The Multimodal Study of Blended Learning Using Mixed Sources: Dataset and Challenges of the SpeakUp Case

María Jesús Rodríguez-Triana[1,2], Luis P. Prieto[2], Adrian Holzer[1], Denis Gillet[1]

[1] École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
[2] Tallinn University, Tallinn, Estonia
maria.rodrigueztriana@epfl.ch, luis.prieto@tlu.ee, adrian.holzer@epfl.ch, denis.gillet@epfl.ch

Abstract. Social media applications have been proposed as a tool to complement students' formal learning experiences, often to increase interactivity and participation. However, evidence regarding the benefits and challenges of such applications is still conflicting. In our latest study to explore this conundrum, we have gathered a multimodal dataset that showcases the teaching and learning processes co-occurring simultaneously on a physical space (face-to-face university lectures) and a digital one (SpeakUp, a social media app). The raw data, provided by different sources and informants, were transformed and analyzed using mixed (quantitative and qualitative) techniques. In this contribution, we describe the multiple pieces that composed our dataset, and the steps we took in the multimodal analyses to explore the learning experience occurring in both the physical and digital spaces. This dataset and analysis pipeline illustrates not only challenges and limitations specific to our study, but also more general ones. Several such challenges and limitations, commonplace in blended learning settings analyzed using mixed (multimodal) methods, are synthesized at the end of our paper.

Keywords: Multimodal learning analytics (MMLA); blended learning; mixed methods; social media

1 Introduction

Technology-enhanced learning (TEL) is, almost invariably, blended in nature: we create new digital spaces and channels to interact and learn – yet we still inhabit the physical world and also learn through it. This inherently blended nature of learning not only has prompted a methodological turn towards mixed methods [5], but also holds great promises for the rise of multimodal learning analytics (MMLA), as researchers strive to understand more deeply the learning processes and outcomes occurring in both kinds of spaces [2].

One example of such research is our ongoing project to study the usage of social media applications to complement formal, co-located learning experiences (e.g., face-to-face university courses). Currently there is no consensus as to whether the benefits that such applications provide in terms of engagement and interaction, outweigh their potential cost as a source of distraction [6, 4, 7]. To help in clarifying these issues, we are performing a case study in an authentic setting, one of our university courses at the École Polytechnique Fédérale de Lausanne (EPFL) in Switzerland.

In this face-to-face course, composed mainly of lectures with more than a hundred students, the usage of a social media tool (SpeakUp[3]) was proposed in order to foster the (otherwise limited) interaction between students and with the instructors. In a typical usage scenario with SpeakUp, teachers create a chatroom that students can join. Inside the chatroom, any user can anonymously post text messages, comment on existing messages, and up/down-vote them. Moreover, SpeakUp had an added value from the analytics perspective: it allows the chatroom creator to download all the traces collected inside of the room, including both actions and posts.

This paper provides an overview of the multimodal dataset gathered and the analyses performed in the study. A more detailed account of the setting, and a partial analysis of the data (including the effects of SpeakUp usage on student engagement, distraction, social interaction and teaching style, as well as their relationship with learning outcomes), are described in [3]. We end our contribution by reflecting on the limitations and challenges that we have faced, as they pertain to those common in the multimodal study of blended learning situations through a mix of quantitative and qualitative analyses.

2 Multimodal Dataset

The dataset for our study was gathered throughout a whole university course on Communication, which involved 6 face-to-face sessions which lasted 90 minutes each. An average of 3 lecturers and 145 students attended per session. We combined quantitative and qualitative data coming from four types of informants (3 teachers, 145 students, 4 assistants, 1 researcher, plus the SpeakUp system itself) using different data gathering techniques, namely: questionnaires, observations, video recordings and system logs. Figure 1 offers an overview of the dataset. In each session, the following data were gathered:

– The researcher video recorded the session (focusing mostly on the front of the class) [R_VID]
– The researcher also wrote down timestamped observations about what was happening during the session (e.g., beginning, start, breaks, activities, topics discussed, interventions, problems with the app, number of students in the room) [R_OBS]
– In parallel, the assistants involved in the course kept track of the students who participated face-to-face in the session (e.g., posing questions or participating in the general discussions) [A_OBS]

[3] http://speakup.info

– SpeakUp logs were used to track the activity of the instructors and students joining the chatroom[4] (e.g., timestamped posted messages, number of likes and dislikes, etc.) [SP_LOG]

Informants	Data gathering technique	Data source	Data Type	Content	Analyses method	Analyses technique
Researcher	Video recording	[R_VID]	Qualitative	Session recording	Mixed	Video coding (Manual)
						Descriptive statistics
	Observations	[R_OBS]	Qualitative	Timestamped classroom events	Quantitative	Descriptive statistics
Assistants	Observations	[A_OBS]	Quantitative	Student f2f interventions	Quantitative	Descriptive statistics
SpeakUp	Logs	[SP_LOG]	Mixed	Timestamped actions & posts	Mixed	Descriptive statistics
						Exploratory computational analyses
						Message coding (Manual)
Teachers	Questionnaire	[T_QUE]	Quantitative	Engagement & attention perception (Likert)	Quantitative	Descriptive statistics
Students	Questionnaire	[S_QUE1]	Quantitative	Predisposition towards social media (Likert)	Quantitative	Descriptive statistics
	Questionnaire	[S_QUE2]	Quantitative	Engagement & attention perception (Likert)	Quantitative	Descriptive statistics
	Test Scores	[S_SCO]	Quantitative	Multiple choice results	Mixed	Content extraction (Manual)
						Descriptive statistics

Fig. 1. Overview of the informants involved in the data collection as well as the data gathering and analyses techniques.

In addition, there were a number of *complementary* data sources collected at the beginning and at the end of the course:

– To understand the student predisposition towards the usage of technology and social media for learning, we conducted a questionnaire at the beginning of the first session (based on 7-point Likert scale questions) [S_QUE1].
– To gather the student and teacher perceptions on how the tool usage affected the engagement and attention, we conducted questionnaires at the end of the first session [S_QUE2] [T_QUE]. While the former was made of 7-point Likert scale questions, the later combined 5-point Likert scale and open questions.

[4] Since in SpeakUp, users join anonymously the chatrooms, there was no way to figure out who was behind each user identifier. Thus, we asked the students to freely reveal their identities just for research purposes. It is noteworthy that SpeakUp logs are multimodal since they contain not only activity traces but also the text posted by the users.

– At the end of the course, students answered a test composed of multiple-choice questions about the topics discussed in the different sessions. The scores [S_SCO] were used for exploring whether the user's actions on SpeakUp during the topic discussions had an impact on the student answers.

It is noteworthy that, although the study was carried out in an university course on Communication, the data sources used and the data gathering techniques applied are not dependant to the content or the educational context. Thus, the same strategies and techniques could be applied in other settings.

3 Multimodal Data Analysis

As detailed in [3], a first partial analysis was performed by transforming and integrating the data from the first session only. This first analysis involved both qualitative analyses (manual coding of the actions recorded in the videos, and content analysis of the messages generated by the users), as well as quantitative ones (descriptive statistics and exploratory computational analyses of system logs). Figure 1 offers an overview of the different analyses applied to the dataset.

Qualitative analyses. We manually coded all the messages and comments generated during the lesson (thus enriching [SP_LOG]) to determine whether they were relevant to learning, and the direction of the interaction (students to teachers, students to students, students to all, and teachers to students). In a similar way, and in order to understand these topics as they occurred in the face-to-face channel of the classroom, the video recording of the lesson was also coded (enriching [R_VID]), according to several categories: which actor was speaking; what topic to appear in the scoring test [S_SCO] was being discussed, if any; what teacher action was being performed at that moment (e.g., presentation/lecturing, asking questions, providing answers, noting technical or other kinds of problems); who was the target of the interaction, if any (e.g., a teacher, students, or all the class); and finally, what supporting resources were being used, if any (e.g., slides, videos, SpeakUp).

Quantitative analyses. The user activity was measured applying descriptive statistical analyses to the actions tracked in the SpeakUp logs (e.g., number of posted messages, number of likes and dislikes, etc.) [SP_LOG], to the face-to-face utterances registered by the researcher and the assistants [R_OBS] [A_OBS], and to the results of the manually-coded video [R_VID] and SpeakUp comments [SP_LOG]. Furthermore, a clustering analysis (using a k-means algorithm) was performed on each student's activity features [SP_LOG] (number of messages, responses, likes/dislikes, etc.), in order to identify usage profiles. Finally, the data from the users' activity [SP_LOG] were triangulated with the teachers' and students' perceptions from the questionnaires [T_QUE] [S_QUE1] [S_QUE2] and the scores obtained by the students in the final test [S_SCO], to understand the impact of such engagement and participation in the learning outcomes.

4 Limitations and Challenges

Despite the richness of the dataset and the usefulness of the analyses described in the previous sections, it presents several limitations, especially apparent in terms of reproducibility and scalability of our approach. While some of these limitations are specific to the particular implementation of our study, others represent widespread challenges in MMLA that tries to study blended learning settings using a mix of quantitative and qualitative techniques:

Manual data gathering. The fact that several of our data sources originate directly from manual work by human actors (e.g., observations by researchers or assistants). The lack of tooling to easily (and consistently) timestamp, label and export such manual data poses limitations to the scaling and diffusion of this kind of efforts.

Data integration. In our study, different units of analysis or measure were used. For instance, while the videos allowed us to measure the length of the face-to-face interventions, the logs informed us only about discrete computer-mediated events, without a duration. This makes merging and comparison difficult, and illustrates a common issue when using multimodal datasets: the heterogeneous nature of the different data sources. Despite the efforts put in order to adopt interoperable standards and specifications (like Caliper or xAPI), this heterogeneity will be hardly avoidable, requiring multiple analysis techniques.

User identification and anonymization. The usage of multiple data sources entails the need to identify a certain user across data sources. Computer-mediated user actions are often easy to trace; however, in video or audio data this can be a challenging (if done automatically) or cumbersome task (if done manually). In our particular study, the situation was even more complicated because both the questionnaires and the system logs provided by SpeakUp were anonymous. This brings up the tension between user traceability and privacy, which will be brought to the forefront by the requirements of the recent European General Data Protection Regulation (GDPR EU 2016/679). This kind of regulation may lead, in the near future, to technological tools that only expose anonymous data (hence hindering learning analytics and interventions that target specific students).

Manual data analysis. As it happens with data gathering, the need for human involvement during the analysis limits the scalability of our approach. In our study, both comments and videos were manually analysed by teachers and/or researchers. Although there are potential solutions that could facilitate the manual content analyses (e.g., crowdsourcing by letting the users tag themselves the comments), others like the video analyses remain still a challenge. We envision that alternative MMLA techniques and approaches (such as speech or text analyses [1]) could ameliorate the aforementioned limitations, and contribute to avert or circumvent the MMLA challenges in the mixed-method study of blended learning phenomena. For example, voice recognition techniques could automatize part of the video analyses, identifying the different speakers participating during the session. In addition, transcriptions could be automatically generated by applying speech recognition to the audio recorded during the sessions. Later on, content analyses could be applied to the transcriptions, the comments post by the students, and the questions of the test, to explore the relations among them. However, it should be noticed that, to put all these ideas in action, further research would be necessary to provide speech recognition and content analyses solutions applicable to different languages.

5 Acknowledgements

This work was partially funded by the European Union in the context of the Go-Lab Integrated Project (FP7-ICT action, grant no. 317601), SiWay (FP7-ICT, IMAILE grant no. 619231-459.76301), the Next-Lab Innovation Action (H2020-ICT, grant no. 731685), and CEITER (H2020-WIDESPREAD-2014-2, CSA action, grant no.669074)

References

1. P. Blikstein and M. Worsley. Multimodal Learning Analytics and Education Data Mining: Using Computational Technologies to Measure Complex Learning Tasks. *Journal of Learning Analytics*, 3(2):220–238, 2016.
2. X. Ochoa and M. Worsley. Editorial: Augmenting Learning Analytics with Multimodal Sensory Data. *Journal of Learning Analytics*, 3(2):213–219, 2016.
3. M. J. Rodríguez-Triana, A. Holzer, L. P. Prieto, and D. Gillet. Examining the effects of social media in co-located classrooms: A case study based on speakup. In K. Verbert, M. Sharples, and T. Klobučar, editors, *11th European Conference on Technology Enhanced Learning: Adaptive and Adaptable Learning*, volume 9891, pages 247–262, Lyon (France), 2016. Springer International Publishing.
4. F. Sana, T. Weston, and N. J. Cepeda. Laptop multitasking hinders classroom learning for both users and nearby peers. *Computers & Education*, 62:24–31, 2013.
5. J.-W. Strijbos and F. Fischer. Methodological challenges for collaborative learning research. *Learning and Instruction*, 17(4):389–393, 2007.
6. Y.-T. Sung, K.-E. Chang, and T.-C. Liu. The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis. *Computers & Education*, 94:252–275, 2016.
7. D. L. Worthington and D. G. Levasseur. To provide or not to provide course powerpoint slides? the impact of instructor-provided slides upon student attendance and performance. *Computers & Education*, 85:14–22, 2015.