Group Coach for Co-located Collaboration

Sambit Praharaj[1 [0000-0003-4039-9696]], Maren Scheffel[1 [0000-0003-4395-4819]], Hendrik Drachsler[1,2,3 [0000-0001-8407-5314]], and Marcus Specht[1 [0000-0002-6086-8480]]

[1] Open Universiteit, Valkenburgerweg 177, 6419AT Heerlen, Netherlands
[2] DIPF, Schloßstr. 29, 60486 Frankfurt am Main, Germany
[3] Goethe Universität, Robert-Mayer-Str. 11-15, 60629 Frankfurt am Main, Germany
{sambit.praharaj, maren.scheffel, hendrik.drachsler, marcus.specht}@ou.nl

Abstract. Collaboration is an important 21st century skill; it can take place in a remote or co-located setting. Co-located collaboration (CC) gives rise to subtle human interactions that can be described with multimodal indicators like gaze, speech and social skills. In this demo paper, we first give a brief overview of related work that has identified indicators during CC. Then, we look briefly at the feedback mechanisms that have been designed based on these indicators to facilitate CC. Using these theoretical insights, we design a prototype to give automated real-time feedback to facilitate CC taking the help of the most abundant modality during CC i.e., audio cues.

Keywords: co-located collaboration · real-time feedback · cscl · collaboration indicators · multimodal learning analytics

1 Introduction

Collaboration is an important skill in the 21st century. It can take place in different settings and for different purposes: collaborative meetings [11,5,8], collaborative project work [3], collaborative programming [4] and collaborative brainstorming [10]. Collaboration can be either co-located (or face-to-face) or in a remote setting. "The requirement of successful collaboration is *complex, multimodal, subtle,* and learned over a lifetime. It involves *discourse, gesture, gaze, cognition, social skills, tacit practices,* etc." [9, p.1-2, emphasis added]. Furthermore, in each context, the indicators of collaboration vary. For example, in collaborative programming pointing to the screen, grabbing the mouse from the partner and synchrony in body posture are relevant indicators for good collaboration [4]; whereas in collaborative meetings gaze direction, body posture, speaking time of group members are relevant indicators for good collaboration quality [11,5]. Different feedback mechanisms are designed based on these indicators of CC to facilitate CC [1,5]. Most of these feedback mechanisms are designed by analyzing the indicators of collaboration in a post-hoc manner instead of a real-time operational design [3]. Some studies which have used real-time feedback suffer from the limitation of employing human observers to drive those feedback mechanisms and others have used simplistic automated feedback mechanisms [7]. For instance, Tausch et al. [10] used human observers to drive the real-time feedback system and Bachour et al. [1] used the total speaking time of each group member to drive a real-time feedback LED table-top display which displayed the amount of total speaking time of each group member by glowing different colour LED lights assigned to each member. So, to address this problem, we seek answer to the following research question:

RQ: How can we design an automated *real-time feedback* system using *audio cues* to facilitate co-located collaboration *in-the-wild*?

2 Related Work

In this section, we will first analyze related work according to the different indicators obtained from audio-based cues during CC; and secondly, we review some of the feedback mechanisms designed using these indicators.

2.1 Audio-based indicators during co-located collaboration

Different types of verbal and non-verbal indicators have been used in the past to measure collaboration quality ranging from tangible interaction, audio-based cues, gesture, posture to gaze and eye interaction [3]. For the scope of this paper, we focus on the most abundantly occurring modality during CC, i.e., audio cues. Lubold and Pon-Barry [6] found that *proximity, convergence* and *synchrony* are different types of coordination (or rapport) cues obtained from the audio features (like intensity, pitch and jitter) of the collaborating dyads.

Bassiou et al. [2] assessed collaboration among students using *non-lexical speech* features. Types of collaboration levels marked are: Good (all 3 members are working together and contributing to the discussion), Cold (only two members are working together), Follow (one leader is not integrating the whole group) and Not (everyone is working independently). This coding was based on two types of engagement: simple (i.e., talking and paying attention) and intellectual (i.e., actively engaged in the conversation). Combination of both the speech-activity features (i.e., *solo duration*, *overlap duration of two persons*, *overlap duration of all three persons*) and speaker-based features (i.e., *spectral*, *temporal*, *prosodic* and *tonal* features of speech) were good predictors of collaboration. Speaking time of each member can also be a good indicator of collaboration [1].

2.2 Feedback based on audio cues during co-located collaboration

Simpler versions of feedback which leverage the audio cues (like speaking time) during collaboration have proved effective in the past. For instance, Bachour et al. [1] reflected back the speaking time of each group member using a real-time feedback during CC by glowing different coloured LED lights on a smart table. According to this study, the real-time feedback helped to maintain the equity of participation. Similarly, Praharaj et al. [8] reflected back the speaking time of each member on the fly by a multicoloured line chart. Tausch et al. [10] used an intuitive metaphorical feedback moderated by human observers during *collaborative brainstorming*. The group members brainstormed on a certain topic and their collaboration was measured by the number of ideas generated from the audio. Then the human observers controlled the public shared display which showed a *metaphorical garden*. This garden comprised of flower plants symbolizing the individual state of a group member and a tree symbolizing the state of the group which was well grown with fruits and flowers when a group had balanced participation.

In summary, most of these studies were in controlled conditions using specialized furniture and devices. Some real-time feedback mechanisms employed human observers (i.e., the non-automated ones), were simplistic (i.e., the automated ones) and acted as a mere reflection for the group to self-regulate instead of an actionable feedback; while others used a post-hoc analysis for the teachers (or facilitators) to reflect on the group activity.

3 Group Coach

Our technological infrastructure[4] is mostly driven by the need to analyze the audio input not disregarding the need for accommodating the input from other modalities in future. So, we looked at different solutions for automated audio analysis in real-time and the possibility of using different real-time feedback mechanisms to our maximum advantage. Finally we decided to use Unity over Audacity, Supercollider and Praat because of the readily-available real-time audio input analysis and game design interface support making it easy to design the automated real-time feedback. For the feedback in unity, we represent each group member as a tree which is nothing but a game object. For the audio input, we use one microphone for each group member and connect this to a single laptop which displayed the feedback using the Unity interface. We use different audio cues such as the speaking time, number of turns and change in loudness and map it to growth of different parts of the tree such as tree trunk, branches, leaves and flowers. Figure 1 and 2 show the growth of the tree at initial and late-mid stages respectively. Inspired from previous works [10], we chose metaphorical feedback which is easier for group members to understand and act as they can associate it with their day-to-day life or surroundings.

4 Use-case for the Group Coach

We tested this prototype design in different types of meetings and continued refining the design further based on the feedback of the stakeholders. We will expand this prototype in future work using video modality in other collaboration settings such as collaborative programming and engineering design in order to test its usability. We will test it further in future meetings with varying numbers of participant ranging between 2 and 6 members for large scale adoption.

[4] https://github.com/sambit2/GroupCoachCC

Fig. 1. Initial tree

Fig. 2. Late-mid tree

5 Conclusions

We succeeded in designing an automated real-time feedback prototype. Its design will be further refined in follow-up studies based on the requirements of the different stakeholders that will be using the system.

References

1. Bachour, K., Kaplan, F., Dillenbourg, P.: An interactive table for supporting participation balance in face-to-face collaborative learning. IEEE Transactions on Learning Technologies **3**(3), 203–213 (2010)
2. Bassiou, N., Tsiartas, A., Smith, J., Bratt, H., Richey, C., Shriberg, E., D'Angelo, C., Alozie, N.: Privacy-preserving speech analytics for automatic assessment of student collaboration. In: INTERSPEECH. pp. 888–892 (2016)
3. Cukurova, M., Luckin, R., Mavrikis, M., Millán, E.: Machine and human observable differences in groups collaborative problem-solving behaviours. In: Proc. of the 12th EC-TEL. pp. 17–29. Springer (2017)
4. Grover, S., Bienkowski, M., Tamrakar, A., Siddiquie, B., Salter, D., Divakaran, A.: Multimodal analytics to study collaborative problem solving in pair programming. In: Proceedings of the 6th Intl. Conf. on LAK. pp. 516–517. ACM (2016)
5. Kim, T., Chang, A., Holland, L., Pentland, A.S.: Meeting mediator: enhancing group collaboration using sociometric feedback. In: Proc. of the 2008 ACM Conf. on CSCW. pp. 457–466. ACM (2008)
6. Lubold, N., Pon-Barry, H.: Acoustic-prosodic entrainment and rapport in collaborative learning dialogues. In: Proc. of the 2014 ACM WS on Multimodal Learning Analytics Workshop and Grand Challenge. pp. 5–12. ACM (2014)
7. Praharaj, S., Scheffel, M., Drachsler, H., Specht, M.: Multifocus: Multimodal learning analytics for co-located collaboration understanding and support. In: EC-TEL (Doctoral Consortium) (2018)
8. Praharaj, S., Scheffel, M., Drachsler, H., Specht, M.: Multimodal analytics for real-time feedback in co-located collaboration. In: European Conference on Technology Enhanced Learning. pp. 187–201. Springer (2018)
9. Stahl, G., Law, N., Hesse, F.: Reigniting cscl flash themes. International Journal of CSCL **8**(4), 369–374 (2013)
10. Tausch, S., Hausen, D., Kosan, I., Raltchev, A., Hussmann, H.: Groupgarden: supporting brainstorming through a metaphorical group mirror on table or wall. In: Proc. of the 8th Nordic Conference on HCI. pp. 541–550. ACM (2014)
11. Terken, J., Sturm, J.: Multimodal support for social dynamics in co-located meetings. Personal and Ubiquitous Computing **14**(8), 703–714 (2010)