Four Challenges in Crafting Multimodal Collaboration Analytics for non-Data experts

Gloria Fernandez Nieto, Kirsty Kitto and Roberto Martinez Maldonado
University of Technology Sydney, Australia
Gloria.m.Fernandez-Nieto@student.uts.edu.au
Roberto.Martinez-Maldonado@uts.edu.au
Kirsty.Kitto@uts.edu.au

Abstract: The Multimodal Learning Analytics (MMLA) community of research is growing fast and considerable effort is being put on creating technological innovations to automatically analyse collocated collaboration and teamwork in physical spaces. However, the MMLA field is new and considerable work has been focused on solving technical problems, such as data collection, storage and synchronisation, with the purpose of performing tailored analysis. Much work still needs to be done for the community in order to start generating educationally meaningful interfaces to support group activity. This poster identifies potential gaps in R&D that the community may want to focus on in the near future. These are presented in the form of challenges that researchers may face while designing MMLA solutions aimed at being used by non-data experts.

Keywords: collocated teamwork, Multimodal Learning Analytics, integration, data collection, multimodal data meaning, interfaces, automatic assessment and feedback.

Introduction
As education providers increasingly focus on teaching 21st century competencies (De Lima & De Souza, 2017; Shum & Crick, 2016), the question of how we might effectively support students in developing the teamwork and collaboration skills that they can apply in authentic settings has become a topic of utmost importance. Capturing digital traces of teamwork and group activity to provide personalised support in online spaces (e.g. from wikis, chats or forums) has received considerable attention (see meta-analysis by Sottilare et al., 2018). By contrast, researchers have paid less attention to teamwork in face-to-face (collocated) scenarios (Chua, Y. H. V., Rajalingam, P., Tan, S. C., Dauwels, 2019). This is in part because it is harder to automatically collect evidence that can be used to model collaborated collaboration. Multimodal Learning Analytics (MMLA) is an emerging area of research that is becoming a promising alternative to understand group processes in situations that have been for a long time considered ephemeral, such as collaboration in the classroom and high effective collocated teamwork in the workplace. Inroads have been made into (i) understanding how team/group behavior is connected to performance and learning outcomes (Richardson & Dale, 2005; B. Schneider & Pea, 2014; B. Schneider et al., 2015); and (ii) exploring the potential of using multimodal learning analytics to find patterns that can be used to personalise instruction or identify potential problems (Blikstein & Worsley, 2016; Martinez-Maldonado, Clayphan, Yacef, & Kay, 2015; Mason, Pluchino, Tornatora, & Ariasi, 2013; B. Schneider & Blikstein, 2014; Sottilare et al., 2018).

However, materialising MMLA solutions for non-data experts (interfaces aimed at being used by teachers or learners who are not necessarily familiarised with interacting with data representations) to automatically provide feedback or assess collocated teamwork brings new challenges. These challenges may not only involve technical aspects, but also human factors (how can people interact with data?), pedagogical aspects (how can teachers appropriate data-intensive solutions into their practice?), learning aspects (how can learners reflect on evidence?) and the characteristics of the learning design (what is the meaning of data in a particular context?). Although the MMLA community is acknowledging some of these challenges, this poster summarises what we consider are the critical challenges to be addressed in order to advance in the design of effective MMLA innovations particularly tailored to support (non-data experts) in the context of collocated teamwork and group activity.
The Challenges

The challenge of creating effective multimodal collaboration interfaces for non-experts
Collocated collaboration and teamwork is a complex process which is, to a great extent, ephemeral and quite invisible for computational analysis (Martinez-Maldonado et al., 2019). This means that collaborators can communicate in multiple (often subtle) modalities that can be critical but hard to capture for algorithms to extract features that could be used for generating personalised actions or feedback. Evidence about collaboration can be captured via different sensors, interactive devices and/or human observations. Yet, the capture of multiple data streams from multiple modalities and group members can result in quite complex interfaces which sometimes require experts to use them (J. Schneider, Di Mitri, Limbu, & Drachsler, 2018). In many cases, some MMLA innovations do not even provide any interface. Most interfaces that provide automated feedback to teamwork are limited to mirroring quantitative information, such as quantifying the amount of speech by each participant in a meeting (MacNeil, Kiefer, Thompson, Takle, & Latulipe, 2019).

MMLA researchers are getting ready to design and build multimodal interfaces to support groups. Yet, more work still needs to be done to tackle the complexity of showing multiple streams of data to non-data experts. There are two recent examples that have address this problem. Echeverria et al. (2019) presented four interfaces that show information automatically captured while teams of nurses participate in training sessions with simulated patients. The four interfaces present information about one modality each, namely conversation patterns, positioning around the patient, arousal peaks and critical actions performed. The challenge here is how to fuse the information into a single interface that can facilitate reflection instead of having one interface per modality. This has been attempted by Ochoa et al. (2018) who presented the multimodal transcript to present automatic feedback to students on basic oral presentation skills analyzing posture, gaze, volume, filled pauses and the slides of the presentation. However, more work needs to be done to assess the complexity of this interface and to investigate alternative ways in which multimodal data can be compressed and visualised. Recent work by Vujovic & Hernández Leo, (2019) is precisely looking at ways in which multiple streams of sensor data can be segmented and visualised minimalistically.

The challenge of multimodal data sensemaking
There is little work in MMLA focused on understanding how non-data experts can make sense of intertwined multimodal streams of data (Echeverria et al., 2019). The use of data in education, and the social sciences in general, requires mix-method methodologies so that multimodal data insights can be associated to constructs that humans can easily interpret (Burke Johnson & Onwuegbuzie, 1963; Onwuegbuzie & Leech, 2005). Some work in MMLA has highlighted the importance of annotating multimodal data to find meaningful patterns (Mitri, Schneider, Klemke, Specht, & Drachsler, 2019). Worsley and Blikstein (2018) suggested the use of the notion of epistemic frames as a way to characterise certain combinations of aspects of the learner (e.g. posture and gaze) that can be tracked using sensors. Echeverria et al. (2019) proposes the need for a careful mapping from multimodal data to High-Order Constructs (HOCs) to give meaning to multimodal data during the design process rather than in-use.

Only the latter of these three examples mentioned above refers explicitly to collaborative settings. Finding new approaches to support non-data experts in making sense and understanding multimodal data about multiple collaborators is a required step towards the creation of new ways to facilitate the provision of feedback and assessment for collocated collaboration and teamwork.

The challenge of multimodal data integration
The data integration is a challenge that needs to be addressed. Nowadays, researchers have adopted their own in-house integration and data fusion approaches. Examples of the latest integration architectures include the Learning hub (J. Schneider et al., 2018) the adoption of the Synchronous Serial Interface protocol (SSI) (Martin et. al., 2019) and EduBrowser (Chua et. el., 2019). Despite the MMLA community is proposing different architectures, pipelines or
other approaches to integrate data coming from different sensors, a critical question remains: How can the community move beyond solving the technical problem of data integration to more rapidly focus on educational problems? Possible solutions may include to partner with industry to adopt widely known integration protocols or all researchers may want to adopt one specific approach. Perhaps, since collocated teamwork is a complex process, there is the possibility that the only solution is for researchers to keep creating tailored solutions. The integration problem (often called interoperability problem) is not unique for MMLA, but it is much more critical for Collaboration Analytics innovations in which multiple sensors are often required to capture information about multiple students.

In sum, MMLA researchers should focus on the educational problems to be solved. However, for this to happen, the technical problem of interoperability still needs to be solved.

The challenge of focusing too much on sensors and multimodal data
It may be fair to say that most of current work in MMLA has adopted a bottom-up approach. This means that the focus has been posed on what can sensors tell us about collaboration or individual learning rather than identifying the higher-order constructs that are important in educational terms and the evidence needed to provide feedback or assess such constructs (Prieto et al., 2017). For the case of collocated teamwork, a bottom-up approach means that researchers may want to collect data about posture or movement using a Kinect sensor, voice or conversation patterns using microphone arrays (e.g. Matrix Voice, ReSpeaker etc.) and physiological data via wristbands. Then, researchers may use machine learning algorithms or statistical analysis approaches to classify behavioral patterns and associated them to group dynamics or learning constructs. In the end, this approach gives the main role to technology and leaves aside stakeholders and educational requirements.

By contrast, a top-down approach would focus on understanding the purpose of the research and prioritising the higher-order educational constructs over technology (Prieto et al., 2017). Then, researchers would invite stakeholder in the definition of the potential MMLA solutions and in identifying the sources of evidence that could help them to assess the important educational constructs. This way, technical aspects, such as sensors, data and software to be used to analyse such data, would be shaped based on the learning design and the actual needs of educators and learners.

If MMLA seeks to make impact on the provision of feedback or assessment for collocated teamwork, we believe that a top-down approach (or a combination of approaches) need to be adopted to consider the actual requirements of different stakeholders and the authentic dynamics of the setting in which the MMLA solution will be used. In this case, educational purposes drive data collection, use of applications, tools and techniques.

Final Remarks
The MMLA field is new and considerable work has been focused on solving technical problems, such as data collection, storage and synchronisation, with the purpose of performing tailored analysis. Much work still needs to be done in order for the community to start generating educationally meaningful interfaces to support group activity. This poster identified potential gaps in research and development that the community may want to focus on in the near future. Four challenges that researchers may face while designing MMLA solutions were presented. These point at the potential future avenues of research if we want to design effective MMLA interfaces aimed at being used by non-data experts (e.g. teachers and learners).

References