Instruction, Student Engagement and Learning Outcomes: A Case Study Using Anonymous Social Media in a Face-to-face Classroom

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 in social media, numerous 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. To investigate these issues, this paper presents a case study on the usage of a social media app (SpeakUp) during a semester-long face-to-face 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. This study also points out that the mere introduction of social media in educational settings does not guarantee a positive impact on learning. To promote an effective use of social media in the classroom (raising engagement while avoiding disaffection), teachers need to design the learning activities aligning the use of social media with the learning goals.**

***Index Terms*—Collaborative learning tools, Computer-assisted instruction, Social networking**

I. Introduction

SOCIAL media platforms such as Facebook, Twitter or WhatsApp enable individuals and communities to share, co-create, discuss, and modify user-generated content through the web [1]. Given the wide adoption of these platforms among students in their personal life, there have been several attempts in the educational research community to reach students through these platforms [2]–[5] or propose alternative, education-oriented tools with similar features (e.g., Edmodo, Twiducate or SpeakUp).

Social media tools have been used especially to foster interaction between teachers and students. While both students [6]–[11] and teachers [12] perceive improvements in participation, there is still conflicting evidence about their impact on learning [13], [14]. As with other educational technologies and innovations, the mere introduction of social media does not guarantee a positive effect on learning [15], [16].

To contribute to the understanding of how social media can be used effectively in educational settings, this paper explores the use of one such tool in face-to-face classrooms, also referred to as 'co-located' settings. More specifically, we aim to address the very real practitioner question of whether using these tools (especially, those that allow anonymous posting of messages) is related, not only to a raise in engagement but *also* of learning outcomes [15], [17]. Apart from being relevant for everyday educational decision-making, it also represents an unsolved research and methodological question (as studies linking individual outcomes with the usage of such anonymous messaging platforms are inherently difficult to perform).

After providing an overview of existing works on student engagement and social media for educational purposes, this paper introduces SpeakUp [18], a mobile app designed to promote participation in face-to-face settings. 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 three teachers during one semester in a lecture hall. Following the model proposed by Skinner et al. [19], 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 behavioral and emotional engagement and disaffection with the learning activity, observed through the tool logs and face-to-face participation in the classroom. 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 [19], [20]. The case study methodology [21] guided the data gathering and analyses, leading us to involve multiple informants, use different data gathering techniques, apply mixed methods analyses, and process evidence from the digital and the physical space.

Manuscript received March 3, 2019; Revised XXX, 2019.(*Corresponding author: María Jesús Rodríguez-Triana.*)

This research has been co-funded by the European Union's FP7 and Horizon 2020 research and innovation programmes (grant agreements no. 317601, 731685 and 669074).

M. J. 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 (e-mail: denis.gillet@epfl.ch)

Digital Object Identifier XX/TLT XXX

II. Related Work

A. Student Engagement

Given the current emphasis on student success and dropout rates in formal education, understanding and promoting student engagement has become an important issue among practitioners and researchers in the educational community [22], considering it an indicator of intrinsic will to learn [23]. As Henrie et al. explain in their review [24], 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' [25]– has later evolved, connecting the time and effort invested by the students to the learning activities and the desired learning outcomes [26].

Although closely related, some authors differentiate engagement and disaffection (or alienation) by looking at the occurrence of behaviors and emotions that reflect maladaptive motivational states [19], [20]. Engagement is considered a strong predictor of student learning, grades, achievement, and school retention [19]. Symmetrically, disaffection has been found to be a strong predictor of poor grades, low achievement in test scores, and eventual dropout.

Several authors proposed the engagement–disaffection dyad as a framework to understand students' relationships to their learning outcome [17], [27], [28]. More concretely, Skinner [28] proposes to look at three elements in those relationships: the *contextual factors* that influence engagement and disaffection (such as the social context, self features, and experience), *student actions* that can be observed in the classroom, and consequences of engagement and disaffection, namely, *learning outcomes* and student success.

Fredricks et al. [29] 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 these three levels [28]: behavioral disaffection includes passivity and withdrawal from participating in learning activities; emotional disaffection spans boredom, anxiety, and frustration states in the classroom; and cognitive disaffection is illustrated by, e.g., aimlessness, resignation, apathy or pressure.

Among the 113 papers studied by Henrie et al. in their review on student engagement in technology-mediated learning settings [24], 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. 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 were followed in the reviewed papers [24]: quantitative self-report (e.g., surveys), qualitative observations (e.g., direct or video observations of students' behavior while learning, interviews, focus groups, and the analysis of digital content) and quantitative observations (e.g., frequency indicators obtained from direct o video-based human and digital observations). Each strategy posed different pros and cons. Self-reporting is useful and frequently used for investigating aspects of student engagement which are not easily observable (especially, around emotional or cognitive engagement). This strategy is in general more scalable than human observations, however, it tends to require significant effort from students, and may be disruptive and intrusive within educational practice. On the other hand, observational methods have the advantage of enabling researchers to measure engagement and disaffection during the learning process without being disruptive. While useful for exploratory studies, qualitative observations are difficult to scale since extensive (human) resources may be needed to collect and analyze the data. On the other hand, quantitative observations, while being more prone to be automatized, may limit the aspects of engagement and disaffection that can be studied. Thus, the combination of both quantitative and qualitative strategies could offer the possibility of using qualitative measures for describing engagement and disaffection, and quantitative ones to establish comparisons between individuals or groups and to see their progress over time [19].

B. 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. Among those practices, fostering social interaction in the classroom is considered by numerous researchers as a *conditio sine qua non* for learning [30], [31]. Since the 80's, when IBM started to experiment with student interaction systems [32], this idea has prompted the apparition of a myriad of digital solutions. 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 [8], [33]–[35]. 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 [11], with which students could send SMSs to the teachers.

More recently, many educational technologies have included social media features (such as social networking, publishing and sharing, collaboration, or content management) [3], enabling, for instance, students to vote and comment on each other's contributions (e.g., ClassCommons [36], Fragmented Social Mirror [7], Pigeonhole Live [37], Backchan.nl [9], or SpeakUp [10]). Mainstream social media, such as Twitter [2], [16], Facebook [4], and Reddit [38], are also popular when attempting to foster interaction in the classroom. 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 [6]–[11]. Furthermore, students often prefer to use a digital channel to interact instead of raising their hand [11].

A survey with 7969 U.S. higher education teachers [12] found that 70.3% of faculty used social media at least once per month in their private life and 41% in the classroom. Teachers saw social media and technology as having a "considerable potential" for learning, and a majority (78.9%) stated that digital communication increases communication with students. However, 56% of teachers also considered that social media in class can be more distracting than helpful.

Beyond student and teacher perceptions, multitasking is receiving increased attention, still with conflicting results. Certain studies suggest that laptop multitasking hinders learning for both users and nearby peers [13]. 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 [39]. 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 [40], [41].

While several studies show significant correlations between the use of educational technology and student engagement [26], [42], [43], many of these studies take place in contexts where the use of technology can almost exclusively be on-tasks (e.g., a university's learning management system or a MOOC provide fewer opportunities for off-task behavior and distraction than other technologies do). Hence, it is not surprising how these studies (e.g., [44]) find that more interactions with the technology are related to better learning outcomes. In contrast, when students interact with social media platforms, both on- and off-task activities are possible (or likely).

The availability of digital traces, has raised the interest in Social Learning Analytics [45]. However, as already noted in Tess' literature review [46], most of the existing research on the utility and effectiveness of social media relies on self-reported data (e.g., surveys, questionnaires) and content analyses (see e.g., WhatsApp [47], Pigeonhole Live [37], Mentimiter [48], Polleverywhere [49], or SpeakUp [10]), which raises the need for further exploring participation attitudes. While we have not found in-depth studies about student engagement and disaffection, Table I provides an overview of similar studies -where social media apps were used to facilitate students anonymously participating in face-to-face classrooms- that report at least some general metrics of engagement and disaffection based on user interactions with the apps.

Thus, when introducing these tools, we cannot assume a direct relationship between engagement and learning outcomes [15], [16]. Further research and practitioner involvement are 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).

C. Assessing the Role of Social Media in the Classroom

Echoing Chapman and Junco's claims [15], [17], there is still a need for understanding the possibilities and limitations of social media, and measuring its impact on learning. In light of such need, this paper aims to contribute to better understanding whether, and under what circumstances, social media usage in the classroom may have a positive impact on learning. To guide us in this effort, we extracted a number of guidelines from the related work presented in the previous sections, shaping the design and methodological decisions of our case study as follows:

- *Choose a social media platform that fits the educational goal* [15], [16]: among the different social media apps available in the market, the lectures involved in the study chose SpeakUp (see Section III) because it 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* [57]: the different use cases for SpeakUp in the classroom were predefined before the study (see Section V-B). 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* [15], [19]: the study spans 6 lessons of a semester-long course, enabling longitudinal analysis of the course evolution (see Section IV-B).
- *Both on-task and off-task activities should be taken into consideration* [16], [19]: as depicted in Figures 2 and 4, our research questions and indicators take 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 others react to them.
- *Gather quantitative and qualitative evidence* [58]: beyond looking at the volume of observable measures, our study includes certain measures of their quality by categorizing comments (see Table II) and mapping student actions 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* [15], [19], [28], [29]: this study extracts indicators related to behavioral and emotional dimensions of engagement and disaffection that go beyond single behavior variables (see Figure 4).
- *Engagement and disaffection happen in a context and generate an outcome* [19], [24], [28], [29]: following Skinner's model [19], our research questions (see Section IV) explore the relation between the engagement and disaffection, the teacher instruction, and learning outcomes.
- *Carefully select which measures of student engagement and disaffection should be used to study the relationship with other variables* [24]: to understand which indicators are more relevant in relation to the context and the

TABLE I
SIMILAR STUDIES WITH SOCIAL MEDIA APPS TO FACILITATE ANONYMOUS PARTICIPATION IN THE CLASSROOM.

Tool & Study	Audience	Duration	Engagement / Disaffection
Fragmented Social Mirror [7]	180 HE students	3 sessions	- In total: 11 questions, 106 messages - 46 off-topic messages
HandsUp [50]	250 HE students	1 session	- In total: 18 questions, 102 votes - No off-topic contribution
Sli.do [51]	34 HE students	2 sessions	- In total: 31 students submitted a comment on demand - No off-topic contribution
Sli.do [52]	110 HE students	6 weeks (1-hour sessions)	- In total: 25.5% students submitted >2 questions, 31.8% 1-2 questions, 42.7% no question - Off-topic contributions detected
Google Docs & SurveyMonkey [53]	154 HE students	22 1-hour sessions	- On average per session: 6 contributions per student, 35 students answered the polls - No off-topic contribution
Hotseat [54]	70-450 HE students per class	Multiple lengths but not specified	- On average per semester: 3-18 posts per student
Engage [55]	<100 HE students	8 45-min sessions	- In total: 54 questions - On average per session: 3-22 questions - Participation fluctuated between sessions without decreasing ostensibly over time (linked with the prompting by the lecturer)
TodaysMeet [56]	41 HE students	7 sessions	- In total: 54 questions - On average per session: 7-21 questions - 4% of the questions were off-topic

learning outcomes, we have run correlational and multiple linear regression analyses. Further details can be found in Sections IV-B and V-D.

- *Look at individual and social forms* [58], [59]: as Section V-D shows, the social aspect is considered both as a contextual factor (in relation to the individual and collaborative activities), and as an analysis dimension (looking both at the individual and at the classroom level).
- *Combine self-reported and observable indicators* [24]: as described in Section IV-A, the study gathers both student perceptions and 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* [24]: in our case, SpeakUp traces are used to monitor the behavioral and emotional engagement as well aw disaffection in the tool.
- *Avoid being disruptive and intrusive* [24]: apart from a poll and pre/post-surveys, observational methods were used to gather evidence while the learning process occurred. Besides, interactions mediated by SpeakUp were automatically collected to reduce the survey and (human) observation workload. A detailed description of the informants and data sources is provided in Section IV-A.

III. SPEAKUP

SpeakUp (www.speakup.info) is a social media tool designed to facilitate and foster participation in physical settings 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).

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 off-task. 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 [15], 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 II, apart from the pedagogical affordances, a crucial reason that led us towards 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

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.

tools, has an outstanding added value for practitioners and researchers willing to understand how the tool has been used.

IV. Methodology

The present study is framed within a wider research effort towards understanding how social media can be used effectively as an additional channel of communication in face-to-face classrooms. Several studies exist about the use of SpeakUp in classrooms, which concluded that the tool was easy to use and motivated students to participate more in lectures [10], [18], [60]. This paper explores how different definitions of student action relate to learning outcomes (i.e., how informative these definitions are to predict academic performance), and to pedagogical aspects of the learning context (e.g., the role teacher instruction plays on student action) in face-to-face learning settings mediated by this technology. To define student action, we use relatively simple and unobtrusive indicators of behavioral and emotional engagement and disaffection already reported in the literature [24].

To achieve a deeper understanding of the usage of this social learning tool in a context as complex as classrooms are, we adopted a case study methodology [21]. This research methodology helps to inform practice by illustrating what has worked well, what has been achieved and what have been the issues or dilemmas that played out in a real-life scenario. Hence, to understand how SpeakUp can be integrated effectively in face-to-face classrooms, this case study addresses three main research questions (see 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 instructional activities, especially regarding SpeakUp) plays on student action?

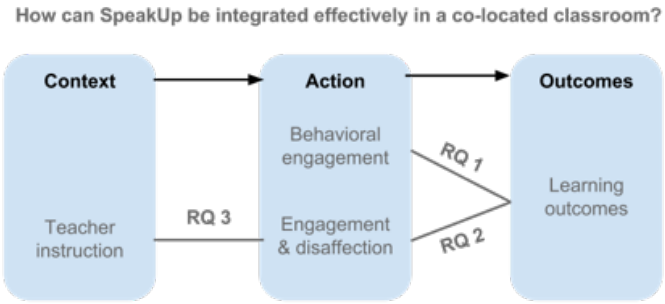


Fig. 2. Diagram representing the relation between the research questions and the theoretical model proposed by Skinner et al. [19].

A. Informants, Data Gathering, and Data sources

Aligned with the trends in technology-enhanced learning research [61] and the case-study methodology [21], we used mixed methods [62], [63] 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) helped us gather quantitative and qualitative data using different techniques such as questionnaires, observations, action logging, and reports.

Our case study followed the use of SpeakUp during a whole semester, in the face-to-face lectures of a blended university course (see Section V-A for further details). During the face-to-face sessions, each seat in the classroom had a code which was used by the observers to keep track of the student participation such as hands-up questions and other oral interventions (see Figure 3). To be able to track the students across lessons, the first day the students were asked to freely choose where they would like to sit and keep the same place during the whole semester. 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 throughout the course. Since due to the anonymity it would not have been possible to relate learning outcomes with the digital traces, the students were asked to voluntarily disclose their anonymous user identifier for the purposes of this research. To gather the student perspective about the usage of SpeakUp, questionnaires were sent to the students before and after using SpeakUp (to map the pre and

TABLE II
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"*

post questionnaires, students provided a non-identifiable ID). 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.

Fig. 3. Picture from one of the lessons of the study. Each student had a card with an identifier to enable the observations about their face-to-face interventions during the sessions.

B. 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* [64] 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 (1182 messages), into three main categories: *on-task* (e.g., questions or comments about the content, course organization, or SpeakUp), *off-task* (e.g., spam or bullying messages) and *neutral* (e.g., greetings and policing messages). Table II shows examples of messages in each category. During the coding 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 (from among the six teaching-learning activity categories above) 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, or when SpeakUp was used in any of the manners/scenarios described in Section V-A ('Backchannel', 'Ask me anything', 'Quiz', or 'Think-pair-share').

To operationalize the different elements of the theoretical model proposed by Skinner et al. [19] (namely, context, action and outcomes), multiple indicators proposed in the technology-mediated learning literature [24] have been used (see Figure 4). While some indicators take into account simple self-reported data about emotional or attitudinal factors (e.g., a single-question poll about the perception of the tool as a distraction, 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 [24]. Hence, we make more intensive use of observational data collected by teachers or tools (e.g., classroom attendance or student actions logged in the system) 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). Considering that the distributions of the student action indicators and learning outcomes were not normal (as tested with Shapiro-Wilk tests), we used the non-parametric and more robust Kendall's correlation coefficient (τ) and tests to evaluate such associations. Since multiple hypotheses are tested (for each indicator), and to keep the false discovery rate under control, p-values in these tests were corrected using the Benjamini-Hochberg method. Further explorations have been made through stepwise multiple linear regression modelling, to understand the relative strength of these trends. We used the 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 indicators and removing the least useful ones), we arrived at models that try balance simplicity and predictive power. When performing these regressions, to avoid the problem of multicollinearity (i.e., several of the proposed indicators were indeed correlated among themselves – like total and average off-task message length) that could

muddle the interpretation of the regression results, several indicators were removed from the modelling *a priori* (using the variance inflation factor –VIF– as an indicator). For similar reasons, exceedingly-influential outliers were removed when building the models. Other regression model assumptions (e.g., homoscedasticity, normality of residuals) were checked using diagnostic plots (e.g., Q-Q plots) and statistical tests (e.g., Anderson-Darling test to evaluate the normality of residuals).

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 kinds of analyses were also made taking the time dimension into account (e.g., aggregated measures for the whole course vs. associations per session). Furthermore, to explore how different kinds of teacher actions were related to student use of the tool (RQ3), we have compared the distributions of actions/indicators across different kinds of teaching/learning activities. Given the aforementioned non-normality of indicators, we performed multiple Kruskal-Wallis rank sum tests (which is a non-parametric test that does not assume/require a particular distribution of the data). This test helped us 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, or the teaching suggesting the use of SpeakUp) seem to be coming from substantially different distributions (e.g., with different median values) or not. Subsequent Dunn tests (with multiple comparison p-values adjusted using the Benjamini-Hochberg method) indicated which pairs of activities were significantly different in terms of the action indicators observed during them.

V. Case Study

A. 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. Figure 3 provides an overview of the setting where the course took place. Three teachers led the course and 149 students (37 females) registered to it. The three 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. Regarding the students, the pre-questionnaire shows that attitudes towards anonymous social media change depending on gender, being females significantly more positive than males to ask more questions with an anonymous tool.

The course introduces first-year undergraduate engineering students to different kinds of communication channels, social media platforms and technology-enhanced learning. A critical goal of the course is to develop self-assessment and critical thinking skills in those settings. Thus, teachers often raise controversial claims often appearing in the media. Due to the scale of the audience and the topics discussed, lecturers in this course often struggle to enable (honest) student participation during the face-to-face sessions.

The course was divided into six face-to-face teacher-led 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. During those sessions, SpeakUp was introduced as a complementary communication channel with students (aside from the classic face-to-face verbal communication), to increase interaction in four different ways:

- *Backchannel.* Throughout the course, the application was used as a digital channel to promote the interaction among participants during the lessons. Students freely posted comments and interacted with their peers by answering and voting each others' comments.
- *Ask me anything.* 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 assess the knowledge or to poll the opinion of the students. This usage of the platform was often coupled with (face-to-face) group discussions.
- *Think-pair-share.* In this kind of activity, the teachers first guided the audience to think individually 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.

B. Action: Attendance, Participation and Usage of SpeakUp

Figure 5 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 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, 128 (85.9%) volunteered to disclose their anonymous user identifier for the purposes of this research (i.e., to relate their learning outcomes with their app behavior). It should be noted that, 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. All in all, we registered 243 different student user identifiers in SpeakUp.

C. Learning Outcomes

As mentioned in Section IV-A, two learning outcomes were taken into consideration: the results of the multi-choice test

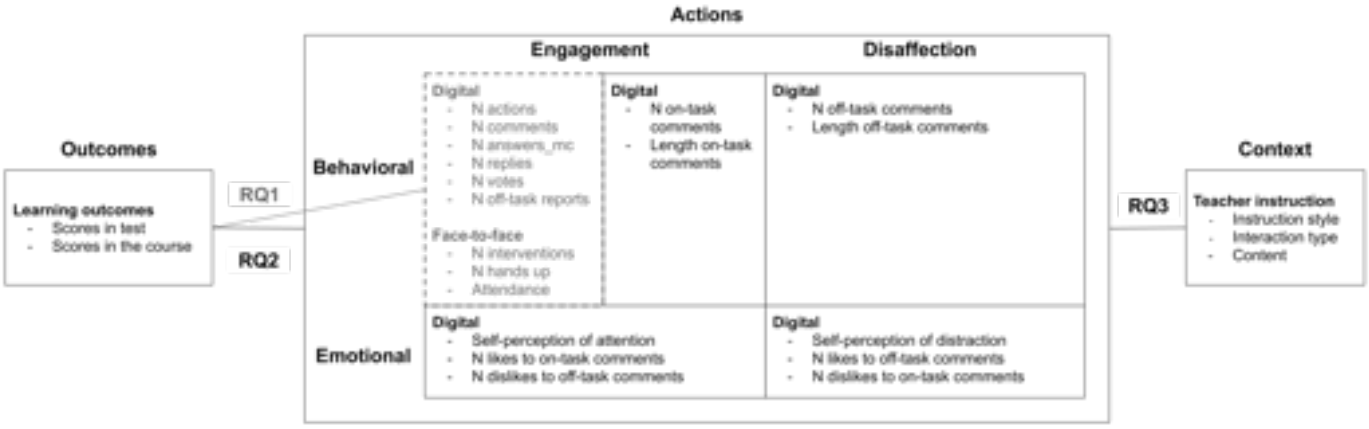


Fig. 4. Indicators extracted in each one of the research questions.

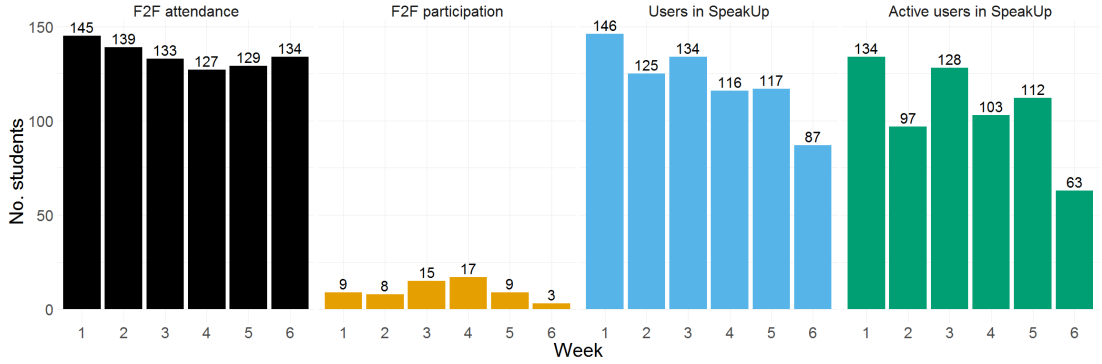


Fig. 5. Overall counts of face-to-face and online participation in the six sessions of the course.

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 [65]. 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).

TABLE III
OVERALL COUNTS OF (SIMPLE, COUNT-BASED) BEHAVIORAL ENGAGEMENT INDICATORS, AND CORRELATIONS OF SUCH INDICATORS (KENDALL'S τ) WITH THE LEARNING OUTCOMES OF EACH STUDENT.

Indicator	Counts	τ
n_actions	22632	0.17*
create_message	928	0.10
reply_comment	254	0.21*
answer_mc	372	0.11
reports_off-task	20	-0.04
vote_comments	8489	0.13
prompted participation	64	0.03
unsolicited hands-up	4	-0.10

Note: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.
(p-values adjusted using the Benjamini-Hochberg method)

D. Results

1) How does a simple definition of student action relate to learning outcomes? (RQ1): As mentioned in Section IV, 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:

Classroom-level analysis. Table III and Figure 6 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 "novelty 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 RQ3 results 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 8 below.

Student-level analysis. Since our goal was to understand

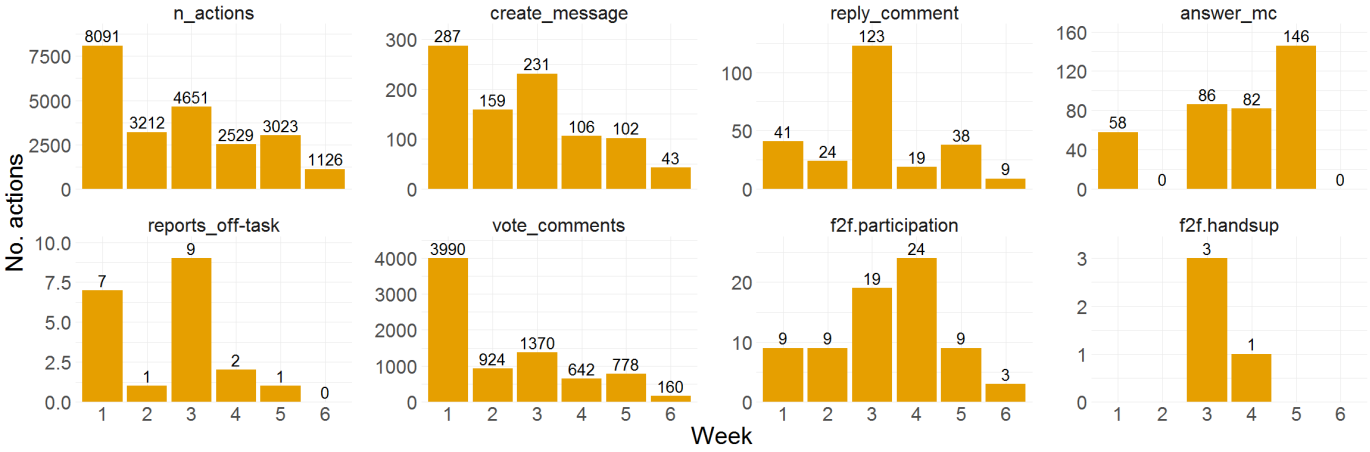


Fig. 6. Overall counts of simple behavioral engagement indicators for each of the sessions in the course.

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 final scores of the 124 students who disclosed their identity and completed the course, we can explore the correlations of the simple behavioral engagement indicators with such learning outcome (Table III). We can observe that many of these indicators bear little correlation with the student learning outcomes. The exceptions are the raw number of actions (of all kinds), and the number of replies to other students' messages, that each student generated in SpeakUp, which seem 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 off-task in SpeakUp, or putting your hand up to ask questions during the classes). However, these are relatively rare events in the dataset, and we should not infer too much from them.

We further explored whether these indicators, or combinations of them, were related to the learning outcomes. Our stepwise multiple linear regression analysis led to the model that is shown in Table IV (top). We can observe that it uses the same two indicators correlated with the learning outcomes above (the total number of SpeakUp actions, replies to others' messages), but neither of those predictors is significant (meaning that we are not certain that the actual value of the coefficient is not zero), and the model 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.

2) How does a more complex view of action relate to learning outcomes (RQ2)?: Classroom-level analysis. Table V and Figure 7 provide an overview of the additional indicators covering behavioral and emotional engagement and disaffection (see Figure 4), 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 off-task messages and their average length) a similar downward trend after the first session (cf. the notion of there being a "novelty 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).

TABLE IV
LINEAR REGRESSION MODELS OF LEARNING OUTCOMES BASED ON DIFFERENT KINDS OF INDICATORS OF ENGAGEMENT AND DISAFFECTION.

Indicators	B	$Std.Err.B$	t	p
Model: Simple Behavioral Engagement ($Adj.R^2$ =0.05 ; p =0.02)				
n_actions	0.07	0.05	1.47	0.14
reply_comment	0.07	0.56	1.43	0.15
Model: Engagement & Disaffection ($Adj.R^2$ = 0.15 ; p = 0.00)				
avg. length on-task messages	0.12**	0.04	2.79	0.01
likes to on-task messages	0.18**	0.06	3.07	0.00
dislikes to off-task messages	-0.10	0.06	-1.89	0.07

Note: All predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

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 V shows the correlations (Kendall's τ) 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 exceptions in this case are the average on-task message length, the number of 'likes' given to other people's on-task messages, and the number of on-task messages posted by oneself. These indicators also make somewhat intuitive sense as predictors of learning outcomes, as they may require a certain level of cognitive processing on the part of the students (to compose long, relevant messages, or read and understand others' relevant messages in order to upvote them).

In a similar manner as we did in the previous section with the simple behavioral engagement indicators, here we

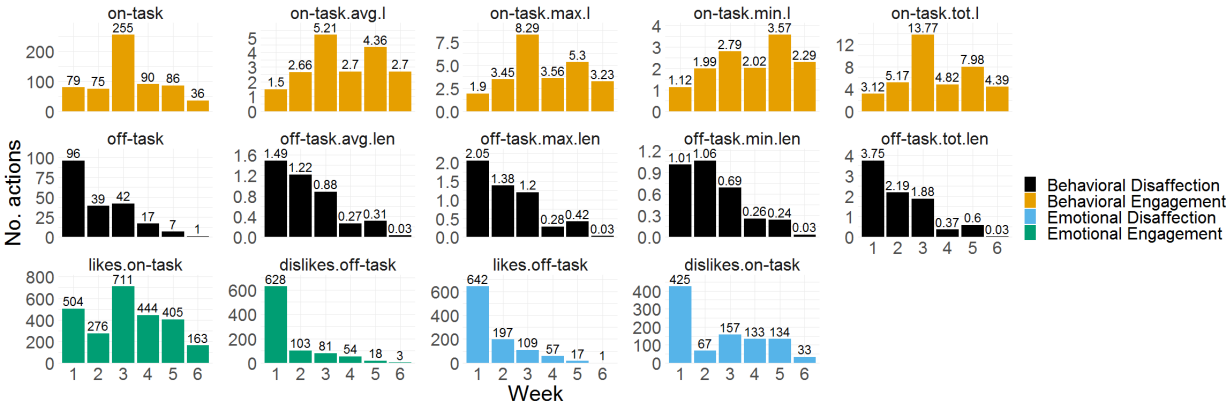


Fig. 7. Overall counts and averages of additional engagement and disaffection indicators for each of the sessions in the course.

TABLE V
OVERALL COUNTS/AVERAGES OF SELECTED ADDITIONAL ENGAGEMENT AND DISAFFECTION INDICATORS, AND CORRELATIONS (KENDALL'S τ) WITH THE LEARNING OUTCOMES OF EACH STUDENT.

Indicator	Category	Counts/Avgs.	τ
off-task messages	Behavioral Disaffection	202	-0.01
off-task message length (avg.)	Behavioral Disaffection	1.71	0.02
on-task messages	Behavioral Engagement	621	0.17*
on-task message length (avg.)	Behavioral Engagement	4.59	0.23**
dislikes to on-task messages	Emotional Disaffection	949	0.10
likes to off-task messages	Emotional Disaffection	1023	0.13
Perception of SpeakUp as useful/distracting (avg. poll, [-2,+2])	Emotional Engagement Disaffection	0.19	0.02
dislikes to off-task messages	Emotional Engagement	887	0.06
likes to on-task messages	Emotional Engagement	2503	0.19*

Note: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.
(p-values adjusted using the Benjamini-Hochberg method)

again used stepwise multiple linear regression to understand the relationships and respective importance of the different indicators. Table IV (bottom) shows the main parameters of the resulting best-fitting model. We can observe that the model performs better in this case (it explains now about 15% of the variance in the outcomes). The model now has two significant predictors, both with a positive coefficient: a) the average length of on-task messages posted; and b) the number of upvotes made to others' on-task messages. These two indicators had already been detected as potentially useful by the correlation analyses above.

It is interesting that the third indicator used in the model, the downvoting (or 'disliking') of off-task messages, has a negative coefficient. While this indicator could be intuitively understood as a marker of positive emotional engagement, the model seems to suggest that, once we control for the other variables in the model, policing others' non-relevant messages in SpeakUp does not lead to better outcomes (maybe rather distracting from engaging with the content). Yet, again, we should not infer too much from this coefficient, as the evidence for its being actually negative is not solid.

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 who want to scaffold learning using these tools. The following section deepens further into the latter.

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 completely) 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 may be 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 8 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 V-A) by the lecturers triggered the sudden rise of both 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

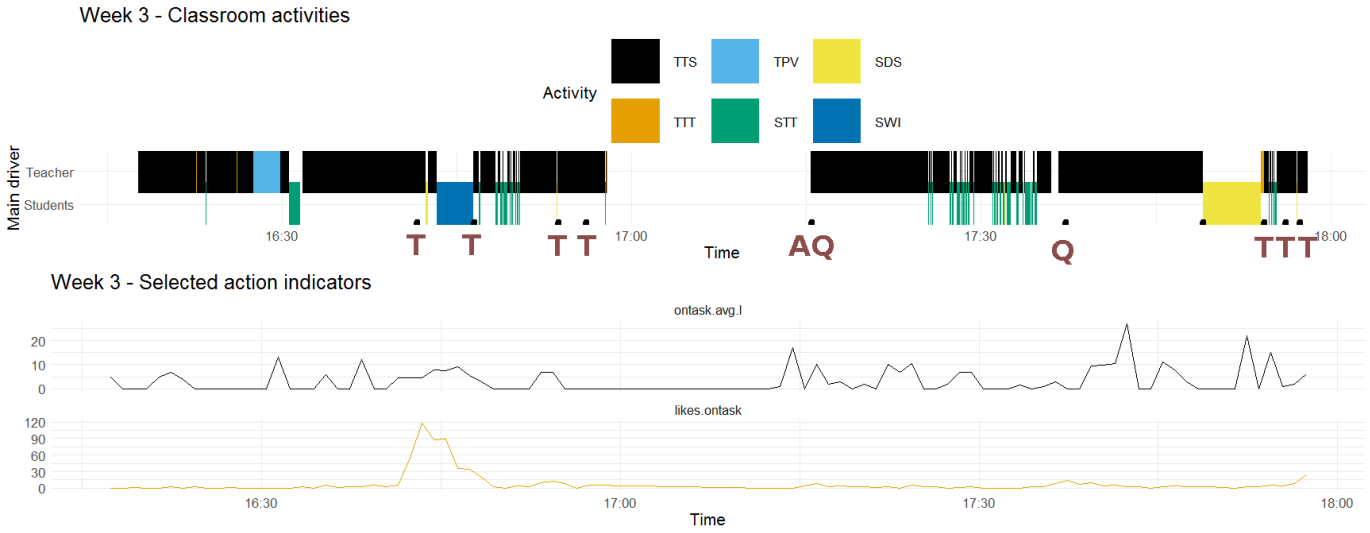


Fig. 8. Graph of selected engagement and disaffection indicators (significant in the linear regression models), compared with the activity happening in the class during session 3. Black dots and letters along the timeline indicate moments in which a lecturer explicitly pointed students to SpeakUp activities (see Section V-A): T=Think-pair-share; A=Ask me anything; Q=Quiz

nevertheless be arduous by mere visual inspection. In this sense, we could ask whether there is an association between the moments the teachers were mentioning or encouraging actively the use of SpeakUp (marked with black dots and letters in Figure 8) and increased amounts of those indicators we have identified as related to learning. Our Kruskal-Wallis rank sum tests seem to indicate that the distribution of likes to on-task messages *is* indeed different (and with higher median value) during those episodes in which the teachers were suggesting to use the tool ($median = 2$ vs. 0, $\chi^2 = 20$, $df = 1$, $p = 0.000$). However, the same tests indicate that there may not be such an association in terms of length of the content-relevant messages ($\chi^2 = 3$, $df = 1$, $p = 0.07$).

We could also ask whether there are any definite patterns in the amounts of action indicators witnessed by the different kinds of classroom activities throughout the course. As Table VI shows, Kruskal-Wallis rank sum tests suggest that the distributions of action values that were associated with learning, 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 average length of on-task messages (which was positively associated with learning outcomes), certain classroom activities like students working independently (SWI) and students asking questions to the teachers (STT) tend to foster longer relevant messages. This makes sense as these kinds of episodes give the student time to think, engage with the content individually, and maybe post relevant questions or comments. This hints at potentially useful pedagogical patterns of use of SpeakUp: lecturers could intersperse such individual reflection or questioning episodes, rather than lecturing non-stop, or fostering student discussions without an individual reflection phase (e.g., using Think-Pair-Share).
- Regarding the number of likes added to on-task messages by others (also positively associated with learning outcomes), we again find activities like students working individually (SWI) or students talking to the teacher (STT) as having more of these behaviors than student discussions (SDS) or teachers playing a video (TPV). Even the classic lecturing episodes (TTS) enable such interaction with other students' content-relevant posts. Again, this makes intuitive sense since student-to-student discussions tend to be highly engaging, redirecting the attention of students away from SpeakUp (and for good reason, probably). This again speaks to having a balanced mix of teaching/learning activities, that enable different ways of engaging with the content, and directing the students' attention into and away from the social apps and other media.

TABLE VI
MEDIAN VALUES OF ENGAGEMENT AND DISAFFECTION METRICS, FOR THE DIFFERENT KINDS OF CLASSROOM ACTIVITIES CODED THROUGHOUT THE COURSE.

Activity	SDS	STT	SWI	TPV	TTS	TTT	χ^2	df	p
on-task message length (avg.) (+)	0	3.0	3.89	0	0	0.5	20	6	0.002
likes to on-task messages (+)	0	2.0	2.5	0	1	2.5	30	6	0.000

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.

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 set up, 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 [10] already points

out, using the tool effectively might require monitoring and guiding students' attention to more productive uses of it.

VI. 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 [6]–[9], [11], [18], [66]. 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 [44]). However, the evidence of a relationship between technology use and academic performance is still unclear in the case of less-controlled technologies like anonymous social apps. In this paper, we contribute to this knowledge base through a case study which followed the use of SpeakUp during a university course.

When looking at the evidence we gathered from different sources, a first question we can ask is whether students engaged at all in the use of the tool. While collaborative learning situations, where the digital tool is expected to be the primary form of communication, can see high digital interactions traffic (e.g., [67] report an average of 15 messages per student in just under one hour of activity), a lecture-style setting where the technology is used rather as a 'backchannel' to complement the main (i.e., verbal) communication, can see very different traffic profiles (see Table I for concrete volumes of engagement indicators). In this sense, the student activity within our study (with an average of about 8 postings and 59 likes per student in the whole semester), is well within the expected range for this kind of applications. This again speaks against taking simplistic measures of engagement (like message counting) at face value, without considering further information about the educational setting being studied.

Another important, inherent factor of the use of this kind of social apps in the classroom comes from the anonymity features themselves (which make it especially difficult to research, and may be related with the predominance of self-report measures in prior studies). To what extent could the fact that only part of the students disclosed their identity, bias our results? We can make several observations and checks to assess this possibility: 1) anonymous accounts, while numerous (46% of the observed student accounts), represent a comparatively low proportion of the traffic in our case study (around 20%); 2) anonymous accounts did not engage in disproportionately different on-/off-task behaviors (e.g., always posting off-task messages), and both disclosed and anonymous accounts had a larger tendency towards on-task messages; 3) students that did not disclose their identity did not have substantially different characteristics, like course outcomes, than those that did (a Kolmogorov-Smirnov test fails to reject the hypothesis that scores from students who disclosed their identities were drawn from the same distribution as anonymous ones, $D = 0.32$, $p = 0.08$). Based on these data, we suggest that anonymous accounts may have introduced moderate amounts of noise in the results (which may be one reason for not observing larger correlations and variance explained by the models in Section V). Indeed, we could consider it remarkable that even with such noisy conditions, the proposed models and indicators still retain some of their predictiveness.

We could also ask whether the observed student behaviors in SpeakUp were affected by students' dispositions or attitudes towards the technology. The pre and post questionnaires were used to reveal attitudes such as feelings of inclusion and attitudes towards the technology that we kept anonymous. This questionnaire shows that students had a positive attitude towards the use of technology in the classroom ("I find the use of technological tools during classes useful" on a 7-level Likert scale $M = 5.90$, $SD = 1.14$). Related to this issue of dispositions, it is worth mentioning that the poll performed via SpeakUp on the first session of the course, about their perception of the tool as distracting, can start revealing potential relationships between student attitudes or dispositions, and their actual behaviors of engagement in SpeakUp. Interestingly, the students' answers to this poll ($n = 92$, including both anonymous and students that revealed their identity) are rather uncorrelated with their later on-/off-task usage of the tool (e.g., correlation with the "relevance ratio" of a student's messages is $r = -0.01$). Future investigations could include more detailed measures of antecedents to engagement, such as predispositions to boredom or procrastination which have been found to be negatively associated with engagement to online learning management systems [68].

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, but probably insufficient to build fitting models of academic performance in the course (Adj. $R^2 = 0.05$), especially compared with other settings like MOOCs.

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 would require manual human intervention) explained a much larger proportion of variance in the academic performance (Adj. $R^2 = 0.15$). 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" [24], [29]. The fact that our models included both positive, expected indicators (like on-task message length), but also somewhat surprising negative ones (disliking an off-task message, which in principle is a good thing), is also in line with the experiences of many practitioners, which consider that social media in class can be sometimes more distracting than helpful [12], and that not all engagement with the tool is necessarily of equal

value. Finally, it is worth comparing these results with studies done on the use of Twitter for learning [16], 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. Our results further detail this notion, indicating that the length of the on-task messages created, and the kind of response they garnered (liking them, and ignoring the noisy, off-task messages), may also matter.

It seems that it is not using social media *per se* which is related to outcomes, but rather how students use (and are guided to use these) tools [15], [46]. 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. Patterns of tool usage and participation (e.g., longer on-task messages during individual work activities) result from our analyses of co-occurrence between the aforementioned indicators of student action (related to learning outcomes) and different kinds of classroom activities. 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 like downvoting an off-task comment). This hints at future work in this area, using advances in multimodal learning analytics [69] to complement the digital indicators of action with physical ones, or to gather automatically the classroom events [70], so that they can be used in analyses such as those presented in this paper.

Our investigation of whether the engagement indicators most related to learning outcomes changed substantially when teachers directed students explicitly to use SpeakUp (within RQ3), also highlights an often-overlooked issue: the importance of teacher actions (in the form of the pedagogical design, or through explicit action) as an influence on the observed engagement behaviors. This influence is ever more present in face-to-face, more teacher-driven scenarios like our case study, as compared to MOOCs or more learner-driven activities. Hence, we posit that it is crucial, when studying engagement, to specify (and gather data about) the pedagogical situation and teacher actions, which may help explain the patterns, ebbs and flows of engagement behaviors. Our study also highlights the need to avoid simplistic or mechanistic views of engagement, or trying to use engagement indicators as the target of interventions (as other technological fields and the industry often do). Rather, our results paint a more nuanced picture in which engagement indicators explain only a part of the variance in outcomes, and whereby having a balanced mix of learning activities (some involving the tool, others directing the attention away from it, or to individual reflection) and different levels of engagement with the technology, is beneficial. Our initial investigation on the impact of teacher actions on engagement indicators also brings to the fore the issue of the differential value or reliability of different indicators across teaching situations (e.g., it seemed that average on-task message length was not greatly impacted by teachers direct references to the tool). Overall, we can conclude that studying engagement and linking it to learning outcomes, in such a complex social setting as a classroom, is an inherently hard problem – and the anonymity afforded by SpeakUp adds further noise to any modelling or prediction efforts. Yet, we also illustrate what methods can be used to start tackling this problem (e.g., adding simple content analysis to the engagement indicators, and descriptions of classroom events), and improve our models of engagement in classroom learning.

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). Indeed, our study followed a primarily quantitative approach, to uncover associations between student behavior, learning outcomes and teacher actions. It did not, however, collect deeper qualitative data (e.g., interviews) that might have helped us delve into the students' experiences of using SpeakUp, and to understand *why* the students used the tool the way they did. Preliminary analyses of the (mostly quantitative) data from our pre-/post-questionnaires reveal intriguing associations between students' self-perceptions and attitudes and their enjoyment of using the tool (e.g., student enjoyment of the tool seems to be correlated with higher sense of safety to express themselves before the study, $\tau = 0.22$*, or higher sense of belonging to the class *after* the study, $\tau = 0.29$**). Yet, those student perceptions (and their relation with different tool usage profiles, such as those uncovered in prior studies with SpeakUp [60]) will need to be further explored in future research studies.

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 [44], [71] (teacher participation, course design, class size, student self-regulated learning characteristics, etc.).

VII. Conclusions

In our way towards understanding how to use social media effectively in the classroom, this paper analyses the use of SpeakUp as an anonymous communication addition in a face-to-face course with three teachers and 149 university students. Following the model proposed by [19], we have explored the context-actions and actions-outcomes relationships. In our case study, the learning context is represented by the teacher instruction, the learning outcomes by the students' scores in the test and the marks in the course, and the learning actions by relatively simple and unobtrusive indicators of behavioral and emotional engagement and disaffection used in the literature [24].

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 [24], [29]. While teachers and students reported in this course that SpeakUp was beneficial in terms of participation [60], our results show that social media in class can be also distracting, as teachers often point out [12]. Thus, in future research, there should be more emphasis on the *quality* of the participation, rather than on the quantity [15].

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 8 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 [15], [16].

While case studies are not envisioned for and do not enable generalizations, the in-depth nature of our case study allowed us to illustrate how the guidelines extracted from the literature can be applied in an authentic setting. Moreover, this case study provides insights for future research and practice. First, it illustrates the limitations of using exclusively behavioral measures of engagement and the added value of combining simple behavioral and emotional engagement and disaffection indicators [15], [19], [28], [29]. Second, the study confirms the need for looking at the quality and not only at the quantity of participation (e.g., to distinguish distraction from positive engagement) [58]. Third, in line with [10], our results suggest that effective use of the app relate to teacher-led usage more than the unstructured backchannel, where there is more room for distraction. On this regard, the extracted patterns connecting teacher instruction and student action may help others effectively integrate social media in the educational practice. These strategies can complement existing advice found in the literature, like allowing messages to be flagged and removed by the lecturer (to avoid spam) [72], or just asking students to behave professionally [53]. Forth, promoting the participation of those students that may require further support (e.g., posing questions) and addressing their needs may help to reduce the gap between students that perform well and those that do poorly.

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 [69], [70]. For example, automatic content analysis could be used to infer emotions from the messages, and audio analysis could automatically detect classroom activity based on the ambient noise.

Acknowledgements

This research has been partially funded by the European Union in the context of Next-Lab, CEITER (Horizon 2020 Research and Innovation Programme, grant agreements no. 731685 and 669074). The authors would like to thank the observers from École Polytechnique Fédérale de Lausanne and the video coders from Tallinn University for their contribution to this study.

References

[1] J. H. Kietzmann, K. Hermkens, I. P. McCarthy, and B. S. Silvestre, "Social media? Get serious! Understanding the functional building blocks of social media," *Bus. Horiz*, vol. 54, no. 3, pp. 241–251, 2011, doi: 10.1016/j.bushor.2011.01.005.

[2] D. Retelny, J. P. Birnholtz, and J. T. Hancock, "Tweeting for class: Using social media to enable student co-construction of lectures." in *Proc. ACM 2012 Conf. Comput. Supported Cooperative Work Companion*, S. E. Poltrock, C. Simone, J. Grudin, G. Mark, and J. Riedl, Eds. ACM, 2012, pp. 203–206, doi:10.1145/2141512.2141578.

[3] J. Mao, "Social media for learning: A mixed methods study on high school students' technology affordances and perspectives," *Comput. Hum. Behav.*, vol. 33, pp. 213–223, 2014, doi:10.1016/j.chb.2014.01.002.

[4] B. Dyson, K. Vickers, J. Turtle, S. Cowan, and A. Tassone, "Evaluating the use of Facebook to increase student engagement and understanding in lecture-based classes," *Higher Educ.*, vol. 69, no. 2, pp. 303–313, Feb 2015, doi:10.1007/s10734-014-9776-3.

[5] C. Greenhow and C. Lewin, "Social media and education: Reconceptualizing the boundaries of formal and informal learning," *Learning, Media and Technol.*, vol. 41, no. 1, pp. 6–30, 2016, doi:10.1080/17439884.2015.1064954.

[6] R. J. Anderson, R. Anderson, T. VanDeGrift, S. Wolfman, and K. Yasuhara, "Promoting interaction in large classes with computer-mediated feedback," in *Designing for change in networked learning environments*. Springer, 2003, pp. 119–123, doi:10.1007/978-94-017-0195-2_16.

[7] T. Bergstrom, A. Harris, and K. Karahalios, "Encouraging initiative in the classroom with anonymous feedback," in *Proc. Int. Conf. Hum.-Comput. Interact.*, ser. Lecture Notes in Computer Science, P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, and M. Winckler, Eds. Springer Berlin Heidelberg, 2011, vol. 6946, pp. 627–642, doi:10.1007/978-3-642-23774-4_49.

[8] C. Fies and J. Marshall, "Classroom response systems: A review of the literature," *J. Sci. Educ.and Technol.*, vol. 15, no. 1, pp. 101–109, 2006, doi: 10.1007/s10956-006-0360-1.

[9] D. Harry, J. Green, and J. Donath, "Backchan.nl: Integrating backchannels with physical space," in *CHI'09*. ACM, 2008, pp. 2751–2756, doi: 10.1145/1518701.1518907.

[10] S. Govaerts, A. Holzer, B. Kocher, A. Vozniuk, B. Garbinato, and D. Gillet, "Blending digital and face-to-face interaction using a co-located social media app in class," *IEEE Trans. Learn. Technol.*, vol. 11, no. 4, pp. 478–492, Oct 2018, doi:10.1109/TLT.2018.2856804.

[11] E. Scornavacca, S. Huff, and S. Marshall, "Mobile phones in the classroom: If you can't beat them, join them," *CACM*, vol. 52, no. 4, pp. 142–146, Apr. 2009, doi: 10.1145/1498765.1498803.

[12] J. Seaman and H. Tinti-Kane, "Social media for teaching and learning," Pearson Learning Solutions, Tech. Rep., 2013.

[13] F. Sana, T. Weston, and N. J. Cepeda, "Laptop multitasking hinders classroom learning for both users and nearby peers," *Comput. Educ.*, vol. 62, pp. 24–31, 2013, doi:10.1016/j.compedu.2012.10.003.

[14] 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," *Comput. & Educ.*, vol. 85, pp. 14–22, 2015, doi:10.1016/j.compedu.2015.02.002.

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

[16] J. H. Kuznekoff, S. Munz, and S. Titsworth, "Mobile phones in the classroom: Examining the effects of texting, Twitter, and message content on student learning," *Communication Education*, vol. 64, no. 3, pp. 344–365, 2015, doi:10.1080/03634523.2015.1038727.

[17] E. Chapman, "Alternative approaches to assessing student engagement rates," *Practical Assessment*, vol. 8, no. 13, pp. 1–7, 2003.

[18] A. Holzer, S. Govaerts, A. Vozniuk, J. Ondrus, D. Rigaud, B. Garbinato, and D. Gillet, "SpeakUp – A mobile app facilitating audience interaction," in *Adv. Web-Based Learn. – ICWL 2013*. Springer, 2013, pp. 11–20, doi: 10.1007/978-3-642-41175-5_2.

[19] E. Skinner, C. Furrer, G. Marchand, and T. Kindermann, "Engagement and disaffection in the classroom: Part of a larger motivational dynamic?" *J. Educ. Psychol.*, vol. 100, no. 4, p. 765, 2008, doi: 10.1037/a0012840.

[20] S. R. Jimerson, E. Campos, and J. L. Greif, "Toward an understanding of definitions and measures of school engagement and related terms," *The California School Psychologist*, vol. 8, no. 1, pp. 7–27, 2003, doi: 10.1007/BF03340893.

[21] R. E. Stake, *The art of case study research*. Thousand Oaks, CA: Sage Publications, 1995.

[22] J. M. Vytasek, A. Patzak, and P. H. Winne, *Analytics for Student Engagement*. Cham: Springer International Publishing, 2020, pp. 23–48. [Online]. Available: https://doi.org/10.1007/978-3-030-13743-4_3

[23] H. Coates, "A model of online and general campus-based student engagement," *Assessment & Eval. Higher Educ.*, vol. 32, no. 2, pp. 121–141, 2007, doi:10.1080/02602930600801878.

[24] C. R. Henrie, L. R. Halverson, and C. R. Graham, "Measuring student engagement in technology-mediated learning: A review," *Comput. & Educ.*, vol. 90, pp. 36–53, 2015, doi:10.1016/j.compedu.2015.09.005.

[25] A. W. Astin, "Student involvement: A developmental theory for higher education," *J. College Student Personnel*, vol. 25, no. 4, pp. 297–308, 1984.

[26] G. D. Kuh, "What student affairs professionals need to know about student engagement," *J. College Student Development*, vol. 50, no. 6, pp. 683–706, 2009.

[27] S. J. Mann, "Alternative perspectives on the student experience: Alienation and engagement," *Studies in higher Educ.*, vol. 26, no. 1, pp. 7–19, 2001, doi:10.1080/03075070020030689.

[28] E. A. Skinner, "Engagement and disaffection as central to processes of motivational resilience and development," *Handbook of motivation at school*, pp. 145–168, 2016.

[29] J. A. Fredricks, P. C. Blumenfeld, and A. H. Paris, "School engagement: Potential of the concept, state of the evidence," *Rev. Educ. Res.*, vol. 74, no. 1, pp. 59–109, 2004, doi:10.3102/00346543074001059.

[30] J. Erickson and K. Siau, "E-ducation," *CACM*, vol. 46(9), pp. 134–140, 2003, doi:10.1145/903893.903928.

[31] K. Kreijns, P. A. Kirschner, and W. Jochems, "Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: A review of the research," *Comput. Hum. Behav.*, vol. 19, no. 3, pp. 335–353, 2003.

[32] H. M. Horowitz, "Student response systems: Interactivity in a classroom environment," in *Proc. 6th Annu. Conf. on Interact. Instruction Delivery*, Salt Lake City, UT, Febr 1988.

[33] E. Blood and R. Neel, "Using student response systems in lecture-based instruction: Does it change student engagement and learning?" *J. Technol. and Teacher Educ.*, vol. 16, no. 3, pp. 375–383, 2008.

[34] J. R. Stowell and J. M. Nelson, "Benefits of electronic audience response systems on student participation, learning, and emotion," *Teaching of Psychology*, vol. 34, no. 4, pp. 253–258, 2007, doi:10.1080/00986280701700391.

[35] A. R. Trees and M. 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 Technol.*, vol. 32, no. 1, pp. 21–40, 2007, doi:doi.org/10.1080/17439880601141179.

[36] H. Du, M. B. Rosson, and J. M. Carroll, "Augmenting classroom participation through public digital backchannels," in *Proc. 17th ACM Int. Conf. on Supporting Group Work*, ser. GROUP '12. New York, NY, USA: ACM, 2012, pp. 155–164, doi:10.1145/2389176.2389201.

[37] G. Grotenbreg and S. B. J. Wong, "Using Pigeonhole® Live to elicit feedback, questions & reinforce learning during lectures," *CDLT Brief*, vol. 16, no. 2, pp. 2–7, 2013.

[38] P. Kumar, A. Gruzd, C. Haythornthwaite, S. Gilbert, M. Esteve del Valle, and D. Paulin, "Learning in the wild: Coding reddit for learning and practice," in *Proc. 51st Hawaii Int. Conf. Syst. Sci.*, 2018, doi:10.1080/17439884.2018.1498356.

[39] 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," *Comput. & Educ.*, vol. 94, pp. 252–275, 2016, doi:10.1016/j.compedu.2015.11.008.

[40] M. B. Kinzie, S. D. Whitaker, and M. J. Hofer, "Instructional uses of instant messaging (IM) during classroom lectures," *Educ. Technol. & Society*, vol. 8, no. 2, pp. 150–160, 2005.

[41] S. Yardi, "The role of the backchannel in collaborative learning environments," in *Proc. 7th Int. Conf. on Learning Sciences*, ser. ICLS '06. Int. Society of the Learning Sciences, 2006, pp. 852–858.

[42] A. W. Chickering and S. C. Ehrmann, "Implementing the seven principles: Technology as lever," *AAHE bulletin*, vol. 49, pp. 3–6, 1996.

[43] P.-S. D. Chen, A. D. Lambert, and K. R. Guidry, "Engaging online learners: The impact of web-based learning technology on college student engagement," *Comput. & Educ.*, vol. 54, no. 4, pp. 1222–1232, 2010, doi:10.1016/j.compedu.2009.11.008.

[44] A. Pardo, F. Han, and R. A. Ellis, "Combining university student self-regulated learning indicators and engagement with online learning events to predict academic performance," *IEEE Trans. Learn. Technol.*, vol. 10, no. 1, pp. 82–92, 2017, doi:10.1109/TLT.2016.2639508.

[45] S. B. Shum and R. Ferguson, "Social learning analytics," *J. of Educ. Technol. & society*, vol. 15, no. 3, pp. 3–26, 2012.

[46] P. A. Tess, "The role of social media in higher education classes (real and virtual) – a literature review," *Comput. Hum. Behav.*, vol. 29, no. 5, pp. A60 – A68, 2013, doi:10.1016/j.chb.2012.12.032.

[47] P. Rambe and A. Bere, "Using mobile instant messaging to leverage learner participation and transform pedagogy at a south african university of technology," *British J. Educ. Technol.*, vol. 44, no. 4, pp. 544–561, 2013.

[48] T. Van Daele, C. Frijns, and J. Lievens, "How do students and lecturers experience the interactive use of handheld technology in large enrolment courses?" *British J. Educ. Technol.*, vol. 48, no. 6, pp. 1318–1329, 2017, doi:10.1111/bjet.12500.

[49] W. M. Kappers and S. L. Cutler, "Polleverywhere! even in the classroom: An investigation into the impact of using polleverywhere in a large-lecture classroom," *The ASEE Comput. in Educ.(CoED) Journal*, vol. 6, no. 2, p. 21, 2015.

[50] H. Jackson, A. Nayyar, P. Denny, A. Luxton-Reilly, and E. Tempero, "Handsup–an in-class question posing tool," in *2018 Int. Conference on Learning and Teaching in Comput. and Engineering (LaTICE)*, April 2018, pp. 24–31, doi:10.1109/LaTICE.2018.00-11.

[51] N. Muthmainnah, "An effort to improve students' activeness at structure class using slido app," *J. of English Educ. Soc.*, vol. 4, no. 1, pp. 1–7, 2019, doi:10.21070/jees.v4i1.1868.

[52] J. Ha, "Using mobile-based slido for effective management of a university english reading class," *Multimedia-Assisted Language Learn.*, vol. 21, no. 4, pp. 37–56, 2018, doi:10.15702/mall.2018.21.4.37.

[53] D. R. George, T. D. Dreibelbis, and B. Aumiller, "How we used two social media tools to enhance aspects of active learning during lectures," *Med. Teacher*, vol. 35, no. 12, pp. 985–988, 2013, doi:10.3109/0142159X.2013.818631.

[54] H. Aagard, K. Bowen, and L. Olesova, "Hotseat: Opening the backchannel in large lectures," *Educause Quart.*, vol. 33, no. 3, p. 2, 2010.

[55] D. Baron, A. Bestbier, J. M. Case, and B. I. Collier-Reed, "Investigating the effects of a backchannel on university classroom interactions: A mixed-method case study," *Comput. & Educat.*, vol. 94, pp. 61–76, 2016, doi:10.1016/j.compedu.2015.11.007.

[56] S. Y. Harunasari and N. Halim, "Digital backchannel: Promoting students' engagement in EFL large class." *Int. J. Emerging Technol. in Learn.*, vol. 14, no. 7, 2019.

[57] V. Trowler, "Student engagement literature review," *Higher Educ. Acad.*, vol. 11, no. 1, pp. 1–15, 2010.

[58] S. Sinha, T. K. Rogat, K. R. Adams-Wiggins, and C. E. Hmelo-Silver, "Collaborative group engagement in a computer-supported inquiry learning environment," *Int. J. Comp.-Supported Collaborative Learn.*, vol. 10, no. 3, pp. 273–307, 2015, doi:10.1007/s11412-015-9218-y.

[59] S. Järvelä and A. F. Hadwin, "New frontiers: Regulating learning in cscl," *Educ. Psychol.*, vol. 48, no. 1, pp. 25–39, 2013, doi:10.1080/00461520.2012.748006.

[60] 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 *European Conf. on Technol. Enhanced Learn.* Springer, 2016, pp. 247–262, doi: 10.1007/978-3-319-45153-4_19.

[61] R. B. Johnson and A. J. Onwuegbuzie, "Mixed methods research: A research paradigm whose time has come," *Educ. Researcher*, vol. 33, no. 7, pp. 14–26, 2004.

[62] J. C. Greene, L. Benjamin, and L. Goodyear, "The merits of mixing methods in evaluation," *Evaluation*, vol. 7, no. 1, pp. 25–44, 2001.

[63] J. W. Creswell, V. L. Plano Clark, M. L. Gutmann, and W. E. Hanson, "Advanced mixed methods research designs," in *Handbook of mixed methods in social and behavioral research*. Sage Publications, 2003, pp. 209–240.

[64] E. G. Guba, "Criteria for assessing the trustworthiness of naturalistic inquiries," *J. of Theory, Research, and Development on Educ. Communic. and Technol.*, vol. 29, no. 2, pp. 75–91, 1981.

[65] A. Holzer, I. V. Cardia, S. Bendahan, A. Berne, L. Bragazza, A. Danalet, A. Fasoli, J. N. Feige, D. Gillet, S. Isaac *et al.*, "Increasing the perspectives of engineering undergraduates on societal issues through an interdisciplinary program," *Int. J. Eng. Educ., 2016*, vol. 32, no. 2, pp. 614–624, 2016.

[66] A. Holzer, S. Govaerts, A. Vozniuk, B. Kocher, and D. Gillet, "SpeakUp in the classroom: Anonymous temporary social media for better interactions," in *Proc. of the Extended Abstracts of the 32Nd Annu. ACM Conf. on Hum. Factors in Comput. Syst.*, ser. CHI EA '14. New York, NY, USA: ACM, 2014, pp. 1171–1176, doi:10.1145/2559206.2581211.

[67] J. Mittelmeier, B. Rienties, D. Tempelaar, G. Hillaire, and D. Whitelock, "The influence of internationalised versus local content on online intercultural collaboration in groups: A randomised control trial study in a statistics course," *Comput. & Educ.*, vol. 118, pp. 82–95, 2018, doi:10.1016/j.compedu.2017.11.003.

[68] D. T. Tempelaar, B. Rienties, and Q. Nguyen, "Towards actionable learning analytics using dispositions," *IEEE Trans. Learn. Technol.*, vol. 10, no. 1, pp. 6–16, 2017, doi10.1109/TLT.2017.2662679.

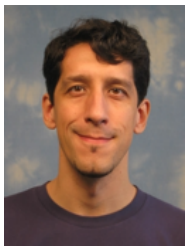
[69] X. Ochoa, "Multimodal learning analytics," *Handbook Learn. Analytics*, pp. 129–141, 2017, doi:10.18608/hla17.011.

[70] L. P. Prieto, K. Sharma, Ł. Kidzinski, M. J. Rodríguez-Triana, and P. Dillenbourg, "Multimodal teaching analytics: Automated extraction of orchestration graphs from wearable sensor data," *J. Comput. Assisted Learn.*, vol. 34, no. 2, pp. 193–203, 2018, doi:10.1111/jcal.12232.

[71] C. Beer, K. Clark, and D. Jones, "Indicators of engagement," *Curriculum, Technol. & transformation for an unknown future. Proceedings ASCILITE, Sydney*, pp. 75–86, 2010.

[72] T. Bergstrom, A. Harris, and K. Karahalios, "Encouraging initiative in the classroom with anonymous feedback," in *IFIP Conference on Human-Computer Interaction*. Springer, 2011, pp. 627–642.

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 focus on learning analytics to support classroom orchestration, teacher inquiry and institutional decision making.



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 École Polytechnique Fédérale de Lausanne, 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 his Ph.D. degree in Information Systems from the École Polytechnique Fédérale de Lausanne in 1995. Currently, Faculty Member at the EPFL School of Engineering and affiliated with the Center for Learning Sciences, head of the REACT multi-disciplinary research group, and co-founder of the Swiss EdTech Collider. He has being the technical coordinator of large-scale European innovation actions for STEM education at schools, and associate editor of the IEEE Transactions on Learning Technologies and of the International Journal of Technology Enhanced Learning. His current research interests include digital education, human computer interaction, humanitarian technology and ICT for development.