Deep Structures of Collaboration

Prerna Chikersal

CMU-RI-TR-17-36

August 2017

School of Computer Science

Carnegie Mellon University

Pittsburgh, PA 15213

Masters Committee:

Laura Dabbish, Chair

Louis-Philippe Morency

Marynel Vázquez

Submitted in partial fulfillment of the requirements

for the degree of Master of Science in Robotics.

Keywords: Physiological synchrony; Behavioral similarity; Collective Intelligence; Distributed/ virtual teams

Abstract

Collective intelligence (CI), a group's capacity to perform a wide variety of tasks, is a key factor in successful collaboration. Group composition, particularly diversity and member social perceptiveness, are consistent predictors of CI, but we have limited knowledge about the mechanisms underlying their effects. To address this gap, we examine how physiological synchrony, as an indicator of coordination and rapport, relates to CI in computer-mediated teams, and if synchrony might serve as a mechanism explaining the effect of group composition on CI. We present results from a laboratory experiment where 120 dyads completed the Test of Collective Intelligence (TCI) together online and rated their group satisfaction, while wearing physiological sensors. The first 60 dyads communicated via video and audio in study 1, while the next 60 dyads communicated via audio only in study 2. In study 1, we find that synchrony in facial expressions and synchrony in standard deviation of loudness in speech (both indicative of shared experience) was associated with CI and synchrony in electrodermal activity (indicative of shared arousal) with group satisfaction. Furthermore, various forms of synchrony mediated the effect of member diversity and social perceptiveness on CI and group satisfaction. In study 2, synchrony in facial expressions no longer had an effect on CI, but synchrony in standard deviation of loudness in speech continued to positively effect CI. Our results have important implications for online collaborations and distributed teams.

Acknowledgments

I would like to thank my adviser, Prof. Laura Dabbish, for giving me the opportunity to work on this project. Her valuable advice and ideas have been useful to me throughout my Masters. In addition, I would like to thank my wonderful collaborators, Dr. Maria Tomprou, Prof. Anita Woolley and Dr. Young Ji Kim, for helping me run the experiments, make better sense of the data and the implications of our research. Without their useful advice, discussions, and efforts, this project would not have been possible. I am also grateful towards Prof. L.P. Morency for his constant support and encouragement, and for sharing his experiences to help me better understand certain technical concepts and workings of the academic world.

Finally, I would like to thank our undergraduate research assistants, Thomas Rasmussen, Mikahla Vicino, Zhong Yu Bing, Sean Tao, Brian Hall, Pavan Gollapalli, and Karen Kim for helping us with data collection.

This project was supported by the National Science Foundation, under awards CNS-1205539, VOSS- 1322278, and VOSS-1322241.

Contents

This work also led to the following publication [1], which was presented at CSCW 2017:

Chikersal, P., Tomprou, M., Kim, Y. J., Woolley, A. W., & Dabbish, L. (2017, February). Deep structures of collaboration: Physiological correlates of collective intelligence and group satisfaction. In *Proceedings of the 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2017)*.

Chapter 1

Introduction

Recent research has demonstrated that groups exhibit “collective intelligence’” (CI) [2] defined as a group’s capacity to perform a wide variety of tasks, and that CI is consistently predictive of future performance [3, 4, 5, 6]. Further, CI has been shown to be heavily influenced by team composition and team structure [6], particularly by team diversity (in terms of sex composition and cognitive diversity; [7]) and inclusion of members with higher average social perceptiveness [2, 4, 8]. These results have been replicated with groups working online [4] and in groups in multiple cultures [3]. In addition to predicting team performance, CI is also associated with teams’ ability to engage in tacit coordination, or coordination without communication [7].

Despite advances in our understanding of CI and its relationship with team performance, we lack understanding of its so-called deep structure, that is how CI develops, and how details of physiological responses and behavior are related to CI and collaboration outcomes. The previous results on CI lead us to ask whether a basic mechanism via which group diversity or composition affects CI may reside in the sensing, and possibly synchronization, of subtle nonverbal physiological signals. Along with CI, we will also explore whether members’ satisfaction with the team, as a measure of how team members “feel” about the interaction, is associated with similar physiological mechanisms.

Our study uses new sensing instrumentation to explore the connections between diversity, physiological synchrony, CI and group satisfaction. Specifically we investigate how synchrony in physiological responses such as in electrodermal activity, heart rate, facial expressions, and speech is a mechanism via which diversity affects collective intelligence and group satisfaction in computer-mediated interaction.

Our results contribute to the field of CSCW by extending our understanding of the mechanisms underlying collective intelligence and group satisfaction. These findings enhance our ability to evaluate the quality of interaction in ongoing computer-mediated teams, advance our ability to model and detect collective intelligence in computer-mediated teams, and inform interventions that could build collective intelligence during early stages of collaboration.

Chapter 2

Background

2.1 Physiological Sensing and Computer Supported Collaborative Work

The previous research in CSCW on sensing group interaction falls primarily in two categories, intelligent meeting systems and group feedback systems. Intelligent meeting rooms use sensors to analyze ongoing group behavior for capture, annotation and review or intelligent intervention [9]. For example, computer vision or spatial microphones are used for speaker identification and automatic camera control rather than assessing interpersonal dynamics (e.g. [10]). More recently sensing is being used to support smart rooms that are aware of participant activity and location for augmented reality interface placement and behavior (e.g. the GravitySpace system tracks users and poses in a smart room to support AR projections that do not intersect with users bodies [11]).

Research on group feedback systems has also used sensors to measure behavior of individuals during group interactions in order to improve interaction quality and collaborative outcomes through real-time or posthoc visualization or feedback displays [12, 13, 14, 15, 16]. Much of this work has focused on equalizing participation in face-to-face meetings. The earliest of these systems SecondMessenger, was a visualization system for reviewing speaker participation patterns in a face-to-face discussion [13]. The goal in group feedback systems research is to capture patterns of interaction dynamics and reflect them to participants in real time or after the fact to alter the dynamics in some way.

Other work beyond these two categories has used physiological sensors to detect engagement or stress levels during collaboration. These sensors have been used to evaluate the quality of collaborative experiences. For example, Mandryk and Inkpen [17] pioneered the use of physiological indicators for indicating engagement levels during collaborative gameplay interactions. In the remote collaboration setting, recent work by Tan et al [18] used physiological sensors to give remote collaborators awareness of their partner's workload. Their results suggested that collaboration supported by physiological feedback provides unobtrusive awareness of confusion or difficulty during a remote assembly task. These results suggest the potential for physiological sensing to provide a more fine-grained understanding of participant experience during collaborative interaction in addition to facilitating awareness among partners. Ultimately this sensing may be able to drive feedback systems and displays that improve collaboration quality.

Our study is distinct from the previous CSCW research on sensing in collaboration in that we are using physiological sensing to examine the underlying dynamics of collaboration, specifically synchrony, and how it is connected to collaborative outcomes. The previous work did not associate synchrony, or physiological sensing at the group level, with collective intelligence or group satisfaction as collaboration outcomes. Thus our results will inform whether and how sensing can provide early indicators of a group's future potential for collaboration.

2.2 Collective Intelligence and Group Satisfaction

Psychologists have repeatedly shown that a single statistical factor called "general intelligence" or "g" emerges from the correlations among individuals' performance on a wide variety of cognitive tasks, and that it predicts an individual's future performance. This general factor is in addition to more task-specific intelligences [19] with the majority of empirical analyses of individual ability supporting the notion of a general intelligence factor. Similarly, researchers recently explored whether such a general intelligence factor exists for groups, by adopting the same approach that psychologists have used in examining general intelligence in individuals [2]. They gave a sample of groups a wide range of different types of tasks, and found that teams that did well on one type of tasks tended to also do well on all of the other tasks. A factor analysis of the groups' scores revealed a single, dominant, general factor explaining a large proportion of the variance in all of the groups' scores. In individuals, this factor is called "general intelligence" or "IQ;" for groups, they call this first factor "collective intelligence." Collective intelligence (CI) was then shown to predict team's future performance on a more complex task [2]. Recent work has replicated these findings with groups working for just one hour on an online battery of tasks [4], in student teams [20], and in groups in multiple cultures [4, 21]. In many of these settings, collective intelligence has been shown to predict future performance, consistent with the prior research done in face-to-face groups.

A consistently puzzling observation in the work on CI is its lack of relationship with various measures gauging the quality of member interpersonal relationships [2, 4, 20] Variables such as group satisfaction or cohesion are generally treated as reliable indicators of the level of rapport in a team, even in online collaborations [22]. Since physiological synchrony has been shown to be an indicator of interpersonal rapport and relationship quality [23], it may be the case that some forms of synchrony will relate more to group satisfaction than CI.

Despite the empirical evidence of collective intelligence and its utility for predicting performance, research on collective intelligence is still in its infancy, leaving many questions unresolved. Prior to the recent studies in human teams, work on collective intelligence originated in other species, where it manifests as large scale coordination with a physiological basis, such as the following of pheromone scent trails by ants [24] or the reaction to visual signals in fish shoals [25] raising plausible research questions as to whether similar effects exist in human interactions. As such here we examine the physiological signals shown to govern various aspects of human social interaction in prior studies to investigate their role in human collective intelligence. We focus on the synchrony of facial myography or what we will refer to facial expressions, electrodermal activity and heart rate as previous experiments have shown synchrony in these signals relate to the quality of social interactions and level of cooperation both online and face to face (e.g., [23, 26, 27, 28]).

2.3 Physiological Synchrony

Across different social environments people often engage in group activities that lead the members to act in synchrony with each other [29]. As a social phenomenon, synchrony promotes affiliation and closeness among members across different teams and groups of individuals (e.g., from close relationships and newly formed teams to armies and group dancers; [29, 30]). Studies consistently find that behavioral synchrony promotes affiliation, establishes rapport and cooperation and supports the pursuit of joint goals [31, 32, 33]. Scholars reason that behavioral synchrony functions as social glue, which is powerful to promote coordinated action and joint outcomes. More recently, researchers have begun to test whether physiological synchrony (manifested in the synchronization of less consciously controllable physiological processes) reveals similar effects.

Recent studies build on advances in sensing technology that provide better capabilities to examine physiological synchrony and how it influences group performance and collaboration. Mitkidis et al. [28] found that trust has a positive effect on heart rate synchrony, and that the degree of heart rate synchrony was predictive of participants' expectations of their partners in a behavioral economics game. Mønster et al. [23] showed that synchronous activation of the zygomaticus major (the smile muscle) was related to team cohesion and members' decisions to adopt a new routine, whereas synchrony in electrodermal activity was related to negative affect and group tension. Interestingly, they found no relationship between physiological synchrony and task performance, but did find a strong relationship between physiological synchrony and the emotional aspects of cooperation (i.e., team cohesion and team tension). During adult-child interactions, synchrony in electrodermal activity is related with child's engagement levels [26] and better emotional attunement [34]. Synchrony in facial expressions also promotes emotional contagion among dyads [27].

In this study, we focus on the physiological synchrony of heart rate, electrodermal activity, facial expressions and speech, which have been shown in prior research to influence cooperation among teams [31, 35]. Electrodermal activity (EDA) or skin conductance is an indicator of emotional arousal and reactivity, both at conscious and unconscious levels [36, 37]. Heart rate is a measure of cardiac activity and also an indicator of arousal [36, 38] that has been used in different studies related to emotional episodes [39]. Both measures capture unconscious physiological processes and variations that link to certain emotional states such as positive affect, anxiety, and boredom as well as cognitive states such as level of engagement, arousal and attention [26, 40, 41, 42, 43, 44, 45, 46]. We interpret HR and EDA to be indicative of generalized arousal [47, 48, 49], however there is some empirical evidence that shows that EDA and HR can alter in opposite ways in certain conditions [45, 50] referred to as directional fractionation[1]. However we do not anticipate those conditions to be relevant to the current study and thus expect the signals of arousal via HR and EDA to be consistent. Further, facial myography can also complement the physiological picture of felt experience [51]. Previous research has shown that facial expressions can reliably detect conscious and unconscious experiences of affect [52] or mimicry [27]. Prosodic characteristics of speech have also been found to be indicators of affect [53], mood [54], and engagement [55].

[1]For example, in some studies, attentiveness caused EDA to increase [41, 42] and HR to decrease [43, 50]. These findings were not consistent though [52], and did not hold when the subject was moving or performing tasks requiring cognitive effort or when two or more cognitive constructs occurred simultaneously (eg: cognitive effort while paying attention) [50, 56] such as TCI). Additionally, this phenomenon is usually related to cognitive constructs like attention or anxiety, and measuring these algorithmically using physiological signals will require validation against participant reported values. Hence, we do not analyze the effects of directional fractionation on CI in this study.

Chapter 3

Hypothesis

Based on the existing findings, we now walk through our predictions for the current study.

3.1 Physiological Synchrony and CI

CI has been associated with teams' ability to engage in tacit coordination, that is mutual adjustment without explicit communication [7, 57]. This relationship was observed in teams of strangers participating in a laboratory study together over the space of a few hours, versus among individuals with a long-standing relationship, suggesting that the level of coordination undergirding CI must exist at a fairly basic, sensory level. Additional evidence demonstrates that CI levels established in a team's first interaction remain relatively stable over time, even following a period of months of regular interaction [58]. This reinforces the notion that rather than being based on relational elements that are developed over time, CI may be rooted in much more basic, instantaneous, and perhaps sensory-based mechanisms, which communicate and perceive all kinds of interpersonal information and expectancies, similar to those that drive thin-slice judgments [59].

However, exactly which forms of physiological synchrony will support CI is unclear, particularly given the lack of relationship demonstrated in prior research between CI and group satisfaction, cohesion, and psychological safety [2, 3, 4, 21]. Many of the existing studies of physiological synchrony support its association with indicators of group member relationships and rapport. For instance, research shows that team members' synchrony in EDA is related to tension and negative affect [23], and that synchrony in HR is related to cooperation [28] Some scholars have also found that spontaneous synchrony in facial expressions are related to cohesion [23] and resulted in increased levels of emotional interaction and liking among dyads (e.g., [27]). Further, synchrony in the prosodic characteristics of speech has also been found to positively correlate with engagement [60] and feelings of empathy [61] among dyads.

Our study will extend these findings in online collaborations and examine whether the association between group satisfaction and physiological synchrony holds for computed-mediated environments. Furthermore, given the differentiation between CI and group satisfaction previously noted, we may find that different patterns of physiological synchrony correspond with CI, deepening our understanding of these two different building blocks of collaboration. Therefore, we propose:

Hypothesis 1. Physiological synchrony is positively related with (a) collective intelligence and (b) group satisfaction.

3.2 The Effect of Social Perceptiveness

We also hypothesize that the effect of group members' social perceptiveness on CI [2, 4] will be explained in part by physiological synchrony. Social perceptiveness is a measure of an individual's ability to infer what others are thinking or feeling based on subtle, nonverbal cues. It is correlated with other aspects of emotional intelligence [62] and consistently related to higher CI [4, 8], and more effective group functioning [63] both in online and face-to-face collaborations [4, 6]. We argue that the underlying reason for the strong relationship of team-level social perceptiveness with CI and potentially with affect-laden group satisfaction is physiological synchrony. People who are high in social perceptiveness are better at communicating as well as coordinating physical movements with others, even in the absence of visual access to their interaction partner

[64]. This occurs because highly socially perceptive team members are more likely to pick up on the subtle nonverbal cues, and we expect that will also enable them to physiologically synchronize with others in a manner that facilitates rapport and coordination. This leads to our second hypothesis:

Hypothesis 2. Group average social perceptiveness will affect (a) collective intelligence and (b) group satisfaction via effects on physiological synchrony.

3.3 The Role of Group Composition

As described previously, collective intelligence has consistently been related to features of group composition and structure [6]. Specifically, linear and curvilinear relationships have been reported between CI and both gender and cognitive diversity [2, 4, 7] and with members' level of social perceptiveness [2, 4, 8]. In addition, some preliminary findings suggest that age diversity serves to disrupt CI [65]. Based on these findings and the existing literature on team diversity and social intelligence, we anticipate we will observe relationships between these various forms of diversity and CI, but furthermore that physiological synchrony may serve as a mechanism.

Typically diversity refers to any attribute that may lead one person to perceive another one as different from self [66]. In practice this may mean any aspect of differentiation with research typically focusing on these aspects that relate to background and social categorization (e.g., gender, education, age, ethnicity and so on). Diversity in work teams has been found to relate to both functional outcomes such as increased information sharing and creativity [67] as well as dysfunctional outcomes such as increased conflict [68, 69]. In our study, we investigate the accessible social categories that people may use to make conclusions, i.e., gender, age, and ethnicity. Ethnicity and age are important variables in team composition research because they are visible characteristics that may be used for social categorization [70] which are typically found to disrupt group relationships and productivity due to stereotyping and associated conflicts [71].

In addition to exploring whether the previously observed relationships between CI and diversity are replicated in dyads working online, we are also interested in exploring the role of physiological synchrony as a mechanism. There is some empirical evidence about the relationship of gender and age with physiological synchrony [72]. In married couples, when male spouses experience negative affect, they are also more likely to demonstrate increased electrodermal activity but the relationship between affect and arousal is absent in wives. In the same study, the researchers found that older couples reported higher positive affect and lower arousal than middle-aged couples [72]. Regarding ethnic composition, Blascovich and colleagues [73] showed that White participants interacting with Black confederates exhibited increased cardiovascular response and performed poorly on a cooperative task compared to participants interacting with White confederates. However, in another experimental setting, Blascovich and colleagues [73] found no effects in heart rate activity of the interaction of White participants with Black confederates. In our experiment, rather than focus on the main effect of ethnicity, sex or age, we focus on the effects of group composition on physiological synchrony. Specifically, we focus on whether dissimilarity in observable variables such as age, ethnicity, and gender disrupts physiological synchrony, and the degree to which that disruption helps to explain their role in collective intelligence and group satisfaction.

Hypothesis 3. Group composition (sex diversity, age diversity (or distance), and ethnic diversity) will affect (a) collective intelligence and (b) group satisfaction via effects on physiological synchrony.

Chapter 4

Method

4.1 Study Design

We investigated our hypotheses in the context of a laboratory study. In the experiment, teams of two completed the Test of Collective Intelligence (TCI). We start with examining our questions in dyads, as construct and measures of CI have been demonstrated to apply to dyads as well as larger groups [2], and focusing on dyads enables us to look at synchrony without the additional statistical and phenomenological complexity of subgroups that may form in larger groups (e.g., [74]). We collected individual measures of social perceptiveness and demographics before the TCI and group satisfaction after its completion. Throughout the TCI, we recorded physiological measures of electrodermal activity and heart rate. All sessions were also video recorded to obtain facial expression data.

4.1.1 Participants

We recruited 120 dyads from the participation pool of a large Northeastern university in the United States with the age range of 18 to 61 years old (M = 26.4, SD = 8.45). All participants were compensated 15 US dollars. We ran the study using both same- and mixed-gender teams (nearly equal male-male, female-female and mixed dyads). We failed to capture physiological signals and video for six dyads due to technical equipment issues.

4.1.2 Procedure

We conducted **two separate studies**. The procedure for the studies was as follows:

Each session lasted approximately 30 minutes. Members of each dyad were seated in different rooms. None of the participant pairs knew each other before the experiment. After completing a pretest survey, they were instructed to wear the E4[1] wristbands on their non-dominant hand and relax for two minutes to obtain a baseline in EDA and heart rate. After that, participants initiated the video conference call with their partner. **In study 1, participants were able to communicate via audio and video, however in study 2, the video window was hidden, such that the participants were only able to communicate via audio.** Participants then logged onto the Platform for Online Group Studies (POGS), a web browser-based platform that supports synchronous multiplayer interaction, to complete the Test of Collective Intelligence (TCI) with the other research participant [2, 3, 4]. The TCI contained six tasks ranging from 2 to 6 minutes each, and instructions were displayed before each task for 15 seconds to 1.5 minutes. At the end of the test, they were instructed to sign off the videoconference and proceed to the post-test survey, which was completed independently. Participants were then compensated and debriefed.

4.2 Measures

Participants provided demographic information and completed the test for social perceptiveness individually prior to working on the TCI with their group. Physiological synchrony was measured during the group work on the TCI. After the group work period was over, group satisfaction was measured at the individual level.

[1] https://www.empatica.com/e4-wristband

4.2.1 Group composition

We examined group composition in terms of three surface-level attributes: sex, age, and ethnicity. For sex, we calculated the number of females in each dyad (male only = 0, mixed = 1, female only = 2). Age diversity was operationalized as the distance between two members' ages in years. Age distance ranged from 0 to 40 years (M = 9.24 years, SD = 10.04). Ethnic diversity was dummy-coded; if participants reported identifying with different ethnic groups, they were considered dissimilar, coded as 1, otherwise, 0. Thirty-four dyads (56.7%) were ethnically dissimilar.

4.2.2 Social perceptiveness

To measure social perceptiveness, we used the Reading the Mind in the Eyes test (RME) developed by Baron-Cohen and colleagues [75]. The test consists of 36 images of the eye region of individual faces. Participants were asked to choose among possible mental states to describe what the person pictured was feeling or thinking. The options were complex mental states (e.g., guilt) rather than simple emotions (e.g., anger). Individual participants' scores were averaged for each dyad.

4.2.3 Collective intelligence

Collective intelligence was measured using the Test of Collective Intelligence (TCI), which was completed by dyads working together. The TCI is an online version of the collective intelligence battery of tests used by Woolley et al. [2], which contains a wide range of group tasks [3, 4]. The TCI was adapted into an online tool to allow researchers to administer the test in a standardized way, even when participants are not co-located. There were a total of six tasks in the version of the TCI used in this study which measured the dyads' ability to collaborate in a variety of ways by having them generate creative ideas, solve word and number puzzles, collectively remember detailed information, and execute detail-oriented tasks quickly and accurately. To obtain collective intelligence scores for all dyads, we first scored each of the six tasks and then standardized the raw task scores. We then computed an unweighted mean of the six standardized scores, a method adapted from prior research on collective intelligence [2].

4.2.4 Group satisfaction

To measure group satisfaction, we used six items that reflect the quality of group collaboration and relationship, adapted from the Team Diagnostic Survey (TDS, [76]) (e.g., "I am very satisfied working with this team"). Participants reported their ratings on a five point Likert-type scale (? = .72, M= 4.12, SD = .42). Since group satisfaction is conceptualized as a group-level construct, we aggregated individual dyad members' mean group satisfaction score to the dyadic level by computing the mean of two ratings (ICC(1) = .64, ICC(2) = .75). The median r*wg(J) [77] was .97. The values ranged from .78 to 1, demonstrating acceptable level of within-group agreement.

4.2.5 Physiological synchrony

We assessed physiological synchrony by recording facial expressions, electrodermal activity (EDA), and heart rate (HR) of each individual in the dyad throughout their interaction during the TCI. Synchrony in facial expressions can capture shared experience or mimicry [52] over time. Synchrony in HR and EDA captures shared arousal during periods of stress, excitement, or high levels of engagement [36, 37, 38]. We processed the individual responses into physiological response signals, sequences of response scores over time in the study, and then calculated distances between the series of scores (or signals) of each individual in a dyad using Dynamic Time Warping. In the rest of this section we describe how we translated sensor data into physiological response signals and calculated interpersonal distances.

4.2.6 Physiological response signals

In order to develop physiological response signals comparable across individuals, we first translated each person's physiological responses into scores over time for each measure. We then reduced noise in the scores and normalized the signals. We did this to account for individual differences that can change the scale of response for signals like EDA and HR that are usually influenced by factors such as age and cardiovascular fitness. For all our physiological signals, we restricted analyses to the task portions of the experiment, trimming out data where participants were not collaborating i.e., while reading instructions. Below, we describe how we transform raw signals from each sensor to physiological response signals, and illustrate the process in figure 4.1.

Facial Expressions

We derived facial expression signals from web conference videos of participants' faces recoded by Evaer[2]. For each CI task, we manually extracted the respective video and used OpenFace [78] to detect Facial Action Units (AUs) [79] in each frame. We coded for two types of expressions in the video, positive (AU12 with or without AU6), and negative (AU15 and AU1 and/or AU4).

We coded smiles of different types as positive expressions. A smile involves pulling the lip corners up (AU12) with or without raising the cheeks (AU6). Early research [79] argued that smiles in which AU12 and AU6 co-occur indicate felt positive emotion, while smiles containing only AU12 are polite, social, or "masking" smiles, and do not indicate felt emotion. However, recent findings [80, 81] show that smiles containing only AU12 can also indicate felt emotions, while smiles containing both AU12 and AU6 can also be feigned. Hence, whenever the system detects the lip corners being pulled upwards (AU12), it labels the expression as a positive one, irrespective of whether the cheeks are raised or not.

We coded activation of a different set of key facial action units as negative expressions. Negative expressions such as worry or displeasure are often conveyed by depression of the lip corners (AU15) and changes to the positioning of eyebrows forming something close to a frown (AU1 and/or AU4). Hence, when the system detects depression of the lip corners, and either raising of inner brows or lowering of brows or both (AU15 and AU1 and/or AU4), it labels the expression as a negative one.

[2]http://www.evaer.com/

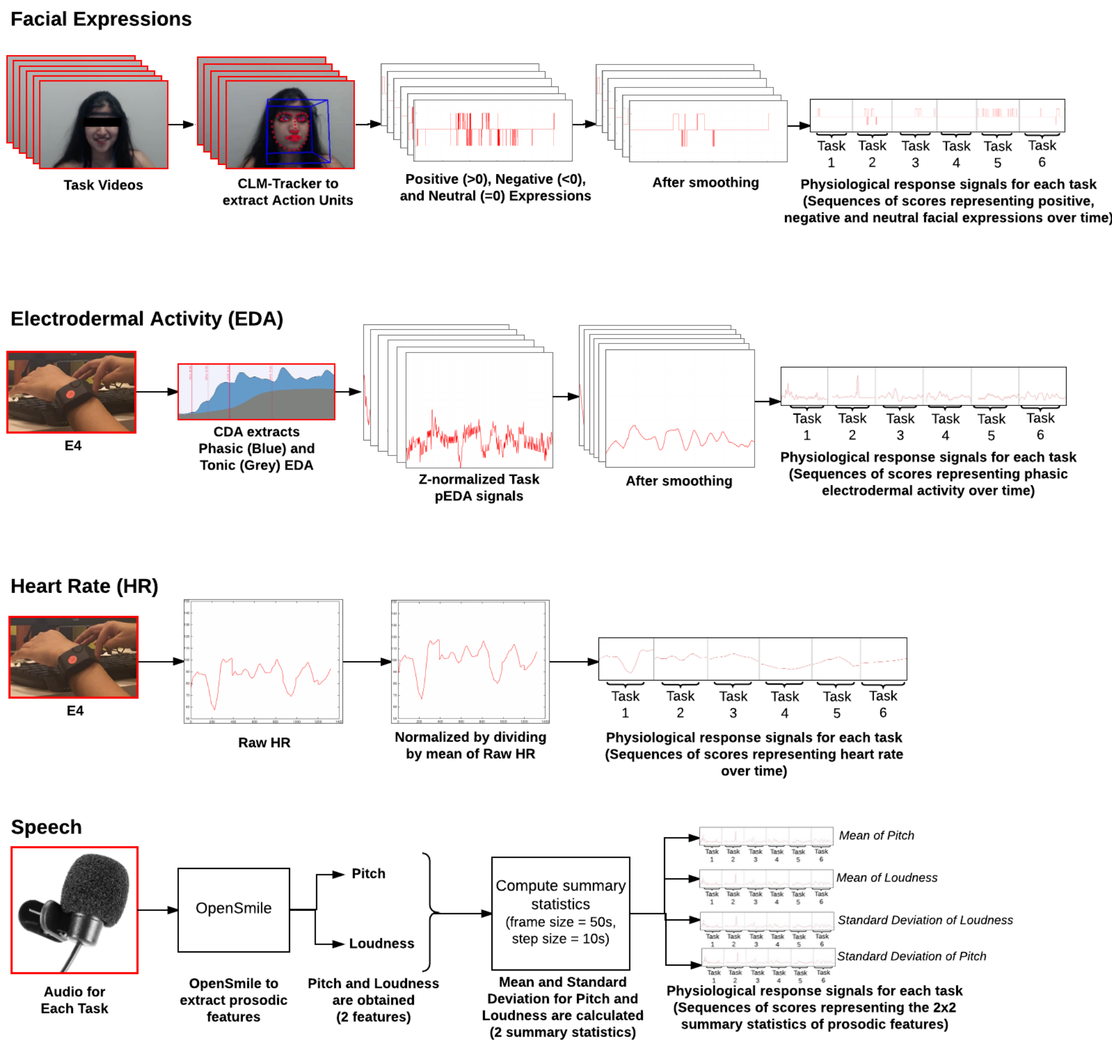

Figure 4.1: Data transformation from raw data to physiological response signals for each measure.

We converted the facial expressions identified to scores for each frame in the video. We assigned a value of "1" to frames containing a positive expression, "-1" to frames containing a negative expression, and "0" for frames containing neither (neutral). The signal obtained is noisy potentially due to jittery facial motion and the use of an automatic tool that can often inaccurately detect expressions. To reduce noise, we smoothed the signal over 29 frames (average frame rate/

approximately 1 second) by applying a Simple Moving Average filter (SMA). SMA allows us to calculate a moving average by adding signal samples over a number of time periods (i.e. 29 samples), and then dividing this total by the number of time periods. Finally, individual scores during each task make up the individual's physiological response signal for that task. Since we have six tasks, we get six physiological response signals for each person. We then calculate the synchrony of these scores between the two interaction partners as described below.

Electrodermal Activity (EDA)

We measured EDA to assess each participant's level of electrodermal arousal during the session. We recorded EDA using the E4 wristband with a sample rate of 4 samples per second (4Hz). The resulting signals have two components - tonic (skin conductance level) and phasic (skin conductance response). The tonic component changes gradually over time, approximates a person's baseline and is not the result of stimuli. The phasic component contains quickly changing peaks that typically occur in response to short-term events or environmental stimuli. We separate phasic EDA from tonic EDA using Continuous Decomposition Analysis [82], and use only phasic EDA (pEDA) henceforth, since we're only interested in participant task responses and not their baselines. We get six task pEDA signals, for each participant. These signals are subjected to the steps described below.

We normalized pEDA signals using z-score of each sample to enable inter-participant comparison. The signals obtained were very noisy and so we applied a Simple Moving Average filter (window size = 5 seconds = 5*4 samples, empirically determined) to smooth the signals, and name these physiological response signals. As in facial expressions, we end up with six physiological response signals (one for each of the six tasks). We then calculate the synchrony of these scores between the two interaction partners as described below.

Heart Rate (HR)

We measured HR to assess each participant's level of cardiovascular arousal during the session. Since different people have different resting heart rates, to enable inter-participant comparison, we normalized the HR data by dividing it by its mean (assuming mean to be an estimate of the baseline) and multiplying it by 100. No smoothing was required for HR signals, since HR is the number of heartbeats averaged or smoothed over a moving window of 1 minute. We obtained six task HR signals that are physiological response signals representing a participant's percentage change in HR over time, from his/her mean HR. We then calculate the synchrony of these scores between the two interaction partners as described below.

Speech

We recorded speech of each participant in the dyad into separate files using Evaer[3]. So, speaker recognition and diarization was not required. For each CI task, we used OpenSMILE [83] to extract 2 features related to pitch and loudness respectively, with a frame step of 10ms. Frame size is 60ms for pitch and 20ms for all other features. For each participant, we computed 2 summary statistics (mean and standard deviation) for each of the 2 features over a moving window of size 50 seconds and step 10 seconds. We aggregate prosodic features over 50 seconds, because they often change rapidly across utterances, and we want to represent affective or communicative states conveyed by prosody and not individual utterances. In the end, we obtained 2x2x6 (2 summary statistics x 2 features x 6 tasks) signals for each participant that are physiological response signals representing prosodic characteristics of speech for that participant.

4.2.7 Computing Synchrony

We use a method called Dynamic Time Warping (DTW) to compute synchrony in facial expressions, electrodermal activity, and heart rate. While, we use Pearsons Correlation to compute

[3] http://www.evaer.com/

synchrony in speech features. This is because to compute synchrony in facial expressions, electrodermal activity, and heart rate, we are trying to align events such as occurrence of facial expressions, increase in heart rate, etc, and need to account for differences in time lag and duration of events across partners . Whereas, speech response signals already represent prosodic features computed over a certain frame size using a moving window, thus there is no need to account for for differences in time lag and duration of events across partners, and Pearons correlation can be used.

Dynamic Time Warping (DTW)

We used Dynamic Time Warping (DTW) [84] to calculate synchrony between facial expressions, heart rate, and EDA of partners in a dyad. DTW is an algorithm for measuring similarity between two temporal sequences that vary in time and speed. Physiological response signals of each participant are also temporal sequences that are computed using the method described above. DTW provides the distance between the partners' physiological response signals for each task, which are then summed across tasks to give the total distance. We operationalize synchrony as similarity of a physiological measure between the partners' response signals of that measure. This similarity is calculated by subtracting the total distance from the total distance of the most different (largest total distance) dyad.

Figure 4.2 shows an example of two signals - Signal A and Signal B that are different in length, time, and speed. DTW warps the time axis of these signals to find corresponding points between the two signals that optimally match. Figure 4.3 shows some of the matched corresponding points in signals A and B. A locality constraint in DTW is the maximum distance allowed between the matched corresponding points in the two sequences. The algorithm used to match the signals illustrated in figure D does not specify a locality constraint, which may or may not be always favorable. Since, our physiological response signals are several minutes long, and we only want to match physiological responses expressed by participants within a few seconds of their partners as synchrony, having a locality constraint is necessary for this analysis. The locality constraint we chose for each measure is approximately 5 seconds, that is 145 samples for facial expressions (since average frame rate is 29Hz), 20 samples for electrodermal activity (frame rate is 4Hz), and 5 samples for heart rate (frame rate is 5Hz). This locality constraint is arbitrarily chosen since it appears to cover roughly one cycle of the signals (see [85]). We use Euclidean distance to calculate distance between any two corresponding points in DTW.

Unlike other methods like sample-wise Euclidean distance and cross-correlation, DTW is able to overcome time lag and flexibility issues in our data. For example, if a participant smiles 4 seconds after his/her partner (time lag issue), or if a person smiles 2 seconds longer than his/her partner (time flexibility issue), DTW matches both these events. Next, we present our findings.

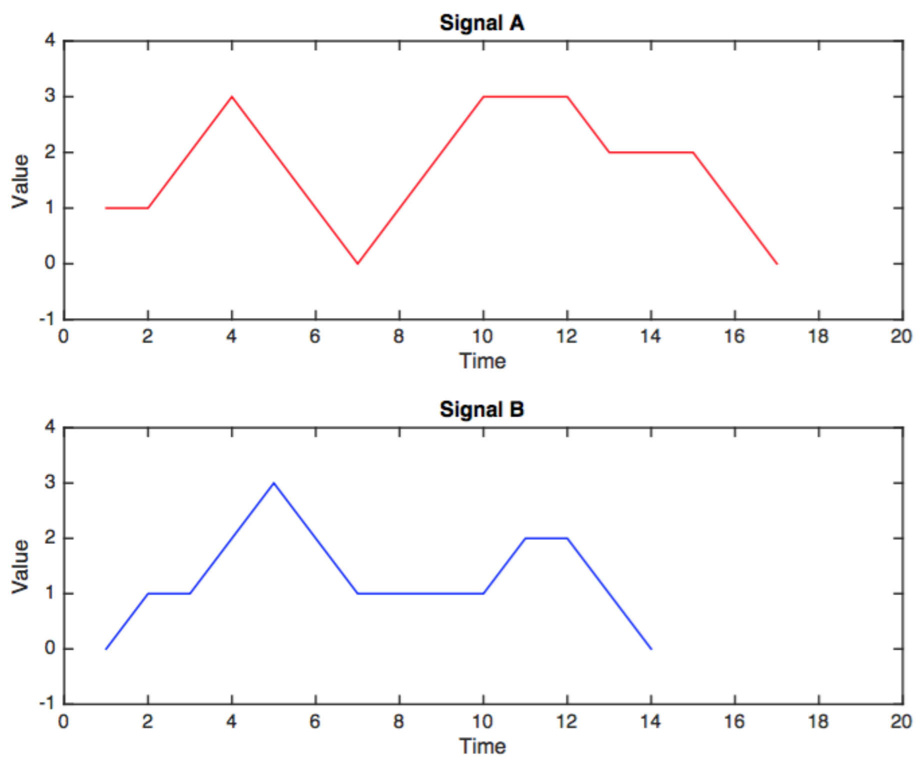

Figure 4.2: Examples of signals A and B of different lengths and varying time and speed.

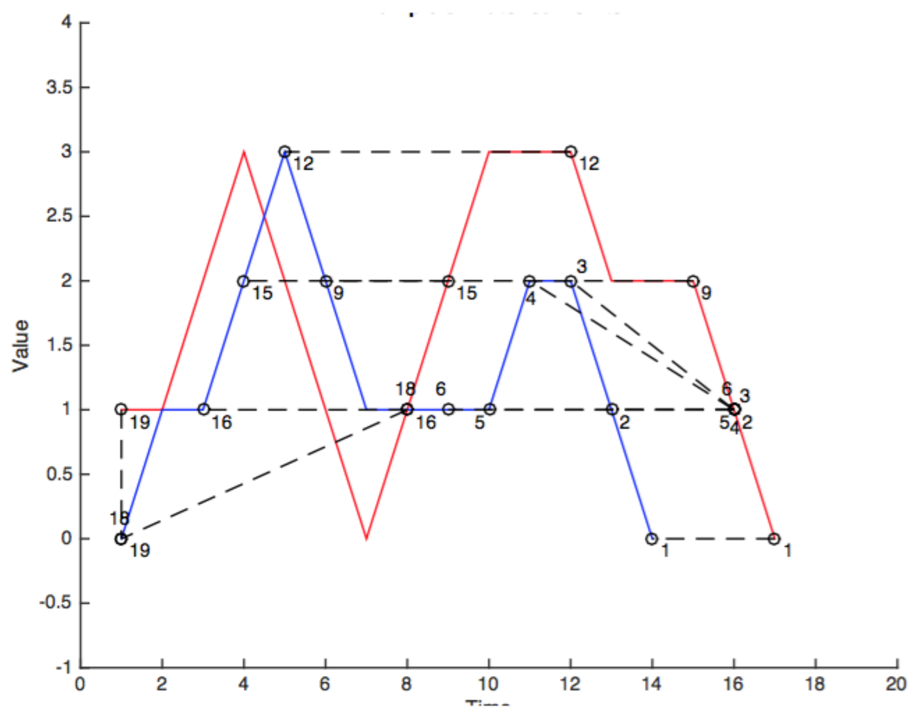

Figure 4.3: Signals A and B matched using DTW. Only some of the corresponding points matched between the two signals are shown for readability purposes.

Pearsons Correlation

For each task, we use Pearsons Correlation to calculate calculate synchrony between the 2x2 speech response signals of each partner, corresponding to mean of pitch, mean of loudness, standard deviation of pitch, and standard deviation of loudness. Mean across all tasks gives us overall synchrony between dyad for each speech response signal.

Pearsons Correlation is calculated using: $\frac{\sum_{i=1}^{n}(a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^{n}(a_i - \bar{a})^2}\sqrt{\sum_{i=1}^{n}(b_i - \bar{b})^2}}$

Where “a” and “b” correspond to the respective response signals of the two dyadic partners.

Chapter 5

Results

5.1 Study 1

Table 5.1 presents zero-order correlations among all variables: group compositional variables (number of females in the dyad, age diversity, ethnic diversity, and social perceptiveness), collective intelligence (CI), group satisfaction, and physiological synchrony variables (facial expressions, electrodermal activity (pEDA), heart rate, and speech).

5.1.1 Physiological Synchrony and Collective Intelligence

In order to test hypothesis 1, we examined the relationship between physiological synchrony and CI, and physiological synchrony and group satisfaction. We found a significant, positive relationship between synchrony in facial expressions and CI ($r = .30$, $p = .01$), and a significant positive relationship between synchrony in the standard deviation of loudness in speech and CI ($r = .294$, $p = .04$). By contrast, CI was neither significantly correlated with synchrony in pEDA nor with synchrony in heart rate.

Interestingly, we found a different pattern with respect to group satisfaction. Group satisfaction was positively associated with high levels of pEDA synchrony ($r = .33$, $p = .04$). In other

Table 5.1: STUDY 1: Correlation coefficient for synchrony strength in facial expressions, electrodermal activity, heart rate, and speech with collective intelligence, group satisfaction and RME.

	Mean	**S.D.**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
1. Age distance	9.24	10.044	1													
2. Number of females	1.01	0.813	0.214	1												
3. Ethnic similarity	1.42	0.498	-0.019	-.271*	1											
4. Education distance	1.51	1.135	.342**	0.083	0.131	1										
5. RME	26.35	2.943	-0.198	-0.24	-0.085	0.05	1									
6. CI	-0.016	0.63	-0.223	0.101	-.295*	-0.041	.390**	1								
7. Group satisfaction	4.12	0.425	0.115	-0.05	.266*	-0.006	0.02	-0.14	1							
8. DTW_EDA	3674.8	411.553	0.114	0.211	-0.085	0.071	0.077	-0.064	.325*	1						
9. DTW_HR	15571	5935.2	0.207	0.047	-0.225	0.072	-0.053	-0.14	-0.002	.340*	1					
10. DTW_FACE	7772.1	3931.1	-0.255	0.168	-0.068	0.152	0.222	.304*	-0.093	-0.057	-0.2	1				
11. Pears_Mean_Pitch	0.36	0.299	-.272	.154	-.056	-.164	-.018	.189	.248	-.275	.124	.116	1			
12. Pears_Mean_Loudness	0.27	0.287	-.070	.258	-.014	-.230	-.134	.216	.134	-.185	-.023	-.015	0.602**	1		
13. Pears_Std_Pitch	0.24	0.428	-.227	.340*	-.067	-.183	.336*	.273	-.048	-.146	.182	.291*	.453**	.288*	1	
14. Pears_Std_Loudness	0.44	0.247	-.145	.253	.168	-.255	-.070	.294*	.150	-.157	.127	.008	.569**	.857**	.379**	1

∗∗ p <.01, ∗ p <.05, N = 60 (in subjective measures), N = 52 (in facial expressions) , N = 53 (in EDA and HR)
Number of females is coded as 0 (male only), 1 (mixed sex), 2 (female only), ethnic similarity coded as 1 (dissimilar), 2 (similar). CI = Collect intelligence (z-score), RME= Read the Mind through the Eyes, DTW = Dynamic Time Warping (measure of similarity or synchrony in dyads), EDA= electrodermal activity, HR= heart rate, FACE = facial expressions, Pears = Pearsons Correlation, Mean_Pitch = mean of pitch in speech, Mean_Loudness = mean of loudness in speech, Std_Pitch = standard deviation of pitch in speech, Std_Loudness = standard deviation of loudness in speech.

words, when both members of a dyad exhibited similar levels of electrodermal activity, they later reported a higher level of satisfaction with the interaction with their partner. However, group satisfaction had no significant relationship with synchrony in facial expressions and heart rate. Finally, pEDA and heart rate were positively related to one another ($r = .34$, $p = .02$). Dyads that had higher synchrony in their heart rate also had higher synchrony in electrodermal activity. Only synchrony in electrodermal activity was significantly associated with group satisfaction.

There was no relationship between CI and group satisfaction, consistent with prior studies [2, 3, 4, 65]. Interestingly, neither synchrony in facial expressions (correlated with CI) nor synchrony in standard deviation of loudness (also correlated with CI) were not significantly correlated with synchrony in pEDA (correlated with group satisfaction), reinforcing the speculation that there might be two separate paths along which collaborative relationships in groups develop. Further, even though both synchrony in facial expressions and synchrony in standard deviation of loudness correlated with CI, they did not correlate with each other. This indicates that there may be more than one mechanisms underlying CI.

5.1.2 Social Perceptiveness

In Hypothesis 2, we predicted that a group's average social perceptiveness, measured by the RME test would positively affect (a) CI and (b) group satisfaction, via effects on physiological synchrony. To test this hypothesis, we ran a series of mediation models using PROCESS macro via SPSS [86]. We tested mediation by first using social perceptiveness to predict CI and then looked at the change in effect of social perceptiveness when each form of synchrony (facial expression vs. pEDA or heart rate or speech) was also included in the model. We ran a similar series of models to test whether both forms of physiological synchrony mediated the relationship between social perceptiveness and group satisfaction (facial expression vs. pEDA). For all tests, we used kappa squared as an index of indirect effect size [87]. Since kappa squared is a ratio, the direction of each relationship is characterized by the coefficient of the indirect effect.

The results showed that the average social perceptiveness had a positive indirect effect on CI via synchrony in facial expressions (kappa squared = .06; 95% bias-corrected 10000 bootstrap confidence intervals ranged from .0038 to .1976). Average social perceptiveness had a significant, positive direct effect on CI, as well ($p = .04$). This suggests that the effect of average social perceptiveness on CI, which had been repeatedly shown in previous research [2, 3], is in part explained by a physiological mechanism, specifically synchrony in facial expressions. Further, average social perceptiveness did not have a significant effect on CI via synchrony in standard deviation of loudness in speech. This suggests that there is another mechanism underlying CI, related to the effect of the standard deviation of loudness on CI that is not associated with social perceptiveness. Average social perceptiveness had a small but positive indirect effect on group satisfaction via synchrony in pEDA (kappa squared = .03; 95% bias-corrected 10000 bootstrap confidence intervals ranged from .0006 to .1130). On the other hand, average social perceptiveness did not have a direct effect on group satisfaction after controlling for pEDA.

5.1.3 Group Composition

Finally, in Hypothesis 3, we predicted that (a) CI and (b) group satisfaction would be influenced by group composition in terms of sex, age, and ethnicity, and that the effects would be mediated by physiological synchrony. Results revealed that the number of females in the dyad had no indirect effect on CI via synchrony in facial expressions or synchrony in standard deviation of loudness in speech. The number of females in the dyad did not have a direct effect on CI controlling for synchrony in facial expressions or synchrony in standard deviation of loudness . On group satisfaction, the number of females did not have a direct effect on group satisfaction, but had a positive indirect effect on group satisfaction via synchrony in pEDA (kappa squared = .03; 95% bias-corrected 10000 bootstrap confidence intervals ranged from .0006 to .1294).

Age diversity, measured by distance, had a negative indirect effect on CI through reduced synchrony in facial expressions (kappa squared = .07; 95% bias-corrected 10000 bootstrap confidence intervals ranged from .0065 to .1726). That is, dyads whose members greatly differ in age were less likely to synchronize in facial expressions, thus having lower CI. However, age distance did not have a significant effect on CI via synchrony in standard deviation of loudness in speech. Age distance did not have a significant direct effect on CI, either. With respect to group satisfaction, age diversity had a positive indirect effect via synchrony in pEDA (kappa squared = .04; 95% bias-corrected 10000 bootstrap confidence intervals ranged from .0027 to .1457). The direct effect of age diversity on group satisfaction controlling for synchrony in pEDA was not significant.

Ethnic diversity had a positive indirect effect on CI through increased synchrony in facial expressions (kappa squared = .03; 95% bias-corrected 10000 bootstrap confidence intervals ranged from .0003 to .0940). Ethnic diversity had a positive direct effect on CI as well, controlling for synchrony in facial expressions. Thus, having ethnically dissimilar members in the dyad seems to increase CI with and without synchrony in facial expressions. No such relationships were observed with synchrony in standard deviation of loudness in speech. For group satisfaction, ethnic diversity had a negative direct effect, controlling for synchrony in pEDA ($p = .01$). However, it had a positive indirect effect on group satisfaction via synchrony in pEDA (kappa squared = .03; 95% bias-corrected 10000 bootstrap confidence intervals ranged from .0008 to .1358). It suggests that ethnically dissimilar dyads are less likely to report satisfaction with their partner; however, ethnically dissimilar dyads who synchronized in electrodermal activity reported higher levels of group satisfaction.

5.2 Study 2

Table 5.2 presents zero-order correlations among all variables: group compositional variables (number of females in the dyad, age diversity, ethnic diversity, and social perceptiveness), collective intelligence (CI), group satisfaction, and physiological synchrony variables (facial expressions, electrodermal activity (pEDA), heart rate, and speech).

Table 5.2: STUDY 2: Correlation coefficient for synchrony strength in facial expressions, electrodermal activity, heart rate, and speech with collective intelligence, group satisfaction and RME.

	Mean	**S.D**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
1. Age distance	5.35	9.159	1													
2. Number of females	1.02	0.792	-.052	1												
3. Ethnic similarity	0.34	0.478	.040	.010	1											
4. Education distance	1.10	1.111	.343**	-.003	.042	1										
5. RME	26.592	2.8611	-.067	.271*	-.244*	-.160	1									
6. CI	0	0.566	-.231*	.120	-.133	-.274*	.321**	1								
7. Group satisfaction	4.227	0.476	.048	-.068	-.230*	-.193	.023	.189	1							
8. DTW_EDA	3865.77	438.978	.209	-.315**	.153	.100	-.151	-.125	-.118	1						
9. DTW_HR	18072.19	6986.151	-.084	-.166	-.051	-.101	.213*	.174	.159	.253*	1					
10. DTW_FACE	7804.69	473.401	-.046	.021	.323*	-.023	-.013	-.130	-.270*	.215	.051	1				
11. Mean_Pitch	0.81	0.161	-.009	.140	-.237	-.220	.106	.242	.190	-.164	.074	-.299*	1			
12. Mean_Loudness	0.70	0.198	.089	.150	-.104	-.130	.103	.334*	-.085	-.123	.020	-.078	.651**	1		
13. Std_Pitch	0.95	0.229	.090	.198	-.151	-.118	.138	.199	.196	-.227	.030	-.362**	.583**	.316*	1	
14. Std_Loudness	0.69	0.177	.076	.204	.023	-.138	-.054	.281*	-.158	.008	.072	-.080	.504**	.852**	.291*	1

$**$ p $<$.01, $*$ p $<$.05, N = 60 (in subjective measures), N = 52 (in facial expressions) , N = 53 (in EDA and HR)

Number of females is coded as 0 (male only), 1 (mixed sex), 2 (female only), ethnic similarity coded as 1 (dissimilar), 2 (similar). CI = Collect intelligence (z-score), RME= Read the Mind through the Eyes, DTW = Dynamic Time Warping (measure of similarity or synchrony in dyads), EDA= electrodermal activity, HR= heart rate, FACE = facial expressions, Pears = Pearsons Correlation, Mean_Pitch = mean of pitch in speech, Mean_Loudness = mean of loudness in speech, Std_Pitch = standard deviation of pitch in speech, Std_Loudness = standard deviation of loudness in speech.

5.2.1 Physiological Synchrony and Collective Intelligence

In order to test hypothesis 1, we examined the relationship between physiological synchrony and CI, and physiological synchrony and group satisfaction. We found a significant positive relationship between synchrony in standard deviation of loudness in speech ($r = .281$, $p = .04$) and synchrony in mean of loudness in speech ($r = .334$, $p = .01$). However, unlike study 1, CI did not correlate with synchrony in facial expressions, and like study 1, CI did not correlate with synchrony in pEDA or with synchrony in heart rate. This suggests that, when participants can no longer see each other, the mechanism that effects CI through synchrony in facial expressions seizes to exist, but the mechanism that effects CI through synchrony in standard deviation of loudness in speech still exists.

Also unlike study 1, no correlation was found between synchrony in pEDA and group satisfaction. We however, found a significant negative correlation between synchrony in facial expressions and group satisfaction ($r = -.270$, $p = .04$). This finding does not make sense, because it is unclear why synchrony in facial expressions would have a negative effect on group satisfaction.

5.2.2 Social Perceptiveness

In Hypothesis 2, we predicted that a group's average social perceptiveness, measured by the RME test would positively affect (a) CI and (b) group satisfaction, via effects on physiological synchrony. To test this hypothesis, we ran a series of mediation models using PROCESS macro via SPSS [86]. We tested mediation by first using social perceptiveness to predict CI and then looked at the change in effect of social perceptiveness when each form of synchrony (facial expression vs. pEDA or heart rate or speech) was also included in the model. We ran a similar series of models to test whether both forms of physiological synchrony mediated the relationship between social perceptiveness and group satisfaction (facial expression vs. pEDA). For all tests, we used kappa squared as an index of indirect effect size [87]. Since kappa squared is a ratio, the direction of each relationship is characterized by the coefficient of the indirect effect.

The results showed that unlike study 1, the average social perceptiveness did not have a significant indirect effect on CI via synchrony in facial expressions. But like study 1,average social perceptiveness had a significant, positive direct effect on CI, as well ($p = .04$). Further, average social perceptiveness did not have a significant effect on CI via synchrony in standard deviation of loudness in speech or synchrony in mean of loudness in speech. Average social perceptiveness had a significant direct effect on CI after controlling for synchrony in standard deviation of loudness in speech ($p = .008$) or controlling for synchrony in mean of loudness in speech ($p = .02$). Average social perceptiveness did not have a significant indirect effect on group satisfaction via synchrony in pEDA or synchrony in facial expressions, or a significant direct effect on group satisfaction after controlling for synchrony in pEDA or synchrony in facial expressions.

5.2.3 Group Composition

Finally, in Hypothesis 3, we predicted that (a) CI and (b) group satisfaction would be influenced by group composition in terms of sex, age, and ethnicity, and that the effects would be mediated by physiological synchrony. Number of females in the dyad had no indirect effect on CI via synchrony in facial expressions or synchrony in standard deviation or mean of loudness in speech. The number of females in the dyad also had no direct effect on CI after controlling for synchrony in facial expressions or synchrony in standard deviation or mean of loudness in speech. On group satisfaction, the number of females in the dyad did not have a indirect effect on group satisfaction via synchrony in pEDA or synchrony in facial expressions. The number of females in the dyad also had no direct effect on group satisfaction after controlling for synchrony in pEDA or synchrony in facial expressions.

Age diversity, measured by distance, had no indirect effect on CI via synchrony in facial expressions or synchrony in standard deviation or mean of loudness in speech. The age distance in the dyad also had no direct effect on CI after controlling for synchrony in facial expressions or synchrony in standard deviation or mean of loudness in speech. On group satisfaction, the age distance in the dyad did not have a indirect effect on group satisfaction via synchrony in pEDA or synchrony in facial expressions. The age distance in the dyad also had no direct effect on group satisfaction after controlling for synchrony in pEDA or synchrony in facial expressions.

Ethnic diversity in the dyad had no indirect effect on CI via synchrony in facial expressions or synchrony in standard deviation or mean of loudness in speech. Ethnic Diversity in the dyad also had no direct effect on CI after controlling for synchrony in facial expressions or synchrony in standard deviation or mean of loudness in speech. On group satisfaction, the ethnic diversity in the dyad did not have a indirect effect on group satisfaction via synchrony in pEDA or synchrony in facial expressions. Ethnic Diversity in the dyad also had no direct effect on group satisfaction after controlling for synchrony in pEDA or synchrony in facial expressions.

Chapter 6

Discussion

In this study, we examined whether collective intelligence of human groups is associated with the deep structures of collaboration manifested by synchronization of physiological responses. In addition, building upon the established teams and organizational literatures on group interpersonal processes, specifically group satisfaction, we further tested whether the interpersonal aspect of group processes is similarly governed by synchrony in physiological signals. Finally, we hypothesized that synchrony in physiological signals is one key mechanism for previously studied effects of group composition on collective intelligence and group satisfaction. All of this was tested in a computer-mediated communication environment, one in which physiological synchrony has not been extensively examined.

To test these hypotheses, we conducted an experiment where 120 dyads interacted in virtual collaborative environments while being measured for physiological signals such as electrodermal activity (EDA), heart rate, and facial expressions. In study 1 (60 dyads), dyadic partners were able to communicate through video and audio using a video conferencing tool. However, in study 2, dyadic partners only communicated through audio as their video window was hidden.

6.1 Study 1

In study 1, we found that collective intelligence was positively correlated with synchrony in facial expression and with synchrony in the standard deviation of loudness in speech, but not with EDA nor heart rate. On the other hand, group satisfaction, which captures the quality of group interaction and relationships, was positively correlated with synchrony in EDA, but not with facial expression synchrony or speech-related synchrony. These findings suggest that the physiological structures of group collaboration are not monolithic, but perhaps comprised of different building blocks. Specifically, similarity in group members' facial expressions and variation in loudness is a symptom of a higher level of attentiveness to other members which may facilitate coordination and collective effort. In contrast, similarity in EDA indicates shared arousal, capturing how the group members feel during the interaction, and thus has effects on members' level of satisfaction. This finding is also consistent with Mnster et al.'s, [23] work that showed physiological synchrony in EDA is related to emotional aspects of the group dynamics but not to task performance.

6.1.1 Social Perceptiveness

We observed that groups with high social perceptiveness on average were more collectively intelligent, consistent with previous research [2, 3] and interestingly, this effect was mediated by synchrony in facial expressions. That is, a group's average social perceptiveness increases collective intelligence because members in such group synchronize their facial expressions more, facilitating coordination. It is important to note that the direct effect of social perceptiveness on collective intelligence, controlling for the mediator, was also significant. A question worth further exploring is what other mechanism explains the positive effect of social perceptiveness on collective intelligence. It is likely that other communication and coordination behaviors manifest in groups with higher levels of social perceptiveness [2, 64], which further enhance collective intelligence.

6.1.2 Group Composition

Physiological synchrony also appeared to be an underlying mechanism for the effect of group composition on collective intelligence and group satisfaction, but to varying degrees. All of the effects on CI operated via their impact on facial expression synchrony, which can be viewed as a gauge of shared attention and concentration [88]. Ethnic diversity indirectly increased collective intelligence via increased synchrony in facial expressions. Ethnicity is a surface-level characteristic, and diversity in such characteristic has been shown to prime heightened levels of mutual attentiveness among group members [89], reinforcing our interpretation of facial expression synchrony as an index of shared attention. Age diversity, on the other hand, negatively affected collective intelligence because the more dissimilar members are in terms of age, the less synchronous in facial expressions, reinforcing the negative effect of age heterogeneity on group performance [90]. Taken together, these findings suggest that ethnic diversity perhaps contributed to a heightened level of attention among members, a favorable condition for collective intelligence; however, age diversity can evoke a sense of hierarchy between members, obstructing the development of collective intelligence [65, 91].

Group satisfaction was also indirectly affected by the group composition variables we examined. Similar to collective intelligence, sex and ethnic diversity had positive indirect effects on group satisfaction, but via synchrony in EDA, a signal of shared arousal in the group. However, age diversity, measured by distance, also had a positive indirect effect on group satisfaction via synchrony in EDA. That is, groups composed of members with a greater age gap demonstrated a high level of synchrony in EDA, which was positively associated with group satisfaction. It is possible that a greater age gap between members created a non-competitive, caring environment for group members; however, without the greater attentiveness engendered via facial expression synchrony, this did not translate into the group's collective intelligence. Finally, social perceptiveness indirectly increased group satisfaction via synchrony in EDA, albeit to a very small degree.

6.2 Study 2

We carried out study 2 after study 1, to understand how the relationships between synchrony and facial expressions and CI, and synchrony in standard deviation of loudness and CI, might change when dyadic partners are unable to see each other. We found that the relationship between synchrony in facial expressions and CI no longer exists in study 2, however synchrony in standard deviation of loudness continues to have a significant positive correlation with CI. Further, unlike study 1, synchrony in pEDA also does not relate to group satisfaction in study 2.

The positive effect of social perceptiveness on CI was mediated by synchrony in facial expressions in study 1, however no form of physiological synchrony mediates the relationship between social perceptiveness and CI in study 2. So while, synchrony in facial expressions can partly explain the relationship between social perceptiveness and CI in study 1, further research is needed to explain the relationship between social perceptiveness and CI in study 2.

In both studies, synchrony in standard deviation of loudness in speech had a positive effect on CI. This effect was independent of social perceptiveness, as synchrony in standard deviation of loudness in speech does not mediate the relationship between social perceptiveness and CI, in either studies.

Qualitative analysis revealed that synchrony in standard deviation of loudness in speech is often higher for dyads that engage in similar thought patterns. For example, they may think out loud together, or may both think without speaking. This suggests that like facial expressions, variation (measured using standard deviation) of loudness in speech is also an indicator of shared experience.

Chapter 7

Conclusion

7.1 Implications

There are a number of important implications of this study and interesting opportunities for future work on CI and CSCW associated with the additional insight that the physiological mechanisms provide.

Different Processes May Drive Cohesion vs Performance

First, it appears that the dissociation between CI and group satisfaction repeatedly observed in prior studies [2, 3, 65] has a parallel in physiological signals. Here facial expression synchrony (related to CI only), speech-related synchrony (related to CI only) and EDA synchrony (relates to group satisfaction only), were also dissociated from each other but related to CI and group satisfaction, respectively. This dissociation harkens back to debates of a few decades ago regarding the cohesion-performance connection or lack thereof [92, 93] and suggests additional mechanisms to gain further insight into that relationship, namely group composition variables and their differential effects of group member physiological response and synchrony. This distinction is important if we want to develop more fine-grained, sensory-driven predictive models of group performance to drive intelligent environments or feedback systems.

Diversity, Social Perceptiveness and Collective Intelligence

While scholars have documented the benefits of diversity for cognition [94, 95] the implications for physiological measures and their independent impact on collective intelligence and group member relationships has only begun to be explored [36, 38]. Our study suggests that diverse groups may engage in fundamentally different interpersonal processes as a function of heightened social perceptiveness. Our study provides only an initial glimpse at answers to questions about diversity and group process, but suggests a host of other relationships to explore. Future work should examine the effects of interventions to regulate physiology as a means of improving working relationships in diverse groups. In addition, it may be possible to improve physiological synchrony and ultimately performance in diverse and homogeneous groups through social perceptiveness training interventions. It would also be interesting and fruitful to look at the relationship of facial synchrony and more detailed process behaviors in diverse groups to unpack exactly how and why it may support CI.

Shared attention, facial synchrony and turn-taking

The positive relationship between collective intelligence and synchrony in facial expressions in our study confirms the importance of visual, nonverbal cues about team members in facilitating collective intelligence, complementing previous work in CSCW. Theory and research in CSCW has long noted the value of nonverbal cues for supporting language understanding and coordinating turn-taking during remote collaboration (e.g. [96]). Our results suggest facial expression synchrony may be a critical aspect of collective intelligence, and that systems may be able to enhance team performance by making faces more visible and salient. Designers may consider, for example, screen layouts for video conferencing systems that could help group members attend to and synchronize other members' facial expressions easily.

Enhancing collective intelligence

One potential application for our results is improving collective intelligence via technological or behavioral intervention. This application raises a host of additional research questions about the relationship between synchrony and collective intelligence and intervention design.

7.1.1 Is synchrony controllable?

Our work raises the question of whether synchrony is a controllable or unconscious process. Can individuals consciously increase facial or physiological synchrony? And by doing so increase collective intelligence or group satisfaction? Synchrony may depend on individual differences in perspective taking or empathic abilities [61]. We found potential evidence for this in group level variations in synchrony.

7.1.2 Intervention design

Next there is a question of how to design interventions for increasing synchrony. Should such interventions be direct or unobtrusive?

Indirect environmental interventions

Unobtrusive interventions could increase the level of physiological synchrony among members through shared activities or other manipulations outside of participants' conscious awareness. Intelligent meeting rooms could attempt to unobtrusively intervene or modify the collaboration environment to enhance group performance in response to sensed levels of synchrony early in a group's life cycle to improve the quality of interaction. For example, an intelligent system that sensed pEDA asynchrony across members could increase room temperature or ambient noise in one location until that member was ?in synch' in terms of their arousal level. It remains to be seen, however, whether synchrony evoked via this kind of unobtrusive background intervention would have a similar positive effect on group satisfaction.

Visual feedback

Alternatively, technological interventions could be more obtrusive or directive, and presented at the individual versus the group level. For example, video conferencing systems could integrate a facial synchronometer showing the level of similarity across participants' facial expression or provide commands to individual group members for increasing synchrony (e.g. "smile more to match your partner"). This kind of on-screen instruction introduces other tradeoffs like mental and visual attention demand. However, previous work in CSCW has successfully applied in-situ visual feedback on group processes such as floor sharing behavior to increase equality of participation and ultimately team performance [13, 14]. Future work should explore the use of directive feedback versus real time or post-hoc visualization of individual and group responses to see whether it is possible to enhance synchrony.

Training

Training is another more direct way to potentially enhance teams synchrony. Social perceptiveness training has been used in other settings effectively to increase individual levels of attention to social cues [97]. If facial synchrony is controllable it may be possible to train team members to better attend to facial expressions and synchronize their own in response. Alternatively, participants can be primed with a pro-social task [98] before they begin collaborating.

7.2 Limitations

As with any study that represents an initial attempt at applying a new methodology to a novel context, our study has some limitations that readers must bear in mind. First, heart rate synchrony turned out to be a less sensitive measure of synchrony of arousal for the present study. A potential reason for this is that the tasks in TCI did not induce strong enough arousal, compared to other studies that show such an effect [10]. On the other hand, existing studies in married couples and infant-mother pairings show heart rate responses to rather subtle changes in activity [88] and thus it is hard to discern the reasons for a lack of relationship observed here. Alternatively, wearable wristbands (like E4) capture HR as averaged over a minute instead of instantaneous HR. ECG sensors are more accurate but also more intrusive for use in lab experiments as these require straps placed close to heart. Future research could explore additional techniques for measuring heart rate, and/or tasks to induce larger changes during the interaction to see if greater effects associated with heart rate synchrony might be observed.

A second potential limitation here is associated with our use of dyads versus larger groups. CI has been examined in dyads in the past [2] and shown to be predicted by the same variables associated with it in larger groups. However, it is unknown in this context whether the more complex form of synchrony that would need to develop in a larger group would (a) develop in larger groups collaborating online, and (b) show the same effects on CI and group satisfaction. Thus future research will need to explore replication of these effects in larger groups.

Finally, one major limitation of this work that could be resolved by additional analysis or future work, is that the reason for the lack of relationship between synchrony in pEDA and group satisfaction in study 2, and the negative effect of synchrony in facial expressions and group satisfaction in study 2, largely remains unclear.

7.3 Future Work

In conclusion, this study represents an initial foray into exploring the role of deep collaboration mechanisms, represented by physiological synchrony, in the development of CI and group satisfaction in computer-mediated teams. We find fairly strong evidence of the role of facial expression synchrony and EDA synchrony in explaining the effects of group diversity and social perceptiveness.

We believe this study is only the beginning of a series of experiments that use sensors to capture the deep structures of collective intelligence. Future research can expand our findings by adopting other sensors such as facial electromyography (EMG), cortisol levels, eye-tracking, motion sensors (for body language), and EEG. Results will further our understanding of the mechanisms underlying group performance and cohesion.

As collaboration becomes ever more dispersed and technology-mediated, we see that these fairly basic, primitive human responses to one another remain. The real challenge will be to develop new and innovative interventions that harness this newly acquired knowledge to enable teams to reach new heights in collective intelligence and satisfaction with their relationships.

References

[1] P. Chikersal, M. Tomprou, Y. J. Kim, A. W. Woolley, and L. Dabbish, "Deep structures of collaboration: Physiological correlates of collective intelligence and group satisfaction," in *Proceedings of the 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2017)*, 2017. (document)

[2] A. W. Woolley, C. F. Chabris, A. Pentland, N. Hashmi, and T. W. Malone, "Evidence for a collective intelligence factor in the performance of human groups," *science*, vol. 330, no. 6004, pp. 686–688, 2010. 1, 2.2, 3.1, 3.2, 3.3, 4.1, 4.1.2, 4.2.3, 5.1.1, 5.1.2, 6.1.1, 7.1, 7.2

[3] D. Engel, A. W. Woolley, I. Aggarwal, C. F. Chabris, M. Takahashi, K. Nemoto, C. Kaiser, Y. J. Kim, and T. W. Malone, "Collective intelligence in computer-mediated collaboration emerges in different contexts and cultures," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2015, pp. 3769–3778. 1, 3.1, 4.1.2, 4.2.3, 5.1.1, 5.1.2, 6.1.1, 7.1

[4] D. Engel, A. W. Woolley, L. X. Jing, C. F. Chabris, and T. W. Malone, "Reading the mind in the eyes or reading between the lines? theory of mind predicts collective intelligence equally well online and face-to-face," *PloS one*, vol. 9, no. 12, p. e115212, 2014. 1, 2.2, 3.1, 3.2, 3.3, 4.1.2, 4.2.3, 5.1.1

[5] Y. J. Kim, D. Engel, A. W. Woolley, J. Lin, N. McArthur, and T. W. Malone, "Work together, play smart: Collective intelligence in league of legends teams," *Proceedings of Collective Intelligence 2015*, 2015. 1

[6] A. W. Woolley, I. Aggarwal, and T. W. Malone, "Collective intelligence and group performance," *Current Directions in Psychological Science*, vol. 24, no. 6, pp. 420–424, 2015. 1, 3.2, 3.3

[7] I. Aggarwal, A. W. Woolley, C. F. Chabris, and T. W. Malone, "Cognitive diversity, collective intelligence, and learning in teams," *Proceedings of Collective Intelligence*, 2015. 1, 3.1, 3.3

[8] N. Meslec, I. Aggarwal, and P. L. Curseu, "The insensitive ruins it all: Compositional and compilational influences of social sensitivity on collective intelligence in groups," *Frontiers in psychology*, vol. 7, 2016. 1, 3.2, 3.3

[9] Z. Yu and Y. Nakamura, "Smart meeting systems: A survey of state-of-the-art and open issues," *ACM Computing Surveys (CSUR)*, vol. 42, no. 2, p. 8, 2010. 2.1

[10] I. Mikic, K. Huang, and M. Trivedi, "Activity monitoring and summarization for an intelligent meeting room," in *Human Motion, 2000. Proceedings. Workshop on*. IEEE, 2000, pp. 107–112. 2.1, 7.2

[11] A. Bränzel, C. Holz, D. Hoffmann, D. Schmidt, M. Knaust, P. Lühne, R. Meusel, S. Richter, and P. Baudisch, "Gravityspace: tracking users and their poses in a smart room using a pressure-sensing floor," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2013, pp. 725–734. 2.1

[12] T. Bergstrom and K. Karahalios, "Conversation clock: Visualizing audio patterns in co-located groups," in *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on*. IEEE, 2007, pp. 78–78. 2.1

[13] J. M. DiMicco, K. J. Hollenbach, A. Pandolfo, and W. Bender, "The impact of increased awareness while face-to-face," *Human–Computer Interaction*, vol. 22, no. 1-2, pp. 47–96, 2007. 2.1, 7.1.2

[14] T. Kim, A. Chang, L. Holland, and A. S. Pentland, "Meeting mediator: enhancing group

collaboration with sociometric feedback," in *CHI'08 extended abstracts on Human factors in computing systems*. ACM, 2008, pp. 3183–3188. 2.1, 7.1.2

[15] G. Leshed, D. Perez, J. T. Hancock, D. Cosley, J. Birnholtz, S. Lee, P. L. McLeod, and G. Gay, "Visualizing real-time language-based feedback on teamwork behavior in computer-mediated groups," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2009, pp. 537–546. 2.1

[16] J. Sturm, O. H.-v. Herwijnen, A. Eyck, and J. Terken, "Influencing social dynamics in meetings through a peripheral display," in *Proceedings of the 9th international conference on Multimodal interfaces*. ACM, 2007, pp. 263–270. 2.1

[17] R. L. Mandryk and K. M. Inkpen, "Physiological indicators for the evaluation of co-located collaborative play," in *Proceedings of the 2004 ACM conference on Computer supported cooperative work*. ACM, 2004, pp. 102–111. 2.1

[18] C. S. S. Tan, J. Schöning, K. Luyten, and K. Coninx, "Investigating the effects of using biofeedback as visual stress indicator during video-mediated collaboration," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2014, pp. 71–80. 2.1

[19] H. Gardner, *Intelligence reframed: Multiple intelligences for the 21st century*. Basic books, 1999. 2.2

[20] A. Williams Woolley and I. Aggarwal, *Collective intelligence and group learning*. Linda Argote and J. M Levine (eds.). Oxford University Press, London, UK. 2.2

[21] E. Glikson, R. Harush, Y. J. Kim, A. W. Woolley, and M. Erez, "Psychological safety and collective intelligence in multicultural globally dispersed teams," in *the 2016 INGRoup Conference. Helsinki, Finland*, 2016. 2.2, 3.1

[22] M. L. Maznevski and K. M. Chudoba, "Bridging space over time: Global virtual team dynamics and effectiveness," *Organization science*, vol. 11, no. 5, pp. 473–492, 2000. 2.2

[23] D. Mønster, D. D. Håkonsson, J. K. Eskildsen, and S. Wallot, "Physiological evidence of interpersonal dynamics in a cooperative production task," *Physiology & behavior*, vol. 156, pp. 24–34, 2016. 2.2, 2.3, 3.1, 6.1

[24] D. M. Gordon, "Collective wisdom of ants," *Scientific American*, vol. 314, no. 2, pp. 44–47, 2016. 2.2

[25] A. Berdahl, C. J. Torney, C. C. Ioannou, J. J. Faria, and I. D. Couzin, "Emergent sensing of complex environments by mobile animal groups," *Science*, vol. 339, no. 6119, pp. 574–576, 2013. 2.2

[26] J. Hernandez, I. Riobo, A. Rozga, G. D. Abowd, and R. W. Picard, "Using electrodermal activity to recognize ease of engagement in children during social interactions," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2014, pp. 307–317. 2.2, 2.3

[27] D. N. McIntosh, "Spontaneous facial mimicry, liking and emotional contagion," *Polish Psychological Bulletin*, vol. 37, no. 1, p. 31, 2006. 2.2, 2.3, 3.1

[28] P. Mitkidis, J. J. McGraw, A. Roepstorff, and S. Wallot, "Building trust: Heart rate synchrony and arousal during joint action increased by public goods game," *Physiology & behavior*, vol. 149, pp. 101–106, 2015. 2.2, 2.3, 3.1

[29] S. S. Wiltermuth and C. Heath, "Synchrony and cooperation," *Psychological science*, vol. 20, no. 1, pp. 1–5, 2009. 2.3

[30] M. J. Hove and J. L. Risen, "It's all in the timing: Interpersonal synchrony increases affiliation," *Social Cognition*, vol. 27, no. 6, pp. 949–960, 2009. 2.3

[31] F. J. Bernieri, "Coordinated movement and rapport in teacher-student interactions," *Journal of Nonverbal behavior*, vol. 12, no. 2, pp. 120–138, 1988. 2.3

[32] L. K. Miles, L. K. Nind, and C. N. Macrae, "The rhythm of rapport: Interpersonal synchrony and social perception," *Journal of experimental social psychology*, vol. 45, no. 3,

pp. 585–589, 2009. 2.3

[33] P. Valdesolo, J. Ouyang, and D. DeSteno, “The rhythm of joint action: Synchrony promotes cooperative ability,” *Journal of Experimental Social Psychology*, vol. 46, no. 4, pp. 693–695, 2010. 2.3

[34] J. K. Baker, R. M. Fenning, M. A. Howland, B. R. Baucom, J. Moffitt, and S. A. Erath, “Brief report: A pilot study of parent–child biobehavioral synchrony in autism spectrum disorder,” *Journal of autism and developmental disorders*, vol. 45, no. 12, pp. 4140–4146, 2015. 2.3

[35] J. Chatel-Goldman, M. Congedo, C. Jutten, and J.-L. Schwartz, “Touch increases autonomic coupling between romantic partners,” *Frontiers in behavioral neuroscience*, vol. 8, p. 95, 2014. 2.3

[36] M. Akinola, “Measuring the pulse of an organization: Integrating physiological measures into the organizational scholar’s toolbox,” *Research in Organizational Behavior*, vol. 30, pp. 203–223, 2010. 2.3, 4.2.5, 7.1

[37] R. Nikula, “Psychological correlates of nonspecific skin conductance responses,” *Psychophysiology*, vol. 28, no. 1, pp. 86–90, 1991. 2.3, 4.2.5

[38] S. J. Peterson, C. S. Reina, D. A. Waldman, and W. J. Becker, “Using physiological methods to study emotions in organizations,” in *New Ways of Studying Emotions in Organizations*. Emerald Group Publishing Limited, 2015, pp. 1–27. 2.3, 4.2.5, 7.1

[39] M. M. Herrald and J. Tomaka, “Patterns of emotion-specific appraisal, coping, and cardiovascular reactivity during an ongoing emotional episode.” *Journal of personality and social psychology*, vol. 83, no. 2, p. 434, 2002. 2.3

[40] C. Merrifield and J. Danckert, “Characterizing the psychophysiological signature of boredom,” *Experimental brain research*, vol. 232, no. 2, pp. 481–491, 2014. 2.3

[41] C. D. Frith and H. A. Allen, “The skin conductance orienting response as an index of

attention," *Biological psychology*, vol. 17, no. 1, pp. 27–39, 1983. 2.3, 1

[42] R. G. O?Connell, M. A. Bellgrove, P. M. Dockree, A. Lau, M. Fitzgerald, and I. H. Robertson, "Self-alert training: volitional modulation of autonomic arousal improves sustained attention," *Neuropsychologia*, vol. 46, no. 5, pp. 1379–1390, 2008. 2.3, 1

[43] M. G. Coles, "Cardiac and respiratory activity during visual search." *Journal of Experimental Psychology*, vol. 96, no. 2, p. 371, 1972. 2.3, 1

[44] J. Papillo and D. Shapiro, "The cardiovascular system in jt cacioppo and lg tassinary (eds.) principles of psychophysiology: Physical, social, and inferential elements (pp. 456-512)," 1990. 2.3

[45] J. I. Lacey, "Psychophysiological approaches to the evaluation of psychotherapeutic process and outcome." in *Research in Psychotherapy, Apr, 1958, Washington, DC; This conference, financed by a grant (M-2031) from the National Institute of Mental Health, US Public Health Service, was held under the auspices of the Division of Clinical Psychology, American Psychological Association, with planning and programming by an Ad Hoc Committee of the Division of Clinical Psychology; Frank Auld, Jr., Morris B. Parloff, Benjamin Pasamanick, George Saslow, Julius Seeman, and Eli A. Rubinstein, Chairman.* American Psychological Association, 1959. 2.3

[46] G. E. Deane, "Human heart rate responses during experimentally induced anxiety." *Journal of Experimental Psychology*, vol. 61, no. 6, p. 489, 1961. 2.3

[47] E. Duffy, "The psychological significance of the concept of" arousal" or" activation."," *Psychological review*, vol. 64, no. 5, p. 265, 1957. 2.3

[48] D. O. Hebb, "Drives and the cns (conceptual nervous system)." *Psychological review*, vol. 62, no. 4, p. 243, 1955. 2.3

[49] R. B. Malmo, "Activation: A neuropsychological dimension." *Psychological review*, vol. 66, no. 6, p. 367, 1959. 2.3

[50] J. I. Lacey and B. C. Lacey, "Some autonomic-central nervous system interrelationships," *Physiological correlates of emotion*, pp. 205–227, 1970. 2.3, 1

[51] P. Winkielman and J. T. Cacioppo, "Mind at ease puts a smile on the face: psychophysiological evidence that processing facilitation elicits positive affect." *Journal of personality and social psychology*, vol. 81, no. 6, p. 989, 2001. 2.3

[52] R. Reisenzein, M. Studtmann, and G. Horstmann, "Coherence between emotion and facial expression: Evidence from laboratory experiments," *Emotion Review*, vol. 5, no. 1, pp. 16–23, 2013. 2.3, 1, 4.2.5

[53] O.-W. Kwon, K. Chan, J. Hao, and T.-W. Lee, "Emotion recognition by speech signals," in *Eighth European Conference on Speech Communication and Technology*, 2003. 2.3

[54] D. J. France, R. G. Shiavi, S. Silverman, M. Silverman, and M. Wilkes, "Acoustical properties of speech as indicators of depression and suicidal risk," *IEEE transactions on Biomedical Engineering*, vol. 47, no. 7, pp. 829–837, 2000. 2.3

[55] C. Yu, P. M. Aoki, and A. Woodruff, "Detecting user engagement in everyday conversations," *arXiv preprint cs/0410027*, 2004. 2.3

[56] D. Kahneman, B. Tursky, D. Shapiro, and A. Crider, "Pupillary, heart rate, and skin resistance changes during a mental task." *Journal of experimental psychology*, vol. 79, no. 1p1, p. 164, 1969. 1

[57] G. M. Wittenbaum, G. Stasser, and C. J. Merry, "Tacit coordination in anticipation of small group task completion," *Journal of Experimental Social Psychology*, vol. 32, no. 2, pp. 129–152, 1996. 3.1

[58] A. W. Woolley and I. Aggarwal, "The mind and the heart of the group: Collective intelligence and relationship quality in task performing teams." 3.1

[59] N. Ambady and R. Rosenthal, "Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis." 1992. 3.1

[60] C. De Looze, C. Oertel, S. Rauzy, and N. Campbell, "Measuring dynamics of mimicry by means of prosodic cues in conversational speech," in *International Conference on Phonetic Sciences (ICPhS). Hong Kong*, 2011, pp. 1294–1297. 3.1

[61] Z. E. Imel, J. S. Barco, H. J. Brown, B. R. Baucom, J. S. Baer, J. C. Kircher, and D. C. Atkins, "The association of therapist empathy and synchrony in vocally encoded arousal." *Journal of counseling psychology*, vol. 61, no. 1, p. 146, 2014. 3.1, 7.1.1

[62] J. D. Mayer, R. D. Roberts, and S. G. Barsade, "Human abilities: Emotional intelligence," *Annu. Rev. Psychol.*, vol. 59, pp. 507–536, 2008. 3.2

[63] H. A. Elfenbein, V. Druskat, F. Sala, and G. Mount, "Team emotional intelligence: What it can mean and how it can affect performance," *Linking emotional intelligence and performance at work: Current research evidence with individuals and groups*, pp. 165–184, 2006. 3.2

[64] B. L. Kirkman, B. Rosen, P. E. Tesluk, and C. B. Gibson, "The impact of team empowerment on virtual team performance: The moderating role of face-to-face interaction," *Academy of Management Journal*, vol. 47, no. 2, pp. 175–192, 2004. 3.2, 6.1.1

[65] A. W. Woolley, R. M. Chow, A. T. Mayo, J. W. Chang, and C. Riedl, "Competition and collective intelligence: Do women always make groups smarter?" 2016. 3.3, 5.1.1, 6.1.2, 7.1

[66] H. C. Triandis, L. L. Kurowski, and M. J. Gelfand, "Workplace diversity." 1994. 3.3

[67] P. L. McLeod, S. A. Lobel, and T. H. Cox Jr, "Ethnic diversity and creativity in small groups," *Small group research*, vol. 27, no. 2, pp. 248–264, 1996. 3.3

[68] C. K. De Dreu and L. R. Weingart, "Task versus relationship conflict, team performance, and team member satisfaction: a meta-analysis." 2003. 3.3

[69] K. A. Jehn, G. B. Northcraft, and M. A. Neale, "Why differences make a difference: A field study of diversity, conflict and performance in workgroups," *Administrative science*

quarterly, vol. 44, no. 4, pp. 741–763, 1999. 3.3

[70] H. Tajfel and J. Turner, "The social identity theory of inter group behavior in s worchel & wg austin (eds) psychology of intergroup relations," *Chicago: Nelson*, 1986. 3.3

[71] D. Van Knippenberg and M. C. Schippers, "Work group diversity," *Annu. Rev. Psychol.*, vol. 58, pp. 515–541, 2007. 3.3

[72] R. W. Levenson, L. L. Carstensen, and J. M. Gottman, "Influence of age and gender on affect, physiology, and their interrelations: A study of long-term marriages." *Journal of personality and social psychology*, vol. 67, no. 1, p. 56, 1994. 3.3

[73] J. Blascovich, W. B. Mendes, and M. D. Seery, "Intergroup encounters and threat," *From Prejudice to Intergroup Emotions, Differentiated Reactions to Social Groups, edited by Diane M. Mackie and Eliot R. Smith*, pp. 89–109, 2002. 3.3

[74] D. C. Lau and J. K. Murnighan, "Demographic diversity and faultlines: The compositional dynamics of organizational groups," *Academy of Management Review*, vol. 23, no. 2, pp. 325–340, 1998. 4.1

[75] S. Baron-Cohen, S. Wheelwright, J. Hill, Y. Raste, and I. Plumb, "The ?reading the mind in the eyes? test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism," *Journal of child psychology and psychiatry*, vol. 42, no. 2, pp. 241–251, 2001. 4.2.2

[76] R. Wageman, J. R. Hackman, and E. Lehman, "Team diagnostic survey: Development of an instrument," *The Journal of Applied Behavioral Science*, vol. 41, no. 4, pp. 373–398, 2005. 4.2.4

[77] M. K. Lindell, C. J. Brandt, and D. J. Whitney, "A revised index of interrater agreement for multi-item ratings of a single target," *Applied Psychological Measurement*, vol. 23, no. 2, pp. 127–135, 1999. 4.2.4

[78] T. Baltrušaitis, M. Mahmoud, and P. Robinson, "Cross-dataset learning and person-specific

normalisation for automatic action unit detection," in *Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on*, vol. 6. IEEE, 2015, pp. 1–6. 4.2.6

[79] P. Ekman and W. V. Friesen, *Manual for the facial action coding system.* Consulting Psychologists Press, 1978. 4.2.6

[80] K. L. Schmidt, Z. Ambadar, J. F. Cohn, and L. I. Reed, "Movement differences between deliberate and spontaneous facial expressions: Zygomaticus major action in smiling," *Journal of Nonverbal Behavior*, vol. 30, no. 1, pp. 37–52, 2006. 4.2.6

[81] E. G. Krumhuber and A. S. Manstead, "Can duchenne smiles be feigned? new evidence on felt and false smiles." *Emotion*, vol. 9, no. 6, p. 807, 2009. 4.2.6

[82] M. Benedek and C. Kaernbach, "A continuous measure of phasic electrodermal activity," *Journal of neuroscience methods*, vol. 190, no. 1, pp. 80–91, 2010. 4.2.6

[83] F. Eyben, F. Weninger, F. Gross, and B. Schuller, "Recent developments in opensmile, the munich open-source multimedia feature extractor," in *Proceedings of the 21st ACM international conference on Multimedia.* ACM, 2013, pp. 835–838. 4.2.6

[84] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series." in *KDD workshop*, vol. 10, no. 16. Seattle, WA, 1994, pp. 359–370. 4.2.7

[85] M. P. McAssey, J. Helm, F. Hsieh, D. A. Sbarra, and E. Ferrer, "Methodological advances for detecting physiological synchrony during dyadic interactions," *Methodology*, 2013. 4.2.7

[86] A. F. Hayes, *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* Guilford Press, 2013. 5.1.2, 5.2.2

[87] K. J. Preacher and K. Kelley, "Effect size measures for mediation models: quantitative strategies for communicating indirect effects." *Psychological methods*, vol. 16, no. 2, p. 93, 2011. 5.1.2, 5.2.2

[88] R. Feldman, "Parent–infant synchrony biological foundations and developmental outcomes," *Current directions in psychological science*, vol. 16, no. 6, pp. 340–345, 2007. 6.1.2, 7.2

[89] K. W. Phillips, "The effects of categorically based expectations on minority influence: The importance of congruence," *Personality and Social Psychology Bulletin*, vol. 29, no. 1, pp. 3–13, 2003. 6.1.2

[90] T. A. Timmerman, "Racial diversity, age diversity, interdependence, and team performance," *Small Group Research*, vol. 31, no. 5, pp. 592–606, 2000. 6.1.2

[91] D. A. Harrison and K. J. Klein, "What's the difference? diversity constructs as separation, variety, or disparity in organizations," *Academy of management review*, vol. 32, no. 4, pp. 1199–1228, 2007. 6.1.2

[92] J. E. Mathieu, M. R. Kukenberger, L. D?Innocenzo, and G. Reilly, "Modeling reciprocal team cohesion–performance relationships, as impacted by shared leadership and members? competence." *Journal of Applied Psychology*, vol. 100, no. 3, p. 713, 2015. 7.1

[93] B. Mullen and C. Copper, "The relation between group cohesiveness and performance: An integration." 1994. 7.1

[94] F. J. Milliken, C. A. Bartel, and T. R. Kurtzberg, "Diversity and creativity in work groups," *Group creativity: Innovation through collaboration*, pp. 32–62, 2003. 7.1

[95] S. Mohammed and E. Ringseis, "Cognitive diversity and consensus in group decision making: The role of inputs, processes, and outcomes," *Organizational behavior and human decision processes*, vol. 85, no. 2, pp. 310–335, 2001. 7.1

[96] M. M. Bekker, J. S. Olson, and G. M. Olson, "Analysis of gestures in face-to-face design teams provides guidance for how to use groupware in design," in *Proceedings of the 1st conference on Designing interactive systems: processes, practices, methods, & techniques*. ACM, 1995, pp. 157–166. 7.1

[97] I. Santiesteban, S. White, J. Cook, S. J. Gilbert, C. Heyes, and G. Bird, “Training social cognition: from imitation to theory of mind,” *Cognition*, vol. 122, no. 2, pp. 228–235, 2012. 7.1.2

[98] J. Leighton, G. Bird, C. Orsini, and C. Heyes, “Social attitudes modulate automatic imitation,” *Journal of Experimental Social Psychology*, vol. 46, no. 6, pp. 905–910, 2010. 7.1.2