Belongingness and Satisfaction Recognition from Physiological Synchrony with A Group-Modulated Attentive BLSTM under Small-group Conversation

Conference Paper · October 2021

DOI: 10.1145/3461615.3485410

3 authors:

Woan-Shiuan Chien
National Tsing Hua University

Huang-Cheng Chou
National Tsing Hua University

Chi-Chun Lee
National Tsing Hua University

Belongingness and Satisfaction Recognition from Physiological Synchrony with A Group-Modulated Attentive BLSTM under Small-group Conversation

Woan-Shiuan Chien, Huang-Cheng Chou, Chi-Chun Lee
wschien@gapp.nthu.edu.tw,hc.chou@gapp.nthu.edu.tw,cclee@ee.nthu.edu.tw
Department of Electrical Engineering, National Tsing Hua University
MOST Joint Research Center for AI Technology and All Vista Healthcare
Taiwan

ABSTRACT

Physiological synchrony is a particular phenomenon of physiological responses during a face-face conversation. However, while many previous studies proposed various physiological synchrony measures between interlocutors in dyadic conversations, very few works on computing physiological synchrony in small groups (three or more people). Besides, belongingness and satisfaction are two critical factors for humans to decide where group they want to stay. Therefore, we want to investigate and reveal the relationship between physiological synchrony and belongingness/satisfaction under group conversation in this preliminary work. We feed the physiology of group members into a designed learnable graph structure with the *group-level* physiological synchrony and heart-related features computed from Photoplethysmography (PPG) signals. We then devise a Group-modulated Attentive Bi-directional Long Short-Term Memory (GGA-BLSTM) model to recognize groups' three levels of belongingness and satisfaction (low, middle, and high). Finally, we evaluate the proposed method on our recently collected multi-modal group interaction corpus (never published before), NTUBA. The results show that (1) the models trained jointly with the group-level physiological synchrony and the conventional heart-related features consistently outperforms the model only trained with the conventional features, and (2) the proposed model with a Graph-structure Group-modulated Attention mechanism (GGA), GGA-BLSTM, performs better than the robust baseline model, the attentive BLSTM. Finally, the GGA-BLSTM achieves a good unweighted average recall (UAR) of 73.3% and 82.1% on group satisfaction and belongingness classification tasks, respectively. In further analyses, we reveal the relationships between physiological synchrony and group satisfaction/belongingness.

CCS CONCEPTS

• **Human-centered computing → Activity centered design**; • **Applied computing → Psychology**; • **Information systems →** *Multimedia information systems.*

ICMI '21 Companion, October 18–22, 2021, Montréal, QC, Canada

ACM ISBN 978-1-4503-8471-1/21/10…$15.00
https://doi.org/10.1145/3461615.3485410

KEYWORDS

small group, satisfaction, belongingness, physiological synchrony, graph attentive BLSTM

ACM Reference Format:
Woan-Shiuan Chien, Huang-Cheng Chou, Chi-Chun Lee. 2021. Belongingness and Satisfaction Recognition from Physiological Synchrony with A Group-Modulated Attentive BLSTM under Small-group Conversation. In *Companion Publication of the 2021 International Conference on Multimodal Interaction (ICMI '21 Companion), October 18–22, 2021, Montréal, QC, Canada.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3461615.3485410

1 INTRODUCTION

Human beings by nature are social animals, who grow and mature by engaging in a series of dyadic, small group, and other group interactions in their lifetime [46]. Small group interaction and co-operation frequently occur in our daily life, especially prevalent in workplace settings. Most of the previous computational studies on a small group or multi-party interaction primarily focus on modeling task-based attributes using behavior dynamics within small groups, such as automating the prediction of group performance or group competence [5, 15, 22]. Fewer computational studies investigate the social-affective aspects (group membership), such as belongingness, satisfaction, emotion, trust, and cohesion. In this work, our goal is to present a computational work on group belongingness and satisfaction prediction.

A psychological construct contains many human behaviors, such as self-esteem, a sense of group belonging, group culture, to name a few. The group belongingness describes that the tendency belongs to a team/group would affect an adolescent's behavior well before he or she is a member of the group [28, 46, 56]. If human perceives a shared sense of belonging, it may lead to negative emotions, and the changes involved in the neural basis [16]. In terms of group satisfaction, Fu et al. [19] have investigated the differences and relationships between group consensus and group satisfaction. Additionally, the findings in [31] suggest that group members who have a relatively high sense of group satisfaction wished to remain within the group, and a sense of group satisfaction is also related to the quality of the teamwork and the mean level of group members' satisfaction [38]. Spehar et al. [55] also have revealed that belongingness has a positive impact on satisfaction at work. These social-affective aspects of a group are critically crucial in small group dynamics and affect the outcome of the task performance.

To understand these important aspects of group memberships, most prior studies in behavior science utilize a questionnaire with a series of questions for quantification [23]. However, this self-reported method is inefficient (non-scalable) and prone to undesired variability (subjectivity and uncontrollable individual factors). Hence, an objective method in modeling social-affective group-level construct is important in continuously advancing our understanding of group dynamics and providing technological solutions. According to social psychologists' studies, people in groups tend to become similar to other group members as they engage in positive and satisfactory interactions. More generally speaking, it has been shown that humans would gradually act synchronously with their interlocutors during face-face interactions [59]. This particular phenomenon, called *synchrony*, can be observed externally in voice [36], facial expression [61], and even internally in physiology [47]. This synchronous acting, *synchrony*, is directly controlled and evoked by mutual changes in autonomic nervous system activity [18, 26, 48]. According to these prior studies, we hypothesize that group members' physiological synchrony is connected to the overall group belongingness and satisfaction.

Physiological synchrony (PS) indicates a similarity of physiological signals between individuals over time. Many previous studies demonstrated the existence of synchrony in dyadic interactions, like maternal-infant (mother-child/mother-adolescent/parent-child), where these intimate social contacts would create an impact on the infant's (child's) physiological systems [40, 47, 67]. Humans do affect the physiological processes of their attached partner through the coordination of acoustic, linguistic, and visual social signals [17, 25, 44]. While there is a large body of prior research, most of these studies are conducted in dyadic interactions and not in the context of small groups (three or more people). Only recently, several researchers have started to investigate the relationships between group performance/creativity/cohesion and PS [15, 22, 26]. For example, Mønster et al. [45] showed that physiological synchronous in Electromyography (EMG) (activation of the smile muscle) was related to group cohesion, and PS in electrodermal activity (EDA) was associated with group tension. Also, PS in the heart rate was correlated to group coordination [21]. Besides, interestingly, both studies [21, 45] consistently found no relationships between perceived group competence/performance and PS. While shedding light on the similarity of physiological signals among individuals, the studies mentioned above mostly ignored the within-session temporal dynamics of physiology and regarded the contributions of group members to the overall group-level construct as equal. In this paper, we compute physiological synchrony over time during small group interaction and devise a graph-attentive mechanism to automatically learn the contributions from individual group members to perform automatic recognition of group belongingness and satisfaction. The proposed method is termed as a Graph-structure Group-modulated Attentive Bi-directional Long Short-Term Memory (GGA-BLSTM) network.

Inspired by [43], they compute physiological synchrony in the first derivatives of electrodermal activity (EDA). While the Photoplethysmography (PPG) signals are different from EDA, we hypothesize that the first derivatives of signals are helpful to capture the linearity between two physiological signals. We can imagine that there is a complete PPG wave. We focus on the section from the on-set to the wave pick and the section from the wave pick to the off-set. Therefore, we use the linear correlation coefficient to calculate the synchrony between two physiological signals. Hence, we compute the physiological synchrony of all members within each group and transform them into group-level physiological synchrony features. Then, we combine them with conventional heart-related features as input to train our network for recognizing group belongingness/satisfaction. The proposed method transforms the individual member's physiological representations into the dynamic *graph-level* concatenation, instead of direct concatenation, and model their temporal dynamics in an attentive BLSTM network. We evaluate our method on two different attributes prediction, i.e., group satisfaction and group belongingness, in our recently collected multi-modal small group interaction database, NTUBA. To compare with the performance of the conventional physiological feature set, we conduct an ablation study on the physiological synchrony computed with PPG and conventional PPG features. The method achieves an excellent unweighted average recall (UAR) of 73.2% and 82.1% on the three levels (low, middle, high) group satisfaction and group belongingness recognition. Moreover, we obtain 4.4% and 16.8% improvements comparing the performance with the conventional PPG features on group satisfaction and group belongingness classifications tasks separately. To sum up, the main contributions of our paper are as below.

- We are one of the first works to propose a *group-level* physiological synchrony feature computed with the first derivatives of *PPG* signals during small group conversations.
- The proposed GGA-BLSTM model can automatically learn the contributions of individuals in group-level physiological synchrony features with a sophisticated attention mechanism to enhance the power of models on group belongingness and group satisfaction predictions.
- We are one of the first computational works in revealing the relationships between the group belongingness/satisfaction and physiological synchrony and introduce a new large collective small-group database.

2 RELATED WORK

2.1 Physiological Synchrony

Physiological synchrony (PS) exists between interlocutors' mutual changes in autonomic nervous system activity. There are variants in measuring PS, as shown below.

- **Pearson correlation coefficient (PCC).** Researchers usually utilized the Pearson correlation coefficient (PCC) to calculate PS in the physiological signals. For instance, the previous studies [26, 34, 43] used PCC to measure PS of EDA signals. Feldmana et al. [18] use PCC to calculate PS in the electrocardiogram (ECG) between mothers and their 3-month old infants during face-face interactions. Chang-Arana et al. [8] estimated PS in EMG of the reactions of the zygomatic major with PCC to understand and analyze the designer's success between users and designers.
- **Spearman rank correlation coefficient (SRCC).** Kaplan et al. [27] have used SRCC to calculate PS in galvanic skin reflex (GSR) to investigate the relationships between PS and affective orientation. Cassani et al. [7] have explored the synchrony of electroencephalography (EEG) spectral features between lead

Table 1: A table summarizes the existing small group databases. "*" represents the dataset is self-collected. Zh, En, and Fr represent Mandarin Chinese, English, and French, respectively.

Database	Language	Groups	Population Composition	Recordings				Questionnaire	Available
				Physiology	Audio	Video	Text	(Group Membership)	
NTUBA	Zh	72	3	✓	✓	✓	✓	✓	✓
UZH* [52]	En	62	4~6	-	-	✓	-	✓	-
ELEA [53]	En/Fr	40	4	-	✓	✓	-	-	✓
GAP [5]	En	13	3~5	-	✓	✓	-	-	✓
UGI [3]	En	22	3~5	-	✓	✓	-	-	✓
NU* [10]	En	58	2	✓	-	✓	-	✓	-
AMI [6]	En	30	4	-	✓	✓	✓	-	✓

dancers and fellow dancers, and Kinreich et al. [30] have computed PS in the EEG with SRCC over the time signal of the Stockwell transform frequency spectrum in two partners.

- **Other Measurements.** Other studies proposed autocorrelation [14] and cross-correlation function [4, 22] to calculate PS for behavior analyses during interactions. Also, Chikersal et al. [9] calculated distances between the series of EDA signals of each individual in a dyad using Dynamic Time Warping (DTW) to compute PS for revealing relationships between PS and dyadic satisfaction. Moreover, there are other PS assessments, such as Single Session Index, Signal Matching, Instantaneous Derivative Matching, Directional Agreement, and Fisher's z-transform [37, 51].

Unlike the studies mentioned above, we propose a new PS measurement by estimating the contemporary trends and changes with the first derivatives of PPG signals and then using both PCC and SRCC to obtain the final PS values.

2.2 Group Satisfaction and Belongingness Recognition

To the best of our knowledge, there are very few computational studies on automatic recognition group satisfaction or belongingness. Only Lai et al. [32] had trained classifiers to automatically recognize group satisfaction in meetings using external behaviors, i.e., acoustic, lexical, and turn-taking features. Moreover, Mønster et al. [45] revealed that PS is an indicator of interpersonal rapport and relationship quality in a group. Also, Chikersal et al. [9] proposed that physiological activation is unconscious and difficult to control with consciousness. In this work, our focus is to predict group belongingness and satisfaction with physiological signals.

2.3 Group-level Graph LSTM

There are various Graph Long Short-Term Memory (Graph LSTM), and researchers modified the structure of inputs to fit their specific graph-like data. For instance, Liang et al. [35] have proposed a Graph LSTM model to capture different degrees of semantic correlation with neighboring nodes on the semantic object parsing task. Peng et al. [50] have designed a particular representation incorporating various intra-sentential and inter-sentential dependencies for a cross-sentence n-ary relation extraction with Graph LSTM model. Zhang et al. [64] have changed the uni-directional LSTM layer of Graph LSTM into bi-directional and add an attention mechanism in their proposed S-LSTM for improving text encoding. Moreover, Tang et al. [57] have proposed the Coherence Constrained Graph LSTM (CCG-LSTM) to effectively recognize group activity by modeling the appropriate motions of individuals while suppressing the irrelevant motions. Shu et al. [54] have introduced a residual LSTM into their model, Graph LSTM-in-LSTM (GLIL), for group activity recognition by modeling the person level actions and the group level activity simultaneously. Zhang et al. [63] have used the graph LSTM model to addresses the limitations of sequential models by converting textual information into a graph and then deploying the message passing operation to ascertain the node representation and the semantic correlation between slot and intent on spoken language understanding task.

However, the studies mentioned above only transform inputs from the individual level into group-level statically, but they did not consider the potential unequal contribution from each individual in deriving group-level inputs. In this work, to learn a more accurate contribution from each input for a group-level input, we slightly modified the gating mechanism in LSTM by adding learnable weights to decide the contributions of group-level input features.

3 METHODOLOGY

3.1 Datasets

3.1.1 ***Small Group Interaction Databases***. There have been several existing small group interaction databases (shown in Table 1). For instance, the ELEA corpus [53] was constructed to analyze developing leadership in freshly arranged groups. The GAP corpus [5] contains thirteen small team conversations in which the subjects achieve the Winter Survival Task for studying perceptions on cohesion, leadership, to name a few. The UGI corpus [3] was collected for fine-grained analysis of the head and body pose and gestures. The AMI corpus [6] was designed to collect for studying performing behavior in small and face-to-face conversations in the IDIAP smart room, and it involves multi-modal sensor data with manually labeled meeting conversations. The database used in [10] gathered from laboratory research where sixty dyads carried through the Test of Collective Intelligence together online and evaluated their group satisfaction while wearing physiological sensors. Also, the database [52] collected the research with student teams in a co-working virtual surrounding.

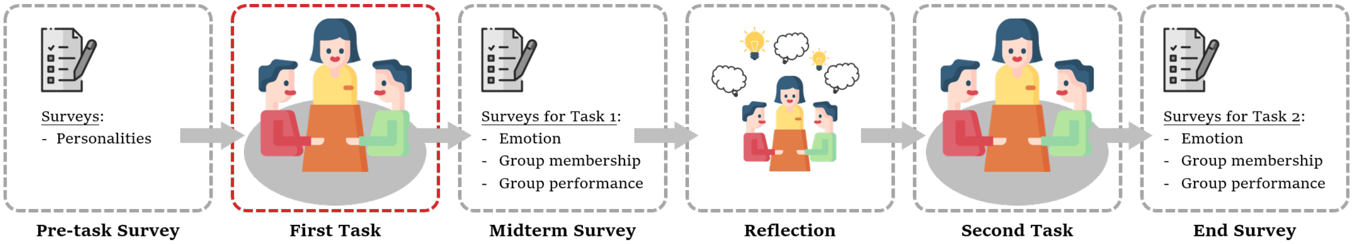

Figure 1: The procedures of collecting NTUBA database.

However, most databases mentioned above only consist of audio and visual recordings and lack physiology for the group membership recognition. In this study, our objective is to use physiology to investigate the group members during the group interaction. Although UZH and NU are suitable for our work, the databases are not released. Therefore, we organized the NTUBA database collected by the College of Management of National Taiwan University (NTU). As Table 1 shows, the NTUBA database contains the most participants and groups and also involves the self-reported questionnaire, including the social-affect aspects of the group. Additionally, the NTUBA database is the largest small-group database in Mandarin Chinese. We will provide the details of the NTUBA database is described in the following section.

3.1.2 ***NTUBA Database.*** The NTUBA is to explores the relationship between group behaviors and group performances. Each group was assigned a shopping task by following [60] of diverse scenarios where they were prompted to discuss with each other and concluded the best solution in a limited 30 minutes. All participants have signed informed consent and been fully informed of all experimental procedures under the approved ethical guidelines (IRB approved). There were 72 three-person groups, who mainly were undergraduate students at NTU, and 7 of the groups were dropped due to signal loss. Hence, this work included 195 participants in 65 groups total. To be noticed, the NTUBA is still collecting, and it is not published before.

The collecting processes have six sessions in total shown in Figure 1. Researchers firstly inquired about prior familiarity between group members and instructed subjects to fill out the self-reported questionnaires. Then, the first task began for 30 minutes. Afterward, the participants completed a midpoint survey about the perceived group cohesion and performance. Furthermore, they were asked to reflect on the task they had just completed and discuss how to perform better at the second task for 10 minutes, and then the second task started for another 30 minutes. Finally, an endpoint survey was reported in self-reports. In this work, we follow [9] to use the data of the first task; compared to the second task, and it would include less confounding factors such as task reflection and increased familiarity between members.

The NTUBA contains audio, transcripts, video, and physiology recordings, which are all simultaneously recorded. In this study, we only used one type of physiological signal, Photoplethysmography (PPG), recorded by the wrist-worn E4 sensor with a 64Hz sample rate, which is widely used in previous studies on physiological synchrony [9, 43]. For the measurement of PPG signals, Fujita et al.

Table 2: A table summarizes the statistics on the NTUBA dataset, including the label distribution and the average and sum timestamps of PPG and ΔPS. PPG and ΔPS represent the PPG features and physiological synchrony features, respectively.

3-class	Statistics		Group Satisfaction	Group Belongingness
Low	Number of Groups		**22**	**33**
	Timestamp (Average/Sum)	PPG	3.318/73	3.848/127
		ΔPS	4.727/104	4.697/155
Middle	Number of Groups		**29**	**24**
	Timestamp (Average/Sum)	PPG	3.448/100	3.167/76
		ΔPS	3.897/113	4.000/96
High	Number of Groups		**14**	**8**
	Timestamp (Average/Sum)	PPG	4.571/64	4.250/34
		ΔPS	5.429/76	5.250/42

[20] measures the sampling rate from 10Hz to 240Hz, and Choi et al. [11] measures from 5 Hz to 10000 Hz. They claimed that 60Hz and 50Hz are the minimum tolerance ranges, respectively, which do not affect the information of PPG signals. Hence, we can point out that our work's 64 Hz PPG signals are sufficient to collect practical information. Besides, the subjects were asked to annotate their subjective perceptions, including group memberships, on a seven-point scale at the end of each task (1 = "highly inaccurate" and seven = "highly accurate"). We list two questions about the degree of the group's satisfaction and belongingness used as learning targets in this work below.

- This question aims to understand your satisfaction with this group. Please indicate your level of agreement with the following narratives: Overall, I am very satisfied with this group? (此部分旨在瞭解您對這個團隊的滿意度，請針對下列敘述句指出您的同意程度: 整體來說，我對這個團隊非常滿意)
- This question aims to understand the relationship between you and the group members. Please indicate your level of agreement for the following narratives: Group members can feel a strong sense of belonging to each other? (此部分旨在瞭解您與團隊成員間的關係，請針對下列敘述回答您的同意程度：團隊成員間彼此可以感受到強烈的歸屬感)

We aggregated the scores of all group members to represent a single group-level score, and the distribution of score is shown in Figure 3. We split group-level values into three-class according to the original

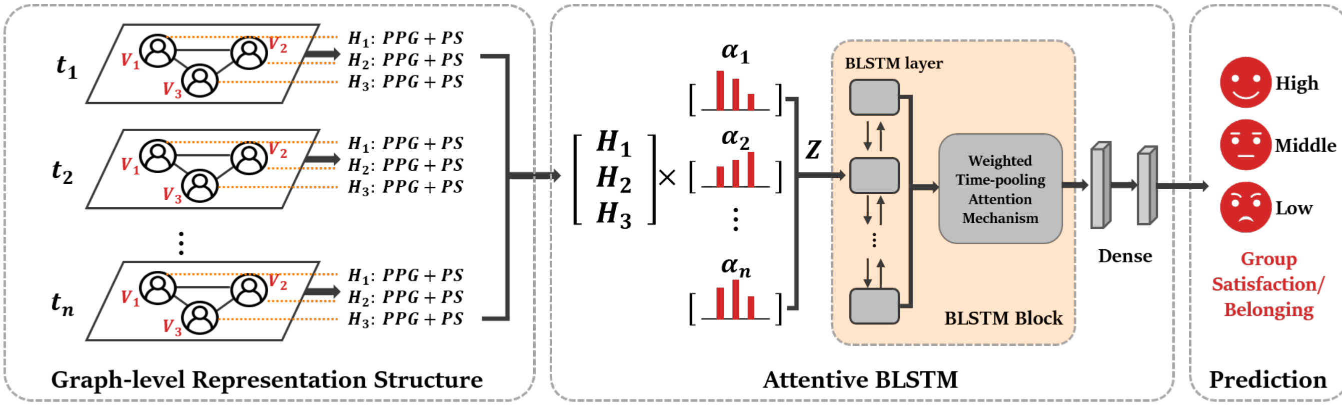

Figure 2: The overview of the proposed Graph-structure Group-modulated Attentive Bi-directional Long Short-Term Memory (GGA-BLSTM) on group satisfaction/belongingness recognition tasks through PPG features (*PPG*) and physiological synchrony features (ΔPS) computed with PPG among group members.

score distribution of each member. By explicitly setting 1-point to 4-point as low, 5-point as middle, and 6-point and 7-point as high, the group-level score can be converted into the following: groups with scores lower than [4,4,5] are divided as low, and groups with scores higher than [5,6,6] are divided to high. Hence, after a simple aggregation, we can obtain the original group-level scores. Then divide the group-level scores into three-class. To make the data distribution of each class balanced, we set the score thresholds that the scores ranging from 3 to 13 are low class, from 14 to 16 is middle class, and from 17 to 21 is high class. We also summarize some statistics in Table 2 including the number of data samples in low, middle, and high categories distribution, the average and sum timestamp of each level.

3.2 Computational Framework

3.2.1 ***Physiological Descriptor Extraction.*** We firstly preprocess individual physiological data with a low-pass filter cut-off at 60Hz on Photoplethysmography (PPG) signals to avoid power frequency noise, then use the consistent FIR filtering from [58] to clean up signals. Also, the first and last 10s of PPG recordings were omitted to avoid artifacts, and then we utilize NeuroKit [13] and HeartPy [58] to extract the standard low-level physiological descriptors (LLDs) widely used in the scientific literature given discrete heart rate signals. There are 35-dimensional features, including time-domain and frequency-domain measures, listed in Table 3. Furthermore, a standard z-normalization is used participant-wise on each feature dimension to ease the effect of individual differences, defined as *PPG*.

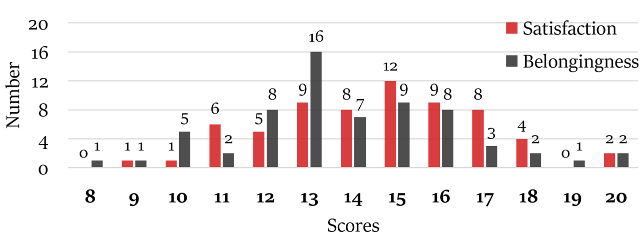

Figure 3: A bar plot summarizes the distribution of the original group-level score, which is aggregated by the scores of all group members of each group. The original group-level score is a list at the bottom. We plot group satisfaction (red) and belongingness (black) in the exact figure and show the total number of each score in our dataset on top of the bars.

Table 3: A table shortens the overview of low-level physiological descriptors extracted from NeuroKit and HeartPy.

Modality	Low-Level Descriptors
PPG(35)	RMSSD, meanNN, sdNN, cvNN, CVSD, medianNN, CD, madNN, mcvNN, pNN50, pNN20, DFA_1, ULF, VLF, LF, HF, VHF, LFn, HFn, LF/HF, LF/P, HF/P, Triang, Sample_Entropy, Entropy_Spectral_HF, Entropy_SVD, Total_Power, FD_Petrosian, FD_Higushi, Shannon_h, Shannon, Fisher_Info, Entropy_Multiscale_AUC, Entropy_Spectral_LF, Entropy_Spectral_VLF

3.2.2 ***Group-based Physiological Synchrony.*** While several different methods have been utilized for assessing physiological synchrony (PS), the simplest and most used technique to assess synchrony is the Pearson correlation coefficient (PCC) [1]. Another simple approach is the Spearman rank correlation coefficient (SRCC) [62]. However, the correlation analysis of continuous human data is vulnerable to spurious conclusions. For instance, when using the Pearson correlation, the data is expected to be independent and stationary; the data has a constant mean and variance over time. Additionally, the Pearson coefficient and Spearman coefficient are both approximately zero when two variables are nonlinear relationships. Therefore, we follow [43] to slightly modify the conventional measurement and calculate *the first-order derivatives* of PPG signals to capture the synchrony trends. Vasundhara et al. [43] also computed the slope value of the EDA signals before calculating PS with PCC. Although the EDA and PPG are different signals, we indeed discover the synchrony trends on *the first-order derivatives* of PPG signals. On the other hand, PCC is better suited for linear relationships in data, whereas SRCC is more accurate for nonlinear correlation and less affected by outliers. Hence, we use both of them to calculate PS in the PPG of individuals in this paper.

Further, since there exist individual differences in the timing of physiological responses, we follow [9] to apply dynamic time warping (DTW) [2] before our synchrony measure. To be noticed, we denote x to be the first-order derivative signal of every group member, which is the reference of every group. Then y is the compared signal, and i is the number of the group member. Afterwards, with the DTW, we allow a maximum warping of 4 seconds, and then calculate synchrony over the entire 30 minutes using the Pearson correlation coefficient ($\boldsymbol{r_p}$) [1] as Eq.(1) and Spearman rank correlation coefficient ($\boldsymbol{r_s}$) [62] as Eq.(2) between the first-order derivative signals of each dyad in this three-person group interaction with 180 seconds (s) as a window size. This parameter is chosen empirically ranging in [60s, 120s, 180s, 240s, 300s, 360s]. The performance using 180s window size is the best among them. Additionally, the use of 50% overlapping size is considered when applying windowing as another type of representation. This operation gives us a time-to-time correlation score revealing the level of synchrony within dyads for the last 180 seconds.

$$\boldsymbol{r_p} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}, \tag{1}$$

$$\boldsymbol{r_s} = 1 - \frac{6\sum(x_i - y_i)^2}{n(n^2-1)}, \tag{2}$$

where $\bar{x}$ represents the mean of the vector x_i, and $\bar{y}$ is the mean of the vector y_i. n represents the number of samples.

Our method returns the max correlation values S when comparing each group member's reference signal to other members in the session. Notice that we only record the values if the p-value is less than or equal to 0.05, else S will be assigned 0 in terms of Pearson and Spearman correlation coefficient defined in Eq.(3) for avoiding capturing noises.

$$\begin{cases} S = \boldsymbol{r}, & p_{\leq} 0.05 \\ S = 0, & else \end{cases}. \tag{3}$$

Since there are three members in a group, we retain all correlation values computed between pairwise combinations, denoted as ΔPS, as a physiological synchrony measure. It includes 4-dimensional features per window.

3.2.3 ***Graph-Structure Group-modulated Attention (GGA) Mechanism***. To better model the dynamics and importance of each group member, we introduce a description of the latent effect of the physiological synchrony arising from each group member within groups shown in Figure 2 (Graph Structures) because group members directly induce the physiological synchrony. To learn more accurate contributions from each input among group-level inputs, we propose a mechanism by adding the learnable weights to decide the contributions of group-level input features. To be more specific, to integrate the information of each other members, at time-stamp t_n (n is ranging from 1 to the maximum length of time-stamp), we construct an undirected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{H}\}$ to bind the 3 members of the groups, and $\mathcal{V}$ represents the set of all graph nodes whose number v is 3 and $\mathcal{H}$ means the node features whose feature dimension is k, and $\alpha \in \mathbb{R}^{k\times v}$ represents a learnable contribution weight vector associated with the set of 3 members nodes. The input feature vector (Z) shown in Figure 2 (Graph-Structure Group-modulated Attentive BLSTM) for any group i is abstracted as follow:

$$Z^i_{t_n} = \sum\nolimits_{s=1}^{v} \alpha_s H_s, \tag{4}$$

where Z is a graph-structure group-based representation as input of following BLSTM Block.

3.2.4 ***BLSTM Block***. The main structure of the BLSTM block is modified from [41] consisting of one BLSTM layer with a weighted time-pooling attention mechanism, one fully-connected layer with Rectified Linear Unit (ReLU) activation function, and then one prediction layer with a softmax activation function. Now, given the output Z, the BLSTM layer then generates an output sequence $y = (y_1, ..., y_t)$. T equals the length timestamp of input features, and t is at each timestamp. The weighted time-pooling attention mechanism is as below. A softmax function is utilized to the results to get a set of final weights for the frames which sum to unity:

$$\alpha_t = \frac{exp(u^T y_t)}{\sum_{t=1}^{T} exp(u^T y_t)}, \tag{5}$$

where u is the attention parameter vector.

The above attention weights are used in a weighted average in sequence to get the output representation:

$$Z_{BLSTM} = \sum\nolimits_{t=1}^{T} \alpha_t y_t. \tag{6}$$

Finally, the Z_{BLSTM} is passed into the following layers, one fully-connected layer with Rectified Linear Unit (ReLU) activation function and one prediction layer with a softmax activation function for prediction.

4 EXPERIMENT

4.1 Experimental Setup

There are two types of group memberships to evaluate our method: **group satisfaction** and **group belongingness**. A group-independent and class-balanced 5-fold cross-validation are used as our evaluation scheme. The BLSTM-based models (Attentive BLSTM and GGA-BLSTM) are trained with a fixed length, and we use the zero-padding to ensure each data sample's time-steps are the same length is less than the maximum timestamp.

Several hyper-parameters as below are grid-searched: learning rate among [0.05, 0.03, 0.01] with adjusting mechanism by multiplying $\frac{1}{\sqrt{1+epoch}}$ per epoch. The number of nodes in the BLSTM layer is fixed as [2, 4, 8]. Batch size is fixed as [16, 32], the max epoch is 1000, and optimizer is ADAMAX [29]. Additionally, we follow [65, 66], which are the closest studies to us, to use an unweighted average recall (UAR) as our final evaluation metric. Zhong et al. [65, 66] modeled the group-level personality composition for group performance classification. Finally, the whole framework is implemented using the Pytorch toolkit [49].

4.2 Model Comparison

We carry out our experiments utilizing SVM and vanilla DNN only with the physiological features or with physiological synchrony features as baseline results for an intact comparison. First,

Table 4: A table summarizes the experimental recognition results in UAR. "*" represents the highest UAR in the task. "x" and "o" indicate the DTW window with overlapping sizes set to 0% and 50%.

Target	Group Satisfaction						Group Belongingness					
Feature Type	*PPG*		ΔPS		*PPG*+ΔPS		*PPG*		ΔPS		*PPG*+ΔPS	
Overlap (50%)	x	o	x	o	x	o	x	o	x	o	x	o
SVM	0.402	0.358	0.359	0.355	0.431	0.397	0.399	0.416	0.374	0.382	0.420	0.468
DNN	0.523	0.506	0.532	0.485	0.594	0.543	0.518	0.569	0.530	0.566	0.569	0.611
Attentive BLSTM	0.672	0.548	0.515	0.467	0.704	0.668	0.599	0.692	0.481	0.505	0.794	0.810
GGA-BLSTM	0.688	0.591	0.611	0.579	**0.732***	**0.682**	0.614	0.653	0.598	0.681	**0.807**	**0.821***

several parameters of SVM are grid-searched: kernel type used $['rbf', 'linear', 'poly']$. The $'coef0'$ is fixed as [1, 10, 100], and the $'gamma'$ of $'rbf'$ kernel is fixed as $[1e-3, 1e-4]$. Second, the architecture of the DNN model includes three dense layers with dimensions [256, 128, 32], and the dropout rates are 0.3, 0.1, and 0.1 in DNN, respectively. For both of these models, we compute 15 statistical functionals[2] on each of the extracted individual short-term *PPG* features and ΔPS features. Afterward, we use 5 statistical functionals[3] to obtain group-wise descriptors. Then we research with both physiological features and physiological synchrony features, and we compare them with the following models to inspect the power of the proposed GGA-BLSTM.

- **Attentive BLSTM.** The baseline Attentive Bi-directional Long Short-Term Memory (Attentive BLSTM) model [41] contains one dense layer in a network with a ReLU activation function, one BLSTM layer with a weighted time-pooling attention mechanism proposed by [42], one more dense layer with ReLU activation function, and then one prediction layer with a softmax activation function. Specifically, we concatenate the *PPG* of three members of each group as the model input.
- **GGA-BLSTM.** The proposed Graph-structure Group-modulated Attentive Bi-directional Long Short-Term Memory (GGA-BLSTM) model is an additional modification from Attentive BLSTM by removing the first dense layer and adding a graph-structure group-modulated attention mechanism to transform the input features into graph-level representations with individual learnable weights. Our objective is to learn better the contributions of each member in each timestamp using group structural information. There are 65 graphs that all the group members would be linked in every timestamp. The specifics of graph structures for group constraints and Graph LSTM have been described in 3.2.3 and 3.2.4.

4.3 Group Satisfaction and Belongingness Recognition Results

Table 4 summarizes the complete recognition results across different methods. The proposed GGA-BLSTM model with an attention mechanism outperforms all comparison methods when using both *PPG* and ΔPS as inputs, which obtains the best UAR 73.2% and 82.1% on group satisfaction and group belongingness classification, respectively. The improvements are the absolute 2.8% and 1.1% on group satisfaction and group belongingness recognition tasks compared to the attentive BLSTM model.

Moreover, there are several observations. First, while the models trained directly with *PPG* features can achieve a relatively high UAR, the models trained with only 4-dimensional ΔPS features can obtain a competitive performance. Second, there exists a large time-series requirement in our tasks. According to the ablation study, the model without the ability to model the temporal relationships perform poorly than the models which can accommodate the temporal information, especially for *PPG* features. Hence, based on the experiments, it is suitable to use a time-series model like BLSTM for modeling *PPG* features. Furthermore, to figure out the effect of the overlap sizes, we conduct the experiments on the same classifiers with the same input features but in different overlap sizes, and the results show that whether taking the overlap or not depends on the learning targets. That is, we still need to investigate the best parameters of overlap sizes according to various tasks.

Furthermore, the proposed GGA-BLSTM model trained with graph structure inputs that link each group member's physiological representations obtains improved robustness results on the group satisfaction and belongingness tasks. The significant difference between BLSTM and GGA-BLSTM is the construction of the features of group members. GGA-BLSTM can learn better the dynamic information contributions of some of the physiological features of each member over time by learning from the representation weights of members with graph strategy. Finally, we provide the additional analyses shown in the following section.

5 ANALYSIS

In this section, to understand relationships between physiological synchrony and group satisfaction/belongingness. We perform a one-way ANOVA test to explore the differences in the physiological synchrony between three levels of group belongingness/satisfaction, respectively.

5.1 One-way ANOVA Significance Test on ΔPS

Having established the presence and characteristics of physiological synchrony in the group belongingness and satisfaction recognition, we are interested in exploring the differences in the physiological synchrony between the three-level group belongingness and satisfaction, respectively. We measure the physiological synchrony representations for each group of every timestamp that represents each level (low, middle, or high) class. While each data sample has

[2] max/min value and respective relative position within input, mean/median value, standard deviation, first percentile, ninety-ninth percentile, the difference between ninety-ninth percentile and first percentile, skewness, kurtosis, quartile 1, quartile 3, and interquartile range

[3] max/min value, mean/median, standard deviation, and differences

Table 5: A table summarizes the one-way ANOVA test on physiological synchrony features. "*F (#,12)*" expresses the $\#^{th}$ representation of all 12 ΔPS features.

Target	**Group Satisfaction**		**Group Belongingness**	
F (#,12)	F statistic	P-value	F statistic	P-value
F (1,12)	4.128	0.017	3.385	0.035
F (3,12)	5.364	0.005	-	-
F (4,12)	3.713	0.026	3.232	0.041
F (7,12)	5.178	0.006	-	-

more than one timestamp, we consider all timestamps as the group in respect to the levels of group belongingness/satisfaction. This procedure results in a total of 293 and 237 pairs corresponding to *PPG* and ΔPS, respectively. Table 2 shows the distribution of timestamp-level pairs. Take ΔPS as an example. There are 155, 96, and 42 pairs on the low-, middle-, and high-level group belongingness. Using this data, we perform a one-way analysis of variance (ANOVA) [24, 39] on each feature of ΔPS with labels of two tasks.

The reporting APA format follows ANOVA test format, and all the results (F statistics and p-values) are shown in Table 5. With group satisfaction target, the significance thresholds (p-value) of four features in ΔPS are lower than 0.05, and three of them including the *F (3,12)*, *F (4,12)*, and *F (7,12)* of ΔPS come from the calculation with Spearman rank correlation coefficient (SRCC). The other one, *F (1,12)*, comes from the Pearson correlation coefficient (PCC). On the other hand, there are 2 significant indicators (*F (1,12)* and *F (4,12)*) on the three-level group belongingness. These findings suggest that we should compute ΔPS with SRCC, which are easier to find the physiological synchrony given two PPG signals of different individuals than PCC.

In the conventional method, most computational studies on calculating physiological synchrony from physiological signals utilize PCC. However, we propose that SRCC can be an alternative measure to estimate physiological synchrony. Besides, we have similar findings that group satisfaction is positively associated with high levels of physiological synchrony as same as [9]. Furthermore, we also do the same significance Test on the *PPG* features, but there are no feature dimensions whose significance threshold (p-value) is smaller than 0.05 on the group belongingness. Instead, there are nine significant indicators among PPG features whose p-values are smaller than 0.05 (CSVD, Entropy_SVD, HF, pNN20, madNN, mcvNN, meanNN, medianNN, CD). Therefore, the synchrony representation (ΔPS) has high potentials for various applications on recognizing other group memberships, such as group cohesion and group emotion.

6 LIMITATION

The work is a preliminary study investigating the relationships between ΔPS and group belongingness/satisfaction. We also propose a group-modulated attention mechanism to learn the contributions of features of each member among groups. However, there are still many factors we did not take into account, such as gender effects in the group composition. Lee [33] has shown evidence that gender composition of groups is related to group cohesion and performance. In this work, we did not consider the gender composition in groups. Additionally, we did not conduct comparative results with physiological synchrony features computed with the *Raw* PPG. Moreover, the proposed approach can not ensure that the motor movement has influenced the physiological synchrony computed with PPG signals recorded by E4-wristband because it has a chance to be affected by movement. That is, recording physiological activity could be a product of motor coordination. We will take video recordings into account to make sure the issues mentioned above in future work. Subsequently, we only use the unweighted average recall (UAR) as the evaluation metric, which cannot be taken as indicative for a good approximation of the actual performance of the proposed system. We will use other metrics (e.g., macro-F1 score) to evaluate our proposed approach in future work.

7 CONCLUSION AND FUTURE WORK

Social-affective aspects of the group, e.g., group belongingness and group satisfaction, significantly impact personal emotional feelings. In this paper, the proposed method, GGA-BLSTM, automatically predicts group satisfaction/belongingness classification with physiological synchrony computed with the slop of PPG (ΔPS) and conventional heart-related features (*PPG*). We design a unique attention mechanism, Graph-Structure Group-modulated Attention (GGA), to learn the contributions of group members. Furthermore, this framework is evaluated on a recently collected larger small group collective database, NTUBA. To be noticed, NTUBA is not published, and this is one of our contributions. Finally, this approach, GGA-BLSTM, achieves a promising UAR of 73.3% and 82.1% on the three-level (low, middle, high) group satisfaction and group belongingness recognition tasks, which get 4.4% and 19.3% improvements comparing with the *PPG* features, respectively.

To the best of our knowledge, this is one of the first studies that have explicitly modeled the physiological synchrony computed with the first derivatives of PPG for predicting group belongingness and satisfaction. Additionally, the ablation study shows that the time-series modeling for physiological features is practical and helpful to improve the performance of two tasks. Also, according to our analyses, we found that Spearman rank correlation (SRCC) is an alternative physiological synchrony measure in PPG, and this type of physiological synchrony helps researchers to quickly capture the contemporary trends of synchrony in group conversations than the conventional method, Pearson correlation coefficient (PCC). In the future work, we will investigate further the contributing factors to the synchrony phenomenon, extend our multimodal fusion framework to combine physiological synchrony computed with *raw PPG signals* and other expressive behaviors (e.g., acoustic behaviors, facial cues, body movements, or conversational temporal dynamics [12]) to enhance the robustness and recognition power.

ACKNOWLEDGMENTS

This work was supported by the MOST under Grants 110-2221-E-007-067-MY3 and 110-2634-F-007-012. We thank the contribution of Yi-Ching Liu from National Taiwan University for collecting the recordings and various annotations in the NTUBA database. We also thank the anonymous reviewers for their valuable comments.

REFERENCES

[1] Jacob Benesty, Jingdong Chen, Yiteng Huang, and Israel Cohen. 2009. Pearson Correlation Coefficient. In *Noise reduction in speech processing*. Springer, 1–4. https://doi.org/10.1007/978-3-642-00296-0_5

[2] Donald J. Berndt and James Clifford. 1994. Using Dynamic Time Warping to Find Patterns in Time Series. In *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining* (Seattle, WA) *(AAAIWS'94)*. AAAI Press, 359–370.

[3] Bhattacharya, Indrani and Foley, Michael and Ku, Christine and Zhang, Ni and Zhang, Tongtao and Mine, Cameron and Li, Manling and Ji, Heng and Riedl, Christoph and Welles, Brooke Foucault and Radke, Richard J. 2019. The Unobtrusive Group Interaction (UGI) Corpus. In *Proceedings of the 10th ACM Multimedia Systems Conference* (Amherst, Massachusetts) *(MMSys '19)*. Association for Computing Machinery, New York, NY, USA, 249–254. https://doi.org/10.1145/3304109.3325816

[4] Andrea Bizzego, Atiqah Azhari, Nicola Campostrini, Anna Truzzi, Li Ying Ng, Giulio Gabrieli, Marc H. Bornstein, Peipei Setoh, and Gianluca Esposito. 2020. Strangers, Friends, and Lovers Show Different Physiological Synchrony in Different Emotional States. *Behavioral Sciences* 10, 1 (2020). https://doi.org/10.3390/bs10010011

[5] McKenzie Braley and Gabriel Murray. 2018. The Group Affect and Performance (GAP) Corpus. In *Proceedings of the Group Interaction Frontiers in Technology* (Boulder, CO, USA) *(GIFT'18)*. Association for Computing Machinery, New York, NY, USA, Article 2, 9 pages. https://doi.org/10.1145/3279981.3279985

[6] Carletta, Jean and Ashby, Simone and Bourban, Sebastien and Flynn, Mike and Guillemot, Mael and Hain, Thomas and Kadlec, Jaroslav and Karaiskos, Vasilis and Kraaij, Wessel and Kronenthal, Melissa and others. 2005. The AMI meeting corpus: A pre-announcement. In *International workshop on machine learning for multimodal interaction*. Springer, 28–39. https://doi.org/10.1007/11677482_3

[7] Raymundo Cassani, Abhishek Tiwari, Ilona Posner, Bruno Afonso, and Tiago H. Falk. 2020. Initial Investigation into Neurophysiological Correlates of Argentine Tango Flow States: a Case Study. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 3478–3483. https://doi.org/10.1109/SMC42975.2020.9282835

[8] Álvaro M Chang-Arana, Matias Piispanen, Tommi Himberg, Antti Surma-aho, Jussi Alho, Mikko Sams, and Katja Hölttä-Otto. 2020. Empathic accuracy in design: Exploring design outcomes through empathic performance and physiology. *Design Science* 6 (2020), e16. https://doi.org/10.1017/dsj.2020.14

[9] Prerna Chikersal, Maria Tomprou, Young Ji Kim, Anita Williams Woolley, and Laura Dabbish. 2017. Deep Structures of Collaboration: Physiological Correlates of Collective Intelligence and Group Satisfaction. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) *(CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 873–888. https://doi.org/10.1145/2998181.2998250

[10] Chikersal, Prerna and Tomprou, Maria and Kim, Young Ji and Woolley, Anita Williams and Dabbish, Laura. 2017. Deep Structures of Collaboration: Physiological Correlates of Collective Intelligence and Group Satisfaction *(CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 873–888. https://doi.org/10.1145/2998181.2998250

[11] A Choi and H Shin. 2017. Photoplethysmography sampling frequency: pilot assessment of how low can we go to analyze pulse rate variability with reliability? *Physiological measurement* 38, 3 (2017), 586.

[12] Huang-Cheng Chou, Yi-Wen Liu, and Chi-Chun Lee. 2019. JOINT LEARNING OF CONVERSATIONAL TEMPORAL DYNAMICS AND ACOUSTIC FEATURES FOR SPEECH DECEPTION DETECTION IN DIALOG GAMES. In *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. 1044–1050. https://doi.org/10.1109/APSIPAASC47483.2019.9023050

[13] Makowski. D. 2016. NeuroKit: A Python Toolbox for Statistics and Neurophysiological Signal Processing (EEG, EDA, ECG, EMG...). Paris, France.

[14] Chad Danyluck and Elizabeth Page-Gould. 2019. Social and Physiological Context can Affect the Meaning of Physiological Synchrony. *Scientific reports* 9, 1 (June 2019), 8222. https://doi.org/10.1038/s41598-019-44667-5

[15] Muhterem Dindar, Sanna Järvelä, and Eetu Haataja. 2020. What does physiological synchrony reveal about metacognitive experiences and group performance? *British Journal of Educational Technology* 51, 5 (2020), 1577–1596. https://doi.org/10.1111/bjet.12981 arXiv:https://bera-journals.onlinelibrary.wiley.com/doi/pdf/10.1111/bjet.12981

[16] Lili Dong and Wei Wang. 2018. Influence of Social Factors on the Group Belongingness of Residents in Small Towns and Its Underlying Neural Base. *NeuroQuantology* 16, 6 (2018).

[17] Bruce P. Doré and Robert R. Morris. 2018. Linguistic Synchrony Predicts the Immediate and Lasting Impact of Text-Based Emotional Support. *Psychological Science* 29, 10 (2018), 1716–1723. https://doi.org/10.1177/0956797618779971 arXiv:https://doi.org/10.1177/0956797618779971 PMID: 30020863.

[18] Ruth Feldman, Romi Magori-Cohen, Giora Galili, Magi Singer, and Yoram Louzoun. 2011. Mother and infant coordinate heart rhythms through episodes of interaction synchrony. *Infant Behavior and Development* 34, 4 (2011), 569 – 577. https://doi.org/10.1016/j.infbeh.2011.06.008

[19] Chao Fu, Wenjun Chang, and Shanlin Yang. 2020. Multiple criteria group decision making based on group satisfaction. *Information Sciences* 518 (2020), 309 – 329. https://doi.org/10.1016/j.ins.2020.01.021

[20] Daisuke Fujita and Arata Suzuki. 2019. Evaluation of the possible use of PPG waveform features measured at low sampling rate. *IEEE Access* 7 (2019), 58361–58367.

[21] Riccardo Fusaroli, Johanne Stege Bjørndahl, Andreas Roepstorff, and Kristian Tylén. 2015. Physiological entrainment and behavioral coordination in a collective, creative construction task.

[22] Ilanit Gordon, Avi Gilboa, Shai Cohen, Nir Milstein, Nir Haimovich, Shay Pinhasi, and Shahar Siegman. 2020. Physiological and Behavioral Synchrony Predict Group Cohesion and Performance. *Scientific reports* 10, 1 (May 2020), 8484. https://doi.org/10.1038/s41598-020-65670-1

[23] Ben D. Harper and Kent L. Norman. 1993. Improving user satisfaction: The questionnaire for user interaction satisfaction version 5.5. In *Proceedings of the 1st Annual Mid-Atlantic Human Factors Conference*. 224–228.

[24] Gary W Heiman. 2001. *Understanding research methods and statistics: An integrated introduction for psychology.* Houghton, Mifflin and Company.

[25] Karine Jospe, Shir Genzer, Nathalie klein Selle, Desmond Ong, Jamil Zaki, and Anat Perry. 2020. The contribution of linguistic and visual cues to physiological synchrony and empathic accuracy. *Cortex* 132 (2020), 296–308. https://doi.org/10.1016/j.cortex.2020.09.001

[26] Eunice Jun, Daniel McDuff, and Mary Czerwinski. 2019. Circadian Rhythms and Physiological Synchrony: Evidence of the Impact of Diversity on Small Group Creativity. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 60 (Nov. 2019), 22 pages. https://doi.org/10.1145/3359162

[27] Howard B Kaplan, Neil R Burch, Samuel W Bloom, and Robert Edelberg. 1963. Affective orientation and physiological activity (GSR) in small peer groups. *Psychosomatic Medicine* 25, 3 (1963), 245–252. https://journals.lww.com/psychosomaticmedicine/Abstract/1963/05000/Affective_Orientation_and_Physiological_Activity.6.aspx

[28] Jeff Kiesner, Mara Cadinu, François Poulin, and Monica Bucci. 2002. Group Identification in Early Adolescence: Its Relation with Peer Adjustment and Its Moderator Effect on Peer Influence. *Child Development* 73, 1 (2002), 196–208. https://doi.org/10.1111/1467-8624.00400 arXiv:https://srcd.onlinelibrary.wiley.com/doi/pdf/10.1111/1467-8624.00400

[29] Diederik P. Kingma and Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, Yoshua Bengio and Yann LeCun (Eds.). http://arxiv.org/abs/1412.6980

[30] Sivan Kinreich, Amir Djalovski, Lior Kraus, Yoram Louzoun, and Ruth Feldman. 2017. Brain-to-Brain Synchrony during Naturalistic Social Interactions. *Scientific reports* 7, 1 (December 2017), 17060. https://doi.org/10.1038/s41598-017-17339-5

[31] RL Krichevsky and MM Smirnova. 1981. Satisfaction with membership in a group in relation to some leadership and management phenomena. *Voprosy Psychologii* (1981).

[32] Catherine Lai and Gabriel Murray. 2018. Predicting Group Satisfaction in Meeting Discussions. In *Proceedings of the Workshop on Modeling Cognitive Processes from Multimodal Data* (Boulder, Colorado) *(MCPMD '18)*. Association for Computing Machinery, New York, NY, USA, Article 1, 8 pages. https://doi.org/10.1145/3279810.3279840

[33] Cynthia Lee and Jiing-Lih Farh. 2004. Joint Effects of Group Efficacy and Gender Diversity on Group Cohesion and Performance. *Applied Psychology* 53, 1 (2004), 136–154. https://doi.org/10.1111/j.1464-0597.2004.00164.x arXiv:https://iaap-journals.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1464-0597.2004.00164.x

[34] Zhi Li, Melissa L. Sturge-Apple, Siwei Liu, and Patrick T. Davies. 2020. Parent-adolescent physiological synchrony: Moderating effects of adolescent emotional insecurity. *Psychophysiology* 57, 9 (2020), e13596. https://doi.org/10.1111/psyp.13596 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/psyp.13596

[35] Xiaodan Liang, Xiaohui Shen, Jiashi Feng, Liang Lin, and Shuicheng Yan. 2016. Semantic Object Parsing with Graph LSTM. In *Computer Vision – ECCV 2016*, Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling (Eds.). Springer International Publishing, Cham, 125–143.

[36] Christy L Ludlow. 2015. Central Nervous System Control of Voice and Swallowing. *Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society* 32, 4 (August 2015), 294—303. https://doi.org/10.1097/wnp.0000000000000186

[37] Carl D Marci, Jacob Ham, Erin Moran, and Scott P Orr. 2007. Physiologic correlates of perceived therapist empathy and social-emotional process during psychotherapy. *The Journal of nervous and mental disease* 195, 2 (February 2007), 103—111. https://doi.org/10.1097/01.nmd.0000253731.71025.fc

[38] Claire M. Mason and Mark A. Griffin. 2003. Identifying Group Task Satisfaction at Work. *Small Group Research* 34, 4 (2003), 413–442. https://doi.org/10.1177/1046496403252153 arXiv:https://doi.org/10.1177/1046496403252153

[39] JH McDonald. 2014. Multiple tests. *Handbook of Biological Statistics. 3rd ed Baltimore, Maryland: Sparky House Publishing* (2014), 233–6.

[40] David H. McFarland, Annie Joëlle Fortin, and Linda Polka. 2020. Physiological measures of mother–infant interactional synchrony. *Developmental Psychobiology* 62, 1 (2020), 50–61. https://doi.org/10.1002/dev.21913 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/dev.21913

[41] Seyedmahdad Mirsamadi, Emad Barsoum, and Cha Zhang. 2017. Automatic speech emotion recognition using recurrent neural networks with local attention. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2227–2231. https://doi.org/10.1109/ICASSP.2017.7952552

[42] S. Mirsamadi, E. Barsoum, and C. Zhang. 2017. Automatic speech emotion recognition using recurrent neural networks with local attention. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2227–2231. https://doi.org/10.1109/ICASSP.2017.7952552

[43] Vasundhara Misal, Surely Akiri, Sanaz Taherzadeh, Hannah McGowan, Gary Williams, J. Lee Jenkins, Helena Mentis, and Andrea Kleinsmith. 2020. Physiological Synchrony, Stress and Communication of Paramedic Trainees During Emergency Response Training. In *Companion Publication of the 2020 International Conference on Multimodal Interaction* (Virtual Event, Netherlands) *(ICMI '20 Companion)*. Association for Computing Machinery, New York, NY, USA, 82–86. https://doi.org/10.1145/3395035.3425250

[44] Sankar Mukherjee, Alessandro D'Ausilio, Noël Nguyen, Luciano Fadiga, and Leonardo Badino. 2017. The Relationship Between F0 Synchrony and Speech Convergence in Dyadic Interaction. In *Proc. Interspeech 2017*. 2341–2345. https://doi.org/10.21437/Interspeech.2017-795

[45] Dan Mønster, Dorthe Døjbak Håkonsson, Jacob Kjær Eskildsen, and Sebastian Wallot. 2016. Physiological evidence of interpersonal dynamics in a cooperative production task. *Physiology & Behavior* 156 (2016), 24–34. https://doi.org/10.1016/j.physbeh.2016.01.004

[46] Barbara M Newman, Brenda J Lohman, and Philip R Newman. 2007. Peer group membership and a sense of belonging: their relationship to adolescent behavior problems. *Adolescence* 42, 166 (2007), 241—263. http://europepmc.org/abstract/MED/17849935

[47] Lauren M. Papp, Patricia Pendry, and Emma K. Adam. 2009. Mother-Adolescent Physiological Synchrony in Naturalistic Settings: Within-Family Cortisol Associations and Moderators. *Journal of Family Psychology* 23, 6 (Dec. 2009), 882–894. https://doi.org/10.1037/a0017147

[48] Sangin Park, Soo Ji Choi, Sungchul Mun, and Mincheol Whang. 2019. Measurement of emotional contagion using synchronization of heart rhythm pattern between two persons: Application to sales managers and sales force synchronization. *Physiology & Behavior* 200 (2019), 148 – 158. https://doi.org/10.1016/j.physbeh.2018.04.022 Consumers' physiological information to measure marketing strategies effectiveness: the power of unconscious responses.

[49] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. 2019. Pytorch: An imperative style, high-performance deep learning library. *arXiv preprint arXiv:1912.01703* (2019).

[50] Nanyun Peng, Hoifung Poon, Chris Quirk, Kristina Toutanova, and Wen-tau Yih. 2017. Cross-sentence n-ary relation extraction with graph lstms. *Transactions of the Association for Computational Linguistics* 5 (2017), 101–115.

[51] Héctor J. Pijeira-Díaz, Hendrik Drachsler, Sanna Järvelä, and Paul A. Kirschner. 2016. Investigating Collaborative Learning Success with Physiological Coupling Indices Based on Electrodermal Activity. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (Edinburgh, United Kingdom) *(LAK '16)*. Association for Computing Machinery, New York, NY, USA, 64–73. https://doi.org/10.1145/2883851.2883897

[52] Rabindra Ratan and Béatrice S Hasler. 2014. Playing well with virtual classmates: relating avatar design to group satisfaction. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing* (Baltimore, Maryland, USA) *(CSCW '14)*. Association for Computing Machinery, New York, NY, USA, 564–573. https://doi.org/10.1145/2531602.2531732

[53] Sanchez-Cortes, Dairazalia and Aran, Oya and Mast, Marianne Schmid and Gatica-Perez, Daniel. 2012. A Nonverbal Behavior Approach to Identify Emergent Leaders in Small Groups. *IEEE Transactions on Multimedia* 14, 3 (2012), 816–832. https://doi.org/10.1109/TMM.2011.2181941

[54] X. Shu, L. Zhang, Y. Sun, and J. Tang. 2021. Host–Parasite: Graph LSTM-in-LSTM for Group Activity Recognition. *IEEE Transactions on Neural Networks and Learning Systems* 32, 2 (2021), 663–674. https://doi.org/10.1109/TNNLS.2020.2978942

[55] Ivan Spehar, Jacques Forest, and Frode Stenseng. 2016. Passion for work, job satisfaction, and the mediating role of belongingness. *Scandinavian Journal of Organizational Psychology* 8, 1 (2016), 17–27.

[56] Margaret R Stone and B Bradford Brown. 1999. Identity claims and projections: descriptions of self and crowds in secondary school. *New directions for child and adolescent development* 84 (1999), 7–20.

[57] J. Tang, X. Shu, R. Yan, and L. Zhang. 2019. Coherence Constrained Graph LSTM for Group Activity Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2019), 1–1. https://doi.org/10.1109/TPAMI.2019.2928540

[58] P Van Gent, H Farah, N Nes, and B van Arem. 2018. Heart rate analysis for human factors: Development and validation of an open source toolkit for noisy naturalistic heart rate data. In *Proceedings of the 6th HUMANIST Conference*. 173–178.

[59] Scott S Wiltermuth and Chip Heath. 2009. Synchrony and cooperation. *Psychological science* 20, 1 (2009), 1–5.

[60] Anita Williams Woolley, Christopher F. Chabris, Alex Pentland, Nada Hashmi, and Thomas W. Malone. 2010. Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science* 330, 6004 (2010), 686–688. https://doi.org/10.1126/science.1193147 arXiv:https://science.sciencemag.org/content/330/6004/686.full.pdf

[61] Kenji Yokotani, Gen Takagi, and Kobun Wakashima. 2020. Nonverbal synchrony of facial movements and expressions predict therapeutic alliance during a structured psychotherapeutic interview. *Journal of Nonverbal Behavior* 44, 1 (2020), 85–116. https://link.springer.com/article/10.1007/s10919-019-00319-w

[62] Jerrold H. Zar. 2005. *Spearman Rank Correlation*. American Cancer Society. https://doi.org/10.1002/0470011815.b2a15150 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/0470011815.b2a15150

[63] Linhao Zhang, Dehong Ma, Xiaodong Zhang, Xiaohui Yan, and Houfeng Wang. 2020. Graph LSTM with Context-Gated Mechanism for Spoken Language Understanding. *Proceedings of the AAAI Conference on Artificial Intelligence* 34, 05 (Apr. 2020), 9539–9546. https://doi.org/10.1609/aaai.v34i05.6499

[64] Yue Zhang, Qi Liu, and Linfeng Song. 2018. Sentence-State LSTM for Text Representation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Melbourne, Australia, 317–327. https://doi.org/10.18653/v1/P18-1030

[65] Shun-Chang Zhong, Yun-Shao Lin, Chun-Min Chang, Yi-Ching Liu, and Chi-Chun Lee. 2019. Predicting Group Performances Using a Personality Composite-Network Architecture During Collaborative Task. In *Proc. Interspeech 2019*. 1676–1680. https://doi.org/10.21437/Interspeech.2019-2087

[66] Shun-Chang Zhong, Bo-Hao Su, Wei Huang, Yi-Ching Liu, and Chi-Chun Lee. 2020. Predicting Collaborative Task Performance Using Graph Interlocutor Acoustic Network in Small Group Interaction. In *Proc. Interspeech 2020*. 3122–3126. https://doi.org/10.21437/Interspeech.2020-1698

[67] Rachel H. Zhuo, Mengyu M. Gao, Julia Yan, Xiaoyi Hu, Wen Zhou, and Xiaomei Li. 2019. Correlates of Parent-Child Physiological Synchrony and Emotional Parenting: Differential Associations in Varying Interactive Contexts. *Journal of Child and Family Studies* 28, 4 (04 2019), 1116–1123. https://webpac.lib.nthu.edu.tw/pds?func=load-login&calling_system=ermg&url=https://www.proquest.com/scholarly-journals/correlates-parent-child-physiological-synchrony/docview/2179175678/se-2?accountid=14427