Assessment of Computer-Supported Collaborative Processes using Interpersonal Physiological and Eye-Movement Coupling

Guillaume Chanel, Mireille Bétrancourt, Thierry Pun
Swiss Center for Affective Sciences
Rue des Battoirs 7
1205 Geneva, Switzerland
guillaume.chanel@unige.ch

Donato Cereghetti, Gaëlle Molinari
Distance Learning University Switzerland
Sierre, Switzerland
Gaelle.molinari@unidistance.ch

***Abstract*—In this paper we propose a method to assess key collaborative processes during computer-supported group work based on physiological signals and eye-movements. Synchronous interpersonal multimodal signals from 30 dyads were recorded while collaborating remotely. Features measuring how much collaborators' eye-movements and physiology are coupled were extracted from the obtained time series and two regression models were trained to assess collaboration. Results show that the two coupling measures can be used to predict collaborative processes such as emotion management and convergence. Assessing those processes is a major step toward the development of remote collaborative interfaces able to adapt to the users' social interactions.**

***Keywords*—physiological signals, dual eye-tracking, interpersonal coupling, regression, computer-supported collaboration**

I. Introduction

Social signal processing and affective computing aim at creating intelligent interfaces able to react and adapt to the social and emotional signals expressed by users [1], [2]. To reach this goal a machine should be able to understand non-verbal behaviors such as smiles and emblematic gestures (e.g. the "peace and love" finger gesture), and to infer user's state of mind. With the development of remote communication environments, intelligent interfaces should not only adapt their behavior to individual user signals; they should also be able to deal with the complexity of social interactions taking place among several users communicating with each other. In the case of Computer-Supported Collaboration (CSC), collaborators are often facing challenging situations which require deploying several collaborative processes for a successful outcome. The collaborators need to reach shared understanding (grounding), to accurately model the partner, and to cope with emotional situations such as conflicts. In remote situations some non-verbal cues can be missing (e.g. a gesture done out of the webcam field of view) which could lead to misunderstandings and to the deterioration of collaborative processes. We believe that processes such as grounding, mutual modeling and emotion management, can be assessed from the analysis of multimodal social signals. This information would then be valuable to design intelligent interfaces able to dynamically adapt to the characteristics of CSC situations. Such an adaptation could be achieved by using group awareness tools [3], [4] and by providing help to manage and guide the collaboration (e.g. advices about how to address conflicts and misunderstandings).

In the situation of human-human communication, social cues can be considered at both the individual (e.g. a gesture from one group member) and group (e.g. a group movement) levels. Current research have mainly studied features describing individual behaviors [1]. However, studies have shown that features describing the behavior of a group are also of high interest to gain insights in group processes [5]. Indexes of inter-individual coupling measure to which extend group members are showing similar behaviors or similar physiological activities [6], [7]. Coupling has been shown to be related to collaborative processes [6], team performance [8], and affective exchanges [9]. Most of the studies in social signal processing have focused on the analyses of speech, facial expression and gestures [1]. To our knowledge no studies have tried to use both eye-movements and physiological signals for the assessment of collaborative processes.

In this work we are interested in developing innovative CSC technologies to improve collaborative outcomes. This could be achieved by providing interfaces able to adapt to the collaborative interactions. The first step toward such interfaces is the development of methods to automatically assess collaborative processes. Considering that the previous research (detailed in Section II) has found relationships between interpersonal couplings and collaborative processes / outcomes, our hypothesis is: collaborative processes that contribute to the success of collaboration can be predicted by physiological and eye-movement couplings between interacting partners.

The goal of the reported study is to test our hypothesis by collecting synchronous interpersonal multimodal data during natural CSC and by building predictive regression models. The collected data corpus is presented in Section III. Section IV presents the methods used to assess collaborative processes from physiological and eye-movement couplings. Finally results are presented and discussed in Section V and conclusions are given in Section VI.

II. Background

A. *Collaborative processes*

Several processes have been identified as playing a key role in successful collaboration such as: sustaining mutual understanding (grounding); building a representation of the collaborative partners (mutual modeling); information pooling; confronting points of view; co-constructing knowledge; building upon partner's reasoning (transactivity); elaborating and resolving conflicts; reaching consensus and convergence; coordinating the joint efforts; regulating the interaction, etc. Such interaction processes are recognized as being positively related to individual outcomes (e.g. learning gains) as well as team performance [10], [11].

The quality of collaboration can be determined based on the analysis of collaborative processes and group performance. Collaborative processes can be assessed through content analysis of transcribed dialogue or, as in the present study, using a self-report questionnaire asking about participants' perceptions of their interaction with their collaborative partner. Compared to performance assessment, collaborative processes assessment helps to find out when, and which type of, adaptive feedback is needed (e.g a grounding problem should not be solved in the same way than a problem due to a low level of transactivity, although they will both impact performance).

B. *Interpersonal physiological and eye-movement coupling*

In the last decade, physiological signals have shown to be reliable indicators of emotions [12], [13]. Despite of these results they have been hardly used in the context of social interfaces involving multiple users. One of the few possible examples of such physiological and social interfaces is [14] where the users could see their own and partner's heart rate during the presentation of movie clips. The results showed that including bio-feedback increases the feeling of co-presence compared to the control condition. When recording the physiological activity of multiple users synchronously, it is possible to measure to what extent the two physiological profiles are dependent from each other, this is named physiological coupling (or physiological linkage / compliance) [15]. Physiological coupling has been shown to be related to several social processes. In [9] the authors showed that physiological coupling was higher for distressed spouses than for non-distressed ones. Physiological coupling was also found to be related to empathy [16]; people who accurately evaluated the negative emotions of others also displayed a high degree of shared physiology. In [17] physiological coupling was found to be correlated with social presence in the context of social games. Studies have also examined the relationships of physiological coupling with group performance. The first of these studies demonstrated that heart rate and skin conductance coupling were positively correlated with task completion time [8]. Unfortunately, these results were not confirmed by a second study using self-reported performance [18]. Despite of all these results, physiological compliance has never been employed in social signal processing for the design of methods able to automatically assess collaborative processes such as grounding and mutual modeling.

Eye-trackers have been employed for two main purposes [19]: to study users' behaviors during their interaction with computers and as a new means for real-time interaction with the machine. Several features of eye movements were found to be related to cognitive processes. For instance, the number of fixations on a given target indicated its importance while long fixations were generally a sign of difficulty in information processing [19]. It has been suggested that sharing gaze position information among collaborators could help them to reach a common ground since they would constantly know where their partner is looking [20]. However results showed that this strategy often brings confusion and does not help collaborators to reach a common ground in several situations. It is thus necessary to find a less disturbing method to improve grounding among collaborators. The use of interpersonal eye-movement coupling was proposed in [6] to study conversation coordination. In this study eye-movement coupling was defined as moments when people are looking at the same place on a shared picture. The results showed that those who share common knowledge (i.e. high level of grounding) had higher eye-movement coupling. Eye-movement coupling is thus a potential predictor of collaborative processes such as mutual grounding. Using eye-movement coupling to assess collaboration was proposed in [21]. In this study the authors proposed to employ features at the individual level (e.g. gaze dispersion) and at the group level (e.g. gaze divergence) to predict collaborative performance on puzzle games. Two classifiers were trained to determine if the games were successfully solved or not. The classifiers successfully classified the games' outcome in 74% of the cases based only on eye-tracker data. However, the proposed features were highly dependent on the type of game played. Furthermore individual and group features were not analyzed separately, and it is thus impossible to know which of the two contributed to classes' discrimination. Finally, only the task outcomes were classified while assessing collaborative processes could also be of interest as explained in Section II.A.

III. Multimodal collaborative experiment

A. *Experimental protocol*

The development of adaptive collaborative systems requires the collection of multimodal data from groups of people collaborating on a given task. In this study, 60 participants were randomly assigned to same gender dyads (16 women dyads and 14 men dyads, mean age 23.5), and were asked to collaborate remotely without seeing each other. After filling consent forms the participants were briefed about the schedule of the experimental protocol and sited in front of a computer (C1 and C2, Fig. 1). They were then trained to the use of two modules of the collaborative environment DREW. The first module was an argumentative map designed to share and argument on each other's ideas. The second module was an emotion awareness tool designed specifically for this experiment and available to only half of the dyads. This module was used to analyze the effects of emotion awareness on collaborative processes and outcomes (not presented in this paper, see [4]).

During the collaboration, the participants had to design a slogan against violence at school in less than 45 min. For this purpose they could communicate orally using headsets and through the DREW software. The task was decomposed in

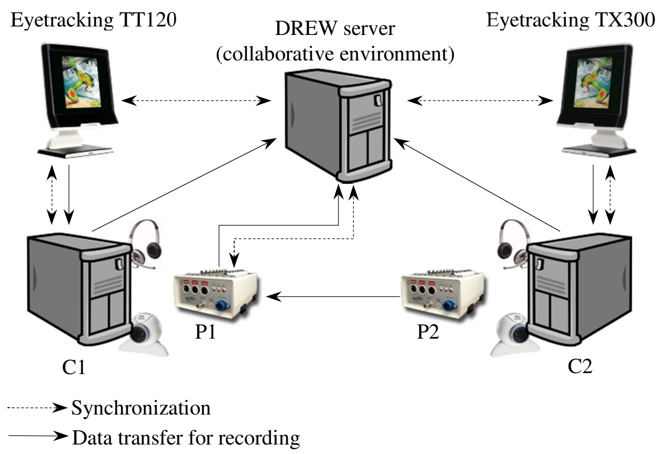

Fig. 1. Framework for synchronous multimodal data acquisition. The two collaborators are interacting through clients C1 and C2. Physiological signals are recorded using Biosemi Active II devices P1 and P2.

three steps. In Step 1, both partners generated as many boxes of slogan ideas as possible in the argumentative map. In Step 2, they were asked to debate and argue slogan ideas. After debating, peers suppressed the less relevant slogan ideas and improved those remaining. In the last step (Step 3), they were asked to negotiate and find a consensus about the best slogan.

B. Synchronous multimodal recordings

The behaviors and physiological activity of the participants were recorded during the collaboration. All the computers and eye-trackers were connected on a 1Gbs network and the devices were synchronized as shown in (Fig. 1.). The DREW software is composed of two clients managing the interaction between the users and a server. The role of the server is to dispatch the information between the clients, so that the two collaborators share a common virtual space for interaction, and to record logs of all events (e.g. creation of an argument).

The eye movements of the collaborators were recorded using two different models of Tobii eye-trackers: a T120 with a sampling frequency of 120 Hz and a TX300 with a sampling frequency of 300 Hz. The DREW client interfaces were displayed on the screens of these two eye-trackers. A webcam pointing at the face of each participant and a headset were also connected to the clients C1 and C2. The eye-positions were recorded synchronously with the face video and the speech recording using Tobii Studio V2.32.

The physiological activity of the collaborators was recorded using two daisy-chained Biosemi Active II systems (P1 and P2, Fig. 1). Peripheral physiological signals were chosen based on their relationship with emotional and cognitive processes. Skin temperature and skin conductance (an index of physiological arousal) were obtained by placing sensors on the proximal phalanges of the hand. Blood volume pulse, a relative measure of blood pressure, was obtained by finger plethysmography. The Biosemi CMS/DRL reference electrodes were positioned on the hand hypothenar surface. All these sensors were attached to the non-dominant hand of the participants to reduce noise contamination due to keyboard typing and mouse movement. A belt was attached around the abdomen of the participants to measure respiration amplitude. Finally, an electro-cardiogram (ECG) was recorded. All physiological signals were recorded at a sampling frequency of 512 Hz.

C. Questionnaires and feedback

After collaboration, the participants answered several questionnaires concerning their perception of the interaction and their emotions. They also annotated a video of the interaction with emotional information. Participants' emotional feedbacks are not used in this study but remain of high interest to the affective computing community, particularly to develop emotion recognition algorithms able to operate in ecological situations, on continuous emotion dimensions and on spontaneous emotional expressions. In this paper we concentrate on the participants' perception of their collaborative experience. This was assessed using a questionnaire which consisted of 4 groups of questions evaluated on 7-point Likert scales. These questions referred to participants' perception of: their relationship with their partner, the time spent on collaborative activities, the frequency with which they themselves displayed a given behavior, the frequency with which their partner displayed a given behavior.

IV. COLLABORATION ASSESSMENT

A. Measuring the degree of perceived collaboration

In order to identify and quantify the different collaborative processes which took place during the interactions, two exploratory factorial analyses were applied to the four groups of questions mentioned in Section III.C. The first factorial analysis was applied on questions from Groups 1 & 2; the second on questions from Groups 3 & 4. The factors were extracted using SPSS 21, the maximum likelihood criteria, a promax rotation and the regression method. With this method a total of 8 standardized factors were obtained and used as ground-truth targets for regression. These factors are briefly presented in Table I while a more detailed explanation can be found in [4].

B. Physiological signal processing

The coupling between the physiological activities of the members of a dyad was used to predict collaboration as defined by the factors listed in Table I. Due to synchronization errors, two recordings were rejected and the physiological signals of 28 dyads out of 30 were analyzed. Two indices of interpersonal physiological coupling R_s and $C_s(F)$ were computed between the pair of physiological signals s_1 and s_2:

$$R_s = \tanh^{-1}(\rho(s_1, s_2)) \; ; \qquad (1)$$

$$C_s(F) = \tanh^{-1}\left(\sqrt{\frac{\sum_{f \in F} Coh_{s_1,s_2}(f)\sqrt{P_{s_1}(f)P_{s_2}(f)}}{\sum_{f \in F}\sqrt{P_{s_1}(f)P_{s_2}(f)}}}\right) ; \qquad (2)$$

where ρ is the Pearson correlation coefficient, $Coh(f)$ and $P(f)$ the coherence and power spectral value at frequency f belonging to the frequency band F. Equation (1) provides an index of coupling in the time domain which follows a normal distribution thanks to Fisher's z-transform (i.e. the application of the inverse hyperbolic tangent). In (2) weighted coherence [8] was used as a measure of coupling in the frequency domain.

TABLE I. FACTORS USED AS A MEASURE OF COLLABORATION

Factor Name	Description of main items related to each factor (loadings in parenthesis)
Grounding & Coordination	Maintaining a shared understanding (.88); managing the progress of the task (.80); managing the quality of the relation (.75); providing/asking for clarification (.73)
Degree of Conflict	Relational conflict (.83); conflict of ideas (.79); competition (.62); emotional tension (.60)
Degree of Convergence	Action synchrony (.77); mutual understanding (.74); conceptual convergence (.72); emotional convergence (.61); symmetry in roles and responsabilities (.68)
Confrontation & Consensus building	Discussing about disagreements (.82); defending and arguing ideas (.80); confronting different points of view (.73); negotiating and finding compromises (.68)
Co-Construction	Building together new ideas (.88); deepening and broadening ideas (.69); co-elaborating of ideas (0.67)
Emotion Management	Communicating on the emotions of others (.88 & .79); communicating on one's own emotions (.68 & .75); adapting to the emotions of others (.50 & .66); partner's effort to understand his/her own emotions (.72); partner's effort to understand emotions in others (.61)
Emotion modeling	Comparing emotions (.90 & .77); imagining reactions to emotions (.83 & .61); participant's effort to understand emotions in others (.61); participants effort to appear able to control his/her own emotions (.66)
Transactivity	Defending and arguing ideas (.74 & .61); understanding the partner's point of view (.57 & .77); providing points of view (.65 & .53); referring and building upon the partner's ideas (.60 & .55)

The transformation proposed by [22] was applied to have a coupling index that follows a normal distribution. Weighted coherence has the advantage of separating the shared variance of two time series among the chosen frequency bands of interest. This is particularly useful for heart rate variability analysis since it is known to be composed of three main periodic components [23]: a high frequency (HF) component [0.15Hz - 0.4 Hz], a low frequency (LF) component [0.05 Hz - 0.15 Hz], and a very low frequency (VLF) component [0.003 Hz - 0.05 Hz]. The Welch's method was employed to compute the power spectral density of the signals with a time window of 5 min and half overlap.

The ECG was first resampled at 256 Hz and a band-pass filter between 0.05 Hz and 40 Hz was applied to remove noise in the signal. The ECG peaks were identified using the Pan-Tompkins algorithm [24] and IBI (Inter-Beat Intervals) were computed by measuring the time elapsed between all pairs of consecutive peaks. From these unevenly spaced time series the final IBI time series were obtained by interpolation at 4 Hz using the algorithm detailed in [25]. This interpolation resulted in two synchronous IBI time series for each dyad (i.e. one for each dyad member) from which the coupling indices R_{IBI}, C_{IBI}(HF), C_{IBI}(LF) and C_{IBI}(VLF) were computed. The respiration signals were filtered between 0.12 Hz and 1.15 Hz to remove noise and the coupling indices R_{Resp} and C_{Resp}(HF) were computed. The HF frequency band was chosen as it corresponds to the standard range of respiratory rate. The indices $R_{RawResp}$ were also computed from the un-filtered respiration signals. This was achieved because the filter significantly reduced the amplitude of signal peaks related to deep respirations which might be of interest for collaboration assessment. The skin conductance and temperature signals were filtered using a moving average window of 0.5 sec. R_{SCR} and R_{Temp} were than computed to measure skin conductance and temperature coupling. The physiological coupling features are summarized in Table II.

C. Eye-movement processing

As for the physiological signals, eye-movement coupling was used to predict the 8 factors presented in Table I. The gaze-position time-series obtained from the eye-trackers were found to miss several samples. The dyads for which one of the two time series contained less than 5 min of eye-movements were considered as irrelevant and rejected from the analysis. It resulted in a total of 13 dyadic recordings. The two gaze-position time series (one recorded at 120Hz and the other at 300Hz) were first resampled by linear interpolation at the sampling frequency of 5Hz.

Cross-recurrence analysis was previously proposed to measure eye-movement coupling [6]. Cross-recurrence analysis is based on the analysis of a matrix R of size NxM which determines when the states of two systems are close to each other [26]:

$$R_{i,j} = \begin{cases} 1 : d(\vec{x}_i^1, \vec{x}_j^2) \le \varepsilon \\ 0 : d(\vec{x}_i^1, \vec{x}_j^2) > \varepsilon \end{cases}; \quad (3)$$

where vectors $\vec{x}_i^p$ represent the state of system p at time i and d is a distance function. To compute the values $R_{i,j}$ the gaze 2D positions of each partner were used as vectors $\vec{x}_i^1$ and $\vec{x}_i^2$, the distance function d was the Euclidean distance, and ε was set to 100 pixels. In this case the R matrix is squared with a size N corresponding to the number of samples in both gaze-position time series. As listed in Table III, several features were extracted from the R matrix to represent the eye-movement coupling of two partners.

D. Regression

Two regression strategies were tested to predict the self-reported measures of collaboration based on the coupling indexes extracted from the physiological and eye-movement

TABLE II. PHYSIOLOGICAL COUPLING FEATURES

Signal	Feat. name	Description
ECG	R_{IBI}	Correlation of the inter-beat intervals (similar to the correlation of heart rates)
	C_{IBI}(HF) C_{IBI}(LF) C_{IBI}(VLF)	Coherence of the inter-beat intervals in the HF [0.15Hz - 0.4 Hz], LF [0.05 Hz - 0.15 Hz], VLF [0.003 Hz - 0.05 Hz] frequency bands.
Respiration	R_{Resp}	Correlation of filtered repiration signals
	$R_{RawResp}$	Correlation of non-filtered respiration signals
	C_{Resp}(HF)	Coherence of filtered repirations signals in the HR [0.15Hz - 0.4 Hz] frequency band.
SCR	R_{SCR}	Correlation of the SCR signals
Temperature	R_{Temp}	Correlation of the temperature signals

TABLE III. EYE-MOVEMENT COUPLING FEATURES

Feat. name	Description	Motivation
RR	$\frac{1}{N}\sum_{i,j=1}^{N} R_{i,j}$	Recurrence rate. Represents how many times the participants did look at the same place, possibly at different time.
RR_{diag}	$\frac{1}{N}\sum_{i=1}^{N} R_{i,i}$	Diagonal recurrence rate. Represents how many times the participants did look at the same place at the same time.
RR_{lag}	$\max_{\tau} \frac{1}{N}\sum_{i=1}^{N} R_{i,i+\tau}$	How many time participants looked at the same place with a delay τ. Only the maximum value is considered. τ values belonged to the interval [-6 6] secs following the results obtained in [6].
Delay	$\arg\max_{\tau} \frac{1}{N}\sum_{i=1}^{N} R_{i,i+\tau}$	Delay for which the participants were more often looking at the same place.
DL_{Mean} DL_{Max}	The average and maximum length of the diagonal lines in the recurrence plot.	Diagonal lines length represents the duration of common gaze trajectories.
VL_{Mean} VL_{Max} HL_{Mean} HL_{Max}	Idem for vertical lines and horizontal lines.	Vertical and horizontal lines length represents the duration for which one participant was looking at the same place than the other at a given time.

signals. The first strategy was to combine a fast-correlation based filter (FCBF) [27] with mean squared linear regression. The FCBF was used for feature selection as we found it to be more efficient than ridge and LASSO regression. The FCBF rejects features for which the correlation with the targets is lower than a given threshold and removes redundant features. The threshold parameter of the FCBF algorithm was set to 0.553 for eye-movement features and 0.374 for physiological features since these values correspond to a p-value of 0.05. The second regression strategy was to train bags of regression trees [28]. This second approach was chosen because it is non-linear. Each bag was configured to contain 2000 trees.

The regression algorithms were trained using the following leave-one-dyad-out cross-validation. For each dyad a regression model was trained using features of the other dyads; the targets were then estimated by applying the trained model on the data of the tested dyad. The performance of each algorithm was evaluated by computing the Root Mean Squared Error (RMSE) and the coefficient of determination R^2 defined by:

$$R^2 = 1 - \frac{SE}{VAR} \quad (3)$$

where SE is the sum of squared errors and VAR the variance of the targets.

When no predictor is available the best possible estimator is the mean of targets. The proposed algorithms were thus compared with a cross-validated mean estimator. Notice that the proposed mean estimator always yields a negative or zero R^2 value since its sum of squared errors is always higher or equal to the variance of the targets. Furthermore, when a leave one out approach is employed the R^2 of the mean estimator is a constant with the same value for all targets.

V. RESULTS

As can be seen from the results in Table IV several targets are predicted with an error inferior to the mean estimator. Based on physiological coupling both least squared regression and bag of trees are predicting the emotion management factor with success. Being able to predict the amount of efforts put in emotion management during collaboration could be very useful for the adaptation of collaborative platforms. For instance when collaborators do not spend time communicating and adjusting to each other's emotions the system could encourage the collaborators to share their emotions. Emotion awareness tools which help collaborators to better understand their partner's emotions have shown to impact collaboration [4]. Being able to assess emotion management during the collaboration might be ideal to determine when to display such emotion awareness information. The most relevant features for regression were analyzed by looking at the FCBF selection. Two features were often selected together: C_{IBI}(VLF) and C_{IBI}(LF). This demonstrates that these two features are not correlated to each other since the FCBF rejects redundant features. The correlation between the two features and emotion management was found to be negative indicating that the participants reported spending less time managing their emotions when IBI variance was shared at the VLF and LF frequencies. Physiological coupling has been shown to increase with emotionally intense situations [9], [17]. One explanation could be that the dyads who reported spending more time managing their emotions regulated (i.e. decreased) the intensity of their emotions which decreased their physiological coupling as opposed to dyads who did not manage their emotions.

The two regression algorithms predict convergence from eye-movement coupling with a low error compared to mean regression. The most relevant features for convergence prediction are RR_{lag} and RR_{diag} both being positively correlated with convergence. Consequently, the more the participants were looking at the same place on the screen at approximately the same time (the delay is of maximum 6 sec.) the more they reported to act synchronously and to have converging ideas and emotions. This result is very coherent since the participants used the DREW argumentative graph to discuss their ideas and to reach a consensus. When doing so they were looking at the same argument boxes represented in the graph which increased the recurrence rate.

TABLE IV. REGRESSION RESULTS FOR PERCEIVED COLLABORATION [a]

	Physiological features						Eye-movement features					
	R^2			RMSE			R^2			RMSE		
Target	***Mean***	***BRT***	***FCBF LS***	***Mean***	***BRT***	***FCBF LS***	***Mean***	***BRT***	***FCBF LS***	***Mean***	***BRT***	***FCBF LS***
Grounding & Coordination		**0.01**	-0.25	0.65	**0.63**	0.70		-0.29	-0.38	0.75	0.79	0.82
Degree of conflict		-0.25	-0.29	0.68	0.73	0.74		-0.20	**0.07**	0.84	0.85	**0.75**
Degree of convergence		**0.06**	-0.17	0.61	**0.57**	0.64		**0.24**	**0.38**	0.60	**0.49**	**0.44**
Confrontation and consensus building	-0.07	**-0.01**	-0.19	0.61	**0.60**	0.65	-0.17	-0.44	-0.82	0.72	0.80	0.89
Co-construction		-0.31	-0.21	0.85	0.94	0.90		-0.38	**0.07**	0.73	0.79	**0.65**
Emotion management		**0.18**	**0.21**	0.75	**0.65**	**0.64**		**-0.02**	-0.85	0.72	**0.67**	0.90
Emotion modeling		-0.49	-0.25	0.76	0.90	0.82		-0.31	-1.15	0.90	0.95	1.21
Transactivity		-0.18	**-0.06**	0.69	0.72	0.69		-0.19	-0.43	0.80	0.80	0.88

[a] Bold values represent regression errors lower (and R^2 values higher) than for mean regression. Undefined acronyms are: Bag of Regression Trees (BRT), Least Squared regression with FCBF (FCBF LS).

A few other variables were predicted with a moderate performance. For instance co-construction was predicted from eye-movement features using least-squared regression while grounding was predicted by physiological features using bags of regression trees. This demonstrates that physiological and eye-tracking measures can be seen as complementary: they allow inferring different collaborative processes. However both the emotion management and convergence factors were assessed, with relative performance, from both the physiological and eye-movement features. This encourages further research on the fusion of the two modalities to check if they can have synergies and improve the prediction of convergence and emotion management

VI. CONCLUSION

This paper proposes an approach to assess several dimensions of collaborative processes based on the analysis of interpersonal physiological and eye-movement coupling. Multimodal synchronous data was collected from dyads collaborating on the design of a slogan. From the participants' self-reports of the quality of collaboration, 8 factors describing the collaborative interactions were extracted and used as targets in a regression analysis. Two sets of features were employed to assess the targets: physiological and eye-movement couplings. Physiological coupling was computed by measuring the correlation and coherence between the physiological activities of the two members of dyads while cross-recurrence analysis was employed to compute eye-movement coupling. Two regression strategies were tested to assess the targets: least squared regression coupled with a FCBF and bags of regression trees. Results showed the importance of coupling features for the assessment of collaborative processes, thus validating our hypothesis. Two targets were particularly well predicted: emotion management and convergence. Those processes are of high importance for successful collaboration, and therefore being able to assess them automatically is a critical step toward the development of interfaces that dynamically adapt to the process of collaboration.

Future work is underway to assess team performance and to analyze indices of the quality of the collaboration process (content analysis of transcribed collaboration dialogue) in addition to self-reported measures. Considering that the data acquisition protocol consisted of three distinct phases, including temporal aspects in the regression should also improve the regression performance. The database constructed for this study is highly multimodal and only two of the available modalities are presented in this paper. The computation of facial expression coupling as well as speech turns and alignment of turns will certainly provide interesting features for a better assessment of collaboration. The multimodal integration of these modalities remains a challenge that once solved will improve the regression performance since both synergies and complementarities were observed among the two modalities analyzed in this study. Finally, a limitation of the current study is the low number of dyads available for eye-movements. The results presented in this paper should thus be validated on larger datasets.

ACKNOWLEDGMENTS

The authors would like to thank Christelle Bozelle Giroud and Konstantina Kalogianni for their help with the data collection. This research was supported by the National Center of Competence in Research (NCCR) Affective sciences financed by the Swiss National Science Foundation (n° 51NF40-104897) and hosted by the University of Geneva.

REFERENCES

[1] A. Vinciarelli et al., "Bridging the Gap Between Social Animal and Unsocial Machine: A Survey of Social Signal Processing," IEEE Transactions on Affective Computing, vol. 3, no. 1, pp. 69–87, 2012.

[2] A. S. Pentland, "Social Signal Processing," IEEE Signal Processing Magazine, no. July, pp. 108–111, 2007.

[3] M. Sangin, G. Molinari, M.-A. Nüssli, and P. Dillenbourg, "Facilitating peer knowledge modeling: Effects of a knowledge awareness tool on collaborative learning outcomes and processes," Computers in Human Behavior, vol. 27, no. 3, pp. 1059–1067, 2011.

[4] G. Molinari, G. Chanel, M. Bétrancourt, T. Pun, and C. Bozelle, "Emotion Feedback during Computer-Mediated Collaboration: Effects on Self-Reported Emotions and Perceived Interaction," in Proceedings of the 10th International Conference on Computer Supported Collaborative Learning, 2013.

[5] A. S. Pentland, "Collective Intelligence," no. August, pp. 9–12, 2006.

[6] D. C. Richardson and R. Dale, “Looking to understand: the coupling between speakers’ and listeners' eye movements and its relationship to discourse comprehension.,” Cognitive science, vol. 29, no. 6, pp. 1045–60, Nov. 2005.

[7] J. M. Gottman, “Detecting Cyclicity in Social-Interaction,” Psychological Bulletin, vol. 86, no. 2, pp. 338–348, 1979.

[8] R. A. Henning, W. Boucsein, and M. C. Gil, “Social-physiological compliance as a determinant of team performance,” International journal of psychophysiology : official journal of the International Organization of Psychophysiology, vol. 40, no. 3, pp. 221–232, 2001.

[9] R. W. Levenson and J. M. Gottman, “Marital interaction: physiological linkage and affective exchange,” Journal of Personality and Social Psychology, vol. 45, no. 3, pp. 587–597, 1983.

[10] A. Weinberger, K. Stegmann, and F. Fischer, “Knowledge convergence in collaborative learning: Concepts and assessment,” Learning and Instruction, vol. 17, no. 4, pp. 416–426, Aug. 2007.

[11] A. Meier, H. Spada, and N. Rummel, “A rating scheme for assessing the quality of computer-supported collaboration processes,” International Journal of Computer-Supported Collaborative Learning, vol. 2, no. 1, pp. 63–86, Feb. 2007.

[12] G. Chanel, J. J. M. Kierkels, M. Soleymani, and T. Pun, “Short-term emotion assessment in a recall paradigm,” International Journal of Human-Computer Studies, vol. 67, no. 8, pp. 607–627, 2009.

[13] H. Hamdi, P. Richard, A. Suteau, and P. Allain, “Emotion assessment for affective computing based on physiological responses,” in IEEE International Conference on Fuzzy Systems, 2012, pp. 1–8.

[14] G. Chanel, S. Pelli, N. Ravaja, and K. Kuikkaniemi, “Social interaction using mobile devices and biofeedback: effects on presence, attraction and emotions,” in BioSPlay Workshop, Fun and Games Conference, 2010.

[15] I. Ekman, G. Chanel, J. M. Kivikangas, M. Salminen, S. Järvelä, and N. Ravaja, “Social Interaction in Games: Measuring Physiological Linkage and Social Presence,” Simulation & Gaming, vol. 43, no. 3, pp. 321–338, 2012.

[16] R. W. Levenson and A. M. Ruef, “Empathy: a physiological substrate,” Journal of Personality and Social Psychology, vol. 63, no. 2, pp. 234–246, 1992.

[17] G. Chanel, J. M. Kivikangas, and N. Ravaja, “Physiological compliance for social gaming analysis: cooperative versus competitive play,” Interacting with Computers, vol. 24, no. 4, pp. 306–3016, 2012.

[18] R. A. Henning, A. G. Armstead, and J. K. Ferris, “Social psychophysiological compliance in a four-person research team,” Applied Ergonomics, vol. 40, no. 6, pp. 1004–1010, 2009.

[19] R. Jacob and K. Karn, “Eye tracking in human-computer interaction and usability research: Ready to deliver the promises,” in The Mind’s Eye: Cognitive and Applied Aspects of Eye Movement Research, no. 1905, J. Hyona, R. Radach, and H. Deubel, Eds. Amsterdam: Elsevier Science, 2003, pp. 573–605.

[20] R. Müller, J. R. Helmert, S. Pannasch, and B. M. Velichkovsky, “Gaze transfer in remote cooperation: Is it always helpful to see what your partner is attending to?,” Quarterly journal of experimental psychology (2006), no. March 2013, pp. 37–41, Nov. 2012.

[21] M. Nüssli, P. Jermann, M. Sangin, and P. Dillenbourg, “Collaboration and abstract representations: towards predictive models based on raw speech and eye-tracking data,” in Proceedings of the 9th international conference on computer supported collaborative learning, 2009, no. 1970, pp. 78–82.

[22] A. M. F. L. Miranda de Sá, D. D. Ferreira, E. W. Dias, E. M. A. M. Mendes, and L. B. Felix, “Coherence estimate between a random and a periodic signal: Bias, variance, analytical critical values, and normalizing transforms,” Journal of the Franklin Institute, vol. 346, no. 9, pp. 841–853, 2009.

[23] G. G. Berntson et al., “Heart rate variability: origins, methods, and interpretive caveats,” Psychophysiology, vol. 34, no. 6, pp. 623–648, 1997.

[24] J. Pan and W. J. Tompkins, “A Real-Time QRS Detection Algorithm,” IEEE Transactions on Biomedical Engineering, vol. 32, no. 3, pp. 230–236, 1985.

[25] R. D. Berger, S. Akselrod, D. Gordon, and R. J. Cohen, “An efficient algorithm for spectral analysis of heart rate variability,” IEEE Trans Biomed Eng, vol. 33, no. 9, pp. 900–904, 1986.

[26] N. Marwan, M. Carmenromano, M. Thiel, and J. Kurths, “Recurrence plots for the analysis of complex systems,” Physics Reports, vol. 438, no. 5–6, pp. 237–329, Jan. 2007.

[27] L. Yu and H. Liu, “Efficient Feature Selection via Analysis of Relevance and Redundancy,” Journal of Machine Learning Research, vol. 5, 2004.

[28] L. Breiman, “Random forests,” Machine learning, vol. 45, pp. 5–32, 2001.