Enhancing Learning Engagement in the Flipped Classroom using a Video-Annotation Tool

Giacomo Cassano, Nicoletta Di Blas and Alessia Mataresi
Departments of Electronics, Information and Bioengineering, Politecnico di Milano, Italy

giacomo.cassano@polimi.it (corresponding author)
nicoletta.diblas@polimi.it

https://doi.org/10.34190/ejel.22.9.3259

Abstract: In the realm of academia, there has been a burgeoning interest in examining student engagement within the context of the flipped classroom approach, particularly at the higher education level. Nevertheless, research pertaining to a pivotal facet of this pedagogical methodology—namely, student engagement during the individualized, at-home preparatory phase—remains limited. Within this void, the primary objective of this study is to delve into the prospective benefits offered by a video annotation tool in augmenting students' level of engagement during this autonomous study phase. Throughout the course of this investigation, a bespoke questionnaire was crafted, grounded in a multifaceted framework for evaluating student engagement, with a specific emphasis on dissecting four distinct dimensions: behavioral, emotional, agentive, and cognitive. This questionnaire is geared towards gauging student engagement within the domain of the Flipped Classroom model that relies on video-based instructional content—a facet heretofore poorly explored within the scholarly literature. Both quantitative and qualitative data were collected by administering the aforementioned questionnaire to a cohort of 68 undergraduate engineering students at Politecnico di Milano (Italy), all within the authentic context of a case study. The outcomes stemming from this empirical inquiry showcase a noteworthy enhancement in student engagement, with particular prominence accorded to the realms of emotional and agentive engagement. Moreover, this study establishes that the interactivity and proactive involvement necessitated by the video annotation tool do not obstruct students' behavioral and cognitive engagement levels. In summation, this research endeavors to illuminate the potentiality of bridging a pivotal juncture in the Flipped Classroom paradigm, specifically the phase characterized by independent at-home study. This bridge is facilitated by the utilization of a video annotation tool designed to heighten student engagement. This transformative approach effectively transmutes the traditionally passive at-home study phase of the Flipped Classroom into an active experience, thereby enhancing the overall efficacy of this pedagogical approach.

Keywords: Flipped classroom, Learning engagement, Student engagement, Video annotation tool, Video-based learning

1. Introduction

The adoption of the Flipped Classroom (FC) approach in higher education has gained momentum as a response to the need for student-centered learning environments (Lai et al., 2021). This approach reverses the traditional teaching sequence, requiring students to independently engage with course content at home, allowing in-class time for activities based on this content. The effectiveness of the in-person session relies heavily on the preparatory phase, where students access materials provided by the teacher. This "at-home" study phase is a crucial component of the Flipped Classroom approach, enabling active and personalized learning and promoting self-regulation skills for class participation, as demonstrated by (Yilmaz and Baydas, 2017). Among the challenges highlighted in the literature, some pertain to the phase of independent study. It is noted not only that students should be generally prepared for the approach (Al-Zahrani, 2015), but also that their motivation to watch pre-recorded video lectures or study course materials outside of class should be stimulated (Zainuddin, 2017). This can be facilitated by providing guidelines on how to prepare for the in-person session (Wanner and Palmer, 2015). Moreover, students do not have the opportunity to ask for help, nor can instructors intervene, as they are unable to monitor the students' progress. Active student engagement during the preparatory phase is thus crucial for the success of the Flipped Classroom (Rahman et al., 2015). The review by (Akçayır and Akçayır, 2018) found that the majority of the challenges of the Flipped Classroom are related to out-of-class activities, with inadequate student preparation prior to class being frequently reported.

Suggestions to improve this phase include providing students with clear instructions on their tasks, equipping them with diverse materials, encouraging note-taking, implementing gamification and using interactive tools (Baig and Yadegaridehkordi, 2023; Divjak et al., 2022). One possible solution is the use of Video Annotation Systems, which foster active learning and have the 'potential to shift the passive experience of watching an instructional video to a more active one' (Mirriahi et al., 2021).

Cite this article: Cassano, G., Di Blas, N. and Mataresi, A. 2024. "Enhancing Learning Engagement in the Flipped Classroom using a Video-Annotation Tool", *Electronic Journal of e-Learning*, 22(9), pp 75-90, https://doi.org/10.34190/ejel.22.9.3259

This study examines how a Video Annotation Tool (VAT) affects student engagement during the home phase of the Flipped Classroom. The tool enables students to express their understanding, ask questions, and share reactions while accessing teacher-provided learning content. The hypothesis is that these tools foster increased engagement by giving students a voice even during this typically independent phase.

Therefore, the research question is: "*Can a Video/Media Annotation tool increase students' learning engagement during the home session of the Flipped Classroom?*"

To answer this question, we analyzed data from a case study conducted in a real-life setting with 378 undergraduate students from Politecnico di Milano, the largest technical university in Italy. Over the course of a semester, the students were provided with the VAT Evoli that allowed them to express not only their level of understanding or confusion about the various topics of the videos and documents their instructors shared with them, but also to ask questions, express doubts, or show interest. The contribution of this study consists in exploring how a VAT can increase engagement during the preparation phase of the FC, investigating which specific types of engagement are affected. The study is based on the hypothesis that making students active through a VAT increases overall engagement, while it remains open as to which specific types of engagement are stimulated.

2. Literature Review

This section introduces the Flipped Classroom model and learning engagement as the two key pillars of the study. It also explores the intersection between these two pillars by reviewing studies that have investigated ways to increase learning engagement during the at-home preparation phase of the Flipped Classroom.

2.1 The Flipped Classroom

In this first section of the state of the art, we present the approach of the flipped classroom, with a particular focus on its component phases. The Flipped Classroom (FC) is an educational approach that reverses the traditional teaching model, with students accessing educational materials before class and using class time for practical activities, discussions and interaction. The teacher facilitates learning, provides support, responds to questions, and fosters participation (Lage, Platt and Teglia, 2000). The FC model can be divided into two phases:

- Independent study phase. In this phase, students are responsible for reviewing the instructional materials on their own, typically accessing online resources such as videos. This phase is focused on individual learning and allows students to work at their own pace and review the materials as many times as needed to understand the concepts.
- In-class learning phase. In this phase, students come to class prepared with a basic understanding of the concepts covered in the instructional materials, and the class time is devoted to interactive activities, case studies application and discussions that reinforce the concepts.

The FC approach has the potential to offer several benefits, including promoting active learning, providing a flexible learning environment, student motivation, and student-teacher interaction (Bryson and Hand, 2007; Halasa et al., 2020; Lage, Platt and Teglia, 2000; Strelan, Osborn and Palmer, 2020). It also poses significant challenges that need to be addressed. These challenges include the need for access to digital technology, which may disadvantage some students, and the importance of self-directed learning, which may be difficult for students who lack support or struggle with time management (Etemi et al., 2024). However, the greatest challenge is the lack of engagement during the preparation phase at home, which is critical for the success of in-class activities (Costley and Lange, 2017). Recent research corroborates the existence of pedagogical challenges experienced by students within this phase. These challenges encompass the absence of explicit guidelines, the constraint on seeking clarifications or aid, and the instructor's limited capacity to oversee students' interaction with educational resources. These inquiries propose that students' self-directed learning preceding classroom sessions can tend to be excessively solitary in nature.

To address this challenge, some studies have focused on the most effective technologies to support independent student preparation, like podcasts, screencasts, annotated notes and captured videos, as well as the use of pre-readings, automated tutoring systems, study guides, and interactive videos. However, these technologies still leave students isolated and compartmentalize the preparation phase from the classroom interaction. To address these issues, it is essential to bridge the gap between these two phases, to increase student engagement and project them towards the in-class session. In the following section, we will examine how the concept of 'learning engagement' has been conceptualized in the literature. Then, in the third and final section of the state of the

art, we will move on to considerations regarding the intersection between the flipped classroom approach and this construct.

2.2 Learning Engagement

Learning engagement refers to the degree of active and meaningful participation that students exhibit in the learning process (Marks, 2000). While there is no single, universally accepted definition of learning engagement in literature, it is widely recognized that it comprises multiple dimensions and components, the number of which varies depending on the perspectives of the authors, with some studies identifying two dimensions and others identifying up to four (Symonds et al., 2021; Fredricks et al., 2004). For instance, Audas & Willms (2001) focused on only two dimensions, namely affective/emotional and cognitive engagement, examining their correlation with school dropout rates. In contrast, Fredricks, Blumenfeld and Paris (2004) introduced a third dimension, "behavioral" engagement. More recently, a fourth dimension, agency, has been added to the construct of learning engagement (Reeve and Tseng, 2011).

The four-dimensional construct, which has been validated multiple times in the literature, has been chosen as the most comprehensive for the purpose of this study (Bergdahl et al., 2020; Wang et al., 2019). The four dimensions are:

1. Behavioral engagement. It refers to the observable actions and behaviors that indicate a student's active participation in learning activities, including attending class, completing assignments, and participating in discussions. Behavioral engagement is closely linked to academic achievement, as it reflects a student's willingness to invest time and effort in their learning (Fredricks, Blumenfeld and Paris, 2004).
2. Emotional engagement. It encompasses students' interest, values, and emotions related to academic activities. Emotional engagement is critical for promoting intrinsic motivation and a sense of personal relevance and value in learning. Students who are emotionally engaged are more likely to persist in their learning and to experience a sense of fulfilment in their academic pursuits (Fredricks, Blumenfeld and Paris, 2004).
3. Cognitive engagement. It relates to the mental effort, strategies, and investment that students bring to their academic tasks. It involves a deep level of thinking and a willingness to engage in challenging and complex tasks, to seek out feedback and support, and to persist in the face of obstacles and setbacks (Fredricks, Blumenfeld and Paris, 2004).
4. Agentic engagement. This fourth dimension focuses on the constructive contribution made by students in their learning process, emphasizing their ability to personalize and proactively enrich the instructional content and learning environment. Agentic engagement involves a sense of ownership and control over one's learning, and a willingness to take initiative for one's academic success (Reeve and Tseng, 2011).

Measuring student engagement presents a challenge due to the pronounced intercorrelation among its various dimensions. Distinct evaluation scales are available for overall engagement and each individual dimension. However, specific components might be attributed to different dimensions depending on how engagement facets are operationalized across various studies. This complexity hinders the ability to draw comparisons and arrive at definitive conclusions about what promotes each type of engagement. However, evidence from various studies suggests a positive relationship between engagement and academic performance (Fredricks, Blumenfeld and Paris, 2004).

2.3 Learning Engagement in the Flipped Classroom

The FC has been widely recognized as a promising approach to enhance teacher-student interaction and potentially increase student engagement (Bergmann and Sams, 2012). However, the effectiveness of this approach in achieving these goals is a controversial topic in literature. Reviews indicate that "the search for evidence of the effectiveness of and improvements engendered by the flipped or inverted classroom approach is becoming frequent", but "knowledge contributions [...] are relatively siloed and fragmented and have yet to stabilize. Academically and socially, the research is quite scattered, and only local evidence and experiences are available. The knowledge contributions within this field of interest seem to be anecdotal rather than systematically researched" (Lundin et al., 2018, p. 2). Some studies seem to conclude that the FC is capable of yielding better learning outcomes and increased learner engagement compared to traditional face-to-face lectures (Elmaadaway, 2018; Kazanidis et al., 2019; Lo and Hew, 2020), while others find substantial equivalence (Subramaniam and Muniandy, 2019). Evidence regarding the effectiveness of the FC in enhancing student

engagement remains limited and sometimes conflicting (Gilboy, Heinerichs and Pazzaglia, 2015). The main reason for this is the absence of a standardized approach to flipping the classroom (Bishop and Verleger, 2013). In practice, since there is no specific methodology or checklist to follow, each teacher applies the FC in a personalized manner and there is significant variation in the way instructors implement it (DeLozier and Rhodes, 2017, p. 142). For instance, instructors may opt for a partial flip, in which only some lectures incorporate a flipped approach (Seery, 2015). Furthermore, they may include a wide range of out-of-class activities beyond lectures, such as readings, homework, and supplemental videos. In-class activities also exhibit considerable diversity, including role-play, debates, quizzes or group presentations (O'Flaherty and Phillips, 2015).

Recent studies address learning engagement in the FC in specific learning environments, such as the fully online FC (Jia et al., 2023) and the FC augmented by Artificial Intelligence (Huang, Lu, Yang, 2023). The former emphasizes the need for instructional design guidelines to make the preparation phase engaging, while the latter highlights the potential of AI-enabled personalized video recommendations to improve learning engagement.

While the effectiveness of the FC in increasing student learning engagement remains inconclusive due to methodological variety, promising findings have been reported in vertical and context-specific studies. Recent attention has been given to the critical first phase of the FC, which our study focuses on as a less explored but highly relevant aspect (Burke and Fedorek, 2017).

3. Method

In order to address our research question, which is whether a VAT can enhance learning engagement in the preparatory phase of the FC, we analyzed a case study in a real life setting in which the students of two undergraduate courses at Politecnico di Milano (Italy) used the Evoli VAT (Cassano and Di Blas, 2024) to support the preparatory phase of a FC. Data was collected through a voluntary and anonymous online questionnaire at the end of the semester, which included closed (on a Likert scale) and open-ended questions. Two voluntary semi-structured interviews were also conducted with students. The numerical data was analyzed using statistical methods, while the qualitative data was analyzed through independent coding by two researchers working on the open-ended questions and the interviews' transcripts. The study covered an entire semester, from February to May 2023.

3.1 The Video/Media Annotation tool

The Video/Media Annotation tool "Evoli", developed by our laboratory, allows instructors to share various types of learning content and gather students' questions and feedback on their level of understanding. The instructor uploads media files, segmented into chapters, and provides students with a unique access code. Students can access the tool anonymously or with their name and surname and, at the end of each chapter, they rate their comprehension and provide a justification. Students can also ask questions and express doubts (Figure 1).

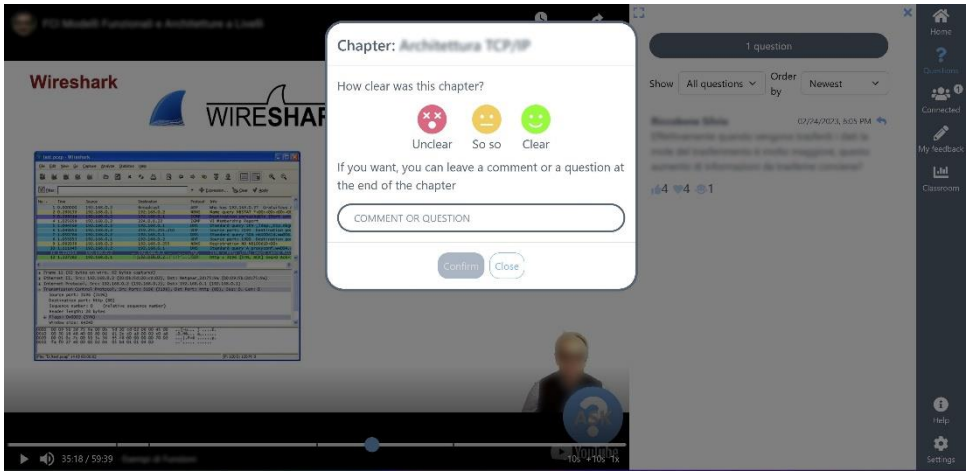

Figure 1: The student interface in Evoli. A modal window appears at the end of each chapter asking the student to rate her level of understanding

Evoli also allows students to see their classmates' questions/comments and shows an overview of their collective level of understanding, promoting a sense of community (Figure 2).

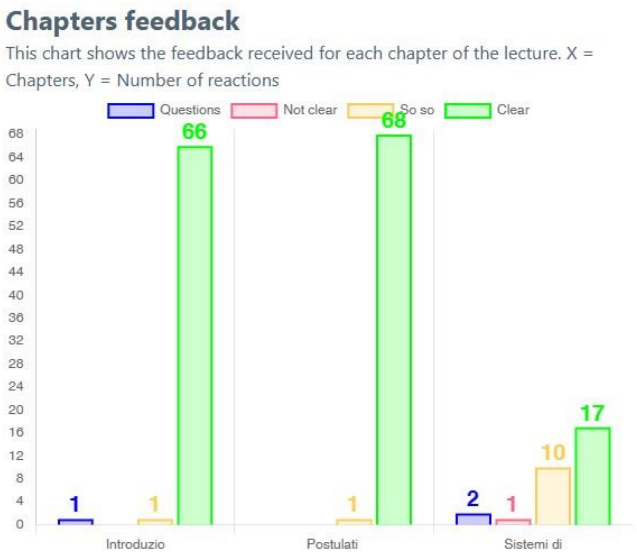

Figure 2: Students can check the understanding of a chapter by their classmates (columns 2, 3 and 4) as well as how many questions they have asked (column 1 - they can also see the questions, in a parallel section)

3.2 Participants and Settings

Two professors and their engineering students at Politecnico di Milano participated in the study. The courses' instructional design involved traditional lectures complemented with various FC sessions. Students prepared before class, and in-class time was used for deeper exploration, practical application, and guided discussion.

The first course was "Fundamentals of Communications and Internet," for undergraduate students in computer science engineering and telecommunications, with 228 students. Four FC sessions were conducted using Evoli. Table 1 presents the various learning units along with their durations and number of chapters.

Table 1: Learning units for "Fundamentals of Communication and Internet" course

Learning unit	Type	Duration	Chapters
Layered architectures	Video	59:40	6
Routing (introduction and algorithms)	Video	1:14:22	2
Routing protocols SDN	Video	1:30:05	2
VPN and IPv6	Video	1:05:23	2

The second course was "Nuclear Physics," attended by students from diverse engineering disciplines. Seven FC sessions were planned, with the aid of the Evoli platform. Table 2 shows the learning units with their durations and number of chapters.

Table 2: Learning units for the "Nuclear Physics" course

Learning unit	Type	Duration	Chapters
Angular momentum theory – part 1	Video	1:35:28	7
Angular momentum theory – part 2	Video	1:21:25	6
Special relativity – part 1	Video	1:48:46	3
Special relativity – part 2	Video	1:34:26	7
Preface nuclear physics texts	Document	20 pages	4
More on the Klein-Gordon equation. Dirac equation	Video	2:01:51	9
Introduction to meson physics	Video	1:12:28	6

3.3 Data collection

At the end of the semester, a questionnaire was administered, requiring participants to provide numerical evaluations indicating their level of agreement with specific statements. In addition, a series of seven open-ended questions was included to capture a more nuanced understanding of the participants' opinions. Furthermore, two semi-structured interviews were conducted, involving one female and one male student. The two courses involved in the study had a total of 378 students, of which 68 (18% approximately) responded to the quantitative part of the questionnaire, and 47 (12% of the total number) also provided answers to the open-ended questions. On the ethical front, it should be noted that the participants took part in the study on a voluntary and completely anonymous basis, without sharing any personal data. The design of our questionnaire was rooted in the questionnaire developed by Subramaniam & Muniandy (2019) for gauging learning engagement. However, some adaptation was necessary as the original questionnaire evaluated the FC approach as a whole, while our study focused specifically on the first phase – the independent study at home. In its turn, the Subramaniam & Muniandy (2019) questionnaire was developed from two earlier instruments, which did not specifically address the FC approach but were concerned with learning engagement in general (Marks, 2000). The Subramaniam & Muniandy (2019) questionnaire includes the concept of students' agency as a fourth dimension of engagement, introduced by Reeve & Tseng (2011).

The questionnaire comprised 11 paired questions designed to elicit participants' assessments of their perceived engagement within two distinct contexts. These contexts were defined as follows: the first context mirrored a 'traditional' educational setting, devoid of technological aids, a setting previously experienced by the students as a result of the university's widespread adoption of the FC methodology across multiple courses. The second context encapsulated the students' experiences during the semester in which Evoli, the Video Annotation tool, was employed This paired design allowed for a direct comparison between the two instructional approaches. Responses to the questions were recorded on a 5-item Likert scale, where 1 indicated "strongly disagree" and 5 "strongly agree". The first three statements were designed to assess behavioral engagement, with a focus on evaluating whether there had been an increase in attention and care for the learning content (statements B1 and B2) or an increase in commitment to the task (statement B3). The next four statements evaluate agentic engagement, aiming to understand if students take actions whenever they encounter difficulties in their studies, such as checking their peers' doubts and questions and contacting them (A1), communicating with the instructor (A2 and A3) or looking for help (A4). The following two statements address emotional engagement by assessing whether students perceive an improvement in their sense of involvement (E1) and in their preparedness for the in-class session of the FC (E2). The last two statements investigate cognitive engagement, specifically examining whether students exert greater effort during the session with Evoli (C1) and whether they can self-regulate their studying more effectively (C2). The questions, along with their scores, are listed in section 4.1.

The questionnaire also included seven open-ended questions designed to gather opinions and details that could shed light on the quantitative data. These questions were based on the themes of the 11 paired questions but approached from a different angle.

- Do you think that the use of Evoli improved your study at home?
- Do you think that the use of Evoli improved the in-class sessions within the course?
- Did you perceive Evoli as useful?
- Did the prompts by Evoli to self-assess your level of understanding help you become more aware of it?
- Was it helpful to see the understanding status and the questions by your classmates?
- In using Evoli, did you feel too much "exposed" to the teacher?
- Did you find using Evoli annoying, with respect to just watching the video / reading a document by yourself?

Finally, the semi-structured interview aimed to investigate the overall experience and the four components of learning engagement.

3.4 Data Analysis

The quantitative data was initially subjected to a descriptive analysis to visualize the mean values of the samples. Subsequently, the appropriate statistical test was applied depending on the normality of the samples. Specifically, the paired t-test was employed when both samples exhibited normal distribution, whereas the non-parametric Wilcoxon signed-rank test was utilized when the normality assumption was violated. To verify the normality assumption, the Shapiro-Wilk test was conducted. For the qualitative data, which included open-

ended responses and two interviews, a thematic analysis was performed. Two independent reviewers tagged all the open-ended responses into six categories: the four types of engagement, comments on the tool, and off-topic comments. They then compared their tagging results to align and group the responses. However, numerical treatments were not performed on these qualitative data due to the insufficient number of respondents. Open-ended responses and interviews were utilized to provide a concrete glimpse of the experiences and perceptions encountered by the students during the activity.

4. Results

To understand whether the cardinality of the sample of students who participated in the study by answering the questionnaire was adequate or not, we conducted a power analysis using G*Power software (version 3.1.9.7). A power analysis for a one-tailed paired-samples t-test indicated that the minimum sample size to yield a statistical power of at least 0.95 with an alpha of 0.05 and a medium effect size (d = 0.5) is 45. Since to perform t-tests the samples must obtain the assumption of normality, we repeated the power analysis for the Wilcoxon signed-rank test (matched pairs) reaching the same results. The students involved in the questionnaire in our case were 68, allowing us to reach a statistical power of 0.9925773.

4.1 Quantitative Data

To evaluate students' engagement, all the responses on a Likert scale from 1 to 5 of all 11 paired questions (FC in normal situation vs. FC with Evoli) of the questionnaire are considered. In particular, for each participant we calculated the mean value of the 11 questions about the engagement in the FC with Evoli and the engagement in the "Traditional FC". The groups are considered to be paired, since the same participants answered both the questions of the FC without and with Evoli. Table 3 shows the mean scores and the questions number for each learning engagement dimensions. Instead, a detail of the distribution plots and box plots can be seen in Figure 3 and Figure 4.

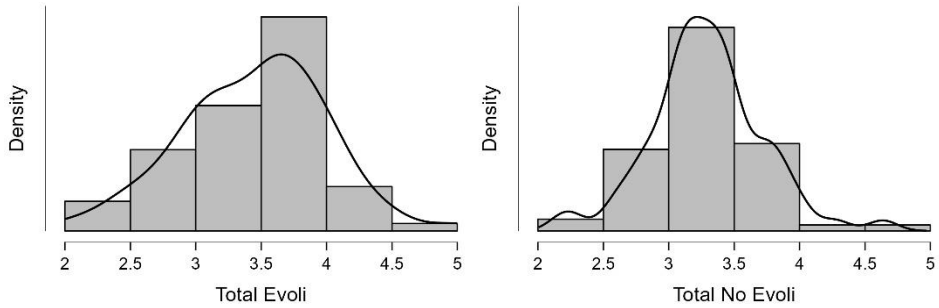

Figure 3: Distribution plots for the overall student engagement with the VAT tool (left) and without (right)

Table 3: Mean score for each dimension (and overall) considering the questions about the use of Evoli and the paired one about the traditional mode (without Evoli)

Dimension	Mean score with Evoli	Mean score without Evoli	Questions number
Behavioral	3.686	3.721	3
Agentic	3.165	2.849	4
Emotional	3.449	3.162	2
Cognitive	3.654	3.647	2
Student engagement	3.448	3.289	11

The descriptive statistics showed that the mean score was higher for the students' engagement for the FC with Evoli (M = 3.448, SD = 0.558) than the "Traditional FC" (M = 3.289, SD = 0.431). In order to generalize the result by a paired t-test, it is necessary that the two paired samples follow a normal distribution. To demonstrate this, we applied the Shapiro-Wilk test to ascertain normality.

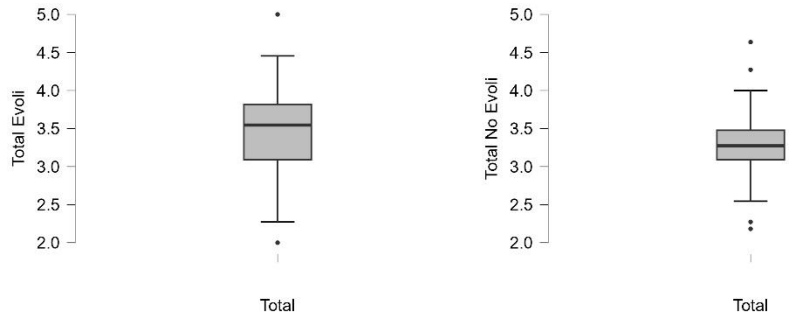

Figure 4: Box plots for the overall student engagement

First, we applied the Shapiro-Wilk test on the FC with Evoli sample with cardinality 68. The test yielded a Shapiro-Wilk statistic of 0.989 and a *p*-value of 0.828. These results indicate that the data in our sample exhibit a minor deviation from the normal distribution. The Shapiro-Wilk statistic, W, being close to 1, suggests that the data are relatively close to a normal distribution. Furthermore, the *p*-value of 0.828, which is well above the commonly used significance level of 0.05, provides strong evidence in support of the null hypothesis that the data follow a normal distribution. Based on these findings, we can conclude that the data in our sample can be considered approximately normally distributed. A similar result was reached applying the Shapiro-Wilk test on the paired "Traditional FC" sample. In this case we reach a value W = 0.981, which suggests that the data are relatively close to a normal distribution and the *p*-value = 0.378 indicates strong evidence in support of the null hypothesis. Thus, it is possible to conclude that both samples follow a normal distribution.

As the assumption of normality was met, a paired samples t-test was conducted to compare the means of the overall students' engagement FC with Evoli sample with respect to the "Traditional FC" paired sample. The null hypothesis of the paired t-test was that the means are not different, while the alternative hypothesis specifies that the mean of the students' engagement in the FC with Evoli sample is greater than the mean of the paired students' engagement in the "Traditional FC" sample. The results show a significant difference between the means of the two samples, t = 3.000 and p = 0.002. Therefore, the overall student engagement was higher in FC with Evoli with respect to the "Traditional FC".

4.1.1 Behavioral engagement

We repeated the analysis restricting only to the questions relative to the behavioral engagement. The descriptive statistics showed that the mean score was slightly higher for the behavioral engagement for the "Traditional FC" (M = 3.721, SD = 0.539) than the "FC with Evoli" (M = 3.686, SD = 0.612). Table 4 shows the related questionnaire items, while a detail of the distribution plots and box plots can be seen in Figure 5 and Figure 6.

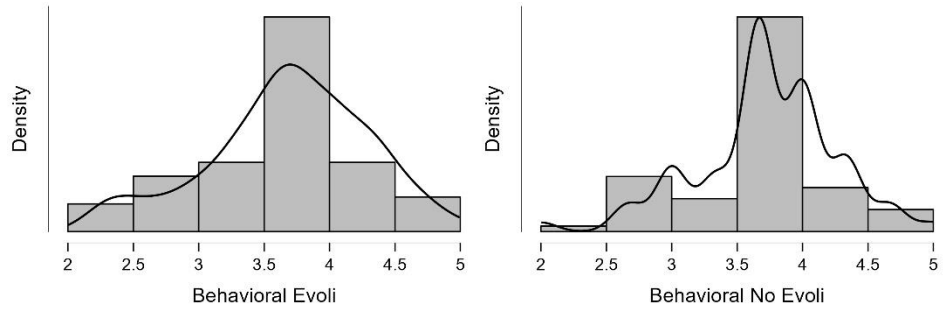

Figure 5: Distribution plots for the behavioral engagement with and without Evoli considering questions from B1 to B3

Table 4: Behavioral engagement questionnaire items with mean score and standard deviation

Item	Statement	Mean score	Std. Dev.
B1a	During the study at home of a traditional FC, I watch the video or read the documents given to me by the teacher very carefully	3.838	0.683
B1b	During the study at home of a FC with Evoli, I watch the video or read the documents given to me by the teacher very carefully	3.824	0.645

Item	Statement	Mean score	Std. Dev.
B2a	During the study at home of a traditional FC, I am able to maintain a constant level of attention	3.324	0.854
B2b	During the study at home of a FC with Evoli, I am able to maintain a constant level of attention	3.279	0.944
B3a	During the study at home of a traditional FC, If I don't understand what I read or see in a video, I go back and read or watch that part of the video all over again	4.000	0.810
B3b	During the study at home of a FC with Evoli, If I don't understand what I read or see in a video, I go back and read or watch that part of the video all over again	3.956	0.888

The normality assumption of the sample was assessed using the Shapiro-Wilk test, yielding a Shapiro-Wilk statistic (W) of 0.953 and a *p*-value of 0.013. These results suggest a moderate departure from normality and provide evidence to reject the null hypothesis of normality. Therefore, the data may not strictly follow a normal distribution. The same result is obtained for the behavioral engagement of the "Traditional FC" data points, since the Shapiro Wilk statistic (W = 0.946) suggests a moderate deviation from a normal distribution with an associated *p*-value (p = 0.005) that is significantly lower than the commonly used significance level of 0.05.

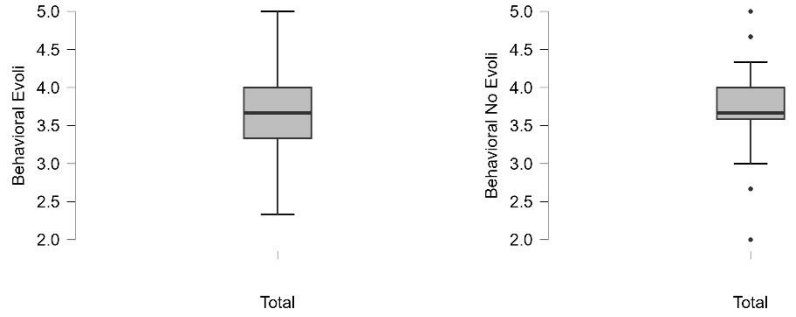

Figure 6: Box plots for the behavioral engagement items

Since the assumption of normality was not met, a paired Wilcoxon signed-rank test is used to compare the mean scores of the behavioral engagement in the FC with Evoli with respect to the paired "Traditional FC". Unlike the overall students' engagement result, the analysis revealed a non-significant difference, indeed despite a slight decrease in behavioral engagement observed with the use of Evoli, the analysis revealed a p-value of 0.556 (W = 411.500, z=0.595), which exceeds the chosen significance level of 0.05. Therefore, the results do not reach statistical significance.

4.1.2 Agentic engagement

Expanding the analysis for the agentic engagement questions (Table 5), it is possible to notice that the descriptive statistics shows that the mean score was higher for the agentic engagement for the "FC with Evoli" (M = 3.165, SD = 0.801) than the "Traditional FC" (M = 2.849, SD = 0.526), while a detail of the distribution plots and box plots can be seen in Figure 7 and Figure 8.

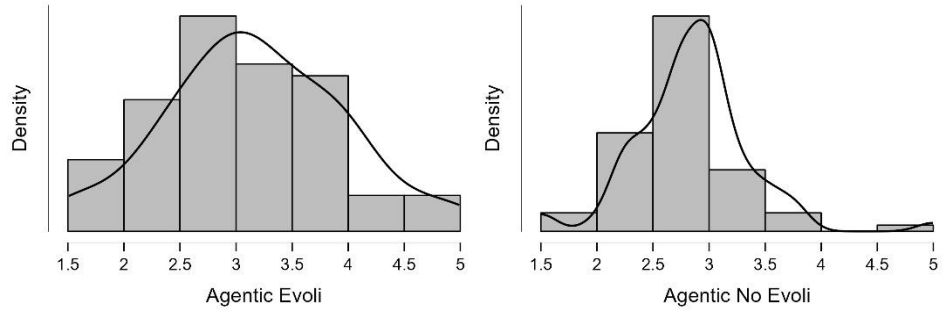

Figure 7: Distribution plots for the agentic engagement items

Table 5: Agentic engagement questionnaire items with mean score and standard deviation

Item	Statement	Mean score	Std. Dev.
A1a	During the study at home of a traditional FC, I check the doubts of my classmates (e.g. by calling them).	2.926	1.124
A1b	During the study at home of a FC with Evoli, I check the doubts of my classmates.	3.382	1.172
A2a	During the study at home of a traditional FC, If something is not clear, I let my teacher know (e.g. via email)	2.471	0.954
A2b	During the study at home of a FC with Evoli, if something is not clear I let my teacher know (e.g. by clicking on the sad emoticon or typing a question)	3.426	1.012
A3a	During the study at home of a traditional FC, I let my teacher know what I am interested in (e.g. via email)	1.912	0.859
A3b	During the study at home of a FC with Evoli, I let my teacher know what I am interested in (e.g. by typing a comment)	2.412	1.096
A4a	During the study at home of a traditional FC, If I don't understand something I look for help (e.g. asking someone or searching the Internet)	4.088	0.787
A4b	During the study at home of a FC with Evoli, If I don't understand something I look for help (e.g. typing a question)	3.441	0.937

This time, the Shapiro-Wilk test showed different results on the two different paired sample. Considering the agentic engagement in the "FC with Evoli" sample we obtained a Shapiro Wilk statistic (W = 0.981) that suggests a minor deviation from the normal distribution with an associated *p*-value (p = 0.376), significantly above the significance threshold of 0.05. Thus, there is no significant evidence to reject the normality assumption. Despite the first group followed an acceptable normal distribution, the second sample of the agentic engagement in the "Traditional FC" obtained a Shapiro-Wilk statistics (W = 0.918) that suggests a significant departure from a normal distribution. The associated *p*-value (p < 0.001) provides strong evidence to reject the null hypothesis of normality. Therefore, based on the Shapiro-Wilk test, the data cannot be considered to follow a normal distribution.

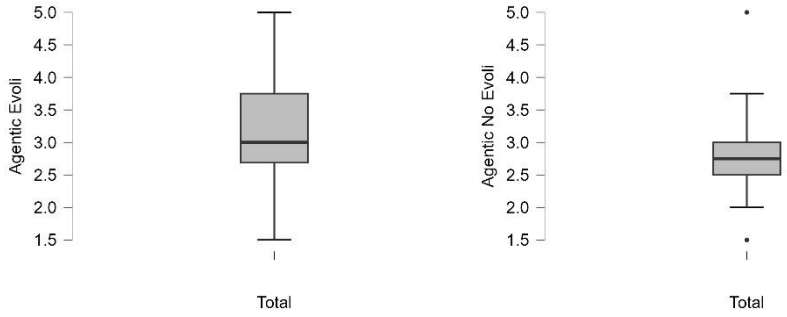

Figure 8: Box plots for the agentic engagement with and without Evoli

To compare the mean values we performed a paired Wilcoxon signed-rank test since the data didn't follow a normal distribution and so the t-test would have been invalid. The analysis revealed a significant difference (W = 1136.500, z = 3.392, p = < 0.001), providing strong evidence to support the alternative hypothesis that the mean score of the agentic engagement in the "FC with Evoli" is higher than the mean score of the agentic engagement in the "Traditional FC".

4.1.3 Emotional engagement

As for the emotional dimension, the process was similar to the agentic engagement analysis. Descriptive statistics showed higher mean scores for emotional engagement in "FC with Evoli" (M = 3.449, SD = 0.739) compared to "Traditional FC" (M = 3.162, SD = 0.720), while a detail of the distribution plots and box plots can be seen in Figure 9 and 10.

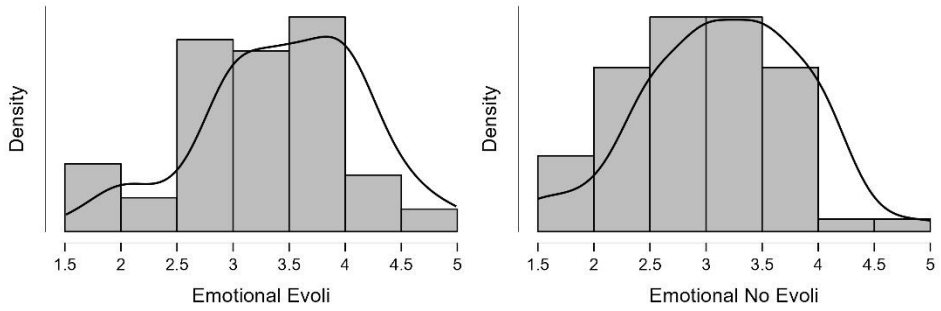

Figure 9: Distribution plots for the emotional engagement with and without Evoli considering questions E1 and E2

Table 6: Emotional engagement questionnaire items with mean score and standard deviation

Item	Statement	Mean score	Std. Dev.
E1a	During the study at home of a traditional FC, I realize that this mode makes me feel involved	2.838	0.908
E1b	During the study at home of a FC with Evoli, I realize that this mode makes me feel involved	3.353	0.842
E2a	During the study at home of a traditional FC, I feel I'm preparing well for the in-person lesson	3.485	0.801
E2b	During the study at home of a FC with Evoli, I feel I'm preparing well for the in-person lesson	3.544	0.836

The Shapiro-Wilk test showed that the samples relative to the two conditions did not follow a normal distribution. Indeed, for the sample about the emotional engagement in the "FC with Evoli" we obtain a Shapiro-Wilk statistic (W = 0.941) that suggests a significant departure from the acceptable normal distribution. The associated *p*-value (p = 0.003) provides strong evidence to reject the null hypothesis of normality. The same result was reached also for the sample relative to the emotional engagement in the "Traditional FC", in which we obtained a Shapiro-Wilk statistic W = 0.947 and a *p*-value p = 0.006. Therefore, based on the Shapiro-Wilk test we were able to compare the mean values of the two samples only with the paired Wilcoxon signed-rank test, and not with a traditional parametric paired t-test.

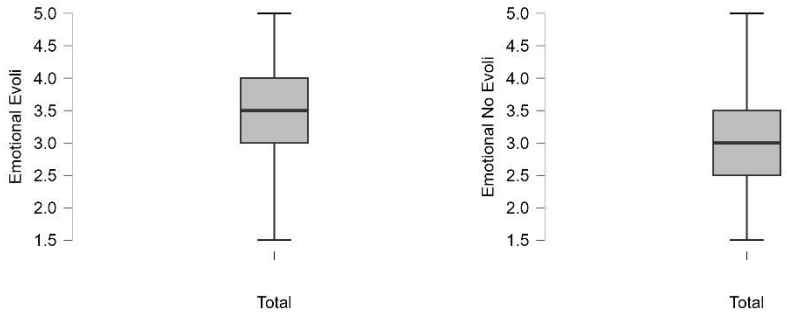

Figure 10: Box plots for the emotional engagement items

The results of the paired Wilcoxon signed-rank test indicate a significant difference between the mean scores of the two paired samples. The test yielded a signed-rank statistic (W) of 674.000 and a z-score of 3.155. Additionally, the associated *p*-value was less than 0.001, indicating strong evidence to reject the null hypothesis and conclude that the mean score of the sample about the emotional engagement in the "FC with Evoli" is significantly higher than the mean score of the emotional engagement in the "Traditional FC".

4.1.4 Cognitive engagement

For the analysis of the cognitive engagement, we considered the last two answers of the questionnaire shown in Table 7. The descriptive statistics showed a very small difference between the mean score of the cognitive engagement in the "FC with Evoli" (M = 3.654, SD = 0.744) and the mean score of the cognitive engagement in the "Traditional FC" (M = 3.647, SD = 0.697).

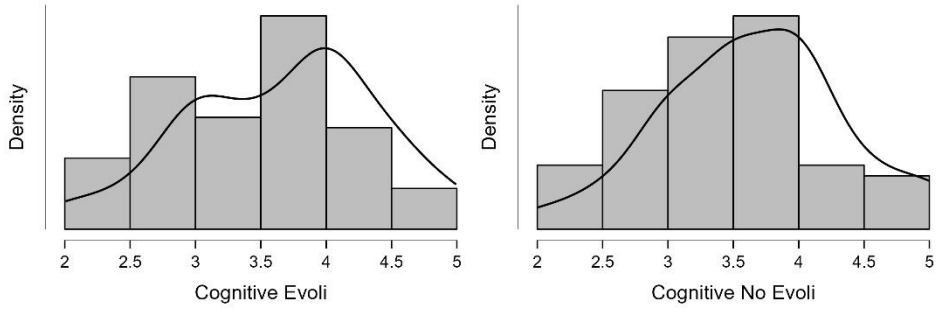

Figure 11: Distribution plots for the cognitive engagement questions

Table 7: Cognitive engagement questionnaire items with mean score and standard deviation

Item	Statement	Mean score	Std. Dev.
C1a	During the study at home of a traditional FC, I feel committed to the task of studying the material	3.618	0.811
C1b	During the study at home of a FC with Evoli, I feel committed to the task of studying the material	3.618	0.847
C2a	During the study at home of a traditional FC, I am able to self-regulate my study	3.676	0.921
C2b	During the study at home of a FC with Evoli, I am able to self-regulate my study	3.691	0.902

Again, the Shapiro-Wilk test showed that the two samples did not follow a normal distribution. For the first sample about the cognitive engagement of the "FC with Evoli," we obtained a Shapiro-Wilk test statistics (W = 0.941) and a p-value (p = 0.003) that confirmed a significant deviation from the normal distribution. The same conclusion can be confirmed for the paired sample, since we obtain a Shapiro-Wilk statistics W = 0.946 and a p-value p = 0.005.

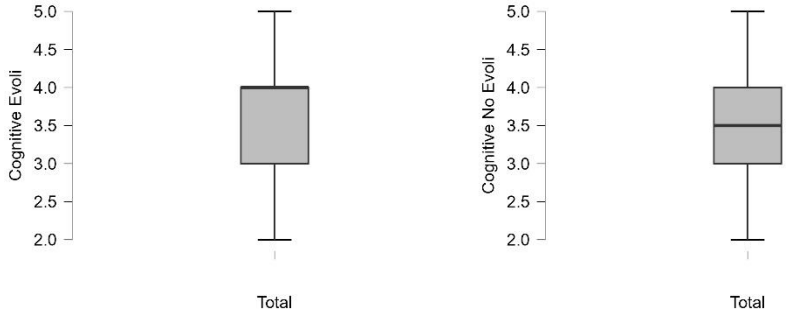

Figure 12: Box plots for the cognitive engagement items

We performed a paired Wilcoxon signed-rank test to examine the difference between the mean values of the two conditions, with Evoli and without Evoli. We chose a paired Wilcoxon signed-rank test because, again, the Shapiro-Wilk test showed that the samples of both conditions were non-normally distributed. The results of the paired Wilcoxon signed-rank test indicate that there is no significant difference between the mean scores of the two paired samples. The test yielded a signed-rank statistic (W) of 156.000 and a z-score of -0.175. Additionally, the associated p-value was 0.576, which is above the commonly used significance level of 0.05. Based on these results, there is insufficient evidence to reject the null hypothesis. Therefore, we can state that the use of Evoli does not result in significant differences in cognitive engagement.

4.2 Qualitative Data

The questionnaire contained seven open-ended questions, which received a total of 47 responses. As previously mentioned, although two independent reviewers tagged them according to six themes (the four types of engagement, comments on the tool, other off-topic comments) we did not conduct any numerical analysis on these data due to the low response rate. However, we present an overview of the students' opinions and include some quotes to provide fresh insight into the quantitative data.

According to the respondents, the behavioral engagement did not show significant improvements, as they felt that their study approach or efficiency remained largely unchanged. However, a subset of participants expressed positive views on the use of Evoli and its influence on their study at home. They valued the opportunity to provide and receive feedback and said that the interactivity provided by Evoli was superior to simply watching videos or reading documents. From the point of view of the agentic engagement, most participants identified the possibility to see their peers' comprehension status and questions as the most engaging feature. It allowed them to understand whether others had similar doubts and difficulties, providing a sense of community (that, consequently, increased the emotional engagement as well). Some participants found it particularly helpful to read other people's questions, as it prompted them to consider different perspectives. In the open questions, emotional engagement was investigated, particularly regarding the risk of students feeling exposed while using Evoli. Responses varied, with some feeling slightly exposed, while others not. Some noticed that asking questions by writing on Evoli instead of speaking up in class would benefit shy students. Concerns were raised about anonymity, with some suggesting an option for anonymous questioning. However, many participants did not feel embarrassed asking questions, emphasizing that seeking clarification is essential to the learning process. From a cognitive perspective, most participants found Evoli's self-assessment reminders to be helpful in increasing their self-awareness of their level of understanding. It is worth quoting *verbatim* the words of one respondent: "[the tool] improved the study experience. It made me aware of what I didn't really understand before the in-class lesson, so it provided me with more time to try and understand the concepts myself or ask the professor." By comparing their own responses to questions with those of their peers and expressing their level of understanding through smiley faces, students heightened their awareness of their preparedness and could better gauge their progress. Additionally, several participants acknowledged that the platform expedited explanations during in-class sessions and alleviated the sense of overwhelm during lessons, making the experience more beneficial overall.

Finally, we present some quotes extracted from the interviews on the four types of engagement. Regarding behavior, one interviewee said: "Using the tool takes away the attitude of 'I'll watch it later' because it gives you more, a sense of something that needs to be done, which motivates a student to watch the video." The two interviewed students also appreciated being invited to be more active, with one remarking: "Even simple questions like 'How well did you understand this topic?' or ''Are there any specific questions?' or seeing 'This is how much others have understood this topic,' are certainly an improvement over the passivity of just sending a YouTube link to the students and ask to watch the video." On the front of agency, one of the interviewed students noted how "with the right input from students, there was a difference, a desire to do more: even people who have no personal interest in telecommunications, for example, were curious and went to learn more about topics."

Regarding emotional impact, two quotes testify to its nature: "I also felt less alone, especially in relation to other students […] I was pleased when, half an hour before the lesson, while I was on the train, I was watching the video, and there were three other students connected. I said to myself, 'okay, I'm not alone,' so much so that we even chatted afterwards to talk about the video." Another participant said: "...I also know that many students are hesitant to ask questions in person, perhaps out of lack of self-confidence. A tool like Evoli also facilitates the exposure of one's doubts and helps the professor understand what the class has understood."

Finally, two quotes testify to the potential on the cognitive front: "The tool made me ask more questions," and "I had to gather my thoughts, ask myself if I had really understood, if I had any questions. I had to stop for a minute or two to collect my thoughts."

5. Limitations

Our study's primary limitation involves assessing the FC experience without a dedicated pedagogical tool relying on the subjective insights from participants' personal experiences, rather than a controlled experiment. However, this limitation is mitigated by the academic context of the university where the research took place, where students are accustomed to video-based FC, giving them a strong foundation in this pedagogical approach. Their extensive exposure to non-technology-assisted FC provides them with a deep understanding of the nuances within the flipped learning paradigm. Considering these factors, despite this inherent limitation, we maintain a high level of confidence in the participants' ability to thoughtfully assess their FC experiences, both with and without a supporting technological tool, based on their experiential depth acquired through prolonged engagement with the non-technology-assisted form of the FC model. Secondly, the voluntary nature of the participation may have introduced bias, as only 68 out of 378 students enrolled in the courses responded to the questionnaire. Thirdly, the study is based on a limited sample of undergraduate students in specific courses,

making the results potentially non-generalizable to other contexts or subjects. Therefore, caution is needed when interpreting the results, and further research with larger samples and in diverse contexts is necessary.

6. Conclusions

In this article, we presented a study aimed at investigating the potential of VATs to increase learning engagement in its four dimensions (behavioral, agentic, emotional, and cognitive) during the independent study phase of the FC. Our research question was: "Can a Video/Media Annotation tool increase students' learning engagement in the home session of the FC?". In a real-world setting, we conducted a case study in a real life setting and collected data through a questionnaire with both closed and open-ended questions, as well as via two interviews. The data collected seems to confirm what the literature already suggests, namely the ability of a VAT to increase engagement in the preparatory phase of the FC approach. Compared to the existing literature, it takes a step forward by specifically highlighting how the different facets of learning engagement are affected, particularly in the agentic and emotional dimensions. Regarding the agentic component, the data confirm what we expected, since a VAT, by its nature, encourages students to do more than passive consumption. The emotional component is increased by the sense of community that the tool creates, by reporting the presence and reactions of all students, and by making students feel already projected towards the interactive in-person session. A delicate emotional aspect, the embarrassment of "raising your hand" and asking questions, appears to be greatly mitigated by the mediation of a tool. Behavioral engagement does not appear to be particularly modified using the tool, although some students, in the open-ended questions, seem to suggest that the tool encourages greater commitment. It should be noted that regarding behavioral engagement, which has received attention in the literature on FC, there is no conclusive data even regarding the difference between a "normal" situation and an inverted situation (Lai et al., 2021), with some authors suggesting that there is an increase in behavioral engagement (Bond, 2020), while others do not find significant differences (Hodgson et al., 2017; Subramaniam and Muniandy, 2019). This topic certainly merits further investigation. Finally, on the cognitive front, the data support that, compared to a normal FC session, there are no appreciable differences, which is a positive element because it means that the demands placed by the tool were not annoying to the point of distracting from the content and increasing cognitive load. Overall, the study demonstrates that the use of a VAT can improve engagement in the FC, not affecting the cognitive load.

The findings of this study have practical implications for educators and designers of educational technology. Teachers are recommended to include VATs that allow for the collection of students' opinions during the preparatory phase of the FC, as this can increase engagement. Designers of educational technology should investigate which characteristics of a tool for collecting student reactions could enhance the behavioral and cognitive components of learning engagement.

Ethics statement

All participants involved were informed about the nature, purpose, and potential outcomes of the study. Data were collected in anonymous form and no personally identifiable information was collected or stored.

AI statement

During the preparation of this work the authors used ChatGPT in order to improve readability and language. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Acknowledgements

This work is part of the dissemination of the results of the project PNRR-ECS-MUSA -Ecosistema dell'innovazione MUSA-Multilayered Urban Sustainability Action - ECS00000037 - CUP: D43C22001410007.

References

Akçayır, G., Akçayır, M., 2018. The flipped classroom: A review of its advantages and challenges. Computers & Education 126, 334–345.

Al-Zahrani, A.M., 2015. From passive to active: The impact of the flipped classroom through social learning platforms on higher education students' creative thinking. British Journal of Educational Technology 46, 1133–1148. https://doi.org/10.1111/bjet.12353

Audas, R., Willms, J.D., 2001. Engagement and dropping out of school: A life-course perspective. Citeseer.

Baig, M.I., Yadegaridehkordi, E., 2023. Flipped classroom in higher education: a systematic literature review and research challenges. Int J Educ Technol High Educ 20, 61. https://doi.org/10.1186/s41239-023-00430-5

Bergdahl, N., Nouri, J., Fors, U., Knutsson, O., 2020. Engagement, disengagement and performance when learning with technologies in upper secondary school. Computers & Education 149, 103783.

Bergmann, J., Sams, A., 2012. Flip your classroom: Reach every student in every class every day. International society for technology in education.

Bishop, J., Verleger, M.A., 2013. The Flipped Classroom: A Survey of the Research. Presented at the 2013 ASEE Annual Conference & Exposition, p. 23.1200.1-23.1200.18.

Bond, M., 2020. Facilitating student engagement through the flipped learning approach in K-12: A systematic review. Computers & Education 151, 103819. https://doi.org/10.1016/j.compedu.2020.103819

Bryson, C., Hand, L., 2007. The role of engagement in inspiring teaching and learning. Innovations in education and teaching international 44, 349–362.

Burke, A.S., Fedorek, B., 2017. Does "flipping" promote engagement?: A comparison of a traditional, online, and flipped class. Active learning in higher education 18, 11–24.

Cassano, G., Di Blas, N., 2024. A Tool to Support Students-to-Teacher Feedback in Asynchronous Online Contexts. IEEE Transactions on Learning Technologies 17, 585–593. https://doi.org/10.1109/TLT.2023.3273109

Costley, J., Lange, C.H., 2017. Video lectures in e-learning: effects of viewership and media diversity on learning, satisfaction, engagement, interest, and future behavioral intention. Interactive Technology and Smart Education 14, 14–30.

DeLozier, S.J., Rhodes, M.G., 2017. Flipped classrooms: A review of key ideas and recommendations for practice. Educational psychology review 29, 141–151.

Divjak, B., Rienties, B., Iniesto, F., Vondra, P., Žižak, M., 2022. Flipped classrooms in higher education during the COVID-19 pandemic: findings and future research recommendations. International Journal of Educational Technology in Higher Education 19, 1–24.

Elmaadaway, M.A.N., 2018. The effects of a flipped classroom approach on class engagement and skill performance in a blackboard course. British Journal of Educational Technology 49, 479–491.

Etemi, B.P., Uzunboylu, H., Latifi, S., Abdigapbarova, U., 2024. The Effect of the Flipped Learning Approach on Engineering Students' Technology Acceptance and Self-Directed Learning Perception. Sustainability 16, 774. https://doi.org/10.3390/su16020774

Fredricks, J.A., Blumenfeld, P.C., Paris, A.H., 2004. School engagement: Potential of the concept, state of the evidence. Review of educational research 74, 59–109.

Gilboy, M.B., Heinerichs, S., Pazzaglia, G., 2015. Enhancing student engagement using the flipped classroom. Journal of nutrition education and behavior 47, 109–114.

Halasa, S., Abusalim, N., Rayyan, M., Constantino, R.E., Nassar, O., Amre, H., Sharab, M., Qadri, I., 2020. Comparing student achievement in traditional learning with a combination of blended and flipped learning. Nursing Open 7, 1129–1138. https://doi.org/10.1002/nop2.492

Hodgson, T.R., Cunningham, A., McGee, D., Kinne, L.J., Murphy, T.J., 2017. Assessing behavioral engagement in flipped and non-flipped mathematics classrooms: teacher abilities and other potential factors. International Journal of Education in Mathematics, Science and Technology 5, 248–261.

Huang, A.Y., Lu, O.H., Yang, S.J., 2023. Effects of artificial Intelligence–Enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. Computers & Education 194, 104684.

Jia, C., Hew, K.F., Jiahui, D., Liuyufeng, L., 2023. Towards a fully online flipped classroom model to support student learning outcomes and engagement: A 2-year design-based study. The Internet and Higher Education 56, 100878. https://doi.org/10.1016/j.iheduc.2022.100878

Kazanidis, I., Pellas, N., Fotaris, P., Tsinakos, A., 2019. Can the flipped classroom model improve students' academic performance and training satisfaction in Higher Education instructional media design courses? British Journal of Educational Technology 50, 2014–2027.

Lage, M.J., Platt, G.J., Treglia, M., 2000. Inverting the Classroom: A Gateway to Creating an Inclusive Learning Environment. The Journal of Economic Education 31, 30–43. https://doi.org/10.2307/1183338

Lai, H.-M., Hsieh, P.-J., Uden, L., Yang, C.-H., 2021. A multilevel investigation of factors influencing university students' behavioral engagement in flipped classrooms. Computers & Education 175, 104318.

Lo, C.K., Hew, K.F., 2020. A comparison of flipped learning with gamification, traditional learning, and online independent study: the effects on students' mathematics achievement and cognitive engagement. Interactive Learning Environments 28, 464–481.

Lundin, M., Bergviken Rensfeldt, A., Hillman, T., Lantz-Andersson, A., Peterson, L., 2018. Higher education dominance and siloed knowledge: a systematic review of flipped classroom research. International Journal of Educational Technology in Higher Education 15, 1–30.

Marks, H.M., 2000. Student Engagement in Instructional Activity: Patterns in the Elementary, Middle, and High School Years. American Educational Research Journal 37, 153–184. https://doi.org/10.3102/00028312037001153

Mirriahi, N., Jovanović, J., Lim, L.-A., Lodge, J.M., 2021. Two sides of the same coin: video annotations and in-video questions for active learning. Education Tech Research Dev 69, 2571–2588. https://doi.org/10.1007/s11423-021-10041-4

O'Flaherty, J., Phillips, C., 2015. The use of flipped classrooms in higher education: A scoping review. The internet and higher education 25, 85–95.

Rahman, A.A., Aris, B., Rosli, M.S., Mohamed, H., Abdullah, Z., Mohd Zaid, N., 2015. Significance of preparedness in flipped classroom. Advanced Science Letters 21, 3388–3390.

Reeve, J., Tseng, C.-M., 2011. Agency as a fourth aspect of students' engagement during learning activities. Contemporary educational psychology 36, 257–267.

Seery, M.K., 2015. Flipped learning in higher education chemistry: emerging trends and potential directions. Chemistry Education Research and Practice 16, 758–768.

Strelan, P., Osborn, A., Palmer, E., 2020. The flipped classroom: A meta-analysis of effects on student performance across disciplines and education levels. Educational Research Review 30, 100314.

Subramaniam, S.R., Muniandy, B., 2019. The effect of flipped classroom on students' engagement. Technology, Knowledge and Learning 24, 355–372.

Symonds, J.E., Kaplan, A., Upadyaya, K., Salmela-Aro, K., Torsney, B.M., Skinner, E., Eccles, J.S., 2021. Momentary engagement as a complex dynamic system.

Wang, M.-T., Fredricks, J., Ye, F., Hofkens, T., Linn, J.S., 2019. Conceptualization and assessment of adolescents' engagement and disengagement in school: A Multidimensional School Engagement Scale. European Journal of Psychological Assessment 35, 592.

Wanner, T., Palmer, E., 2015. Personalising learning: Exploring student and teacher perceptions about flexible learning and assessment in a flipped university course. Computers & Education 88, 354–369. https://doi.org/10.1016/j.compedu.2015.07.008

Yilmaz, R.M., Baydas, O., 2017. An examination of undergraduates' metacognitive strategies in pre-class asynchronous activity in a flipped classroom. Education Tech Research Dev 65, 1547–1567. https://doi.org/10.1007/s11423-017-9534-1

Zainuddin, Z., 2017. First-Year College Students' Experiences in the EFL Flipped Classroom: A Case Study in Indonesia. International Journal of Instruction 10, 133–150.

