Instruction, student engagement and learning outcomes: A case study using social media in co-located classrooms

María Jesús Rodríguez-Triana, Luis P. Prieto, Adrian Holzer, and Denis Gillet *Senior Member, IEEE*

Abstract—With the wide availability of mobile devices and the growing interest about social media platforms, numerous mobile applications have emerged to support student engagement in the classroom. There is conflicting evidence, however, on whether the engagement benefits of such applications outweigh their potential cost as a source of disaffection. Social media apps have been especially questioned in this regard. To investigate these issues, this paper presents a case study on the usage of a social media app (SpeakUp) during a semester-long co-located university course, and its relations with the context and the learning outcomes. In this mixed-methods study, we gathered data from multiple sources (video recordings of the lessons, SpeakUp logs and messages, student questionnaires and course assessments) in order to extract self-reported and observable behavioral and emotional indicators. Our findings reveal that simple measures of behavioral engagement were insufficient to predict academic performance. Nevertheless, our models significantly improved using relatively simple and unobtrusive indicators of both behavioral and emotional engagement and disaffection. These results emphasize the role that teachers play in the effective use of social media in the classroom, and how important it is to design learning activities that raise engagement while avoiding disaffection.

Index Terms—Social networking, Collaborative Learning Tools, Computer-assisted instruction

1 Introduction

The usage of social media apps such as Facebook, Twitter or WhatsApp, is widely adopted among students in their personal life. Seeing the momentum, there have been several attempts in the educational research community to reach students through these platforms [1], [2] or propose alternative, education-oriented tools with similar features (e.g., Edmodo, Twiducate or SpeakUp).

Social media tools have been used inside and outside the classroom, especially to foster interaction between teachers and students. While both students [3], [4], [5], [6], [7], [8] and teachers [9] perceive improvements in participation, there is still conflicting evidence about their impact on learning [10], [11]. Additionally, as it happens with other educational technologies and innovations, the mere introduction of social media in the teaching practice does not guarantee a positive effect [12], [13]. Thus, to determine the impact that social media has on learning outcomes, and to understand how instructional practices facilitate effective engagement, we need useful measures of student engagement.

To contribute to the understanding of how social media can be used effectively in the classroom, this paper explores how one such tool can be integrated effectively in a co-located classroom. SpeakUp [14] is a mobile app designed to promote student participation in face-to-face sessions. In SpeakUp, students can anonymously join chatrooms, post messages, comment, like or dislike them, as well as answer polls (multiple choice questions) set by the teachers.

This paper presents a case study where SpeakUp was used in an authentic learning scenario carried out with 149 university students and 3 teachers during one semester. Following the model proposed by [15], this case study explores the relation between context and actions, as well as actions and outcomes. In our case, the context refers to the teacher instruction (in terms of instruction style, interaction type and content) that frames the learning activity. The actions represent the student engagement and disaffection with the learning activity (both at behavioral and emotional level), observed through the tool logs and face-to-face during the lessons. Finally, the outcomes are the student answers in the exam and the final score in the course. As part of our analysis, we explore whether simple indicators of behavioral engagement predict reliably learning outcomes, or rather we need more complex constructs that distinguish behavioral and emotional engagement and disaffection [15], [16].

The case study methodology [17] guided the data gathering and analyses, leading us to use multiple informants (students, teachers, researchers, face-to-face observers, video coders, and the technology used), different data gathering techniques (observations, questionnaires, video recording, SpeakUp logs, and user comments in the app), and mixed methods analyses, to process evidence from the digital and the physical space (e.g., attendance to the session, teacher and student participation, face-to-face and via SpeakUp, content and length of the comments, as well as teacher and student perceptions about the impact on engagement).

- *María Jesús Rodríguez-Triana and L.P. Prieto are with Tallinn University (Estonia). E-mail: mjrt@tlu.ee, lprisan@tlu.ee*
- *A. Holzer is with university of Neuchâtel (Switzerland). E-mail: adrian.holzer@unine.ch*
- *D. Gillet is with École Polytechnique Fédérale de Lausanne (Switzerland). Email: denis.gillet@epfl.ch*

Abstract submitted in March 2019

2 Related Work

2.1 Student engagement

Given the current emphasis on student success and dropout rates in formal education [18], [19], understanding and promoting student engagement has become an important issue among practitioners and researchers in the educational community, considering it an indicator of intrinsic will to learn [20]. As Henrie et al. explain in their review [21], several definitions have been provided about engagement in the literature. The early definition provided by Astin –who conceived it as 'the amount of physical and psychological energy that the student devotes to the academic experience' [22]– has later evolved, connecting the time and effort invested by the students to the learning activities and the desired learning outcomes [23].

Although closely related, some authors distinguish engagement from disaffection (also known as alienation) [15], [16], looking at the occurrence of behaviors and emotions that reflect maladaptive motivational states. Engagement is considered a strong predictor of student learning, grades, achievement, and school retention. Symmetrically, disaffection has been found to be a strong predictor of poor grades, low achievement in test scores, and eventual dropout. Thus, several authors proposed the engagement–disaffection dyad as a more useful framework to understand students' relationships to their learning outcome [24], [25], [26].

Fredricks et al. [27] described three main types of engagement: behavioral, emotional, and cognitive. According to these authors, behavioral engagement refers to students' effort, attention, and persistence during the learning activities. Emotional engagement includes both feelings learners have about their learning experience, such as interest, frustration, or boredom, and their social connection with others at school. Cognitive engagement is the focused effort learners give to effectively understand what is being taught (e.g., student's involvement in planning, monitoring, and evaluation when accomplishing tasks). Also disaffection presents both behavioral, emotional, and cognitive components [26]. Behavioral disaffection includes passivity and withdrawal from participating in learning activities. Emotional disaffection spans boredom, anxiety, and frustration states in the classroom. Finally, cognitive disaffection is illustrated by, e.g., aimlessness, resignation, apathy or pressure.

Around the concepts of engagement and disaffection, Skinner [26] distinguish among *facilitators* (i.e., factors that influence engagement and disaffection such as the social context, self features, and experience), *indicators* (i.e., student actions that can be observed in the classroom), and *consequences* of engagement and disaffection (namely, learning outcomes and student success).

Among the 113 reviewed papers studied by Henrie et al. in their review [21], 77% operationalized engagement from a behavioral perspective, using indicators such as participation, attendance, assignments completed, or time logged in order to measure on-task behaviors. Technology-mediated learning settings offer the chance of measuring additional behavioral engagement indicators based on the functionalities available in the system (e.g., completing or submitting assignments, logging-in/out, posting, responding, voting or viewing), the learning material available, and the time devoted using the aforementioned functionalities and resources. Cognitive engagement was measured in 43.4% of the cases using qualitative measures inferred from the student-created artifacts. Emotional engagement indicators (in 40.7% of papers) included positive or negative emotions towards learning or the context. Out of those 113 papers, 43%, 36% and 21% of the papers measured one, two or three types of engagement, respectively. Regarding data collection, three main strategies have been followed [21]:

Quantitative self-report. Surveys are useful and frequently used for investigating aspects of student engagement which are not easily observable (especially, around emotional or cognitive engagement) [20], [28], [29], [30]. While this approach is in general more scalable than human observations, surveys tend to require significant effort from students, and may be disruptive and intrusive within educational practice.

Qualitative observations. These methods include direct or video observations of students' behavior while learning. Interviews, focus groups, and the analysis of digital content extracted from the tools used (to extract behaviors and themes), are also included here. While useful for exploratory studies, qualitative measures are difficult to scale since extensive (human) resources may be needed to collect and analyze the data. To overcome this limitation, Multimodal Learning Analytics solutions could be used to automatize part of this process [31], [32], e.g., contributing to the detection of emotions.

Quantitative observations. Frequency indicators have been obtained from direct human and digital observations, or based on video recordings. Only 10 papers in Henrie's review collected data automatically. Observational methods, which include both frequency (number of on/off-task behaviors) and qualitative measures, have the advantage of enabling researchers to measure engagement and disaffection during the learning process without being disruptive, making them especially suitable for observations at the activity level. However, quantitative measures may limit the aspects of engagement and disaffection that can be studied.

The combination of both quantitative and qualitative strategies could offer the possibility of using qualitative measures for describing engagement and disaffection, and quantitative measures to establish comparisons between individuals or groups and to see their progress over time [15]. Along the same lines, Chapmann's overview provides teachers with a list of instruments that they can use to assess engagement on a class-wide basis [25].

2.2 Student engagement and social media

ICT researchers and other providers have proposed technology-mediated innovative practices that may have a positive impact on student engagement. Indeed, several studies show significant correlations between the use of educational technology and student engagement [23], [33], [34]. It is worth noting that many of these studies take place in contexts where the students' use of technology can almost exclusively be on-tasks (e.g., a university's learning management system provides less opportunities for off-task behavior and distraction than other technologies do). Hence, it is not surprising how these studies find that more interactions with the technology are related to better learning outcomes (e.g., [35]).

Among those practices, fostering social interaction in the classroom is considered by numerous researchers as a *conditio sine qua non* for learning [36], [37]. Since the 80's, when IBM started to experiment with student interaction systems [38], this idea has prompted the apparition of a myriad of digital solutions to support student interaction. Many of these systems are based on reactive interaction where teachers can conduct live polls by asking multiple-choice questions and students answer by pressing a button on a clicker. Studies on clickers show that they can foster more participation in the classroom, and that students generally have a positive attitude towards them (e.g., [5], [39], [40], [41]). On top of such a reactive channel, some systems provide a proactive channel, where students can post questions and comments. With the rise of mobile devices, systems also started relying on the students' own devices. An early effort in this direction was the TXT-2-LRN mobile system [8], with which students could send free-form SMSs to the teachers.

More recently, many other educational technologies also have included a social media layer, where students can vote and comment on each other's contributions (e.g., Class-Commons [42], Fragmented Social Mirror [4], Pigeonhole Live [43], Backchan.nl [6], or SpeakUp [7]). Mainstream social media, such as Twitter [2], [13], [44], Facebook [1], and Reddit [45], are also popular when attempting to foster interaction between speakers and their audience in both conferences and classrooms. Research investigating the use of such social media applications in the classroom generally concludes that students perceive such systems positively, and that they feel it increases interactivity [3], [4], [5], [6], [7], [8]. Furthermore, students often prefer to use a digital channel to interact instead of raising their hand [8]. Often, these studies rely on questionnaire instruments to gather data about the student perception (e.g., the National Survey of Student Engagement – NSSE [46]), students' comments about their interventions, and relate them with the learning outcomes (e.g., with the scores in the exams) [1], [13].

The Pearson education service company conducted a survey with 7969 U.S. higher education teachers to better understand the bigger picture of the social media usage by teachers [9]. The survey found that teachers are generally aware of social media and they are using it in their private lives (70.3% of faculty used it at least once per month). Even if the use of social media in the classroom lags behind the usage in their personal lives (41%), it is increasing every year. Teachers see social media and technology as having a "considerable potential" for learning, and a majority (78.9%) states that digital communication has increased communication with students. However, 56% of teachers also consider that social media in class can be more distracting than helpful.

Beyond student and teacher perceptions, the issue of multitasking is receiving increased attention, with conflicting results in the educational research community. Certain studies suggest that laptop multitasking hinders learning for both users and nearby peers [10]. On the other hand, a recent meta-analysis on the use of mobile devices in the classroom nuances these claims, and shows moderate positive learning effects [47]. Other researchers argue that it is possible to take advantage of social media in the classroom by embracing multitasking, which students seem to be able to do effectively in the classroom [48], [49].

In contrast with what normally happens in a Learning Management System (LMS), when students interact with social media platforms, both on- and off-task activities are possible (or likely). Thus, when introducing these tools, we cannot assume a direct relationship between engagement and learning outcomes. Several authors agree that, despite the affordances of social media, simple incorporation of these solutions in the classroom may not always yield desired outcomes [12], [13], and that further research and involvement of practitioners is needed to untangle the complex relationships between engagement, distraction, and learning outcomes. The following section synthesizes advice from the literature on how to perform such research on student engagement and social media (which helped us define the design and methods of our study).

2.3 Assessing the role of social media in co-located classrooms

Echoing Chapman and Junco's claims in relation to the adoption of social media tools in educational settings [12], [25], we believe there is a need for (1) understanding the possibilities and limitations, and (2) measuring the impact, of social media on (formal) learning. This information would enable practitioners to properly introduce social media in their educational context, and to critically report their experience. In light of such need, this paper aims to contribute to the better understanding of whether, and under what circumstances, social media usage in the classroom may have a positive impact on learning. To guide this assessment, we extract below a number of guidelines from the related work presented in the previous sections:

- *Choose a social media platform that fits the educational goal* [12], [13]: among the different social media apps available in the market, the case study presented in this paper focuses on SpeakUp (see Section 3). This tool provides the functionality required for the learning activities (access without registration to an anonymous chatroom compatible with phone, tablets and laptops) and offers the possibility to download the data of all digital traces and content from the chatroom for later analysis.
- *Design the associated learning activities with specific outcomes in mind, which can be assessed* [50]: the different use cases for SpeakUp in the classroom were predefined before the study (see Section 5.2). Also, a multiple-choice exam was used at the end of the course to assess what students knew about the different topics presented in each of the lessons.
- *Conduct interventions utilizing social media over longer periods of time* [12], [15]: this paper presents the study carried out during a whole semester, so that the temporal dimension can be considered as part of the analysis (see Section 4.2), not only at different time granularity levels, but also longitudinally to understand the evolution of the course.
- *Both on-task and off-task activities should be taken into consideration* [13], [15]: the study takes into account both engagement and disaffection of the students by looking at the on-task and off-task comments

they have produced, and analyzing how they have reacted to others' (on- and off-task) comments. As depicted in Figures 2 and 3, our research questions and indicators are structured according to the model proposed by [15], which considers both aspects.

- *Gather quantitative and qualitative evidence* [51]: in relation to the previous point, this study takes into account not only the volume of observable measures, but also certain measures of their quality. To that aim, the comments were categorized, as illustrated in Table 1, and the student actions were mapped to the different types of engagement and disaffection, depending on whether they were on-task or not.
- *Engagement and disaffection are multidimensional constructs that require thorough analysis* [12], [15], [26], [27]: as it is shown in Figure 3, this study extracts indicators related to behavioral and emotional dimensions of engagement and disaffection that go beyond single behavior variables.
- *Engagement and disaffection happen in a context and generate an outcome* [15], [21], [26], [27]: Skinner's model [15] frames or our research questions (see Section 4). The current study explores the relation between the engagement and disaffection, contextual facilitators (in our case, the teacher instruction and the social level of the learning activity), and learning outcomes. Besides, to better understand the relation with the context and the learning outcomes, we evaluate different types of engagement and disaffection.
- *Carefully select which measures of student engagement and disaffection should be used to study the relationship with other variables* [21]: to better understand which indicators are relevant in relation to the context and the learning outcomes, we have run correlational and multiple linear regression analyses. Further details can be found in Sections 4.2 and 5.4.
- *Combine individual and social forms* [51], [52]: as Section 5.4 shows, the social aspect is considered both as a contextual factor (in relation to the individual and collaborative activities), and as a analysis dimension (looking both at the individual and at the classroom level).
- *Combine self-reported and observable indicators* [21]: as described in Section 4.1, the study combines student perceptions with observations made by researchers, observers and digital tools.
- *More research using computer-generated data should be done to better understand its value for studying student engagement and disaffection* [21]: In our case, SpeakUp traces are used to monitor the behavioral and emotional engagement and disaffection in the tool.
- *Avoid being disruptive and intrusive* [21]: apart from a poll and pre/post-surveys, observational methods were used to gather evidence while the learning process occurred. Besides, as mentioned before, learning activities happening in SpeakUp, were automatically collected to reduce the survey and (human) observation workload. A detailed description about the informants and data sources is provided in Section 4.1. Indeed, recent research shows that even the manual human labor involved in our analyses could be somehow automated (through natural language processing, machine learning and multimodal learning analytics) [53], [54], thus opening the door to engagement measurements for practitioners, without the need for a team of researchers in the room.

3 SpeakUp

SpeakUp[1] is a social media tool designed to foster participation in co-located situations where face-to-face interaction is difficult either within the audience or between the speaker and the audience (e.g., a university lecture with a large number of students, or a conference).

In a typical educational scenario with SpeakUp, teachers create a chatroom that students can join by typing its number, as shown in Figure 1.1. Inside the chatroom, teachers and students can post text messages, comment on existing messages, vote them (like or dislike, see Figure 1.2) or report them as spam. Each message has a relevance score, which shows the difference between the number of likes and dislikes. For instance, the top message in Figure 1.2 has a relevance score of 16 (a total of 20 votes, 18 likes and 2 dislikes) and the bottom message a score of 4. The chatroom creator (i.e., the teacher, in our case) can create multiple choice questions by pressing the '+' button on the bottom-left part of the screen, which leads to the question creation screen depicted in Figure 1.3. Teachers can give a title to their question and customize the number of choices with several other settings, such as if the results are displayed directly after students answer, or if teachers can actively show or hide results, and open or close the poll. Figure 1.4 shows how the results of the poll are displayed.

To use Speakup, participants simply need access to a device (phone or computer) connected to the Internet. No registration is required from either teachers or students, enabling an immediate use of the tool. Furthermore, aligned with Junco's view about anonymity in social media [12], in a chatroom all users are either anonymous or pseudonymous (depending on how the teacher configured it), fostering the expression of more uninhibited points of view. This implies that users interact, not directly with one another, but rather on the basis of the content posted by the different (pseudo-)anonymous users.

The participation in a lesson supported with SpeakUp can occur face-to-face (i.e., teachers and students interacting orally), as well as along the digital channel (i.e., posting comments and voting on SpeakUp). Moreover, the usage of the tool can be either spontaneous (e.g., students posing questions) or guided by the teacher. For instance, a teacher can instruct students to answer a poll on SpeakUp, or ask them to write down what they think about a certain topic.

As it was mentioned in Section 2, a crucial reason that led us towards the choosing SpeakUp was the access to the logged data. The owner of the room can download all actions and messages in a single CSV file. This feature, not so often available in social media tools, has an outstanding added value for practitioners and researchers willing to understand how the tool has been used.

1. SpeakUp: www.speakup.info

Fig. 1. Screenshots of the SpeakUp mobile app. (1) Joining a chatroom. (2) Viewing messages in the chatroom ordered by time or score. (3) Creating a multiple choice question in the chatroom. (4) Visualizing the results of a multiple choice question.

4 Methodology

The present study is framed within a wider research effort towards understanding how social media can be used effectively in the classroom. Several studies have been performed in the past on the use of SpeakUp in classrooms [7], [14], [55], pointing out that the tool was easy to use and motivated students to participate more in lectures. In turn, this paper explores how to measure engagement and disaffection in co-located settings mediated by this technology, and how these two variables are related to the context where the learning activity takes place, and the learning outcomes.

Our aim to achieve a deeper understanding of the usage of a social learning tool in a context as complex as classrooms are, led us to adopt a case study methodology [17]. This research methodology helps to inform practice by illustrating what has worked well, what has been achieved and what have been the multiple issues or dilemmas in a real-life scenario.

In order to understand how SpeakUp can be integrated effectively in a co-located classroom, this case study addresses three main research questions (as depicted in Figure 2):

- *RQ1*: How does a simple definition of action (as student behavioral engagement) relate to learning outcomes?
- *RQ2*: How does a more complex view of action (as behavioral and emotional engagement and disaffection) relate to learning outcomes?
- *RQ3*: What is the role that teacher instruction (in terms of instruction style, interaction type and content) plays on student action?

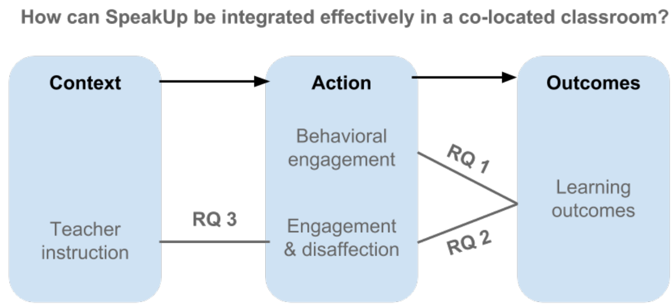

Fig. 2. Diagram representing the relation between the research questions and the theoretical model proposed by Skinner et al. [15].

4.1 Informants, data gathering, and data sources

Aligned with the trends in technology-enhanced learning research [56] and with the case-study methodology [17], we used mixed methods [57], [58] in order to look at our questions from different perspectives. More concretely, six types of informants (2 teachers, 145 students, 1 researcher, 4 observers during the face-to-face sessions, 7 post-hoc video coders, and SpeakUp itself) gathered quantitative and qualitative data using different techniques such as questionnaires, observations, action logging, and reports.

During the face-to-face sessions, students had allocated seats and the observers kept track of their participation including hands-up questions and other oral interventions. Moreover, one researcher video recorded the teachers and made observations about the teacher instruction and the general dynamics of the classroom. On the digital side, SpeakUp registered all the actions and contributions made by teachers and students along the course. To gather the user perspective about the usage of SpeakUp, questionnaires were sent to teachers and students. Finally, to measure learning outcomes, apart from the results of a multiple-choice test on different course topics, we used the overall course score, which included the test but also student presentations and projects.

4.2 Data analysis

Different quantitative (descriptive statistics and exploratory computational analyses) and qualitative analyses (manual coding of the messages generated by the users and video coding of the observations) have been performed on the data. Then, the results from these analyses were *triangulated* [59] to increase the trustworthiness of our findings.

In order to better understand the relevance of the comments posted by the students, one teacher and one researcher manually coded all of them (1226 messages), into three main categories: Messages that are *on-task* (e.g., questions or comments about the content, course organization, or SpeakUp), *off-task* (e.g., spam or bullying messages)

TABLE 1
Examples of SpeakUp message categories.

Category	Number	Examples
On-task	644	*"The course made me think of this TED talk: https://www.ted.com/..."* *"Looking for a group"* *"This app is ruining my battery"*
Off-task	208	*"If I get 100 likes, I will take out my clothes"* *"answer of a blond"*
Neutral	374	*"Hello"* *"LOVE NCC"*

and neutral (e.g., greetings and policing messages). Table 1 shows examples of messages in each category. During the codification process, both teacher and researcher had to agree on the category of each comment.

In a similar way, and in order to understand how was the teacher instruction structured, the video recording of each lesson was also coded, according to the following categories:

- TTS - Teacher talks to students (with slides/web/...)
- TPV - Teacher plays video to students
- TTT - Teacher talks to other teachers
- STT - Student talks to teacher (e.g., question)
- SWI - Students work individually
- SDS - Students discuss with each other

Each video was coded iteratively by at least 2 people. First, each coder analyzed individually the videos. Then, all coders and two researchers discussed the discrepancies found in the inter-reliability analysis. After clarifying the discrepancies and ambiguous situations, each coder went again through the videos adapting the codification according to the feedback received. Again, a sequence of inter-reliability analysis, discussion and review process took place (final mean Krippendorff's $\alpha = 0.85$). Any (small) remaining discrepancies were decided through majority voting, leading to the assignment of a code to each second within the videos. Additionally, one teacher annotated those moments where the contents of the questions in the multi-choice test were mentioned during the sessions, as well as when SpeakUp was used in any of the scenarios described in Section 5.1 ('Backchannel', 'Ask me anything', 'Quiz', or 'Think-pair-share').

To operationalize the different elements of the theoretical model proposed by Skinner et al. [15] (namely, context, action and outcomes), multiple indicators have been used (Figure 3). While some indicators take into account simple self-reported data of emotional factors (a single-question poll taken during one of the sessions in the middle of the course), intensive self-report measures have been avoided. This is due to their being tedious and intrusive over longer periods of time [21]. Hence, we make more intensive use of observational data collected by teachers or tools (e.g., classroom attendance or logged actions) to build our indicators.

To understand the relationships between these elements of context (teacher instruction), action (engagement and disaffection) and outcomes (learning outcomes), basic descriptive and exploratory statistics have been used (e.g., correlation analyses). Further explorations have been made through multiple linear regression modelling, to start understanding the relative strength of these trends.

For each research question, two levels of analysis are presented: a descriptive overview of what happened with these indicators at the classroom-level, and a second analysis at the student-level (e.g., on the relation between engagement and disaffection measures for each student, and their learning outcomes). These two analyses are also discussed taking the time dimension into account (e.g., aggregated measures for the whole course vs. per session).

5 Case Study

5.1 Context of the study

The case study took place in a Communication course at the École Polytechnique Fédérale de Lausanne (Switzerland) in 2016, during the spring semester. Three teachers led the course and 149 students (37 females) registered to it. The course was divided into six sessions of 105 minutes (from 16:15 to 18:00 GMT) with a break in the middle. After these lectures (spread over two months) followed a period of student work, classroom presentations and final exams. In this study, we focus on the aforementioned six sessions, which represent the teaching period of the course.

The course discusses different kinds of communication channels, social media platforms and technology-enhanced learning. This Communication course is part of the Global Issues program [60], which aims at introducing first-year undergraduate engineering students to interdisciplinary topics and soft skills. Due to the scale of the audience, lecturers in these courses often struggle to enable student participation during the face-to-face sessions.

The teaching team was composed of three lecturers with expertise in social media, information systems, behavioral sciences and management. The lecturers were familiar with the usage of social media in the classroom, as they had already used social media apps (such as Twitter or SpeakUp) in their practice.

During this course, SpeakUp was introduced as a communication channel with students to increase interaction. Further, the use of the app had another pedagogical purpose: since the course deals with communication channels, social media and TEL, SpeakUp would provide students with hands-on experience of many of the subjects studied in class. During the course, SpeakUp was used in four different ways:

- *Backchannel.* During the whole course, the application was used as a digital channel to promote the interaction among participants during the lessons. Students posted comments and interacted with their peers by answering and voting each others' comments. This use did not require any particular preparation from the teachers beyond the creation of the room.
- *Ask me anything.* During the course, the application was also used to help students post questions or problems. In each lesson, the group of instructors checked periodically the tool and answered whenever needed, either orally or digitally.
- *Quiz.* In specific situations, the instructors posed multiple-choice questions to the audience, in order

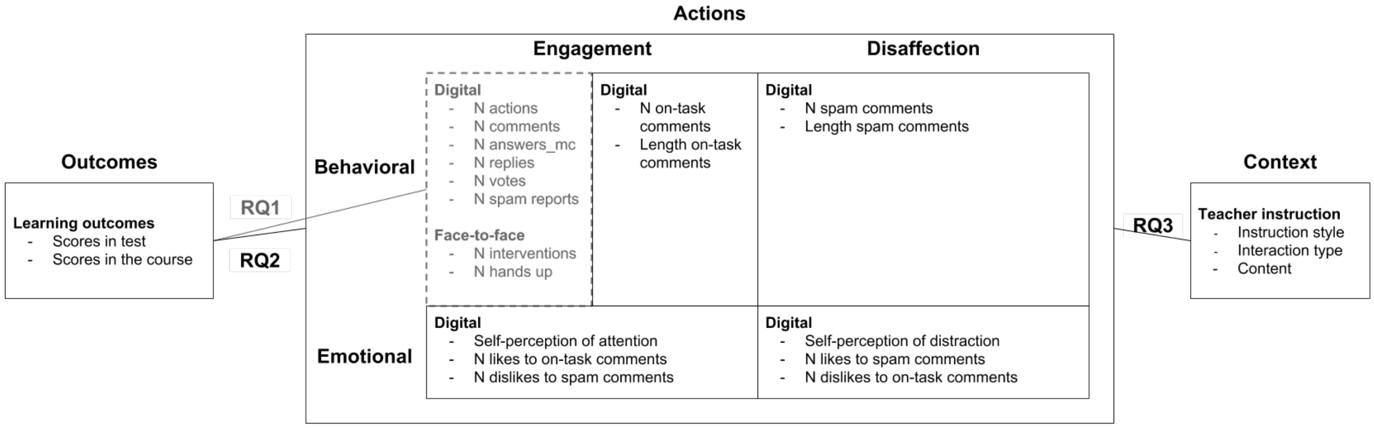

Fig. 3. Indicators extracted in each one of the research questions.

to assess their knowledge or to poll their opinion. In this kind of scenarios, the instructors had to prepare questions and answers in advance to avoid stopping the flow of the activity. This usage of the platform was often coupled with (face-to-face) group discussions.

- *Think-pair-share.* In this scenario, first the teachers guided the audience to think about an open question, ask the students to discuss it in pairs, and post an answer on the application. Later, students were asked to read the answers of others, and express their agreement or disagreement by commenting or voting other people's comments. Finally, the teachers discussed orally with the students on the comments generated during the activity. This scenario allowed teachers to get an overview of the audience's opinions on a subject. While not needing preparation in advance, this usage deviated more of teachers' attention to the tool (as opposed to the face-to-face interaction).

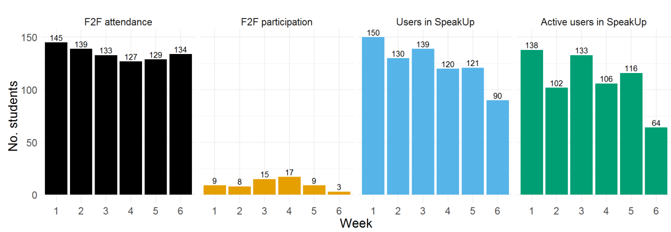

Fig. 4. Overall counts of face-to-face and online participation in the six sessions of the course.

5.2 Action: Attendance, participation and usage of SpeakUp

Figure 4 represents the attendance to the six face-to-face sessions of the course (left-hand side), as well as the face-to-face interventions, the number of student users in SpeakUp[2] and the number of active users in the app (i.e., not only opened it, but also created messages, voted, etc.). We can observe how, despite the fact that physical attendance remained quite constant throughout the course, overall SpeakUp usage showed certain signs of declining (maybe due to the "novelty effect" of using the tool gradually subsiding). Out of the 149 students registered for the course, 132 (88%) volunteered to disclose their anonymous user identifier for the purposes of this research (i.e., to relate their learning outcomes with their app behavior). All in all, we registered 243 different student user identifiers in SpeakUp.

2. Since a student can access the app from multiple devices and web browsers at the same time, and users are anonymous by default, it can be the case that the number of SpeakUp student users is higher than that of the students physically in the classroom – as it happens in the first session.

5.3 Learning outcomes

As mentioned in Section 4.1, two learning outcomes were taken into consideration: the results of the multi-choice test and the overall course score (calculated as a weighted mean of the multi-choice test, student presentations and projects during the course). The multi-choice test was composed of twenty questions, and the resulting student scores ranged from 44–95/100, with an average value of 78. The course scores followed the Swiss scoring standards that range from 1–6. In our course, the scores were bell-shaped, but also had a limited range (the lowest score was 3.6, and the highest, 6.0, with median and average values of 5.1).

For most of the analyses described below, we used the overall course score as the main indicator of learning outcomes, since it tracks more than just factual content knowledge, an important aspect of the learning experience in this kind of inter-disciplinary courses [60]. Only in the case of more fine-grained analyses we used the multi-choice question responses (e.g., to understand the action profiles of different moments where test questions are mentioned during the lectures).

5.4 Results

5.4.1 How does a simple definition of student action relate to learning outcomes? (RQ1)

As mentioned in Section 4, we have explored the relationship between simple, count-based indicators of behavioral engagement (both from SpeakUp and face-to-face events), and the student learning outcomes:

TABLE 2
Overall counts of (simple, count-based) behavioral engagement indicators, and correlations (r) with the learning outcomes of each student. *Note:* *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Indicator	**Counts**	r
n_actions (SpeakUp)	23538	0.18*
create_message (SpeakUp)	950	0.00
reply_comment (SpeakUp)	271	0.11
answer_mc (SpeakUp)	382	0.12
reports_spam (SpeakUp)	21	-0.09
vote_comments (SpeakUp)	7753	0.13
prompted participation (face-to-face)	64	0.04
unsolicited hands-up (face-to-face)	4	-0.13

Classroom-level analysis. Table 2 and Figure 5 provide an overview of these eight indicators, for the whole course and for each of the six sessions that made up the course. We can observe than in many of the indicators (e.g., total number of actions in SpeakUp, messages created, votes) there is a clear downward trend after the first session (again, indicating the presence of a "novely effect"). We can also notice in many of the indicators that some sessions (e.g., session 3) saw a more intense engagement of students, both in and outside the app. This probably was motivated by the strategies used by the lecturers in those sessions. Please, refer to Section 5.4.3 for a deeper discussion of this issue. An even more fine-grained temporal view of the behavioral engagement as portrayed by these indicators is illustrated in Figure 7 below.

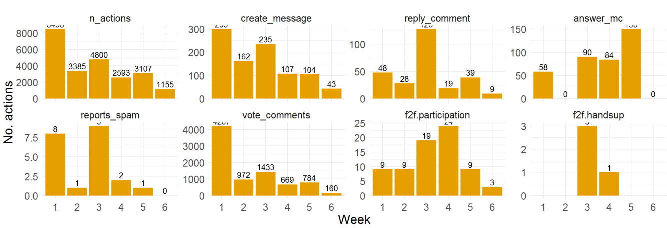

Fig. 5. Overall counts of simple behavioral engagement indicators for each of the sessions in the course.

Student-level analysis. Since our goal was to understand the relationship between the aforementioned measures of engagement and student learning outcomes, we have performed different kinds of analyses to establish the nature and strength of such relationships. Using the course outcomes (i.e., scores) of the 132 students who disclosed their identity, we can explore the correlations of the simple behavioral engagement indicators with such learning outcome (Table 2). We can observe that, in general, these indicators bear almost no correlation with the student learning outcomes. The exception is the raw number of actions (of all kinds) that each student generated in SpeakUp, which seems to have a mild (but significant) correlation with course outcomes. Interestingly, some of the indicators have negative correlations with outcomes, even if intuitively one could consider them as a good thing (e.g., reporting a message as spam in SpeakUp, or putting your hand up to ask questions during the classes). However, these are relatively rare events in the dataset, and hence it is hard to make solid inferences about them.

We further explored whether these indicators, or combinations of them, were related to the learning outcomes. A step-wise multiple linear regression analysis was performed, using Akaike Information Criterion (AIC) for model selection. By performing this model selection, both forwards (adding predictors successively from a model with no predictors) and through backwards elimination (starting with the full model with all possible predictors and removing the least useful ones), we arrived to the model that is shown in Table 3 (top). We can observe that, despite having a significant predictor (the total number of SpeakUp actions), it only explains about 5% of the variance in learning outcomes.

We can hence conclude that these eight measures of behavioral engagement, taken from both the SpeakUp logs and simple face-to-face observations, seem insufficient to predict reliably the academic performance of the students in our course. To try to address the limitations of this simplistic view of engagement as it pertains to learning, we explore further indicators in the next section.

5.4.2 How does a more complex view of action relate to learning outcomes (RQ2)?

Classroom-level analysis. Table 4 and Figure 6 provide an overview of the additional indicators covering behavioral and emotional engagement and disaffection (Figure 3), for the whole course and for each of the six sessions that made up the course. We can observe in many of these indicators (e.g., number of spam messages and their average length) a similar downward trend after the first session (cf. the notion of there being a "novely effect"). We can also again notice that session 3 saw a more intense behavioral engagement of students for many indicators (which is not present in the behavioral or emotional disaffection measures).

Student-level analysis. Once again, to explore the relationship between these additional indicators of engagement/disaffection and the student learning outcomes, we performed correlation analyses. Table 4 shows the correlations (r) between the additional emotional and behavioral indicators and the overall course score of each student. We can observe that many of these indicators also present low correlation with student outcomes. The main exceptions in this case are the average on-task message length and the number of 'likes' on such on-task messages. The fact that this kind of events are also more commonplace in the dataset, offers some initial hope that they will carry some power in terms of predicting the students' performance.

In a similar manner as we did in the previous section with the simple behavioral engagement indicators, here we again used the AIC for model selection in a stepwise multiple linear regression. Table 3 (bottom) shows the main parameters of the resulting best-fitting model. We can observe that the model in this case performs better (it explains now about 17% of the variance in the outcomes). The model now has four significant predictors, two positive and two negative: a) the number of messages created (of any kind) (-); b) the total number of votes made (regardless of them being like or dislike) (+); c) the total number of on-task messages created (+), and d) the number of dislikes assigned to spam messages (-).

It is worth noting that the two negative predictors are somewhat surprising, since they could be intuitively un-

TABLE 3
Linear regression models of learning outcomes based on different kinds of indicators of engagement and disaffection. *Note: All predictors are mean-centered and scaled by 1 standard deviation. Standard errors are heteroskedasticity robust.* $^{****}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Indicators	B	$Std.Err.B$	t	p	**Adj.** R^2	p
Model: Simple Behavioral Engagement					0.05	0.02
n_actions (SpeakUp)	0.11*	0.05	2.10	0.04		
reports_spam (SpeakUp)	-0.08	0.06	-1.24	0.22		
unsolicited hands-up (face-to-face)	-0.33	0.26	-1.27	0.21		
Model: Engagement & Disaffection					0.17	0.00
create_message (SpeakUp)	-0.36*	0.15	-2.38	0.02		
vote_comments (SpeakUp)	0.26*	0.10	2.57	0.01		
unsolicited hands-up (face-to-face)	-0.51	0.27	-1.89	0.06		
number of on-task messages (SpeakUp)	0.30*	0.11	2.82	0.01		
avg. length on-task messages (SpeakUp)	0.10	0.05	1.85	0.07		
max. length of spam messages (SpeakUp)	-1.26	0.79	-1.60	0.11		
min. length of spam messages (SpeakUp)	-0.78	0.70	-1.12	0.27		
total length of spam messages (SpeakUp)	0.18	0.28	0.64	0.53		
avg. length of spam messages (SpeakUp)	1.90	1.41	1.35	0.18		
dislikes of spam messages (SpeakUp)	-0.22*	0.10	-2.16	0.03		

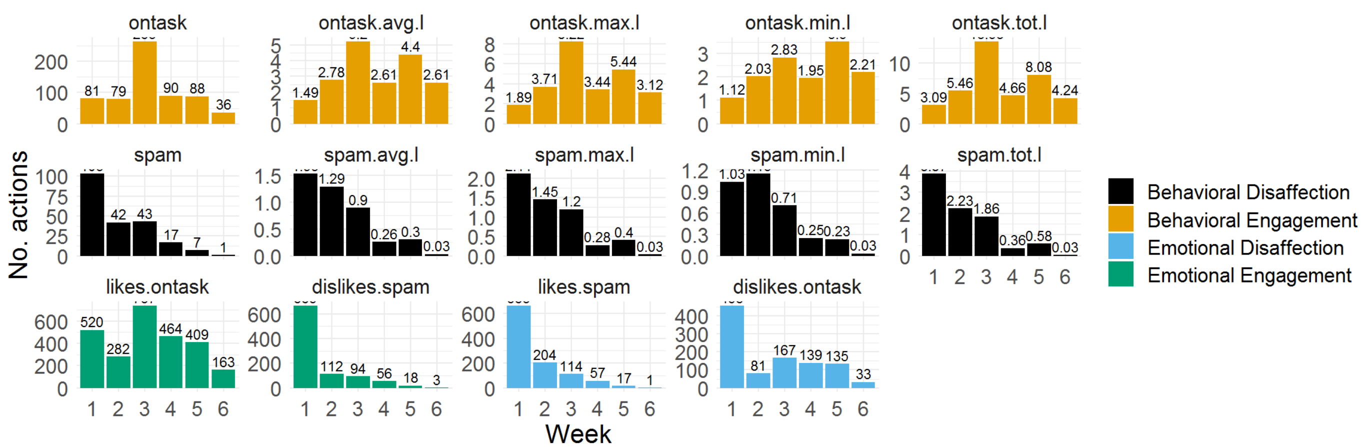

Fig. 6. Overall counts of additional engagement and disaffection indicators for each of the sessions in the course.

derstood as markers of positive behavioral and emotional engagement. The statistical model seems to suggest that, once we control for the other variables in the model, creating more messages in SpeakUp does not lead to better outcomes (and neither does disliking a 'spammy' comment). We can conjecture that such events, while positive in principle, have relatively low value, and may be distracting from deeper work or engagement with the course contents.

This kind of regression modeling thus seems to suggest that taking into account a wider range of engagement and disaffection metrics, can provide enhanced predictive power about student performance, at least for the particular case of using anonymous, potentially distracting social media technologies in lectures. It also starts providing ideas of a certain hierarchy of actions in the system, and their value for learning. These insights can be useful both for learning technology designers and practitioners that want to scaffold learning using these tools. We will deepen further into the latter, in the following section.

5.4.3 What is the role that teacher instruction plays on student action? (RQ3)

As we have already hinted at in some of our results above, the fact that certain sessions, or certain moments within a session, saw increased rates of student activity, is bound to have been influenced (but probably not determined directly) by what the lecturers asked students to do at that moment, and what resources and pedagogical moves they were using. The fact that now we have evidence that certain markers of engagement and disaffection are related to learning outcomes (see RQ1, RQ2) allows us to start disentangling the relationship between what the classroom activity was and the aspects of student action that seem to relate with outcomes.

Figure 7 exemplifies these relationships, in the context of session 3, by representing the classroom activities (as coded by the research team from the video recordings), along with several engagement indicators. Around 16:45 and 17:55 (according to the observations) two think-pair-share activities (as described in Section 5.1) by the lecturers triggered the sudden rise of all the action indicators related to learning outcomes (with different amounts of lag among signals). Other variations in these indicators, smaller and less coordinated across signals, were due to less structured activities related to SpeakUp (e.g., around 17:15 and 17:37 there were whole-class reflections about comments previously posted in SpeakUp) or spontaneous activity of the students related to the topics discussed in the classroom.

Localizing and interpreting such relationships between the face-to-face data (from the videos) and the SpeakUp logs can nevertheless be arduous by mere visual inspection. Are there any definite patterns in the amounts of action indicators witnessed by the different kinds of classroom activities witnessed throughout the six sessions of the course?

A Shapiro-Wilk test revealed that our different indicators do not take a normal distribution. We have thus performed multiple Kruskal-Wallis rank sum tests, to understand whether the distribution of action indicators (e.g., raw number of SpeakUp actions, or number of on-task messages) co-occurring with different kinds of classroom activities (e.g., the teacher playing a video, or the students working individually) seem to be coming from different distributions (with different median values). As we can see in Table 5, the distributions of action values for different teaching activities are indeed different. Subsequent Dunn tests (with multiple comparison p-values adjusted using the Benjamini-Hochberg method) indicated which pairs of classroom activities were significantly different in terms of their action indicators. The combinations of comparisons are too numerous to be detailed here, but certain noteworthy trends should be mentioned:

- For the raw number of SpeakUp actions (n_actions, which was positively associated with outcomes), we can observe that certain classroom activities (the teacher playing videos to students – TPV, and the students discussing among themselves – SDS) have much lower median values than other activities (TTT, STT). This seems logical, since these are activities that have a tighter "grip on students' attention", leaving less for parallel actions like sending messages around. Indeed, SDS and TPV have low median indicators across the board. While this may be interpreted as moments of low behavioral engagement, it can also be interpreted as the opportunity for intense cognitive engagement with the content *outside* the realm of the digital.
- The number of created messages (create_message, which was negatively correlated with outcomes) seems to be higher (in terms of medians) for STT, SWI, TTT. In these moments, the lecturer's attention tends to go away from the students, and seems to be a trigger for using the backchannel that SpeakUp represents (for better and for worse).
- The number of votes (vote_comments, positively associated with outcomes) seems to be also higher in TTT, STT and SWI activities. It seems that in these activities, not only students create more new messages in SpeakUp (see previous point); also, as new messages arrive, people rush to vote them. We see how a certain activity may have both upsides and downsides, as it is associated with both productive and unproductive. It is then the challenge for the teacher to keep students' attention productively directed during these times especially.
- For the number of on-task messages (another positively associated indicator), again STT and SWI seem to have substantially larger medians, suggesting a potential for these activities to become good 'learning moments' with positive action patterns.
- Regarding the emotional engagement metric that appeared in our regression models (dislikes.spam, negatively related to outcomes) only STT seems to have higher activate this kind of indicator. It seems that the students talking to teachers (e.g., posing questions) is a moment that the rest of students may take advantage to engage in the backchannel (SpeakUp). However, as we saw in RQ2 above, while this kind of activity seems positive, maybe it has relatively low value, and distracts the rest of the students.

All in all, these results seem to indicate that SpeakUp can indeed enable active, unstructured participation by the students. While the tool is easy and quick to setup, and can significantly increase classroom interactions (to the level of thousands of them during a whole course), it is also the ground for off-task and distracting behavior. As [7] already points out, using the tool effectively might require a

TABLE 4
Overall counts/means of selected additional engagement and disaffection indicators, and correlations (r) with the learning outcomes of each student. *Note:* $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Indicator	**Category**	**Counts/ Averages**	r
spam messages (SpeakUp)	Behavioral Disaffection	213	-0.04
spam message length (avg.) (SpeakUp)	Behavioral Disaffection	1.8	0.10
on-task messages (SpeakUp)	Behavioral Engagement	637	0.13
on-task message length (avg.) (SpeakUp)	Behavioral Engagement	4.7	0.28**
dislikes on on-task messages (SpeakUp)	Emotional Disaffection	1011	0.03
likes on spam messages (SpeakUp)	Emotional Disaffection	1054	0.08
Perception of SpeakUp as useful/distracting (poll)	Emotional Engagement Disaffection	0.21	0.01
dislikes on spam messages (SpeakUp)	Emotional Engagement	946	0.02
likes on on-task messages (SpeakUp)	Emotional Engagement	2575	0.24**

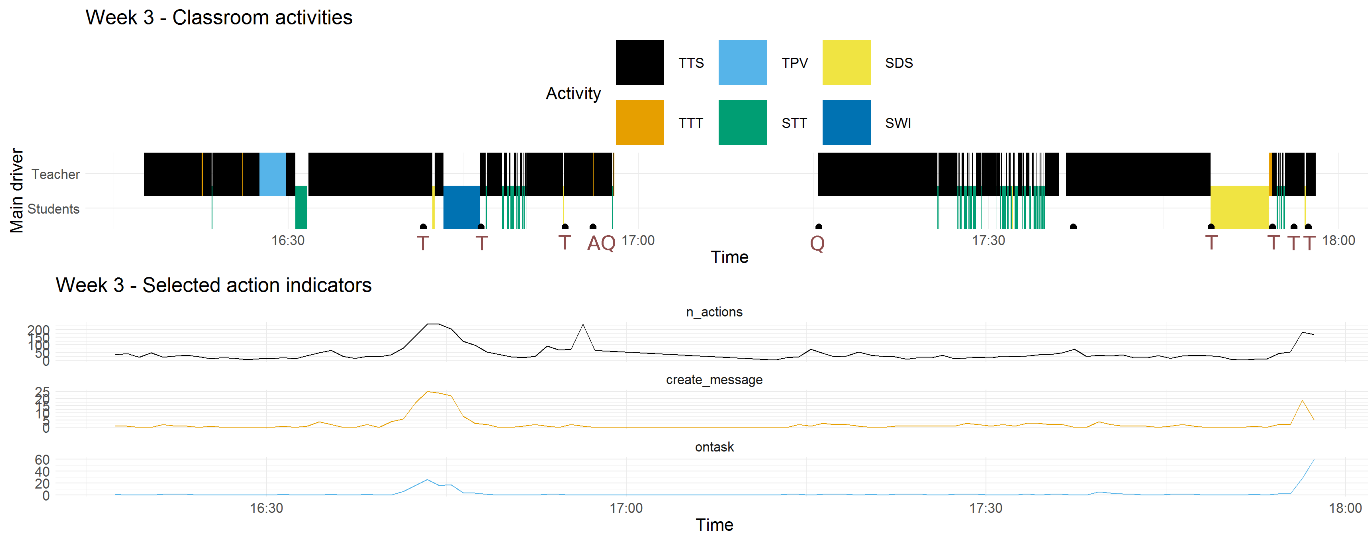

Fig. 7. Graph of selected engagement and disaffection indicators (significant in the linear regression models), compared with the activity happening in the class during session 3. Black dots and letters along the timeline indicate moments in which a lecturer explicitly pointed students to SpeakUp activities (see Section 5.1): T=Think-pair-share; A=Ask me anything; Q=Quiz

TABLE 5
Median values of engagement and disaffection metrics, for the different kinds of classroom activities coded throughout the course. *Note:* χ^2, df and p *from Kruskal-Wallis rank sum test against the hypothesis that the data from each kind of activity comes from the same distribution.* $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Activity	SDS	STT	SWI	TPV	TTS	TTT	χ^2	df	p
n_actions (+)	0	29	18	0	21	66	60	6	0.000***
create_message (-)	0	3	2	0	1	3	50	6	0.000***
vote_comments (+)	0	14	7	0	5	44	50	6	0.000***
on-task (+)	0	1	2	0	0	0.5	30	6	0.000***
dislikes.spam (-)	0	1	0	0	0	0	20	6	0.001**

certain amount of monitoring, and the guiding of students' attention to more productive uses of it.

6 Discussion

Research investigating the use of social media applications in the classroom generally concludes that students perceive such systems as positive and that they feel it increases interactivity [3], [4], [5], [6], [8], [14], [61]. Higher engagement and interactivity tend to have advantages for learning, as has been proven time and again in formal settings where the technology has been designed for learning (i.e., in an LMS, see for example [35]). However, the evidence of clear relationship between technology use and academic performance in the case of less-controlled technologies like (anonymous) social media, is still an emergent field. In this paper we contribute to this knowledge base through a case study which followed the use of SpeakUp during a university course.

The evidence gathered across multiple data sources (logs, observations, video recordings, questionnaires, academic assessments, etc.) paints a nuanced picture of how student action, both face-to-face and through SpeakUp, relates to learning outcomes. Regarding our first research question (*How does a simple definition of action (as student behavioral engagement) relate to learning outcomes?*), we found that simple measures of behavioral engagement, based on simple log counts, are somewhat informative (e.g., the fact that n_actions was a significant predictor in our first linear regression model), but insufficient to build fitting models of academic performance in the course (Adj. $R^2 = 0.05$).

Through our second research question (*How does a more complex view of action relate to learning outcomes?*), we investigated the advantages of adding multiple (but still relatively simple and unobtrusive) indicators that cover both behavioral and emotional engagement and disaffection. Our models based on these potentially-automatable indicators (as opposed to, e.g., deeper content analysis that requires human intervention) explained a much larger proportion of variance in the academic performance (Adj. $R^2 = 0.17$). In this sense, our findings are in line with previous research stating that "measuring engagement across more than one indicator may produce the most productive information for researchers, instructional designers, and educators" [21], [27]. The fact that our models included both positive and negative significant indicators, is also in line with the experiences of many practitioners, which consider that social media in class can be sometimes more distracting than helpful [9]. Finally, it is worth comparing these results with studies done on the use of Twitter for learning [13], which found that message content (related vs. unrelated to the class lecture) and message creation (responding to or creating a message) seemed to impact student learning. In our case, the amount of messages created was itself a negative predictor of performance (maybe hinting at the distraction that composing a message represents), offset by the creation of *on-task* messages related to the course contents (which was a positive, significant predictor).

It seems that it is not using social media *per se* which is related to outcomes, but rather how students use these tools [12]. This places the innovative practitioner in the difficult position of using a tool (social media) which might as well have adverse consequences if not used effectively. The third research question of our case study (*What is the role that teacher instruction plays on student action?*) investigated this issue. The results of our analyses of co-occurrence between the aforementioned indicators of student action (related to learning outcomes) and different kinds of classroom activities, show patterns of tool usage and participation (e.g., more on-task message creation during individual work activities, if the tool is explicitly integrated in the activity). They also represent a warning against conceptions of engagement that are too tool-centered: in our data, students discussing was accompanied by low values in all action indicators (since students were too busy discussing to post messages or vote on them), which should not be necessarily interpreted as detrimental (quite the opposite, as deep concentration in discussing the course contents may leave no spare attention for lower-value tool uses or any distractions). This hints at future work in this area, using advances in multimodal learning analytics [62] to complement the digital indicators of action with physical ones, or to gather automatically the classroom events [54], so that they can be used in analyses such as those presented in this paper.

The findings presented in this case study, however, should also be considered in light of a number of limitations. There are obvious limits to the generalizability of the findings, which stem from the fact that the study was conducted on a single course, at an institution and by a set of lecturers that may not be necessarily representative of other settings. Similarly, the course being targeted at first-year undergraduate students (which may have influenced the predisposition and behavior of the students regarding the use of SpeakUp). Future research should try to randomize across more classes from different backgrounds.

Certain limitations in the dataset should also be acknowledged, such as the fact that the academic scores, both of the multi-choice test and overall (which had a limited range of values, and did not include a baseline pre-test before the course) were used as the main proxy for learning outcomes. The study design and operationalization tried to strike a balance between simplicity (in terms of technology and indicators used – to enable easier transfer of results) and predictive power. In this sense, even if our study required a high investment in terms of human labor for content analysis and video coding, both labeling tasks were kept intentionally simplistic, to make them feasible to be automated in the near future. This, of course, limits the depth of the insights that we can take from the analysis of the context and the cognitive engagement of students. Further studies can also explore other sets of indicators of engagement from existing work in formal learning settings using LMSs [35], [63] (teacher participation, course design, class size, student self-regulated learning characteristics, etc.).

7 Conclusions

In our way towards understanding how to use social media effectively in the classroom, this paper analyses the use of SpeakUp in a co-located course with three teachers and 149 university students. Following the model proposed by [15], we have explored the context-actions and actions-outcomes relationships.

Regarding the relation between student actions and outcomes, our study reveals that simple measures of behavioral engagement were somewhat informative, but insufficient to predict academic performance. On the other hand, adding multiple (but still relatively simple and unobtrusive) indicators that cover both behavioral and emotional engagement and disaffection, our models performed better, as other authors also reported [21], [27]. While teachers and students reported in this course that SpeakUp was beneficial in terms of participation [55], our results also show that social media in class can be also distracting, as teachers often point out [9]. Thus, there should be more emphasis on the *quality* of the participation, rather than on the quantity [12].

The results seems to show that it is not the best students asking questions face-to-face (as there is no or even negative correlation between the level of interaction and the learning outcomes), but rather students requiring more support. So, this additional opportunity to interact is mainly useful for students with difficulties. Thus, digital interventions may help to reduce the gap between students that perform well and those that do poorly (by pushing up the bottom of the cohort). However, this hypothesis cannot be assessed with the currently available data, and studies with a control group (i.e., without digital intervention) would be necessary.

While exploring the relation between context and action, we have been able to identify patterns that connect teacher instruction and student action (as behavioral and emotional engagement and disaffection). The orchestration diagram in Figure 7 shows that the highest points of on-task message activity match those moments when teachers guided the use of SpeakUp (either proposing polls or open-questions, or organizing think-pair-share activities). These results emphasize the role that teachers play in the effective use of social media in the classroom, and how important is to design accordingly the learning activities [12], [13].

To increase the transfer and scale of our study methods into educational practice, in our future work we plan to automate the current qualitative data codification, and expand our catalogue of indicators through multimodal data gathering and analytics techniques [54], [62]. For example, automatic content analysis could be use to infer emotions from the messages, and audio analysis could automatically detect classroom activity based on the ambient noise.

Acknowledgements

This research has been co-funded by the European Union's FP7 and Horizon 2020 research and innovation programmes in the context of the Go-Lab Integrated Project (grant agreement no. 317601), the Next-Lab Innovation Action (grant agreement no. 731685), and CEITER (grant agreement no.

669074). The authors would like to thank the team of observers from École Polytechnique Fédérale de Lausanne and the video coders from Tallinn University for their contribution to this study.

REFERENCES

[1] Benjamin Dyson, Kristin Vickers, John Turtle, Sara Cowan, and Adrianna Tassone. Evaluating the use of facebook to increase student engagement and understanding in lecture-based classes. *Higher Education*, 69(2):303–313, Feb 2015.

[2] Daniela Retelny, Jeremy P. Birnholtz, and Jeffrey T. Hancock. Tweeting for class: using social media to enable student co-construction of lectures. In Steven E. Poltrock, Carla Simone, Jonathan Grudin, Gloria Mark, and John Riedl, editors, *CSCW (Companion)*, pages 203–206. ACM, 2012.

[3] Richard J. Anderson, Ruth Anderson, Tammy VanDeGrift, Steven Wolfman, and Ken Yasuhara. Promoting interaction in large classes with computer-mediated feedback. In *Designing for change in networked learning environments*, pages 119–123. Springer, 2003.

[4] Tony Bergstrom, Andrew Harris, and Karrie Karahalios. Encouraging initiative in the classroom with anonymous feedback. In Pedro Campos, Nicholas Graham, Joaquim Jorge, Nuno Nunes, Philippe Palanque, and Marco Winckler, editors, *Human-Computer Interaction – INTERACT 2011*, volume 6946 of *Lecture Notes in Computer Science*, pages 627–642. Springer Berlin Heidelberg, 2011.

[5] Carmen Fies and Jill Marshall. Classroom response systems: A review of the literature. *Journal of Science Education and Technology*, 15(1):101–109, 2006.

[6] Drew Harry, Joshua Green, and Judith Donath. Backchan.nl: integrating backchannels with physical space. In *CHI'09*, pages 2751–2756. ACM, 2008.

[7] Sten Govaerts, Adrian Holzer, Bruno Kocher, Andrii Vozniuk, Benot Garbinato, and Denis Gillet. Blending digital and face-to-face interaction using a co-located social media app in class. *IEEE Transactions on Learning Technologies*, 11(4):478–492, Oct 2018.

[8] Eusebio Scornavacca, Sid Huff, and Stephen Marshall. Mobile phones in the classroom: if you can't beat them, join them. *CACM*, 52(4):142–146, April 2009.

[9] Jeff Seaman and Hester Tinti-Kane. Social media for teaching and learning. Technical report, Pearson Learning Solutions, 2013.

[10] Faria Sana, Tina Weston, and Nicholas J. Cepeda. Laptop multitasking hinders classroom learning for both users and nearby peers. *Comput. Educ.*, 62:24–31, 2013.

[11] Debra L Worthington and David 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.

[12] Reynol Junco. *Engaging students through social media: Evidence-based practices for use in student affairs*. John Wiley & Sons, 2014.

[13] Jeffrey H. Kuznekoff, Stevie Munz, and Scott Titsworth. Mobile Phones in the Classroom: Examining the Effects of Texting, Twitter, and Message Content on Student Learning. *Communication Education*, 64(3):344–365, 2015.

[14] Adrian Holzer, Sten Govaerts, Andrii Vozniuk, Jan Ondrus, David Rigaud, Benoît Garbinato, and Denis Gillet. SpeakUp – a mobile app facilitating audience interaction. In *ICWL'13*, pages 11–20. Springer, 2013.

[15] Ellen Skinner, Carrie Furrer, Gwen Marchand, and Thomas Kindermann. Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of educational psychology*, 100(4):765, 2008.

[16] Shane R Jimerson, Emily Campos, and Jennifer L Greif. Toward an understanding of definitions and measures of school engagement and related terms. *The California School Psychologist*, 8(1):7–27, 2003.

[17] Robert E. Stake. *The art of case study research*. Sage Publications, Thousand Oaks, CA, 1995.

[18] Johan J. Vossensteyn, Andrea Kottmann, Benjamin W.A. Jongbloed, Franciscus Kaiser, Leon Cremonini, Bjorn Stensaker, Elisabeth Hovdhaugen, and Sabine Wollscheid. Dropout and completion in higher education in europe: Main report, 2015.

[19] Larry Johnson, Samantha Adams Becker, Michele Cummins, Victoria Estrada, Alex Freeman, and Courtney Hall. *NMC horizon report: 2016 higher education edition*. The New Media Consortium, 2016.

[20] Hamish Coates. A model of online and general campus-based student engagement. *Assessment & Evaluation in Higher Education*, 32(2):121–141, 2007.

[21] Curtis R. Henrie, Lisa R. Halverson, and Charles R. Graham. Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, 90:36–53, 2015.

[22] Alexander W Astin. Student involvement: A developmental theory for higher education. *Journal of college student personnel*, 25(4):297–308, 1984.

[23] George D. Kuh. What student affairs professionals need to know about student engagement. *Journal of college student development*, 50(6):683–706, 2009.

[24] Sarah J. Mann. Alternative perspectives on the student experience: Alienation and engagement. *Studies in higher education*, 26(1):7–19, 2001.

[25] Elaine Chapman. Alternative approaches to assessing student engagement rates. *Practical Assessment*, 8(13):1–7, 2003.

[26] Ellen A. Skinner. Engagement and disaffection as central to processes of motivational resilience and development. *Handbook of motivation at school*, pages 145–168, 2016.

[27] Jennifer A Fredricks, Phyllis C Blumenfeld, and Alison H Paris. School engagement: Potential of the concept, state of the evidence. *Review of educational research*, 74(1):59–109, 2004.

[28] Paula Witkowski and Thomas Cornell. An investigation into student engagement in higher education classrooms. *InSight: A Journal of Scholarly Teaching*, 10:56–67, 2015.

[29] Xi Zhang and Yao Meng. Students' engagement in collaborative learning group supported by communication tools: an empirical study. In *PACIS*, page 55, 2016.

[30] Helen Patrick, Allison M Ryan, and Avi Kaplan. Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *Journal of educational psychology*, 99(1):83, 2007.

[31] Sidney D'Mello. Emotional learning analytics. *Handbook of learning analytics*, pages 115–127, 2017.

[32] Daniele Di Mitri, Jan Schneider, Marcus Specht, and Hendrik Drachsler. From signals to knowledge: A conceptual model for multimodal learning analytics. *Journal of Computer Assisted Learning*, 34(4):338–349, 2018.

[33] Arthur W Chickering and Stephen C Ehrmann. Implementing the seven principles: Technology as lever. *AAHE bulletin*, 49:3–6, 1996.

[34] Pu-Shih Daniel Chen, Amber D Lambert, and Kevin R Guidry. Engaging online learners: The impact of web-based learning technology on college student engagement. *Computers & Education*, 54(4):1222–1232, 2010.

[35] Abelardo Pardo, Feifei Han, and Robert A Ellis. Combining university student self-regulated learning indicators and engagement with online learning events to predict academic performance. *IEEE Transactions on Learning Technologies*, 10(1):82–92, 2017.

[36] John Erickson and Keng Siau. E-ducation. *CACM*, 46(9):134–140, 2003.

[37] Karel Kreijns, Paul A. Kirschner, and Wim Jochems. Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in human behavior*, 19(3):335–353, 2003.

[38] Harold M. Horowitz. Student response systems: Interactivity in a classroom environment. In *Proceedings of the Sixth Annual Conference on Interactive Instruction Delivery*, Salt Lake City, UT, Febr 1988.

[39] Erika Blood and Richard Neel. Using student response systems in lecture-based instruction: Does it change student engagement and learning? *Journal of Technology and Teacher Education*, 16(3):375–383, 2008.

[40] Jeffrey R. Stowell and Jason M. Nelson. Benefits of electronic audience response systems on student participation, learning, and emotion. *Teaching of Psychology*, 34(4):253–258, 2007.

[41] April R. Trees and Michele H. Jackson. The learning environment in clicker classrooms: student processes of learning and involvement in large university-level courses using student response systems. *Learning, Media and Technology*, 32(1):21–40, 2007.

[42] Honglu Du, Mary Beth Rosson, and John M. Carroll. Augmenting classroom participation through public digital backchannels. In *Proceedings of the 17th ACM International Conference on Supporting Group Work*, GROUP '12, pages 155–164, New York, NY, USA, 2012. ACM.

[43] Gijsbert Grotenbreg and Soon Boon Justin Wong. Using Pigeonhole® Live to elicit feedback, questions & reinforce learning during lectures. *CDLT Brief*, 16(2):2–7, 2013.

[44] W. Reinhardt, M. Ebner, G. Beham, and C. Costa. How people are using twitter during conferences. In *Proc. of 5th EduMedia Conf.*, page 145, 2009.

[45] Priya Kumar, Anatoliy Gruzd, Caroline Haythornthwaite, Sarah Gilbert, Marc Esteve del Valle, and Drew Paulin. Learning in the wild: Coding reddit for learning and practice. In *Proceedings of the 51st Hawaii International Conference on System Sciences*, 2018.

[46] National Survey of Student Engagement (NSSE). Engagement insights: Survey findings on the quality of undergraduate education. anual resutls 2018, 2018.

[47] Yao-Ting Sung, Kuo-En Chang, and Tzu-Chien 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.

[48] Mable B. Kinzie, Stephen D. Whitaker, and Mark J. Hofer. Instructional uses of instant messaging (IM) during classroom lectures. *Educational Technology & Society*, 8(2):150–160, 2005.

[49] Sarita Yardi. The role of the backchannel in collaborative learning environments. In *Proceedings of the 7th International Conference on Learning Sciences*, ICLS '06, pages 852–858. International Society of the Learning Sciences, 2006.

[50] Vicki Trowler. Student engagement literature review. *The higher education academy*, 11(1):1–15, 2010.

[51] Suparna Sinha, Toni Kempler Rogat, Karlyn R. Adams-Wiggins, and Cindy E. Hmelo-Silver. Collaborative group engagement in a computer-supported inquiry learning environment. *International Journal of Computer-Supported Collaborative Learning*, 10(3):273–307, 2015.

[52] Sanna Järvelä and Allyson F. Hadwin. New frontiers: Regulating learning in cscl. *Educational Psychologist*, 48(1):25–39, 2013.

[53] Cristina Rădulescu, Mihaela Dinsoreanu, and Rodica Potolea. Identification of spam comments using natural language processing techniques. In *2014 IEEE 10th International Conference on Intelligent Computer Communication and Processing (ICCP)*, pages 29–35. IEEE, 2014.

[54] Luis P. Prieto, Kshitij Sharma, Łt Kidzinski, María Jesús Rodríguez-Triana, and Pierre Dillenbourg. Multimodal teaching analytics: Automated extraction of orchestration graphs from wearable sensor data. *Journal of computer assisted learning*, 34(2):193–203, 2018.

[55] María Jesús Rodríguez-Triana, Adrian Holzer, Luis P. Prieto, and Denis Gillet. Examining the effects of social media in co-located classrooms: A case study based on SpeakUp. In *European Conference on Technology Enhanced Learning*, pages 247–262. Springer, 2016.

[56] R. Burke Johnson and Anthony J. Onwuegbuzie. Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7):14–26, 2004.

[57] Jennifer C. Greene, Lehn Benjamin, and Leslie Goodyear. The merits of mixing methods in evaluation. *Evaluation*, 7(1):25–44, 2001.

[58] John W. Creswell, Vicki L. Plano Clark, Michelle L. Gutmann, and William E. Hanson. Advanced mixed methods research designs. In *Handbook of mixed methods in social and behavioral research*, pages 209–240. Sage Publications, 2003.

[59] Egon G. Guba. Criteria for assessing the trustworthiness of naturalistic inquiries. *Journal of Theory, Research, and Development on Educational Communication and Technology*, 29(2):75–91, 1981.

[60] Adrian Holzer, Isabelle Vonèche Cardia, Samuel Bendahan, Alexis Berne, Luca Bragazza, Antonin Danalet, Ambrogio Fasoli, Jerome N Feige, Denis Gillet, Siara Isaac, et al. Increasing the perspectives of engineering undergraduates on societal issues through an interdisciplinary program. *International Journal of Engineering Education, 2016*, 32(2):614–624, 2016.

[61] Adrian Holzer, Sten Govaerts, Andrii Vozniuk, Bruno Kocher, and Denis Gillet. Speakup in the classroom: Anonymous temporary social media for better interactions. In *Proceedings of the Extended Abstracts of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI EA '14, pages 1171–1176, New York, NY, USA, 2014. ACM.

[62] Xavier Ochoa. Multimodal learning analytics. *The Handbook of Learning Analytics*, pages 129–141, 2017.

[63] Colin Beer, Ken Clark, and David Jones. Indicators of engagement. *Curriculum, technology & transformation for an unknown future. Proceedings ascilite Sydney*, pages 75–86, 2010.

María Jesús Rodríguez-Triana received her PhD in Information and Communication Technologies from the University of Valladolid (Spain) in 2014, joining the École Polytechnique Fédérale de Lausanne (Switzerland) as a Postdoctoral Fellow. She is a Senior Researcher at the School of Digital technologies in Tallinn University (Estonia). Her research interests have a strong focus on classroom orchestration and teacher inquiry supported via learning analytics, including aspects such as pedagogically-grounded and multimodal analytics, ethics and privacy.

Luis P. Prieto is a Senior Research Fellow at the School of Educational Sciences in Tallinn University (Estonia). He received his Ph.D. from the University of Valladolid (Spain) in 2012, and was a Marie Curie Fellow at the École Polytechnique Fédérale de Lausanne (Switzerland). His research interests include learning analytics, especially multimodal learning and teaching analytics, the study of teacher orchestration, and their application for teacher professional development.

Adrian Holzer is a Professor of Management Information Systems at the University of Neuchâtel. He holds a PhD in Information Systems from the University of Lausanne. He was a research associate at EPFL, the co-head of the interdisciplinary platform at the University of Lausanne, and an SNF research fellow at Polytechnique Montréal. His research interests cover digitalisation in learning and humanitarian contexts.

Denis Gillet received Ph.D. degree in Information Systems from the Swiss Federal Institute of Technology in Lausanne (EPFL) in 1995. He is currently a Faculty Member at the EPFL School of Engineering, where he leads the REACT multi-disciplinary research group. He is also affiliated with the EPFL Center for Learning Sciences. Dr. Gillet was Associate Editor of the IEEE Transactions on Learning Technologies (TLT) and of the International Journal of Technology Enhanced Learning. His current research interests include Digital Education, Human Computer Interaction (HCI), as well as Humanitarian Technology and ICT for Development. Denis Gillet is also co-founder of the Swiss EdTech Collider, an accelerator for startups in digital education. Finally, he is the technical coordinator of large-scale European innovation actions for STEM education at schools.