony we adapted an existing, standard video-based paradigm [39] to a more naturalistic situation. In our paradigm two naïve participants took turns being demonstrator and observer, the demonstrator undergoing a direct conditioning procedure and the observer learning from the demonstrator's reactions (see Fig. 1, Methods). The experiment consisted of four blocks each consisting of two phases, and participants switched roles between the second and third block (see Methods). The first phase was a learning phase where the observer watched the demonstrator receive probabilistic shocks to one of two visual images serving as conditioned stimulus images (CS+), and never to the other (CS-). The learning phase was followed by a testing phase where both CSs were repeatedly presented again to the observer under threat of shock. To measure social learning no shocks were administered to the observer following CS+ presentations during this phase (except for the last presentation, see Methods). Both participants' electrodermal activity was continuously recorded throughout the experiment allowing for synchrony to be calculated during the learning phase. Threat learning was measured as CS differentiation - stronger skin conductance responses to CS+ compared to CS- images in the testing phase.

Results

We first evaluated the threat learning procedure and confirmed that observers readily acquired threat information from demonstrators. We tested 69 same-sex dyads and analyzed observers' skin conductance responses to the onset of the CS images. We included a factor coding for if a trial consisted of a CS+ or a CS- presentation (CS). Additionally, since our experiment consisted of four blocks and participants were observers either in the first two or final two blocks of the experiment, we included variables coding for these effects in the same regression model (Block and Role respectively). This allowed us to test if learning was equally effective for each observer regardless if this was their first or second time observing the demonstrator and if there was any difference in learning depending on the role reversal. All variables were deviation coded (see Methods).

We found a robust learned threat response indicating CS differentiation (Bayesian multilevel regression; b = 0.15, SE = 0.016, CrI (95% Credible Interval) = [0.12, 0.19], BF_{10} (Bayes Factor) $> 10^6$, see Fig. 1). Importantly, we found no interaction between CS status and Block (b = 0.03, SE = 0.019, CrI = [-0.007, 0.065], BF_{10} = 1.43), meaning that observers, on average, learned equally well in their first and second blocks as observers. Likewise, we found no interaction between CS status and Role (b = 0.036,

SE = 0.028, CrI = [-0.020, 0.092], BF_{10} = 1.31), meaning that prior experience as a demonstrator did not affect observers ability to learn socially through observation. Together, this indicates that learning was equally effective during all stages of the experiment and we conclude from this that our procedure translates standard video-based observational learning paradigms [8, 39] into the more realistic situation, involving two live participants, tested here.

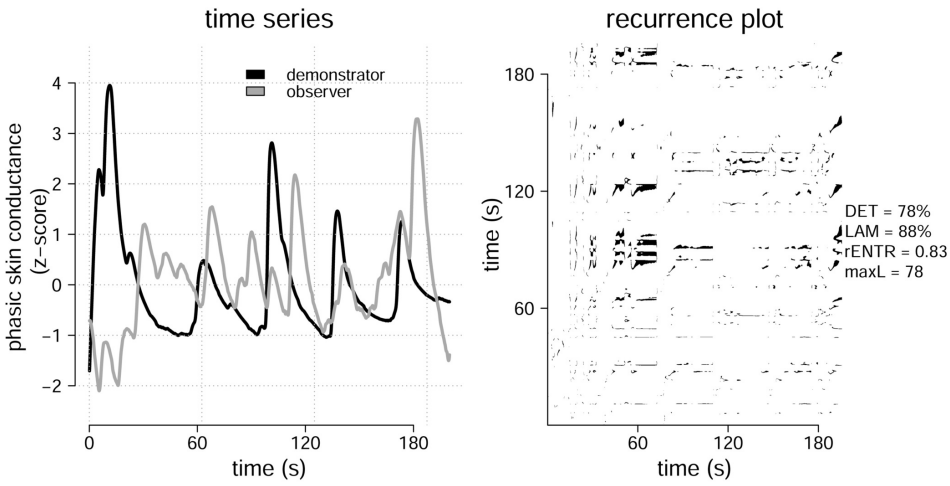

Figure 2: Left panel, example of one observer's and demonstrator's phasic skin conductance time series from a single learning phase, z-scored to facilitate comparison. No testing phase data shown. Right panel, the resulting recurrence plot following cross-recurrence quantification analysis (CRQA) on the time series in the left panel. From each recurrence plot, four standard metrics capturing synchrony were computed; ratio of points in diagonal lines to all points (determinism, DET), maximal diagonal line length (maxL), entropy of diagonal line length distribution (rENTR), and the ratio of points in vertical lines to all points (laminarity, LAM). Resulting metrics computed from the left panel example series displayed. Both panels are displayed for illustrative purposes depicting only one representative time-series and its resulting cross-recurrence plot.

Synchrony predicts observational threat responses

Next, we tested the main hypothesis: that demonstrator-observer synchrony of physiological arousal during the learning phase would predict the strength of the observer's threat responses in the test phase. To quantify synchrony we performed a cross-recurrence quantification analysis (CRQA; see Fig. 2, Methods). CRQA quantifies the similarity of two signals and is suitable for complex non-stationary signals where non-linear dyanamics may exist

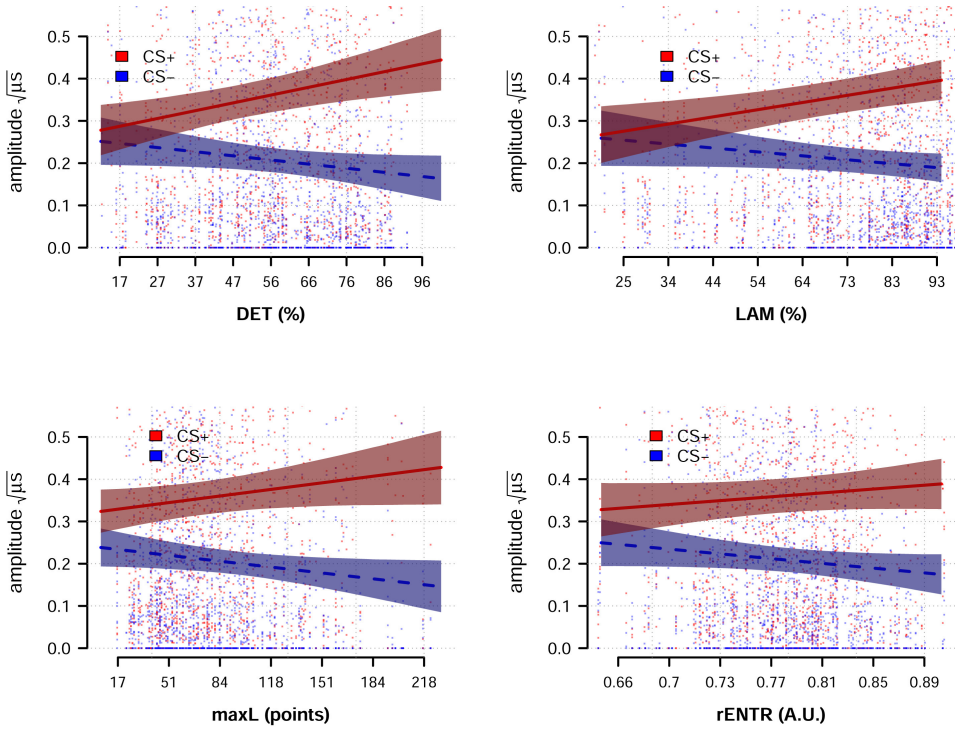

Figure 3: Posterior predictions from regression models predicting CS+ (red, solid line) and CS- (blue, dashed line) responses as a function of each of the four metrics of synchrony computed via CRQA. Points represent individual data points from separate CS presentations. Shaded regions indicates 95% posterior predictive interval.

[27, 34, 40, 41]. We first constructed cross-recurrence plots for each dyad's phasic skin conductance time series from the learning phase portion of each block (see Fig. 2 for example cross-recurrence plot). From these plots four standard metrics of CRQA were computed that capture predictability (% DETerminism), maximum strength of coupling (maxLine) and complexity (rENTRopy) of the relationship between the time series.

We regressed each of four CRQA metrics separately on each observer's skin conductance responses to each CS image presentation from the corresponding testing phase, together with a variable indicating CS status. Interactions between the CRQA metrics and the CS variable indicate support for our hypothesis.

Our hypothesis was supported across all four metrics considered. We found strongest evidence for a link between DET and CS differentiation (b = 0.056, SE = 0.014, CrI = [0.029, 0.083], BF_{10} = 780) and between LAM and CS differentiation (b = 0.052, SE = 0.014, CrI = [0.026, 0.080], BF_{10} = 181); see Fig. 3). We also found strong evidence for maxL (b = 0.027, SE

= 0.014, CrI = [0.013, 0.067], BF_{10} = 15.4) predicting CS differentiation but only weak evidence for rENTR (b = 0.032, SE = 0.012, CrI = [0.004, 0.050], BF_{10} = 3.1).

While the preceding analyses indicated strong support for our hypothesis that interpersonal synchrony during observational learning predicts later threat responses, it does not show that this effect is specific to the actual demonstrator-observer dyads from our experiment. To address this issue we followed existing recommendations [31, 35, 40], and created random permutations of our data by pairing participants' across dyad boundaries. These pairings can be thought of as pseudo-dyads. We completed 500 such permutations and computed the four CRQA metrics for all the resulting data sets. We performed a regression analysis analogous to the one presented above on each dataset and extracted the model evidence for the critical interaction between the CRQA metric from the pseudo-dyads and the CS indicator variable. The results from this analysis can be seen in Fig. 4. The results demonstrate that the predictive effect of the CRQA metrics on threat responses does not arise between pseudo-dyads and is specific to actual demonstrator-observer pairings.

Single component of CRQA metrics captures effect of synchrony on threat responses

Since the four CRQA metrics were highly correlated in our sample (r = .33 to r = .74) we followed prior work and reduced these measures into a single value using principal components analysis [35]. The factor loadings resulting from the analysis and variance explained of each factor are displayed in Supplementary Table 1. The loadings suggested that the first component, which loaded roughly equally, and positively, on all metrics and captured 63% of the variance, should represent synchrony between participants best.

In line with this interpretation and our earlier analysis on the separate metrics, the first component we extracted (PC1), capturing synchrony, positively predicted CS differentiation (b = 0.039, SE = 0.009, CrI = [0.021, 0.058], $BF_{10} > 10^6$, see Fig. 5). The remaining principal components did not reliably predict CS differentiation (all b's < 0.012, all $BF_{10} < 0.58$). In the remainder of the paper we restrict our analyses to this synchrony component.

Stability of synchrony across trials, blocks and participant roles

Analogously to our evaluation of overall threat learning, we tested if the predictive effect of synchrony on the observer's threat responses differed depending on if the observer started the experiment in that Role or, instead, as demonstrator and if it was the first or second Block in that role. It is possible that additional experience with the experiment might in some way

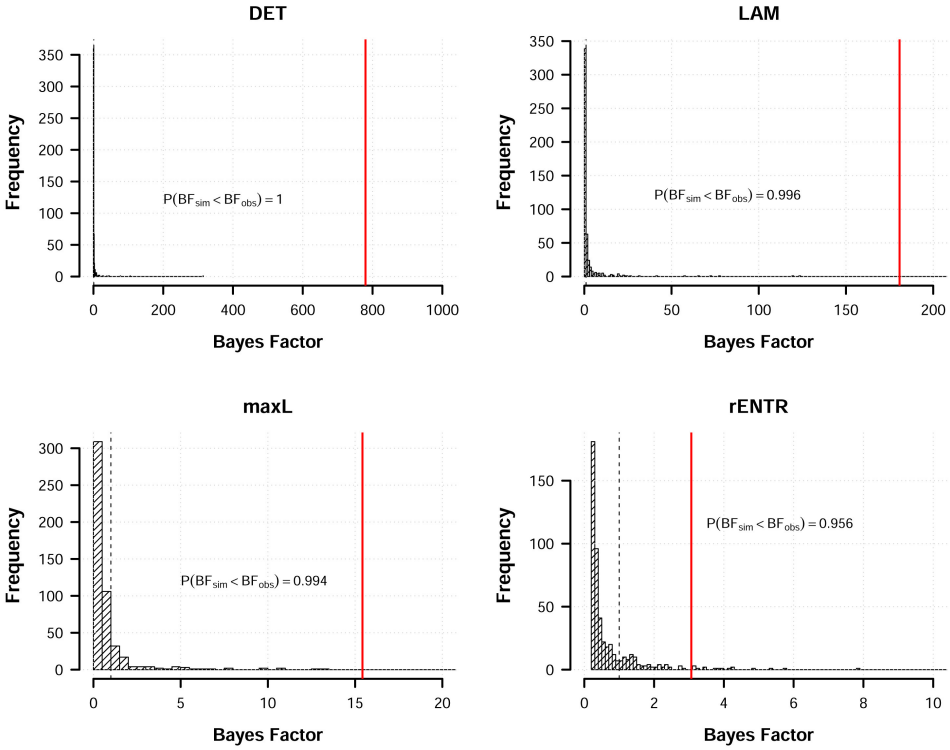

Figure 4: Histograms of model evidence (Bayes Factors) in favor of an effects of each of the four CRQA metrics on CS differentiation, computed from creating pseudo-dyads through permutation of the real participants pairings in our dataset. The vertical red line indicates the observed Bayes Factor and the dashed vertical line shows BF = 1.

affect how observers and demonstrators synchronize and that our effects were only present in parts of the data. Our analyses indicated that this was not the case. We found no evidence for moderation in the effect of synchrony on CS differentiation dependent on initial Role (b = -0.007, SE = 0.0017, CrI = [-0.040, 0.026], BF_{10} = 0.38) or by Block (b = -0.005, SE = 0.015, CrI = [-0.033, 0.024], BF_{10} = 0.30). In other words, the relationship between synchrony and strength of conditioned responses was not affected by participants switching roles or by their experience with the observational learning procedure.

We next tested if the effect of synchrony (PC1) changed over the series of consecutive CS image presentations as the participants' responses extinguished (cf. Fig. 1) . We added a Trial variable to a model with the CS status and the synchrony component predictors, together with all their interactions. Unsurprisingly there was both a main effect of trial (b = -0.034, SE = 0.003, CrI = [-0.040, -0.029], $BF_{10} > 10^6$) and a weak interaction with CS status possibly indicating marginally faster extinction of responses

to the CS+ (b = -0.013, SE = 0.005, CrI = [-0.023, -0.003], BF_{10} = 2.51). However, of critical interest, the relationship between synchrony and CS differentiation did not interact with trial, with strong evidence for null effect (b = 0.002, SE = 0.003, CrI = [-0.004, 0.008], BF_{10} = 0.068). The relationship between CS differentiation and synchrony was stable across all trials in the testing phase.

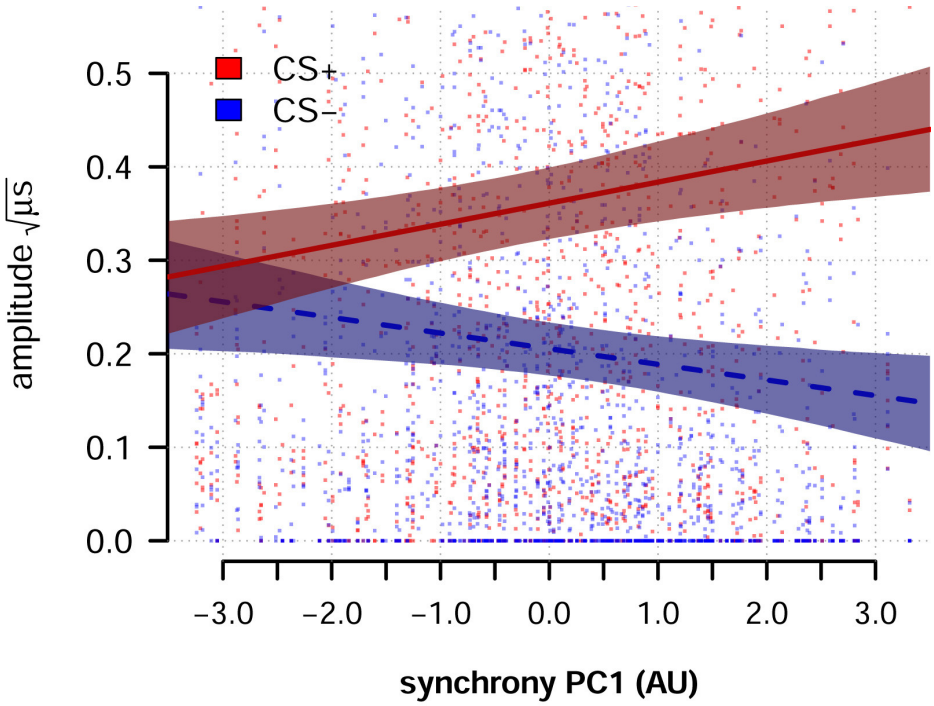

Figure 5: Posterior prediction of the observer's CS+ (red, solid line) and CS- (blue, dashed line) responses during the testing phase, as a function of synchrony during the learning phase as captured by the first principal component of the CRQA measures. Points represent individual data points from separate CS presentations. Shaded region indicates 95% posterior predictive interval.

Specificity of synchrony as predictor of threat learning

To rule out non-synchrony based mechanisms explaining our findings we identified three measures based on the observer's arousal in the learning phase. Each measure could plausibly capture relevant aspects of the observer's learning process. The first measure was the average strength of the observer's skin conductance response to the social UCS (the demonstrator receiving shocks) during the learning phase. During direct conditioning it is generally accepted that the strength of the UCS predicts the strength of

later conditioned responses [42–45], and the social UCS is considered to play a similar role in observational learning [5, 6]. We therefore hypothesized that the observer's reactions to the social UCS might be indicative of its perceived strength. As such the UCS response might be capturing similar empathic processes as the synchrony component. The second measure was the average difference of the observer's responses to the CS+ over the CS- during the learning phase. While the observers were under no direct threat, it is possible that some observers began to develop responses to the CS+ anticipating the future shocks to the demonstrators. It is therefore possible that this early CS differentiation during the learning phase could be the source of the observed synchrony in that phase and account for the effect of synchrony we observed. Third, we included a time-lagged correlation between the demonstrators' and observers' skin conductance time-series during the learning phase. Correlations capture a direct linear relationship between the signals and, unlike the previous two measures, take the whole time-series into account.

The three measures outlined above exhibited low correlations with each other (r = .14 to r = .20), so we jointly regressed each of them and the synchrony component identified earlier together with their interactions with CS status on observer's skin conductance responses from the testing phase. We found that while observers' average UCS responses, during the learning phase, were positively related to the strength of their average skin conductance responses, during the testing phase, (b = 0.045, SE = 0.012, CrI = [0.022, 0.069], BF_{10} = 220), they did not interact with CS status (b = 0.020, SE = 0.017, CrI = [-0.013, 0.053], BF_{10} = 0.64). Similarly, neither learning phase CS differentiation (b = -0.020, SE = 0.045, CrI = [-0.105, 0.069], BF_{10} = 0.99) nor lagged correlations (b = 0.005, SE = 0.013, CrI = [-0.020, 0.031], BF_{10} = 0.28) interacted with CS status. Importantly, the synchrony component continued to robustly predict CS differentiation in the testing phase in this model (b = 0.035, SE = 0.009, CrI = [0.017, 0.053], BF_{10} = 1484), even when accounting for these additional measures capturing other aspects of the observer's skin conductance signals during the learning phase. Finally, we found no evidence that any of the three measures moderated the effect of synchrony on CS differentiation (all b's < 0.01, all $BF_{10} < 0.64$).

Together these analyses show that synchrony is a specific predictor of observational threat learning and that the findings are robust to several plausible alternative predictors also derived from electrodermal activity during the learning phase.

Self-reported empathy does not account for threat learning

Next we considered if individual differences in self-reported trait empathy (see Supplementary Table 2 for descriptive statistics), as measured by the four sub-scales of the Interpersonal Reactivity Index [46], predicted obser-

vational threat learning and if this could explain the effects of synchrony. We regressed the four sub-scales together with CS status and the synchrony component. We found no interactions between any of the sub-scales and CS status nor with the synchrony component (all b's < 0.01, all $BF_{10} < 0.43$).

Similarly, all observers rated their perceptions of the demonstrator: how much pain the demonstrator appeared to be in, their quality as a learning model, how much compassion the observer felt for the demonstrator and how similar to the observer the demonstrator appeared to be. Again none of these measures interacted with CS status or with the synchrony component (all b's < 0.028, all $BF_{10} < 1.18$).

These results suggest the momentary physiological coupling between observers and demonstrators occurs beyond participants' introspective abilities and that synchrony might constitute a more fundamental feature of empathic learning than captured by trait scales.

Discussion

We investigated if spontaneous synchrony between an observer's and a demonstrator's arousal states, as indexed by their electrodermal activity, during observational threat learning predicted the strength of the observer's conditioned responses in a later testing phase. We found that the first principal component assessed from four common metrics of synchrony, calculated using cross-recurrence quantification analysis (CRQA), robustly predicted conditioned responses. Indeed, our findings indicated that at low levels of observer-demonstrator synchrony almost no differentiation in responses to threat-relevant versus threat-irrelevant stimuli was expected (Fig. 5). Together, this suggests a critical and previously undocumented facilitating role of synchrony in observational threat learning. We discuss the interpretation and implications of our findings below.

Synchrony, as we measured it using CRQA, reflects similarity between the observer's and demonstrator's electrodermal activity during the learning phase. This similarity was quantified using four commonly used metrics calculated on the resulting cross-recurrence plots from each learning phase recorded in our experiment (see Fig 2). These metrics capture salient patterns in how the patterns of similarity between observers and demonstrators evolve. We found particularly strong evidence for determinism (DET) and laminarity (LAM) as predictors of later conditioned responses even if all metrics conformed to our hypothesis. More determinism between two signals indicates that a larger proportion of the recurrent time points form diagonal lines in the cross-recurrence plot. This implies a stronger coupling between the two signals. Laminarity indicates vertical segments in the cross-recurrence plots, and suggest sustained, smooth periods in the signal's mutual evolution. We also found that all our metrics were highly correlated,

as has been previously reported [35], and we therefore relied on composite measure of all four metrics captured by their first principal component. Across all our analyses, the more similar demonstrator and observers were in their electrodermal activity during the observational learning phase, the stronger the observer's CS differentiation was during the testing phase.

How should synchrony be interpreted beyond the similarity of the physical properties of the two electrodermal signals? In this study, we argue, synchrony likely reflects the observer mirroring the demonstrator's autonomic nervous system trajectories as they react to and learn the associations between the CS images and shocks. This experience sharing in turn facilitates the observer learning of the CS contingencies absent direct experience with the shocks. Psychophysiologists have hypothesized that sharing another person's autonomic nervous system state is the physiological substrate of empathy [1, 47, 48], part of what empathy researchers will refer to as 'experience sharing' [49] or 'empathic distress' [50]. Individual arousal levels from viewing another person in pain, as indexed by electrodermal activity, have been shown to correlate with later costly helping, which provides indirect evidence for a link between empathy and matches in arousal states [51]. The sharing hypothesis has also been supported by classic findings that empathic accuracy is greatest during periods of synchronized physiology [48] and more recently that empathic accuracy moderates synchrony in a pain task [52]. Both behavioral and neural synchrony has been linked to mirror neurons responsible for representing the actions and intentions of social partners [19, 21, 24, 25], and several accounts have attempted to extend these mechanisms to cover empathic responses [19, 25, 53]. However, to the best of our knowledge, the results reported here represent the first direct experimental evidence that coupling of autonomic nervous system trajectories might serve as a marker of empathic sharing in a learning context.

The current work advances previous indirect evidence for the experience sharing hypothesis in social learning. For example, past research has found that individuals high in psychopathic traits exhibited impaired conditioned responses to a demonstrator getting shocks compared to normal controls [17] and that facial mimicry in response to watching a demonstrator getting shocks reflected empathic experience sharing [16]. In humans, the anterior insula and the anterior cingulate cortex are part of a network responding to observational threat learning [12], and these areas are known to facilitate empathy and emotion sharing in humans [50, 54]. In rodents, brain regions homologous to those supporting empathy and emotion sharing in humans have been shown to be necessary for successful observational threat learning [14, 15, 55]. However, these past findings concerned single individuals and not shared neural patterns between observers and demonstrators. In the current study we have a direct measure of two individual's autonomic nervous system activity. A central question for future work is how this synchrony arises in the brain and how it interacts with neural systems known

to be involved in threat learning. It might be possible to use hyperscanning techniques in humans using EEG or fMRI [36, 56], to address this question using the experimental paradigm established here. Another promising method would be to begin with investigating if similar synchrony can be found in rodent models, where neural recordings of sub-cortical structures implicated in arousal are more readily available.

Apart from experience sharing, other forms of empathy have also been suggested to facilitate observational threat learning [10, 18]. For example, one study found that instructing observers to take the perspective of the demonstrator increased conditioned responses for observers who were also high in trait empathy [10]. This appraisal instruction engages mentalizing aspects of empathy that involve making explicit inferences about a partner's internal states [20, 49]. In the results reported here we found no relation between trait empathy, as captured by the interpersonal reactivity index, and conditioned responses. This suggests that we should interpret previous links between trait mentalizing empathy and observational threat learning with some caution, especially given the large sample size of the present study. There are several differences in method between the current study and studies mentioned above, most notably the absence of appraisal instructions and the use of live versus video-filmed demonstrators. In this study, trait empathy also didn't moderate the relationship between synchrony and CS differentiation. This is consistent with a recent study on empathy for pain where empathic accuracy, but not trait empathy, was a moderator for synchrony [52], even if synchrony was operationalized differently in that study compared to ours. Our findings indicate that interpersonal synchrony affects learning independently of trait empathy and this is consistent with our interpretation of synchrony reflecting experience sharing [20] or empathic distress [50].

Similarly, we found no relationship between the observer's ratings of the demonstrator, including compassion, and their conditioned responses in the testing phase. As with the trait measure, these ratings also didn't moderate the effect of synchrony. One explanation for this might be that these ratings reflect retrospective recollections of the state observers were in during the observational learning and therefore are not accurate, by contrast synchrony will be a direct state measurement. Alternatively, interpersonal synchrony might reflect empathic processes which are more difficult to introspect on. Further work is necessary to fully understand the contributions of multiple empathic systems on processing social stimuli during threat learning.

Our findings are specific to shared patterns of arousal arising from the dynamics of the two participants' skin conductance signals. We demonstrated that competing predictors derived from the observers' electrodermal activity during learning showed no relation to the strength of the observers' conditioned responses and that controlling for these, did not remove the relationship between synchrony and conditioned responses. Importantly, this allows us to rule out that synchrony can be explained away as being caused by the

observer responding particularly strongly to the CS+ or to the demonstrator receiving shocks.

We found that the observers' responses to the social UCS predict their general level of arousal during the testing phase, but not the strength of their later conditioned response. This was surprising since responses to the social UCS have typically been taken to index the strength of observer's empathic responses to the demonstrator and should translate into stronger threat memories available for later recall, analogous to how a stronger UCS works during direct conditioning. In our data, observers who react strongly to seeing the demonstrator being shocked have higher response amplitudes to both the CS- and the CS+, perhaps reflecting an anxious or fear-like state during the testing phase. This suggests that the direct response to the social UCS might be less important for the observer's differential learning than previously theorized, especially when compared to mirroring and directly sharing the dynamics of the demonstrator's arousal states and might represent an uncharted difference to direct learning.

Outstanding questions arising from this study concern the factors that can affect the degree of demonstrator-observer coupling, and how these can be manipulated. Past work has demonstrated that people tend to synchronize more with people they are more positively disposed towards [57, 58] or closer to [34]. Similarly, observers learn about threats better from demonstrators who are similar to them [59]. Hence, an important avenue for future work is to manipulate the relationships between participants, for example by minimal group induction or by using natural covariates such as friendship [38], to investigate if and how this affects synchrony's role in learning.

A better understanding of the conditions during which synchrony emerges is important for establishing whether synchrony affects phenomena related to observational learning, like social buffering and social safety learning [60, 61]. In those paradigms observational learning protects the individual from experiencing strong threat responses and provides a safe route towards extinction of previously learned threat associations. If synchrony attunes the observer to the demonstrators experiences then it is possible that similarly synchrony will serve to aid vicarious extinction, which in turn has clinical relevance as a model for understanding phobias and other anxiety disorders.

Methods

Participants

We recruited a total of 138 participants who formed 69 unique demonstrator-observer dyads. Dyads were matched by gender (24 male, 45 female). Average age was 25 years (SD = 4.1). Participants were recruited from the student population at Karolinska Institutet and the surrounding local community. We ensured that participants did not already know each other prior

to participating in the experiment. Participants were screened from having previously partaken in conditioning experiments. Participants were given two cinema ticket vouchers as thanks for their participation. The experimental procedure was approved by the local ethics committee (2015/2115-32).

Equipment

Electrodermal activity was recorded using a BIOPAC MP150 system and AcqKnowledge software as a skin conductance signal measured in microsiemens (μS). Recording was at 1000Hz from the distal phalanges of the middle and index fingers from the hand belonging to the arm not receiving shocks. Ag-AgCl electrodes were used together with a conducting gel, following standard recommendations [62]. Electric shocks, consisting of a single 100ms DC pulse, were administered using a Biopac STM200 module (Biopac Systems Inc.) applied to the lower forearm. The strength of the electric shocks was individually calibrated so that participants experienced the shocks as being "unpleasant but not painful". Stimulus presentation and shock administration was controlled by scripts programmed in PsychoPy [63].

Procedure

Participants first read and signed consent forms. They were then given separate written instructions about their initial role during the experiment; one participant read instructions pertaining to the role of "observer" and the other participant read instructions to role of "demonstrator". Both participants were taken to the recording room and sat on chairs positioned so that they would afford a good view both of a computer screen and the other participant (see Supplementary Fig. 1). Both participants were asked to place their respective instruction sheets in front of them as a reminder of their current role. Each participant's instructions also emphasized that all communication between them was forbidden for the duration of the experiment.

The experiment was divided into four blocks and followed established protocols for video-based observational learning paradigms [39].

Shock electrodes were attached to the arm of each participant closest to the stimulus monitor, to maximize visibility of the arm to the other participant. Electrodes for recording electrodermal activity were attached to the distal phalanges of the middle and index fingers on the side not receiving shocks. The observer also received shock electrodes, as they would later be exposed to the threat of shocks during the testing phase, as is standard procedure in our observational learning protocol [39].

Each participant, in turn, went through an individual work-up procedure to calibrate the appropriate level of electrical stimulation in the experiment. Participants were instructed to choose a level of stimulation which they experienced as "uncomfortable, but not painful". Past research has shown that

an observer receiving information about a (confederate) demonstrator's level of stimulation does not affect observational learning [6]. Both participants undergoing work-up implies that the observer underwent work-up prior to their first learning phase. This is a difference in our procedure to previous experiments and was included to facilitate the multi-block design of the present experiment (see below). Since the observers know the shock level they have set for themselves, they will have an accurate idea of the threat the shocks pose. In our previous, video-based work, observers had to guess what shocks would be like from the (exaggerated) reactions of the demonstrators. Hence, any effect of this prior shock exposure would likely have been to dampen the expression of the CS contingencies.

Once the participants were connected to the equipment the experiment commenced. The experimenter proceeded to give final verbal instructions to the participants. The demonstrator was reminded that the only way the observer could understand the pattern between the images that are shown on the computer screen and them receiving the shocks is by looking at them and therefore they are asked to act naturally and expressively, but without attempting to exaggerate or fake their responses. The demonstrator was additionally instructed to keep their gaze focused on the screen and to not look towards the observer. The observer was reminded that they may look wherever they wanted - to the demonstrator and the screen - to learn which stimulus predicts shocks and which doesn't. Finally, both participants were reminded that direct communication was forbidden. The experimenter remained in a corner of the room monitoring the participants, but also to ensure that participants were easily able to cancel their participation should they so wish.

Each block consisted of a learning phase and was followed immediately by a testing phase. In the learning phase the observer attempted to learn the CS contingencies by watching the demonstrators reactions to the CS images. The learning phase consisted of six alternating presentations of each CS+ and CS- image. Each CS was shown for six seconds. There was a variable length 10-16 second inter-trial interval between each CS presentation. Four of the six CS+ presentations, randomly determined, terminated with a shock to the demonstrator. Importantly, the observer received no shocks during this phase nor any instructions about which image was the CS+ (shock predicting) and which was the CS-. The only way to learn this contingency for the observer was through observation.

Following the learning phase, on-screen instructions informed the observer that they would view the same two CS images again and *now* receive shocks to the *same* image that they had observed the demonstrator previously receive shocks to. During this phase the demonstrator was instructed to close their eyes and a screen was placed between demonstrator and observer occluding the observers' view of the demonstrator (Supplementary Fig. 2). These steps were taken to ensure that the observer would not be

able to pick up any cues during the testing phase about the valence of the CS images. Hence, any expression of heightened electrodermal activity to the CS+ image compared to the CS- image would *only* reflect associations observationally formed during the previous learning phase. During the testing phase each CS image was shown seven times. Unbeknownst to the observer, only the final CS+ presentation actually terminated with a shock. This final shock was given to ensure that the observers would consider the threat of shock credible also in the next block. The whole procedure was repeated the following block. Critically, each block featured *unique* stimulus images, hence there was no possibility of carry-over of learning between blocks.

After the second block had completed, the observer was asked to rate the demonstrator on four metrics: on how much pain the demonstrator seemed to be in, how much compassion the observer felt for the demonstrator, the extent to which the demonstrator was a good model and helped them learning and to how similar to them the demonstrator appeared to be. All ratings were completed using a 16cm visual-analog scale.

Once the observer had completed rating the demonstrator, the participants were asked to switch instruction sheets with one another. The experiment continued for two more blocks with the participants in reversed roles. Participants were unaware that this role reversal would occur. The role reversal was implemented to record more blocks from each recruited dyad. Ethical considerations limited us from exposing the demonstrator to more than two phases of direct conditioning. The role reversal means that each participant is observer for two of the four blocks of the experiment, and demonstrator for the other two. Half the participants begin the experiment with two blocks as observer while the other half of the participants become observers only after having been demonstrators first for two blocks.

Once all four blocks had completed, the current observer rated the demonstrator on the same four metrics as introduced above. Both participant then completed an Interpersonal Reactivity Index (IRI). The IRI consists of 28 items answered on a 5 point Likert scale ranging from "Does not describe me well" to "Describes me very well" and contains 4 subscales: Perspective taking, Fantasy, Empathic Concern and Personal Distress [46]. Finally, participants were thanked and debriefed.

Signal processing

Electrodermal activity was continuously measured from both participants throughout the experiment. All processing of the electrodermal recordings was done in AcqKnowledge 4.1 (Biopac Systems Inc.). The raw signal from each participant was filtered offline with a low-pass filter (1Hz) and a high-pass filter (0.05Hz). Using CS onset and shock delivery as event markers and following established protocols [39], skin conductance responses (SCRs) were measured as the largest peak-to-peak amplitude difference in the phasic skin

conductance signal in the 0.5 to 4.5 second window following stimulus onset. Responses below $0.02\mu S$ were scored as zero. Scoring was first done using AcqKnowledge's automated scoring algorithm, and then manually checked by an experimenter. SCRs were square-root transformed prior to analysis. The full phasic skin conductance signal from both participants was also exported for use in the cross-recurrence quantification analysis (see below).

Cross-Recurrence Quantification Analysis

Cross-Recurrence Quantification Analysis (CRQA) is the bivariate extension of Recurrence Quantification Analysis (RQA) a method to analyze the structure and revisitation (recurrence) patterns in non-linear time-series [64]. RQA is based on the analysis of recurrence plots [65]. In a recurrence plot (see 1), each dot marks a point of recurrence in a reconstructed phase space of the signal. The phase space is constructed using time-delay embedding. Points are considered to be recurrent if they are within some radius of one another in the resulting high dimensional phase space. Hence, three parameters need to be set to compute a recurrence plot from a time-series: time delay, number of embedding dimensions and radius (see [65] for a rigorous treatment). CRQA works analogously but where the patterns of revisitation are compared between two signals [27, 66]. CRQA yields cross-recurrence plots, analogous to regular recurrence plots.

We used the *crqa* package [41] implemented in the R statistical language to construct the cross-recurrence plots. Each cross-recurrence plot was based on the phasic skin conductance signal from the demonstrator and observer from each learning phase. The signals were down-sampled to 8Hz and then z-scored. Optimal parameters for the CRQA analysis (delay, embedding dimensions and radius) were determined individually for each pair of signals so that they would yield an average recurrence rate between 2% and 4% [34, 65]. From each resulting cross-recurrence plot various metrics can be computed that capture the dynamics of the system being analyzed [27, 41, 65]. Here we computed four metrics: DETerminism, LAMinarity, maximum line (maxL) and relative Entropy (rENTR). DET represents the relative amount of recurrent points forming diagonal segments, as such DET measures the predictability of the time-series as they evolve over time. LAM is analogous to DET but instead represents recurrent points forming vertical line segments, which can be thought of capturing relative stability in the system. maxL is length of the longest diagonal sequence of recurrent points, capturing the maximal strength of coupling between the two time series. rENTR calculates the Shannon entropy of the histogram of the deterministic (diagonal) sequences and indexes the complexity of the relationship between the time series.

Analysis

All analyses were performed in the R statistical language using the *brms* package [67]. We analyzed the data using Bayesian multi-level regression including varying intercepts and slopes by participant and between intercept and slope correlations. All categorical regressors were deviation coded (0.5/-0.5) and all continuous regressors were standardized. Weakly informative priors were used for all analyses to regularize estimates. These were Normal(0,0.5) for the intercept, Normal(0,0.2) for the slope of CS, and Normal(0,0.05) for all other slopes which usually contained the parameters of interest. LKJ(2) priors were placed on the correlation matrix and Normal(0,1) priors on the group standard deviations and the model residual standard deviation. For all estimates 95% credible intervals were computed as well as Bayes Factors based on the Savage-Dickey ratio for the parameter at the value 0. We rely on Bayes Factors to make inferences about effects. We interpret Bayes Factors above 10 to constitute strong evidence for an effect, and Bayes Factors between 3 and 10 to constitute weak evidence for an effect.

Data availability

All processed data are available from the Open Science Framework (`https://osf.io/mkv8c/?view_only=5789aa6170984416b9c9fb6ae4626fbe`), and raw data are available from the authors upon request.

Acknowledgements

This research was supported by the European Research Council (Independent Starting Grant 284366; Emotional Learning in Social Interaction) and the Knut and Alice Wallenberg Foundation (KAW 2014.0237) to A. Olsson. We thank E.C. Nook, A. Golkar and B. Lindström for helpful comments.

References

[1] William McDougall. *An introduction to social psychology.* Meuthen, London, 1908.

[2] Floyd Henry Allport. *Social Psychology.* Houghton Mifflin Company, Boston, 1924.

[3] Seymour M Berger. Incidental learning through vicarious reinforcement. *Psychological Reports*, 9(3):477–491, 1961.

[4] Seymour M Berger. Conditioning through vicarious instigation. *Psychological review*, 69(5):450, 1962.

[5] Albert Bandura and Theodore L Rosenthal. Vicarious classical conditioning as a function of arousal level. *Journal of personality and social psychology*, 3(1):54, 1966.

[6] Staffan Hygge. Information about the model's unconditioned stimulus and response in vicarious classical conditioning. *Journal of Personality and Social Psychology*, 33(6):764, 1976.

[7] Susan Mineka, Mark Davidson, Michael Cook, and Richard Keir. Observational conditioning of snake fear in rhesus monkeys. *Journal of abnormal psychology*, 93(4):355, 1984.

[8] Andreas Olsson, Katherine I Nearing, and Elizabeth A Phelps. Learning fears by observing others: the neural systems of social fear transmission. *Social cognitive and affective neuroscience*, 2(1):3–11, 2007.

[9] Andreas Olsson and Elizabeth A Phelps. Social learning of fear. *Nature neuroscience*, 10(9):1095, 2007.

[10] Andreas Olsson, Kibby McMahon, Goran Papenberg, Jamil Zaki, Niall Bolger, and Kevin N Ochsner. Vicarious fear learning depends on empathic appraisals and trait empathy. *Psychological science*, 27(1):25–33, 2016.

[11] Jacek Debiec and Andreas Olsson. Social fear learning: from animal models to human function. *Trends in cognitive sciences*, 21(7):546–555, 2017.

[12] Björn Lindström, Jan Haaker, and Andreas Olsson. A common neural network differentially mediates direct and social fear learning. *NeuroImage*, 167:121–129, 2018.

[13] Björn Lindström, Ida Selbing, and Andreas Olsson. Co-evolution of social learning and evolutionary preparedness in dangerous environments. *PloS one*, 11(8):e0160245, 2016.

[14] KZ Meyza, I Ben-Ami Bartal, MH Monfils, JB Panksepp, and E Knapska. The roots of empathy: through the lens of rodent models. *Neuroscience & Biobehavioral Reviews*, 76:216–234, 2017.

[15] Daejong Jeon, Sangwoo Kim, Mattu Chetana, Daewoong Jo, H Earl Ruley, Shih-Yao Lin, Dania Rabah, Jean-Pierre Kinet, and Hee-Sup Shin. Observational fear learning involves affective pain system and ca v 1.2 ca 2+ channels in acc. *Nature neuroscience*, 13(4):482–488, 2010.

[16] Katherine Burns Vaughan and John T Lanzetta. Vicarious instigation and conditioning of facial expressive and autonomic responses to a model's expressive display of pain. *Journal of Personality and Social Psychology*, 38(6):909, 1980.

[17] Albert S Aniskiewicz. Autonomic components of vicarious conditioning and psychopathy. *Journal of Clinical Psychology*, 35(1):60–67, 1979.

[18] Johan L Kleberg, Ida Selbing, Daniel Lundqvist, Björn Hofvander, and Andreas Olsson. Spontaneous eye movements and trait empathy predict vicarious learning of fear. *International Journal of Psychophysiology*, 98(3):577–583, 2015.

[19] Vittorio Gallese. The roots of empathy: the shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology*, 36(4):171–180, 2003.

[20] Jamil Zaki, Jochen Weber, Niall Bolger, and Kevin Ochsner. The neural bases of empathic accuracy. *Proceedings of the National Academy of Sciences*, 106(27):11382–11387, 2009.

[21] Uri Hasson and Chris D Frith. Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. *Phil. Trans. R. Soc. B*, 371(1693):20150366, 2016.

[22] JA Scott Kelso. *Dynamic patterns: The self-organization of brain and behavior*. MIT press, 1997.

[23] Hermann Haken, JA Scott Kelso, and Heinz Bunz. A theoretical model of phase transitions in human hand movements. *Biological cybernetics*, 51(5):347–356, 1985.

[24] Thalia Wheatley, Olivia Kang, Carolyn Parkinson, and Christine E Looser. From mind perception to mental connection: Synchrony as a mechanism for social understanding. *Social and Personality Psychology Compass*, 6(8):589–606, 2012.

[25] Gün R Semin and John T Cacioppo. Grounding social cognition: Synchronization, coordination, and co-regulation. In G. R. Semin and E. R. Smith, editors, *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches*, pages 119–147. Cambridge University Press, New York, NY, 2008.

[26] Michael J Hove and Jane L Risen. It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, 27(6):949–960, 2009.

[27] Kevin Shockley. Cross recurrence quantification of interpersonal postural activity. In M. A. Riley and G. C. Van Orden, editors, *Tutorials in contemporary nonlinear methods for the behavioral sciences*, chapter 4, pages 142–177. 2005.

[28] Emilie Delaherche, Mohamed Chetouani, Ammar Mahdhaoui, Catherine Saint-Georges, Sylvie Viaux, and David Cohen. Interpersonal synchrony: A survey of evaluation methods across disciplines. *IEEE Transactions on Affective Computing*, 3(3):349–365, 2012.

[29] Kyongsik Yun, Katsumi Watanabe, and Shinsuke Shimojo. Interpersonal body and neural synchronization as a marker of implicit social interaction. *Scientific reports*, 2:959, 2012.

[30] Zoltán Néda, Erzsébet Ravasz, Yves Brechet, Tamás Vicsek, and A-L Barabási. Self-organizing processes: The sound of many hands clapping. *Nature*, 403(6772):849, 2000.

[31] Daniel C Richardson and Rick Dale. Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive science*, 29(6):1045–1060, 2005.

[32] Olivia Kang and Thalia Wheatley. Pupil dilation patterns spontaneously synchronize across individuals during shared attention. *Journal of Experimental Psychology: General*, 146(4):569, 2017.

[33] Kevin Shockley, Marie-Vee Santana, and Carol A Fowler. Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2):326, 2003.

[34] Ivana Konvalinka, Dimitris Xygalatas, Joseph Bulbulia, Uffe Schjødt, Else-Marie Jegindø, Sebastian Wallot, Guy Van Orden, and Andreas Roepstorff. Synchronized arousal between performers and related spectators in a fire-walking ritual. *Proceedings of the National Academy of Sciences*, 108(20):8514–8519, 2011.

[35] Dan Mønster, Dorthe Døjbak Håkonsson, Jacob Kjær Eskildsen, and Sebastian Wallot. Physiological evidence of interpersonal dynamics in a cooperative production task. *Physiology & behavior*, 156:24–34, 2016.

[36] Uri Hasson, Yuval Nir, Ifat Levy, Galit Fuhrmann, and Rafael Malach. Intersubject synchronization of cortical activity during natural vision. *science*, 303(5664):1634–1640, 2004.

[37] Greg J Stephens, Lauren J Silbert, and Uri Hasson. Speaker–listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences*, 107(32):14425–14430, 2010.

[38] Carolyn Parkinson, Adam M Kleinbaum, and Thalia Wheatley. Similar neural responses predict friendship. *Nature communications*, 9(1):332, 2018.

[39] Jan Haaker, Armita Golkar, Ida Selbing, and Andreas Olsson. Assessment of social transmission of threats in humans using observational fear conditioning. *nature protocols*, 12(7):1378, 2017.

[40] Sebastian Wallot and Giuseppe Leonardi. Analyzing multivariate dynamics using cross-recurrence quantification analysis (crqa), diagonal-cross-recurrence profiles (dcrp), and multidimensional recurrence quantification analysis (mdrqa)–a tutorial in r. *Frontiers in psychology*, 9:2232, 2018.

[41] Moreno I Coco and Rick Dale. Cross-recurrence quantification analysis of categorical and continuous time series: an r package. *Frontiers in psychology*, 5:510, 2014.

[42] Delos D Wickens and Gordon B Harding. Effect of ucs strength on gsr conditioning: A within-subject design. *Journal of experimental psychology*, 70(2):151, 1965.

[43] Richard W Morris and Mark E Bouton. Effect of unconditioned stimulus magnitude on the emergence of conditioned responding. *Journal of Experimental Psychology: Animal Behavior Processes*, 32(4):371, 2006.

[44] Zoltan Annau and Leon J Kamin. The conditioned emotional response as a function of intensity of the us. *Journal of comparative and physiological psychology*, 54(4):428, 1961.

[45] T Rao Laxmi, Oliver Stork, and Hans-Christian Pape. Generalisation of conditioned fear and its behavioural expression in mice. *Behavioural brain research*, 145(1-2):89–98, 2003.

[46] Mark H Davis et al. A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, (10), 1980.

[47] Albert F Ax. Goals and methods of psychophysiology. *Psychophysiology*, 1(1):8–25, 1964.

[48] Robert W Levenson and Anna M Ruef. Empathy: A physiological substrate. *Journal of personality and social psychology*, 63(2):234, 1992.

[49] Jamil Zaki and Kevin N Ochsner. The neuroscience of empathy: progress, pitfalls and promise. *Nature neuroscience*, 15(5):675, 2012.

[50] Tania Singer and Olga M Klimecki. Empathy and compassion. *Current Biology*, 24(18):R875–R878, 2014.

[51] Grit Hein, Claus Lamm, Christian Brodbeck, and Tania Singer. Skin conductance response to the pain of others predicts later costly helping. *PloS one*, 6(8):e22759, 2011.

[52] Pavel Goldstein, Irit Weissman-Fogel, and Simone G Shamay-Tsoory. The role of touch in regulating inter-partner physiological coupling during empathy for pain. *Scientific reports*, 7(1):3252, 2017.

[53] Laurie Carr, Marco Iacoboni, Marie-Charlotte Dubeau, John C Mazziotta, and Gian Luigi Lenzi. Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proceedings of the national Academy of Sciences*, 100(9):5497–5502, 2003.

[54] Claus Lamm, Jean Decety, and Tania Singer. Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage*, 54(3):2492–2502, 2011.

[55] Stephen A Allsop, Romy Wichmann, Fergil Mills, Anthony Burgos-Robles, Chia-Jung Chang, Ada C Felix-Ortiz, Alienor Vienne, Anna Beyeler, Ehsan M Izadmehr, Gordon Glober, et al. Corticoamygdala transfer of socially derived information gates observational learning. *Cell*, 173(6):1329–1342, 2018.

[56] Pavel Goldstein, Irit Weissman-Fogel, Guillaume Dumas, and Simone G Shamay-Tsoory. Brain-to-brain coupling during handholding is associated with pain reduction. *Proceedings of the national academy of sciences*, 115(11):E2528–E2537, 2018.

[57] Zhong Zhao, Robin N Salesse, Ludovic Marin, Mathieu Gueugnon, and Benoît G Bardy. Likability's effect on interpersonal motor coordination: Exploring natural gaze direction. *Frontiers in psychology*, 8:1864, 2017.

[58] Alexandra Paxton and Rick Dale. Argument disrupts interpersonal synchrony, 2013.

[59] Armita Golkar, Vasco Castro, and Andreas Olsson. Social learning of fear and safety is determined by the demonstrator's racial group. *Biology letters*, 11(1):20140817, 2015.

[60] Yasushi Kiyokawa, Yukari Takeuchi, and Yuji Mori. Two types of social buffering differentially mitigate conditioned fear responses. *European Journal of Neuroscience*, 26(12):3606–3613, 2007.

[61] Armita Golkar, Cathelijn Tjaden, and Merel Kindt. Vicarious extinction learning during reconsolidation neutralizes fear memory. *Behaviour research and therapy*, 92:87–93, 2017.

[62] Wolfram Boucsein, Don C Fowles, Sverre Grimnes, Gershon Ben-Shakhar, Walton T Roth, Michael E Dawson, and Diane L Filion. Society for psychophysiological research ad hoc committee on electroder-

mal measures. publication recommendations for electrodermal measurements. *Psychophysiology*, 49(8):1017–1034, 2012.

[63] Jonathan W Peirce. Psychopy—psychophysics software in python. *Journal of neuroscience methods*, 162(1-2):8–13, 2007.

[64] J-P Eckmann, S Oliffson Kamphorst, and David Ruelle. Recurrence plots of dynamical systems. *EPL (Europhysics Letters)*, 4(9):973, 1987.

[65] Norbert Marwan, M Carmen Romano, Marco Thiel, and Jürgen Kurths. Recurrence plots for the analysis of complex systems. *Physics reports*, 438(5-6):237–329, 2007.

[66] Norbert Marwan and Jürgen Kurths. Nonlinear analysis of bivariate data with cross recurrence plots. *Physics Letters A*, 302(5-6):299–307, 2002.

[67] Paul-Christian Bürkner. brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software*, 80(1):1–28, 2016.