JOURNAL OF LEARNING ANALYTICS

Volume 9(3), 125–151. https://doi.org/10.18608/jla.2022.7555

Tcherly: A Teacher-Facing Dashboard for Online Video Lectures

Pankaj Chavan[1] and Ritayan Mitra[2]

Abstract

The use of online video lectures in universities, primarily for content delivery and learning, is on the rise. Instructors' ability to recognize and understand student learning experiences with online video lectures, identify particularly difficult or disengaging content and thereby assess overall lecture quality can inform their instructional practice related to such lectures. This paper introduces Tcherly, a teacher-facing dashboard that presents class-level aggregated time series data on students' self-reported cognitive-affective states they experienced during a lecture. Instructors can use the dashboard to evaluate and improve their instructional practice related to video lectures. We report the detailed iterative prototyping design process of the Tcherly Dashboard involving two stakeholders (instructors and designers) that informed various design decisions of the dashboard, and also provide usability and usefulness data. We demonstrate, with real-life examples of Tcherly Dashboard use generated by the researchers based on data collected from six courses and 11 lectures, how the dashboard can assist instructors in understanding their students' learning experiences and evaluating the associated instructional materials.

Notes for Practice

- Learning analytics for investigating student interactions and experiences with video lectures (e.g., clickstream) have certain limitations, such as a lack of understanding of students' reasons behind their clicks (or actions) and a lack of practical implications for instructors to inform their instructional design and practice, as well as for designers of video lectures.
- Experience sampling of students affords the generation of context-specific learning analytics data. We propose a system that uses anonymous experience sampling of students' voluntary (as and when they want to report without any prompt), real-time self-reports of their cognitive-affective states, and presents aggregated time series of these data on a teacher-facing dashboard. The dashboard can assist instructors in understanding student learning experiences, evaluate lecture materials, and support instructional decision-making.
- This paper demonstrates how stakeholders (instructors and designers) can be involved in the design process of such a dashboard to inform microlevel design decisions such as visualization, the format of dashboard elements, and supports required by instructors to make sense of the presented information (analytics). The evolution of the dashboard design through iterative prototyping with instructors and designers is demonstrated along with usability and usefulness evaluation results.
- We present guidelines for instructors to use the dashboard based on data gathered from six courses and 11 lectures. The real-life examples of dashboard use demonstrate how to use dashboard features and visualizations in tandem to understand student learning experiences and evaluate the associated instructional materials.

Keywords

Video lecture, learning analytics, teacher dashboard, student feedback

Submitted: 24/07/2021 — **Accepted:** 02/08/2022 — **Published:** 24/10/2022

Corresponding author [1] Email: pankajchavan@iitb.ac.in Address: IDP in Educational Technology, IIT Bombay, Mumbai – 400076, India. ORCID: 0000-0002-4538-6731

[2] Email: rmitra@iitb.ac.in Address: IDP in Educational Technology, IIT Bombay, Mumbai – 400076, India. ORCID: 0000-0001-5182-8789

1. Introduction

The increasing number and diversity of students have driven interest in increasing flexibility in higher education (Preston et al., 2010). Thus, video-based learning has been used in universities due to its advantages in terms of cost and accessibility (Yoon et al., 2021). The extensive use of videos started prominently in virtual environments (such as MOOCs) to ensure instructor presence (Scagnoli et al., 2017) by imitating the long-standing tradition of lecturing (Morris, 2009). Pre-recorded video lectures soon became a core component of pedagogical approaches such as the flipped classroom for content delivery and learning in universities (Mirriahi & Vigentini, 2017). Video lectures are used in different scenarios (Montrieux et al., 2015), e.g., as an out-of-class activity in a flipped classroom, as supplementary material to broaden or deepen student understanding, or as problem-solving or worked example videos. Lecture captures (video captures of classroom lectures) are another form of video lectures that are provided to students for the initial viewing and/or reviewing of content (Brooks et al., 2009). However, with video lectures, learning occurs in an environment where students and instructors are physically separated. Hence, it is critical for instructors to be aware of how students interact with and learn from video lectures. For example, in the most commonly adopted blended learning approach in universities, i.e., the flipped classroom, it is important for instructors to know how students learn from and experience video lectures to effectively integrate in-class activities with out-of-class activities.

Typical approaches to investigate student experiences with video lectures include using post hoc questionnaires to collect data on students' subjective video preferences (Cross et al., 2013), affect, mental effort, and perceived learning (Kizilcec et al., 2015, 2014). Although these instruments can effectively capture overall feelings and feedback, they do not provide detailed insights into the specific parts of lectures and their effect on student experiences during learning due to a lack of granularity. Other approaches include video learning analytics, using clickstream data such as pause, play, and seek, to identify which specific parts of lectures students are watching, skipping, rewatching, or speeding through. Recent works employing such an approach have investigated video interaction patterns (Atapattu & Falkner, 2018) and effect of behavioural patterns on learning achievement (Yoon et al., 2021). Brooks and colleagues (2011) used fine-grained user tracking features (e.g., clicking on a button, seeking a new time position) along with student surveys to analyze student behaviours, such as learning approaches and perceptions of learning from lecture captures, while using lecture captures in the Recollect system (Brooks et al., 2009). Other research has investigated the effect of video content (i.e., visual and verbal) on interaction peaks (Gajos et al., 2014; Kim et al., 2014) and the association between video interactions and lecturers' video discourse (Atapattu & Falkner, 2018). Visual learning analytics of students' interaction with a lecture capture has also been used to provide insights into students' usage and investigate differences in the usage patterns of different groups (e.g., high and low achievers; Brooks et al., 2013). They have investigated different ways of visualizing learners' interaction data to provide insights into learning and teaching, such as visualizations of students' rewatching behaviour (e.g., viewing behaviours, such as regular or pauser re-watcher), temporal inter-video usage (i.e., watching behaviour of learners over the span of a course), and usage across different groups of students (e.g., viewership activity of high vs. low achievers). However, the approaches discussed above face an important limitation: they are unable to capture the students' reasons behind the clicks made while watching a video lecture (Srivastava et al., 2021, 2019). For instance, the number of views or watching behaviour could indicate different underlying reasons, such as perceived importance, confusion, or engagement. Likewise, there may be multiple underlying reasons for pausing and replaying a video that may not be immediately obvious from clickstream data alone.

Furthermore, the practical implications of such approaches for the design of video lectures and teaching practice are limited because these approaches do not provide instructors with any direct indicators regarding the instructional materials (e.g., the elements of instruction that are working or not working and why) or changes required in instructional materials, and they do not inform instructor actions (e.g., changing or adapting in-class activities in a flipped classroom). Kay's (2012) comprehensive review that explored the use of videos in education called for further research on the characteristics of video lectures that affect learning (e.g., quality of visuals used, pace, type of explanations offered, cognitive load, engagement) and their impact on instructors' pedagogical decisions (e.g., revising and refining the design of video lectures). Some video learning analytics approaches discussed above used video interaction data (such as pause, forward, or backward seek) to explore the relationship between video interaction and perceived video difficulty (Li et al., 2015), whereas others combined interaction data and content-based analysis (Gajos et al., 2014; Kim et al., 2014) to associate the interaction peaks (a significantly large number of students showing a similar interaction pattern) with the characteristics of the video lecture, such as visual transitions in the video or important remarks by the instructor. Even with advancements in video analytics, most of the areas highlighted by Kay (2012) — except engagement, perceived video difficulty, and the effect of visuals or verbal actions — are not extensively explored due to the limitations of data captured in informing instructors why certain patterns exist and about the effects of instructional materials on student learning experiences and the changes required. Mirriahi and Vigentini (2017) also called for further research to investigate the following questions: *What type of intervention can be applied to inform instructors of changes required in video lectures? How is it integrated into a course?*

To address some of the limitations of existing approaches in video lecture analytics, we propose a system to capture students' voluntary (as and when they want to report), real-time self-reports of their cognitive-affective states throughout a video lecture, supplement that information with students' self-reported reasons behind those states, and finally present classroom aggregated data to teachers on the Tcherly teacher-facing dashboard to inform their instructional practice. The goal of this paper is twofold: *first*, to demonstrate and describe in detail the teacher-centred design process behind the dashboard (visualizations, features, support for instructors, etc.) along with a brief usability and usefulness evaluation by instructors and, *second*, to identify ways in which instructors could use the dashboard as a guide by using real data captured from video lectures. We first present the background of our work, describing current approaches in video learning analytics, the challenges associated with it, and the role of student feedback in teacher awareness and reflection (section 2). We then describe the data related to Tcherly Feedback in section 3. Subsequently, we present the detailed process and results of the teacher dashboard design study and its evaluation with instructors in section 4. In section 5, we discuss the implementation and evaluation of Tcherly in a real teaching–learning setting. Using real-life examples of dashboard use generated by researchers based on the real data, we demonstrate how instructors can use the dashboard to inform their teaching practices and student learning. Specifically, we show how different dashboard features could help instructors 1) retrospectively understand how students learn from video lectures, 2) identify the features of instructional materials or events in instruction that are working or not working for students, and 3) become aware of student perceptions of why particular lecture sections worked or did not work for them. The solution we present here was inspired by our previous studies on face-to-face lectures (Chavan et al., 2018; Chavan & Mitra, 2019; Mitra & Chavan, 2019). We propose that it is equally applicable to video lectures since they imitate the long-standing tradition of lecturing (Morris, 2009) and can be considered an instructional medium capable of providing a learning experience equivalent to traditional instructor-led lectures (Stöhr et al., 2020; Yoon et al., 2021).

2. Background

2.1. Online Video Lectures and State-of-the-art LA: Challenges and Limitations

Video learning analytics approaches using explicit factors (macrolevel data), such as views, video hits, and annotations (de Barba et al., 2016; Namuddu & Watts, 2020), provide insights into video usage (e.g., number of views, time spent on a video) and students' questions/comments related to content, but they do not provide collective fine-grained information about when students are engaged, confused, or distracted while interacting with a video. Although user-defined annotations are helpful (e.g., MRAS; Bargeron et al., 2002), analyzing and interpreting them to inform instruction is challenging. This limitation could be addressed by the use of implicit factors (microlevel data), such as pause, seek, play, and stop, in combination with video content analysis (i.e., visual, text, verbal; Gajos et al., 2014; Kim et al., 2014). However, video interaction data have certain inherent limitations:

1) They require large-scale data. The central assumption of such analysis is that a large number of students is watching a video (Gajos et al., 2014), making this approach impractical for university classrooms where the maximum class size is usually less than a hundred.

2) Although clickstreams or logs are an objective measure of student activity (e.g., pause, play, and seek) in an online learning environment, we know only what buttons students are clicking (Ellis et al., 2017; Scheffel et al., 2014) and not why they are clicking them (e.g., pausing and replaying a video many times due to content difficulty, rewinding due to insufficient explanation).

3) Due to this insufficient understanding of students' reasons behind the analytics, we do not have a sufficient context to interpret what the numbers mean (Boyd & Crawford, 2012).

For example, Li et al. (2015) found that patterns such as replay and more frequent pausing indicate higher perceived video difficulty. However, the question of why such patterns are observed in a specific segment or an entire video remains unanswered. Similarly, Gajos et al. (2014) and Kim et al. (2014), investigating how video interaction patterns (i.e., peaks) shown by a large number of students can be explained by visual/nonvisual content, found that 61% of interaction peaks were associated with visual transitions (e.g., changing from a whiteboard explanation to a talking head, changing slides to present a new concept, presenting an interesting segment (e.g., demo) with visual transitions before and after) and verbal transitions caused 39% of interaction peaks (e.g., instructor introducing an important concept or re-emphasizing what has already been covered visually). Such an analysis is claimed to reveal points of interest or confusion for instructors and students. However, it is challenging to understand whether a video's presentation quality or storyline generates a particular pattern (e.g., students returning to watch the video) or what a point of interest or confusion means without knowing the students' *reasons behind their actions*. Hence, the practical implications of such analytics to inform teaching practice are limited (Mirriahi & Vigentini, 2017) because these analytics do not provide instructors any indicators to address the following questions: *Is the design of instructional tasks working? What changes/additions are required in the instructional materials? How likely are students to* *follow the lecture's discourse and become engaged in different parts of the lecture? Which parts are students not able to understand and why?*

However, there are some notable exceptions, such as the CLAS system (Risko et al., 2012), the gaze-based metric "with-me-ness direction" (Sharma et al., 2016; Srivastava et al., 2021), and real-time difficulty self-reports (Srivastava et al., 2019). Using the CLAS system, students can indicate important points in a video lecture by clicking a button indicating *"for this user, something important happened at this point in the lecture."* The instructors are then provided with analytics to evaluate the efficacy with which they have communicated important points in their lectures. However, this system does not help instructors determine *why* these points were important. Moreover, research has focused more on students' adoption of the tool (Gašević et al., 2017) and less on instructors' use of the tool. The gaze-based metric "with-me-ness direction" measures co-attention between learners' gaze and instructors' dialogue in a video lecture, helping instructors understand students' attention state. However, drawing actionable insights from this information is challenging for instructors. The details of how this information is presented to instructors and how they can use it to investigate the effects of instructional materials on students' learning experiences and inform their pedagogical actions are largely missing. The approach of using real-time difficulty self-reports addresses some of the above-mentioned limitations, but it collects data on a single cognitive construct, i.e., difficulty. It thus highlights only problematic parts of a lecture, overlooking positive aspects. Hence, to address these limitations, we propose collecting data on multiple constructs (cognitive-affective states), and we design a teacher dashboard presenting data to provide teachers with actionable insights, assist them in understanding and evaluating the student learning experiences associated with the instructional materials, and inform their pedagogical actions.

Another strand of work focused on inferring learners' cognitive and affective states in mobile-based MOOCs captures photoplethysmography (PPG) signals and facial expression data (Pham & Wang, 2017, 2018a, 2018b; Xiao & Wang, 2015, 2016). AttentiveLearner uses on-lens finger gestures on a back camera for video control (i.e., a video plays when the camera lens is covered and pauses if it is uncovered). It collects PPG signals through fingertip transparency changes and extracts learners' heart rates to predict their mind wandering (zoning out) events in MOOC sessions (Pham & Wang, 2015). AttentiveLearner has also been used to infer other states, such as interest/boredom and confusion (Xiao & Wang, 2015). AttentiveLearner was proposed to be useful in enabling adaptive tutoring features (e.g., alerting learners when they are zoned out or providing more relevant exercises) as well as in assisting instructors with fine-grained, aggregated visualization of learners' cognitive and affective states (such as confusion or attention) synchronized with the learning materials to identify and reflect upon areas within lectures needing improvement (e.g., confusing or boring segments of a lecture) to refine instructional content for the future (Pham & Wang, 2015; Xiao & Wang, 2015). The initial results related to how instructors can identify confusing topics for learners based on overall feedback (i.e., aggregated cognitive states inferred from PPG signals) were demonstrated. The advanced version of the system, AttentiveLearner^2, uses both back and front cameras to monitor PPG signals and track learners' facial expressions, respectively, to infer learners' cognitive-affective states in real time (Pham & Wang, 2018b). The two channels are used to achieve more robust emotion detection. AttentiveLearner^2 also proposed exploring visualization techniques to help instructors identify difficulties among learners and opportunities for improvements in learning materials. However, these systems were designed primarily to provide feedback to students about their own learning and lack a teacher-facing equivalent.

2.2. Teacher Reflection and Student Feedback

According to Schön (1987), alternative perspectives about their teaching are a critical aspect of teachers' professional development. The author claims that student perspectives can be valuable sources of data for a teacher's personal reflection (Bell & Aldridge, 2014). The importance of seeking student perspectives has been well recognized in the literature (Aldridge & Fraser, 2008; Bustingorry, 2008; Hoban & Hastings, 2006; Rhine, 1998), which suggests that student perspectives can help teachers question their assumptions and view their practice through the eyes of others. Moreover, both of these aspects are acknowledged as essential for changing teaching practice. Research has shown the potential of using student perception data as a basis for reflection, evaluation, and improvement in teaching (Bell & Aldridge, 2014; Fraser, 2012; Hoban & Hastings, 2006). Using student feedback consisting of student perceptions of the learning environment as a tool for reflection and change has been shown to be effective in improving the classroom environment (Aldridge et al., 2010; Bell & Aldridge, 2014).

Hattie & Timperley (2007) conceptualized feedback as *"information provided by an agent (e.g., teacher, student, peer, self, experience) regarding aspects of one's performance or understanding."* Such feedback can support the recipient in changing her actions and improving her performance. Using student feedback to gather information about teaching to evaluate and improve teaching quality has always been a focus of research in education (Bell & Aldridge, 2014; Chen & Hoshower, 2003; Hoban & Hastings, 2006; Richardson, 2005). Surveys or questionnaires such as Students' Evaluations of Educational Quality (SEEQ), the Course Perception/Experience Questionnaire (CPQ/CEQ), and the Constructivist-Oriented Learning Environment Survey (COLES) are the most common forms of student feedback (Bell & Aldridge, 2014; Richardson, 2005). Richardson (2005) argues that such surveys or questionnaires can provide necessary evidence for assessing and supporting attempts to improve teaching quality. Marsh (1987) and Marsh & Roche (1993) also noted that student ratings are used to provide formative feedback to instructors about their teaching efficacy. Extensive and empirical studies investigating how teachers use student evaluations of teaching (SETs) to improve their teaching are rare. Some studies evaluating the effects of feedback from SETs indicate that the process of consultation and counselling on the received feedback improved SET scores (Roche & Marsh, 2002). In a survey by Jacobs (1987), the majority of instructors (70%) reported that SETs had helped them improve their teaching. However, 63% of instructors indicated that even when they were able to interpret their ratings, they often did not know what to do to improve their teaching, indicating the importance of consultation and counselling. Consultation and counselling sessions can be elaborate, wherein a consultant discusses the student ratings and student-written comments and works with the teacher to develop or suggest strategies for improving the particular areas selected by teachers based on the student ratings (Marsh & Roche, 1993). Other, less commonly used forms of student feedback include student interviews, student learning logs, and observation schedules completed by students (Hoban & Hastings, 2006). Specifically, in the context of video lectures, post hoc questionnaires are used to collect student feedback on their video preferences (Cross et al., 2013), affect, mental effort, and perceived learning (Kizilcec et al., 2015, 2014). The research on the use of student feedback to inform teaching is based on an underlying assumption that teachers value student opinions about teaching enough to change their practices to improve their instruction.

3. Tcherly Feedback: Data Collection from Students

The data that inform the teacher-facing dashboard are collected through the Tcherly website, which presents students with a feedback interface called Tcherly Feedback (Figure 1a). This is separate from the teacher-facing dashboard that is the focus of this study. DEBE (pronounced as Debbie) is an acronym for difficult, easy, boring, and engaging, which are the four types of feedback buttons provided in the interface. Students are expected to click on any of the buttons based on their perception while watching a lecture.

3.1. Theoretical Background of Tcherly Feedback

Learning is a multidimensional process where cognition, emotion, and motivation produce the end result. As such, any data and analytical framework that work toward improving learning should consider as many of these factors as possible. Several theories/frameworks indicate how such factors combine to produce learning, such as the cognitive disequilibrium framework (CDF; Graesser & D’Mello, 2012), control-value theory (Pekrun, 2006), and flow theory (Csikszentmihalyi, 1990). In addition, the literature on emotions in learning suggests that cognition and emotion are tightly coupled and play a vital role in student learning (Graesser & D’Mello, 2012; Pekrun, 2006). Complex learning (i.e., *learning of difficult materials*) has been shown to trigger emotions such as *engagement*, *confusion, frustration*, and *boredom* (Graesser & D’Mello, 2012), which influence the cognitive resources of students during learning, such as their attention in a given academic activity (Tze et al., 2016). On the other hand, Csikszentmihalyi’s (1990) flow theory suggests that *boredom* is experienced when the *difficulty of a task is too low* compared to an individual’s skills. Thus, we focus on harvesting information about students’ cognitive-affective states when they are listening to lectures. The choice of the four states, i.e., difficult, easy, engaging, and boring, was informed by the theories/frameworks above and some practical considerations. According to Graesser & D’Mello (2012), boredom is only the end state of the affective trajectory of a student, one that should never be reached. It might seem more appropriate to solicit data on intermediary affective states such as confusion and frustration. While that is entirely possible, we must keep in mind that it might be difficult for students to differentiate between the exact nature of the negative valence they might be experiencing, e.g., confusion vs. frustration. Moreover, it is impossible to gather data with very high construct accuracy, as students have less time for judgment. We suggest that the states engaging and boring are easily interpretable and meaningful for students (feedback provider) to judge and report as well as for teachers (feedback receiver) to gather quick and useful insights from the data regarding affective experiences.

Some initial findings related to the temporal dynamics of students’ affective states in video-based MOOC learning (Xiao et al., 2017) indicated that engagement was the most frequent affect, followed by boredom and then confusion; other affective states, such as frustration, delight, and surprise, were less frequent. Moreover, a strong engagement-to-boredom transition was observed in the study as opposed to a weak confusion-to-engagement transition. Such reasons, pertaining to the theories/frameworks of cognition and affect and those findings related to the temporal dynamics of students’ affective states, supported our selection of the cognitive-affective states.

3.2. DEBE Methodology

We use a variant of the experience-sampling data acquisition methodology (Larson & Csikszentmihalyi, 2014) to collect real-time data from students by allowing them to report on aspects of their learning in face-to-face or online video lectures in real time. Self-report measures are a practical approach, as they provide a quick and easy subjective measure of cognition and affect. Experience sampling methodology (ESM) is well-established for understanding the content of people's thoughts, e.g., their cognitive, emotional, and motivational states, through self-reports. This aspect of ESM aligns with our decision to capture learners' real-time cognitive-affective states through self-reports. However, we further the methodology by adding the element of voluntary reporting (as and when the student wants to report without any prompt) of cognitive-affective experiences. Our approach also shares a commonality with the CLAS system (Risko et al., 2012), which includes a system-defined annotation strategy (i.e., the annotation has the same meaning for all users). We also build upon the LIM App that captures real-time comprehension levels from the audience in mass lectures (Rivera-Pelayo et al., 2013) and the approach of collecting real-time difficulty self-reports in video lectures (Srivastava et al., 2019). We incorporate the multidimensional nature of the learning process, proposing data collection on multiple constructs that reflect cognition and affect in the learning process. The student feedback interface (Figure 1a) has four buttons on it: Difficult, Easy, Boring, and Engaging (right). Students can provide voluntary feedback during lectures about their cognitive-affective states by clicking on these four buttons. A video automatically pauses when a button (e.g., Difficult) is clicked, and a detailed feedback options view (Figure 1b) that includes a list of possible reasons for the provided feedback (e.g., 'lack of explanation' or 'unclear or cluttered presentation') appears. Students can either choose from the list or enter a reason by clicking on Other. If students do not want to provide reasons for their feedback, they can click on Close and restart the lecture. All student feedback is then aggregated to generate class feedback for teachers (Figure 1c). The design considerations that inform the student feedback interface (e.g., using complementary cognitive-affective states, using multiple buttons as opposed to a dial or scale to keep the interface as unobtrusive as possible for students, and using anonymous feedback) and design decisions such as incorporating a pop-up with a list of annotations to capture reasons for feedback are detailed in a previous publication (Chavan & Mitra, 2019).

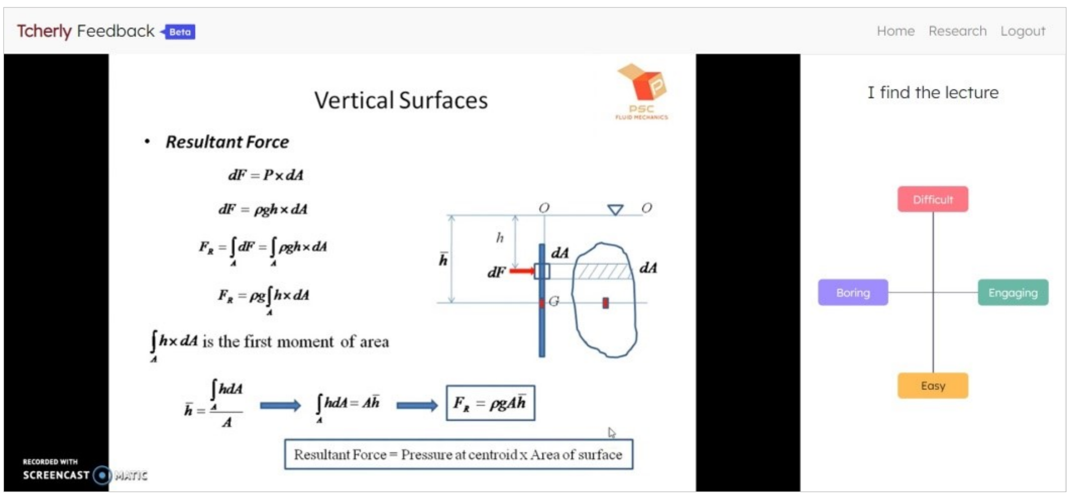


Figure 1a. Student feedback interface (Tcherly Feedback)

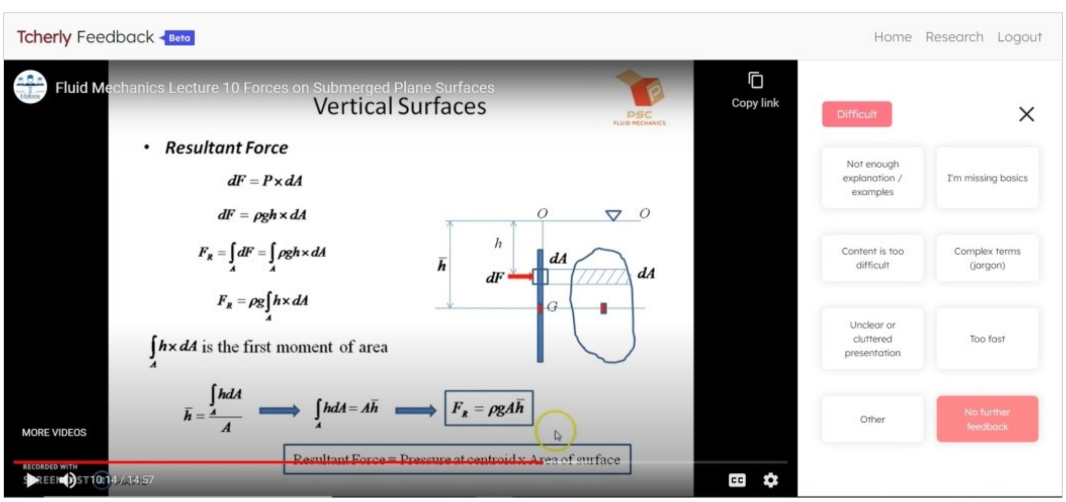


Figure 1b. The detailed feedback options view that appears after clicking on a feedback button

Figure 1c. Time series of class-level aggregation of student feedback. Data collected from individual students with the Tcherly Feedback system, along with a timestamp for each response, are aggregated to generate a time series of feedback during a lecture (x-axis = time in minutes; y-axis = number of students). The red line shows net difficulty (number of students clicking difficulty minus the number of students clicking easy), and the green line shows net engagement (number of students clicking engaging minus the number of students clicking boring). Positive values indicate that the lecture was difficult (red) or engaging (green). Negative values indicate that the lecture was easy (red) or boring (green)

4. Design of the Teacher-Facing Dashboard of Tcherly

The data collected from the student feedback interface (Tcherly Feedback) is reported to teachers through a teacher-facing dashboard (Tcherly Dashboard) to assist them in evaluating their lectures and materials, understanding students' learning experiences, and informing their instructional decision-making, e.g., refining and revising their lecture materials or changing, adapting and improving their instructional activities formatively. Our approach aligns with the data-assisted approach suggested by Brooks and colleagues (2014), which leverages the sense-making process of instructional experts (e.g., instructors, instructional designers) by providing them with summaries of learner activities/interactions (e.g., learner responses to a survey or navigation through video content) in technology-enhanced learning environments to gain insights into their activities and choose instructional interventions to enhance students' learning experiences. One of the approaches suggested to aggregate and present learners' data to instructors is information visualization, and we follow a similar approach to present Tcherly Feedback to teachers.

Research on learning analytics for teachers has focused on designing, developing, and evaluating visual analytics methods and tools for teachers to understand teaching and learning processes and support their teaching practice (Vatrapu, 2012). Such methods and tools enable teachers to adapt their teaching and teaching practices through awareness and reflection on data about learners and learning processes (Verbert et al., 2013). The most commonly used method to report analytics is a learning analytics dashboard (LAD; Verbert et al., 2014). Schwendimann et al. (2016) define a LAD as *"a single display that aggregates different indicators about learner(s), learning process(es) and/or learning context(s) into one or multiple visualizations."* Dashboards have received considerable attention as tools that can provide users with relevant insight, prompt user reflection, and potentially inform interventions aimed at optimizing learning and the quality of the student experience. A review of dashboard applications by Verbert et al. (2013, 2014) indicates its widespread use to report data to teachers captured from a variety of learning environments, such as face-to-face lectures, face-to-face group work, and online or blended learning, to support teacher awareness, reflection, sensemaking, and behaviour changes. We adopt the method of reporting analytics through a teacher-facing LAD that reports student feedback data to teachers using different information visualization techniques to give a macro view (e.g., aggregate student feedback and associated metrics) as well as a meso-view of students (e.g., a subgroup that provided both difficult and engaging feedback for a specific part of the lecture or distribution of student reasons for their feedback in a lecture segment). We also leverage the affordances of using video-based approaches to support teacher reflection. A large body of work has focused on recording and analyzing video lessons to support reflection (Santagata et al., 2007; van Es et al., 2015). In the teacher-facing dashboard, student feedback is synched with a video lecture to facilitate reflection that is precise and embedded in everyday practice (i.e., such feedback provides teachers with an opportunity to reflect on specific aspects of their instruction).

We implemented one of the user-centred design techniques, i.e., iterative prototyping with stakeholders, to develop the dashboard design based on instructors' analytics needs from Tcherly Feedback data, such as how to visualize the feedback data, what interactions to include, how the analysis is saved, and what supports are required on the dashboard. The importance of focusing on users' needs in analytics to provide them with rich and interpretable information has been highlighted (Corrin & De Barba, 2015; Demmans Epp et al., 2019). LA researchers have used a range of design techniques, such as interviewing end users about their needs, contextual inquiry interviews to understand teacher practice (Michos et al., 2020; Xhakaj et al., 2016), scoping sessions with practitioners to identify and prioritize needs, obtaining their feedback on designs (Fiorini et al., 2018), and engaging teachers throughout the prototyping process itself (Holstein et al., 2018). Demmans Epp et al. (2019) reported a user-centred iterative teacher dashboard design conducted with instructors and information visualization or human–computer interaction (HCI) experts. The iterations were performed based on a review of the analytics research (e.g., design considerations, information visualization principles, usability) to identify areas for improvement and interviews with instructors to understand their analytics needs and workflow. This was followed by the creation of a low-fidelity prototype of the dashboard, which was evaluated by researchers with a background in information visualization or HCI. We followed a similar approach of involving instructors and designers (information visualization and HCI experts) in the dashboard prototyping process and obtaining their continuous feedback on the design. Our approach differed slightly in that instructors and designers were involved throughout the prototyping process (e.g., what visualizations to use and what features to include to support instructors' informational needs). We involved instructors and designers to develop visualization ideas to represent feedback data and different dashboard designs, refine the ideas and designs in the process, and finalize the dashboard design. Finally, we evaluated the usability and usefulness of the dashboard with a different cohort of instructors that used a high-fidelity dashboard prototype. In section 4.1, we give an overview of the design study participants, followed by a brief explanation of the three phases of the design process. The results section (4.2) demonstrates how instructors and designers developed a range of visualization ideas to represent feedback data (section 4.2.1) and evaluated and finalized a set of visualizations to generate low-, mid- and high-fidelity prototypes (section 4.2.2). The usability and usefulness evaluation results of the final dashboard design (high-fidelity prototype) are presented in section 4.2.3.

4.1. Methods

We conducted a design study with five instructors (all males with an engineering background and 6–15 years of teaching experience) and two designers (experts in information visualization, 1 male and 1 female) followed by a preliminary usability and usefulness evaluation of the high-fidelity dashboard prototype with 20 instructors (12 males and 8 females with 4–18 years of teaching experience in the following backgrounds: 17 in engineering, two in science, and one other) to evaluate the expected practicality of the dashboard. The five instructors who participated in the design study were not involved in the usability and usefulness evaluation.

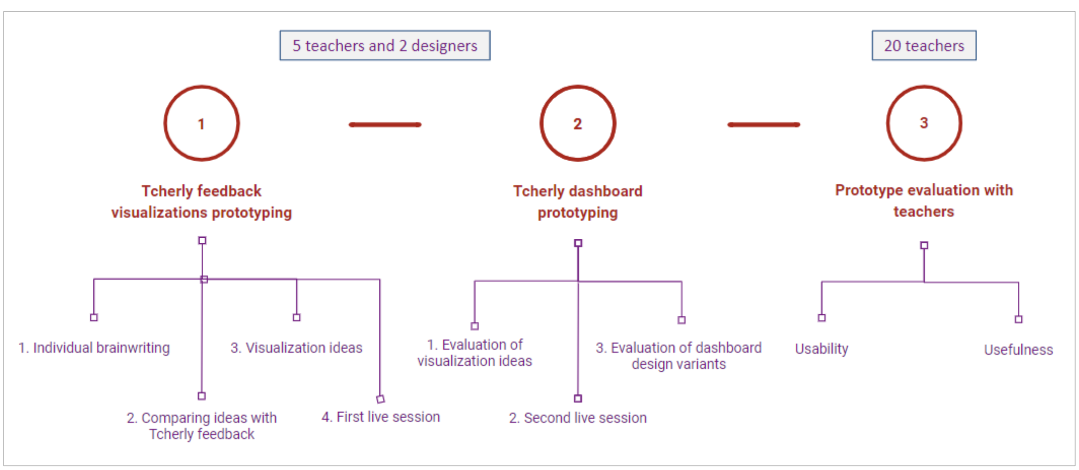


Figure 2. Design study phases

The study was conducted in three phases (Figure 2). All phases were conducted online (a combination of synchronous and asynchronous activities). There were six asynchronous activities, which included tasks such as writing, prototyping, and filling out survey forms, and two live (synchronous) sessions. The total study duration was approximately 1.5 weeks. The primary goal of **Phase 1** (Tcherly Feedback visualization prototyping, days 1–3) was to generate as many visualization ideas as possible from the instructors and designers, collaboratively improvise visualization ideas through discussion, and develop preliminary dashboard designs (low-fidelity prototypes). **Phase 2** (Tcherly Dashboard prototyping, days 4–5) was targeted at identifying potential visualization ideas, improving the design, and developing a set of dashboard designs (mid-fidelity prototypes) through discussion and finalizing the dashboard design through an evaluation of different variants. In **Phase 3** (Tcherly Dashboard prototype evaluation with instructors, days 6–10), a high-fidelity dashboard prototype was developed in Figma, and usability and the usefulness testing of the prototype was conducted using the System Usability Scale (SUS; Brooke, 1996) and the Technology Acceptance Model 2 (TAM2; Venkatesh & Davis, 2000), respectively. Usability and usefulness testing was followed by a closing remarks session where we asked instructors for specific comments and suggestions related to the dashboard based on their interaction experience. The names DEBE Feedback and DEBE Teacher Dashboard were used during the design study instead of Tcherly Feedback and Tcherly Dashboard, respectively, because the name *Tcherly* was conceived afterward. To avoid confusion, we use the names Tcherly Feedback and Tcherly Dashboard throughout the paper.

4.2. Results

4.2.1. Phase 1: Tcherly Feedback Visualization Prototyping

This phase comprised three tasks (Table 1 and Figure 2) followed by the first live (synchronous) session. Individual brainwriting (Task 1) was an icebreaker activity. The goal was to trigger instructors' thinking about the context and prepare them for the next task instead of directly introducing the Tcherly Feedback. The next task, comparing ideas with Tcherly Feedback (Task 2), asked the instructors to compare the ideas they had come up with in Task 1 with what the Tcherly Feedback system does. For this task, the details of how Tcherly Feedback works, along with a schematic of the feedback interface, was shown to the participants. The objective of this task was to familiarize them with the Tcherly Feedback data and establish the relevance for their teaching practice. The two designers were included in the process only after Task 2 completion. They were apprised of the basics of Tcherly Feedback data and the first two tasks given to the instructors. Finally, the visualization ideas task (Task 3) required the instructors and designers to develop visualization ideas (graphs, charts, and any other form of representation) for Tcherly Feedback. All three tasks were conducted asynchronously. The questions and activities were sent sequentially for each of the three tasks. These asynchronous activities were followed by a live (synchronous) session where the participants first presented their visualization ideas and then discussed the ideas. This was followed by a discussion of the preliminary layouts of the dashboard (low-fidelity prototypes) prepared by the researchers and designers based on some of the visualization ideas submitted by instructors and designers.

Table 1. Tasks in Phase 1 of the Design Study

Task 1: Individual brainwriting	Task 2: Comparing ideas with Tcherly Feedback	Task 3: Visualization ideas
Goal: Icebreaker activity to trigger instructors' thinking about the context	**Goal:** Familiarizing instructors and gradually involving them in the design process	**Goal:** Generating as many visualization ideas as possible
Questions: (*only for* instructors) What feedback about your lecture and lecture materials would you like to know from students and why? What feedback about students' feelings about your lecture can be helpful for you?	**Questions:** (*only for* instructors) Identify what is common in the ideas you have come up with and Tcherly Feedback. Does Tcherly Feedback do something that you did not identify? If yes, please explain those aspects briefly. Which aspects of Tcherly Feedback will be useful to evaluate and improve your teaching and teaching materials?	**Activity:** (*for* instructors and designers) Student feedback data (minute-by-minute data) in the excel sheet was provided. Instructors and designers were free to choose to create either a paper or digital prototype of visualization ideas. They were asked to think about the following specific scenario: *If you want to analyze a specific part of the lecture (identify parts in the data provided to you) based on the student feedback data, how do you want to proceed with that?*

In the first live (synchronous) session, the instructors, designers, and researchers first presented their visualization ideas (Figure 3). After each participant's presentation, the pros and cons of different visualizations were discussed. During the discussion, suggestions for improving some of the presented visualization ideas emerged, e.g., allowing instructors to plot variables of their choice on a line chart. The utility of such an interactive line chart was cited by one of the instructors: *"As a teacher, after getting feedback, we just have to click on this [difficult, easy, boring, engaging] checkbox. If I want to have two parameters, difficult and easy, I want to plot only difficult and easy against each other. Then I will be able to do this."*

Moreover, new visualization ideas emerged (e.g., a participation indicator to indicate student participation in a feedback-giving activity). While talking about the importance of the participation indicator, one instructor mentioned the significance of such visualization: *"This parameter is the total count of students. If I have 500 students, then out of 500 students, 300 students have given some feedback."*

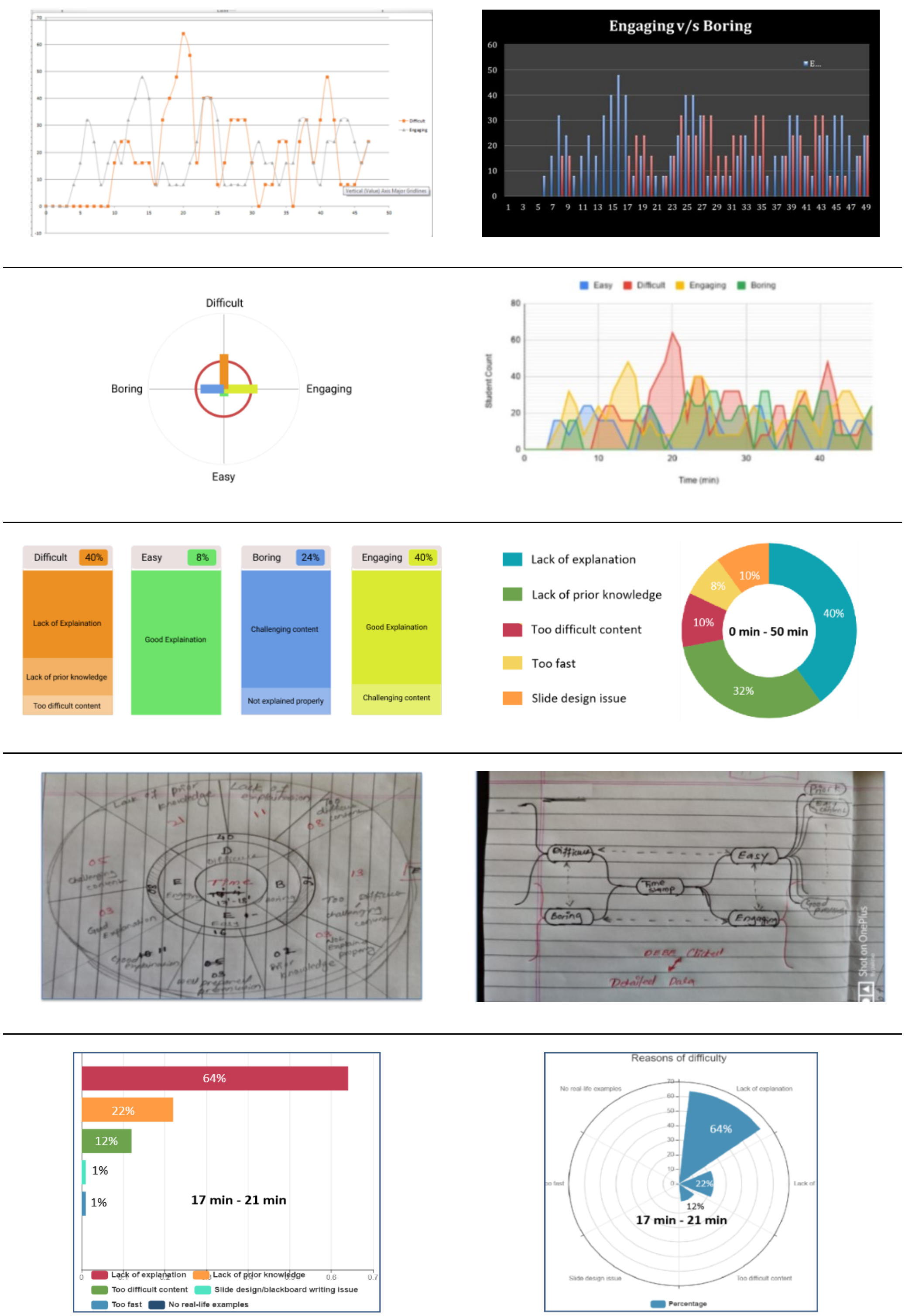


Figure 3. Visualization Ideas Generated in Task 3

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

After the visualization idea discussion, preliminary layouts of the dashboard were demonstrated to the instructors to facilitate an initial discussion on the dashboard elements and their positioning. Figures 4a and 4b show sample dashboard designs that emerged from the visualization ideas generated in Task 3. The preliminary discussion of different aspects of the dashboard, such as the importance of having a separate section to obtain an overall picture of the feedback and selecting a specific segment of the lecture for analysis, resulted in the possible changes/additions required in the dashboard. To give one example, one new feature called *"Bookmark"* was introduced by one of the instructors at this point (*"As a teacher, I am interested in that portion … which topic they (students) felt was difficult. So, here, I can make a bookmark section for my lecture time"*).

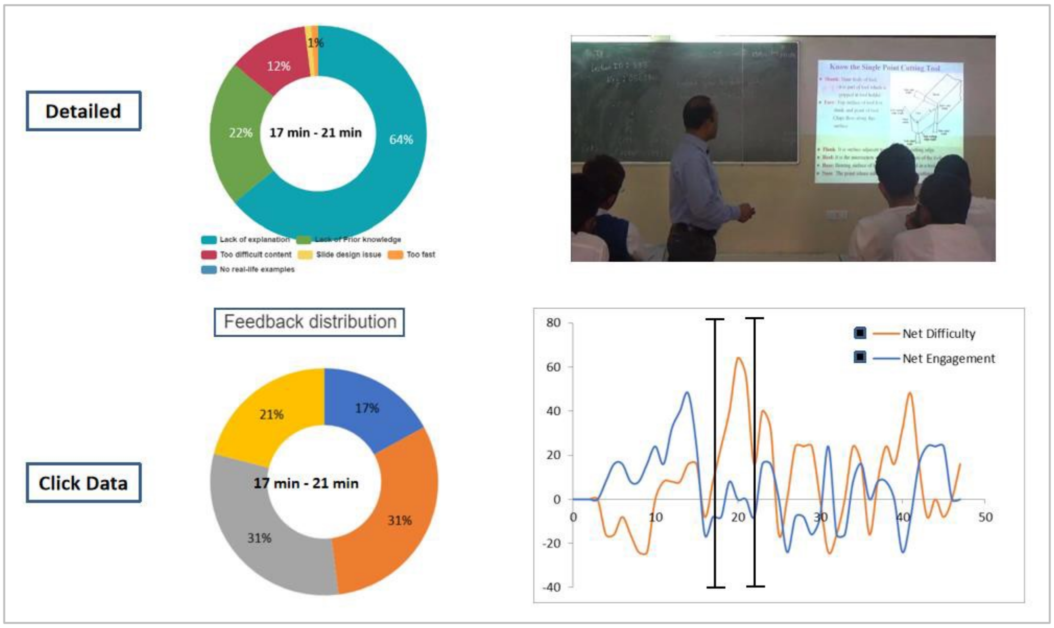


Figure 4a. Initial sample dashboard design (low-fidelity PowerPoint prototype).

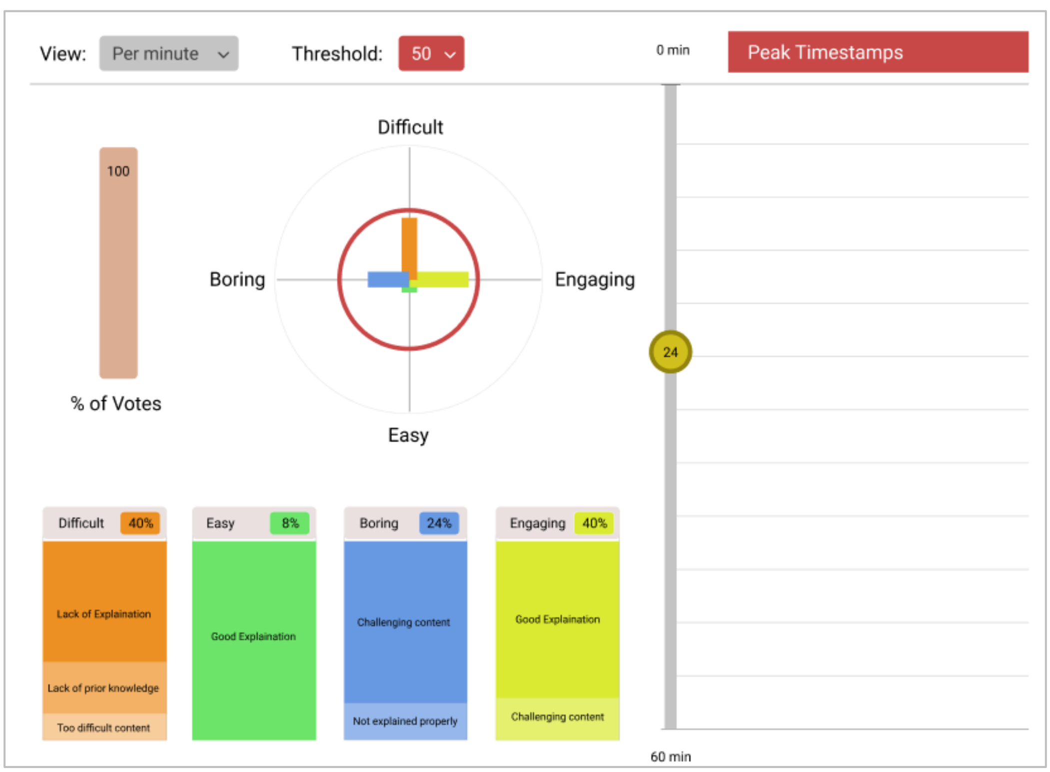
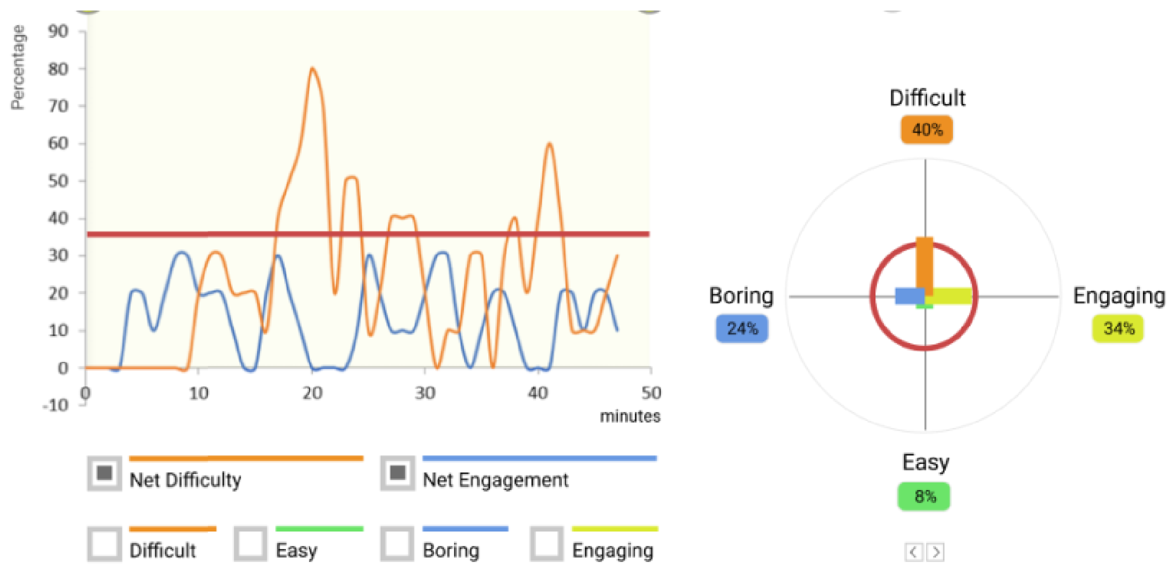


Figure 4b. Initial sample dashboard design (low-fidelity Figma prototype)

4.2.2. Phase 2: Tcherly Dashboard Prototyping

After the first live session, the survey forms to evaluate the visualization ideas were sent to the instructors and designers. The participants were asked to evaluate the visualization ideas on a scale of 1 to 10 (where 1–2 = not at all useful, 3–4 = useful to a very small extent, 5–6 = moderately useful, 7–8 = useful, 9–10 = very useful) based on the following aspects: *Ease of use* (ease of interpretation, clarity in information presentation), and *Time required* to interpret data. Along with providing a rating, participants were asked to indicate the *Strengths* and *Weaknesses* of each visualization idea to gather their subjective evaluations. Table 2 shows instructors' and designers' evaluations of a visualization idea that scored highest.

Table 2. Instructors' and Designers' Evaluations of Visualization Ideas



Average Score	9.00
Strengths	- It is easy to understand the differences between criteria [parameters of DEBE] - It is concise and easy to interpret, and decisive actions can be taken. - Radial charts can be easily understood and are easy to interpret. - The radial column chart represents the DEBE input mechanism, maintaining consistency if it were also used by students in the future. - It is easy to compare overall difficulty in the left figure and actual difficulty in the right figure. - The design looks minimal while conveying all necessary information.
Weaknesses	- The right chart uses a limited area to represent the percentage, but this might be an advantage if we consider the overall dashboard to minimize the colour clutter. - It takes some time to comprehend.

In the second live (synchronous) session, the researchers and designers created and presented a new set of dashboard designs based on the suggestions and inputs received from all participants during Phase 1. This was followed by a discussion of different aspects of the dashboard design, such as the interactions required by the instructor while analyzing feedback, the positioning of different elements of the dashboard (e.g., video, line chart, detailed feedback chart), and which combination of visualizations could provide the required information while working on the analysis task. The features suggested during the first live session, such as a student participation indicator, a bookmark section, and a summary report for the instructors were discussed in detail and finalized. Figures 5a and 5b show the intermediate sample dashboard designs.

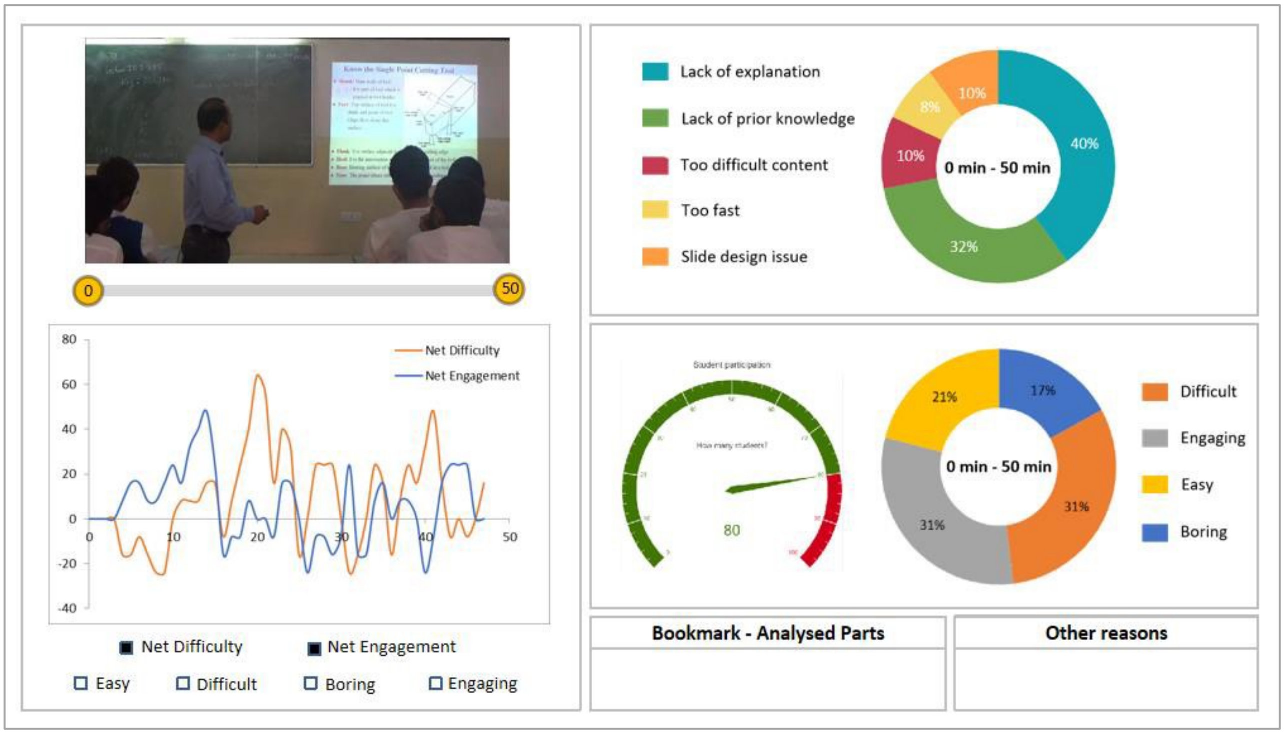


Figure 5a. Intermediate sample dashboard design (mid-fidelity PowerPoint prototype)

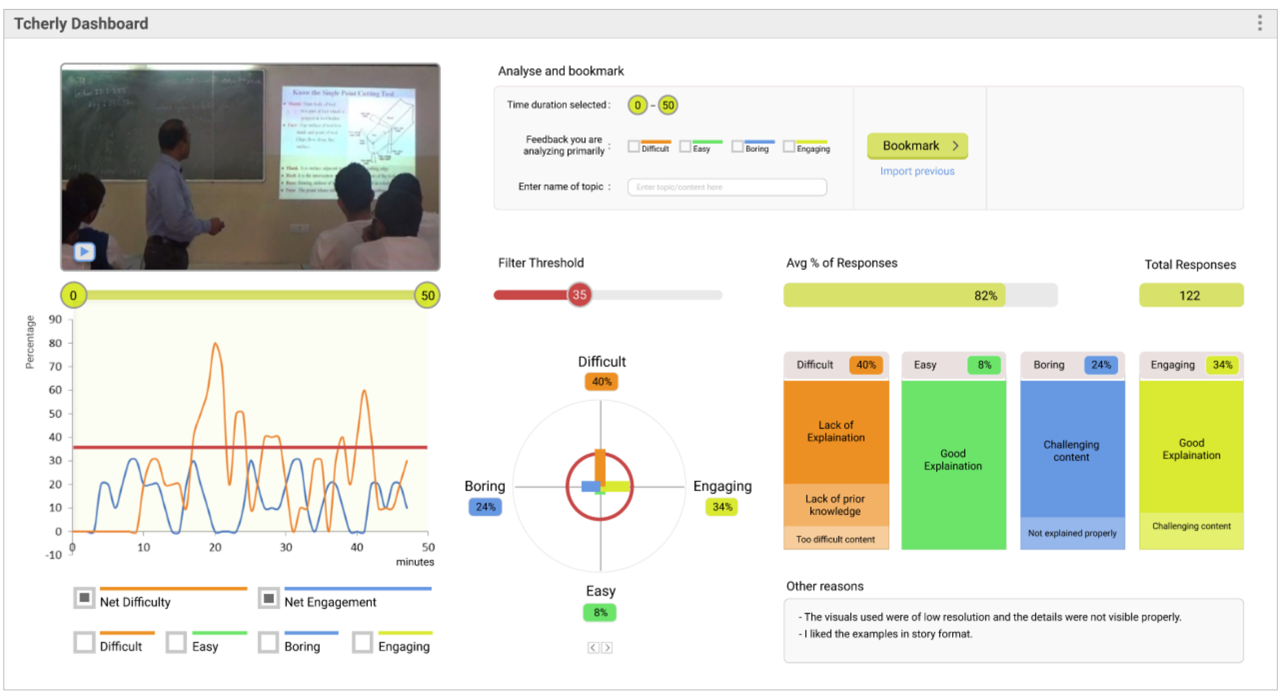
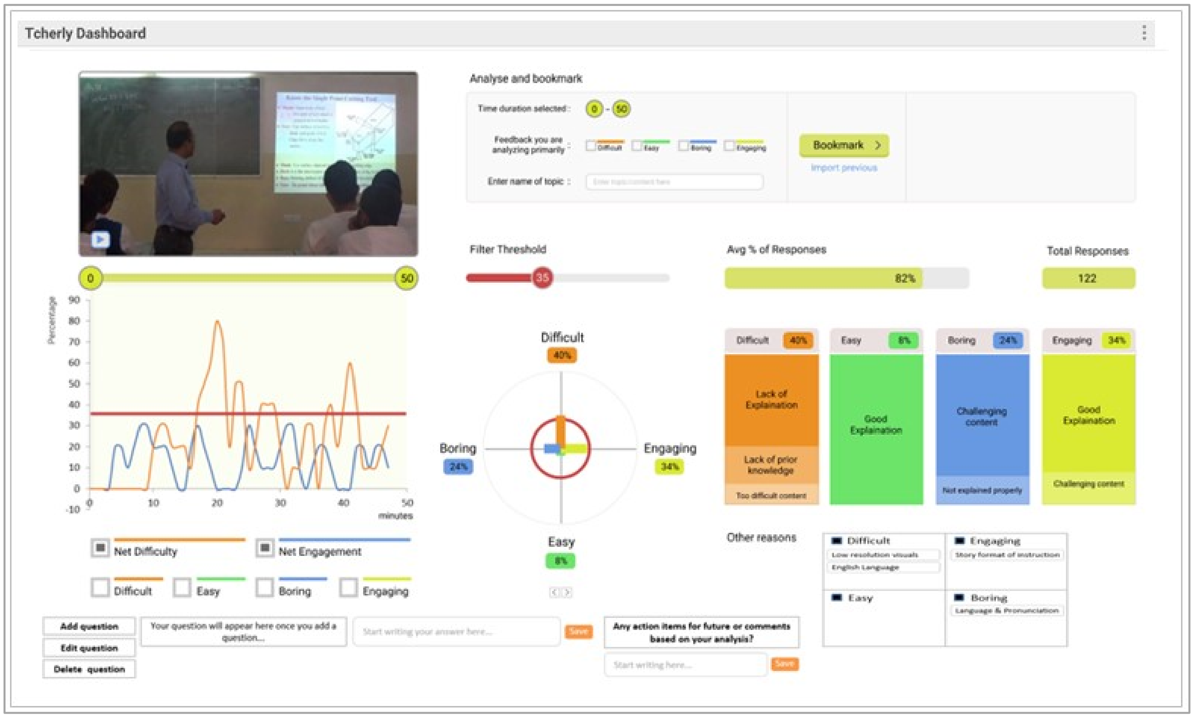


Figure 5b. Intermediate sample dashboard design (mid-fidelity Figma prototype)

Two new sections were introduced to participants during the second live session, namely, *Question generator* and *Action tracker*. The purpose of these sections aligned with suggestions provided by Wise & Jung (2019) on dashboard analytics, namely, teacher-facing analytics dashboards should support teachers in the generation and maintenance of questions related to the analytics and possibly answers. Thus, these sections were retained for the final prototype despite not emerging from the discussions. The instructors' overall perceptions of the usefulness of those sections were positive. Based on the insights from the two live sessions, four dashboard design alternatives were created. They varied in some aspects, such as the position of different elements (video, line chart, DEBE Reasons, bookmark section, etc.) and a set of visualizations. A dashboard design evaluation survey was then sent to the participants to score the four dashboard design alternatives on a scale of 1 to 10, where 1–2 = very poor, 3–4 = poor, 5–6 = average, 7–8 = good, 9–10 = excellent. We asked them to briefly write the *Strengths* and *Weaknesses* of each dashboard design. Table 3 shows the results of the instructors' and designers' evaluations of the dashboard design that scored the highest.

Table 3. Instructors' and Designers' Evaluations of Dashboard Design



Average Score	9.17
Strengths	- The bookmarks in the top right of the layout, being the most important part of the whole analysis activity, gain prominence there and also suggest the steps to follow using this tool in the form of questions right at the top. - This looks [like] a clean dashboard with fewer graphical representations to interpret. - It focuses on important information. It has a good design and is easy to understand. - The bookmark setting on the top right will help the user see it whenever she takes a ride backward into the archives.
Weaknesses	- The detailed other reasons need not have a separate box with a heading again; instead, it could be an expandable vertical rectangle on hover using coding. - The dashboard does not have demarcation presently, which makes it slightly difficult to understand how to observe and operate the dashboard. - The radial chart's representation as a doughnut chart or pie chart is not that impactful, but the user would find it easy.

Based on the strengths and weaknesses reported by the participants, some changes were made in the design, e.g., using clear demarcation between different sections of the dashboard, making other reasons the part of the heatmap, and structuring the question generator and action tracker. Finally, Venn diagrams were added to the final prototype to understand the reporting behaviour of the class (i.e., whether any individuals reported multiple states over a specific time range and is there any relationship between them) in more detail (refer to "Analysis of lecture segments with line chart and Venn diagrams" in section 5.1.2). Figure 6 shows the final dashboard design.

Figure 6. Final dashboard design (high-fidelity prototype, Figma)

4.2.3. Phase 3: Tcherly Dashboard Prototype Evaluation with Instructors

In the last phase of the design, the usability and usefulness evaluation of the high-fidelity dashboard prototype (Figure 6) was performed by a different cohort of 20 instructors. Instructors interacted with the dashboard prototype in three stages. The **first stage** was the *dashboard walkthrough,* in which instructors took an interactive tour to understand the following aspects of each section of the dashboard: Purpose, How to use, and Features. In the **second stage** *(feature interaction)*, the instructors explored and interacted with different features of the dashboard with a focus on understanding their functions and connections between different sections (e.g., how the selection of a specific part of the lecture on the line chart using sliders results in changes in the other sections). In the **third stage,** the instructors *analyzed specific parts of the lecture* (engaging and difficult segments). A demo was given to the instructors for all stages before they interacted with the dashboard prototype. The feedback data they analyzed on the dashboard were from one of our previous studies rather than from their courses. At the end of the third stage, the instructors were given usability and usefulness evaluation surveys. This was followed by a closing remarks session, where instructors' specific comments and suggestions/feedback on the dashboard based on their interaction experience were gathered.

We used the SUS (Brooke, 1996) and TAM2 (Venkatesh & Davis, 2000) surveys to evaluate the perceived usability and usefulness of the dashboard, respectively. The SUS was used to evaluate the user's ability to successfully carry out the feedback analysis tasks, whereas the TAM2 survey was used to test the acceptance of the dashboard, i.e., whether the instructors perceived the dashboard to be useful and relevant for their teaching context and their intent to use it. Hence, we investigated the following four constructs from TAM2: *Intention to use, Perceived usefulness, Perceived ease of use, and Job relevance*. We used only these four constructs from the TAM2 survey because our goal was to evaluate the *expected practicality* of the dashboard. There were 12 items from the TAM2 survey.

The mean SUS score was 86.60 (SD = 9.91; n = 20), reflecting that the dashboard is an acceptable system (Bangor et al., 2008). The participants mentioned the following: *"Good visual interpretation makes the design more user-friendly, which aptly follows the eye ergonomics"*; *"Of course, anyone using it the first time will need a little guidance and some effort. But once you get the hang of it, it should be very easy to understand"*; *"The interface design is simple to understand, but work has to be done on its aesthetic appearance"*; *"Appearance of the dashboard seems to be good, and it is well managed with the consideration of everyone's visualization"*; *"I believe that this dashboard is very intuitive"; and "The Filter Threshold was confusing and not sure whether it is useful as we can easily identify points of interest visually."* Some instructors also mentioned that *the Venn diagrams were crowded*.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

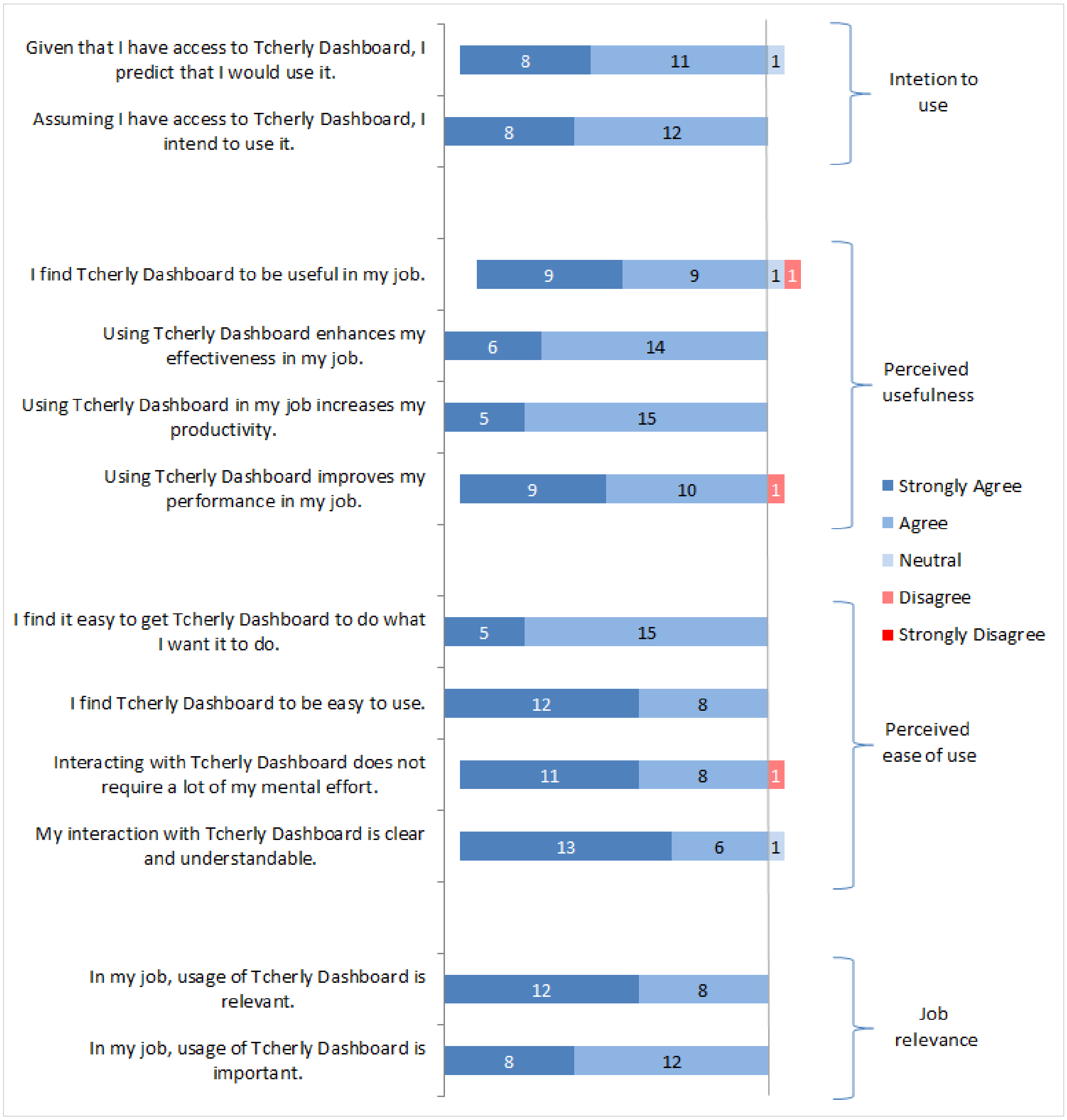


Figure 7. TAM2 survey results. TAM2 items had 7-point Likert scale responses (1-Strongly Disagree, 2-Moderately Disagree, 3-Somewhat Disagree, 4-Neutral, 5-Somewhat Agree, 6-Moderately Agree, 7-Strongly Agree), which were categorized into five groups: Strongly Agree (response 7), Agree (combining responses 5 and 6), Neutral (response 4), Disagree (combining responses 2 and 3), and Strongly Disagree (response 1)

Figure 7 shows the distribution of TAM2 survey responses. Overall, the results showed positive perceptions, with few exceptions. Forty percent of instructors (8 out of 20) reported a strong intention, whereas the remaining 60% (12 out of 20) showed a moderate (or slight) intention to use the dashboard. The closing remarks session also revealed instructors' perceptions of the usefulness of the dashboard.

The teacher dashboard is helpful to improve teaching

The dashboard was perceived as a tool that can help instructors improve their teaching practice: *"I feel that the opportunities to improve my own teaching methodology are endless with this dashboard"*; *"It would greatly benefit the teachers who are ready to adapt and improve their teaching methodology based on the student feedback on video lectures"; "Teachers have a scope for development"*; *"The real-time feedback taken from students will definitely help to improve lectures"; "We can easily see the point where the students are getting confused or when they are experiencing difficulty in understanding"*; *"It is a very well-thought-of dashboard, and the teacher should use it very honestly and with an open mind* *so that they can improve their teaching"*; and *"Real-time suggestions sound more authentic and helpful to upgrade the course and teaching techniques."*

The teacher dashboard is helpful to build a student–teacher relationship

The dashboard was perceived as useful to enhance the student–teacher relationship: *"In times of enhancing student–teacher bond, to inculcate student-friendly teaching skills for a better understanding of concepts and curriculum."* One participant called it a *"student-centric system."*

The evaluation of the high-fidelity prototype indicated the overall usability of the dashboard and instructors' interest and intention to use the dashboard. Some participants appreciated the simplicity and intuitiveness of the design. During this preliminary evaluation study, some instructors pointed out that the interface was easy to use. However, improvements were suggested for the aesthetic appearance of the dashboard. Moreover, some instructors mentioned that the *Filter Threshold* feature might not be required and that the Venn diagrams were crowded. Hence, during the development of the system (Tcherly Feedback and Tcherly Dashboard), an expert designer was involved in incorporating the necessary design changes considering ease of use, colour scheme, interaction elements, etc. Figure 8 shows the Tcherly Dashboard (the details of different sections and features of the dashboard are provided in Appendix A). The system is available at https://www.tcherly.com/.

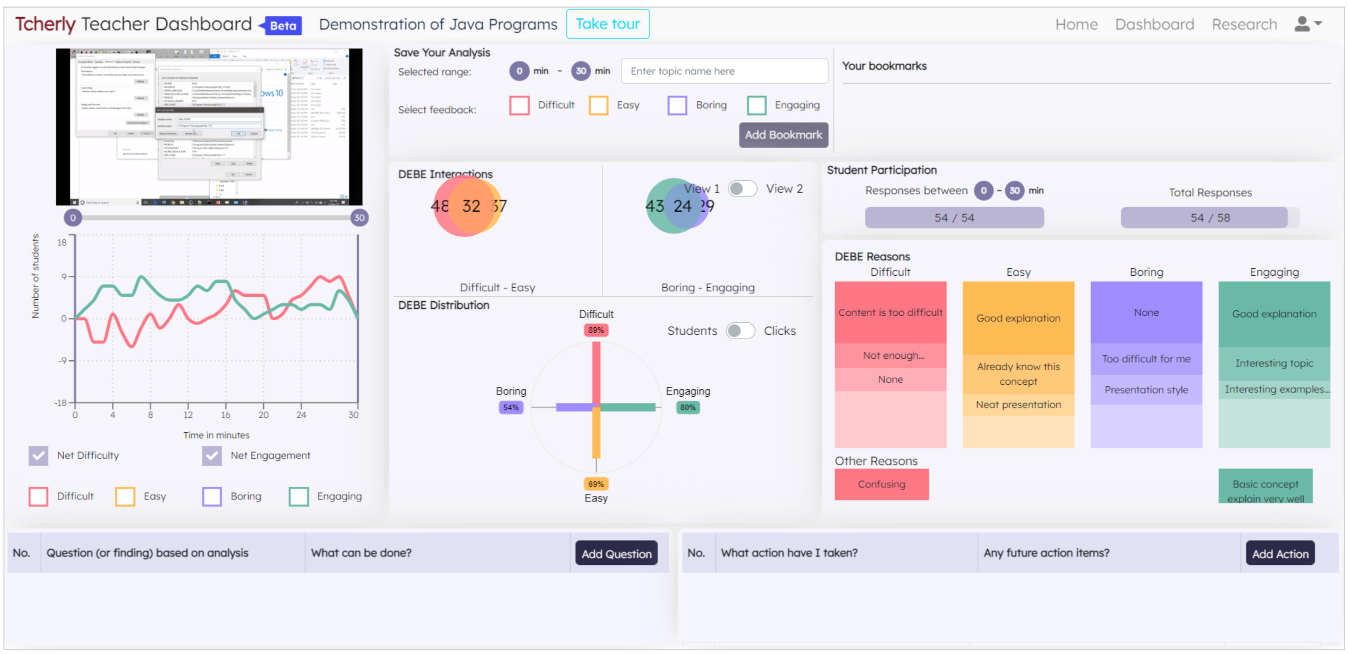


Figure 8. Tcherly - A teacher-facing dashboard

5. Lessons from the Implementation of Tcherly

After the teacher dashboard prototype showed positive perceptions of usability and usefulness, we brought Tcherly into the real teaching–learning setting. Due to the COVID-19 pandemic, all classes were being conducted online. We approached instructors who used video lectures as instructional materials or intended to use them along with live lectures. We implemented Tcherly in six courses across three institutions and collected data from 11 lectures. A total of 273 students and six instructors participated in the study. Table 4 shows the details of the courses and lectures in which the data were collected. Informed consent was obtained from the students. We collected instructors' perceptions of the usability and usefulness of the teacher dashboard using the SUS and TAM2 surveys, respectively. The instructors also responded to the Evaluation Framework for Learning Analytics (EFLA) survey (consisting of the three dimensions: Data, Awareness & Reflection, and Impact) used to measure the overall quality of the dashboard (Scheffel, 2017). Although the usability and usefulness evaluation was conducted with a cohort of 20 instructors in the design phase, it was conducted with the high-fidelity dashboard prototype that did not present actual course data to instructors. Hence, instructors may have rated the dashboard higher than they would have if real data from their classrooms had been presented. Thus, we evaluated the dashboard again when it was implemented in real classrooms. We also collected student responses to the SUS to evaluate their perceptions of the usability of the student feedback interface (section 3).

Table 4. Study Details

Name of course	No. of instructors	No. of students	No. of lectures
Analysis of Algorithms	2	53*	3
Microprocessor	1	35*	2
Operating Systems	1	96*	1
Subtotal:		**96**	
Java Programming	1	54	2
SPCC	1#	51	1
Analysis of Algorithms	1#	72	2
Total:	**6**	**273**	**11**

*Students from the same institute and same year of study; #same instructor.

All six instructors completed the SUS, TAM2, and EFLA surveys. The mean SUS score for the teacher dashboard was 95.0 (SD = 2.74, n = 6), and the responses of the TAM2 survey were also overall positive. The overall EFLA score for the dashboard was 90/100 (Figure 9). The Data dimension outperformed the other dimensions with a score of 99/100. The scores for the Awareness & Reflection and Impact dimensions were 85/100 and 87/100, respectively, indicating that the dashboard stimulated *"teachers' awareness and reflection"* and enabled them *"to teach more efficiently and effectively"* with video lectures.

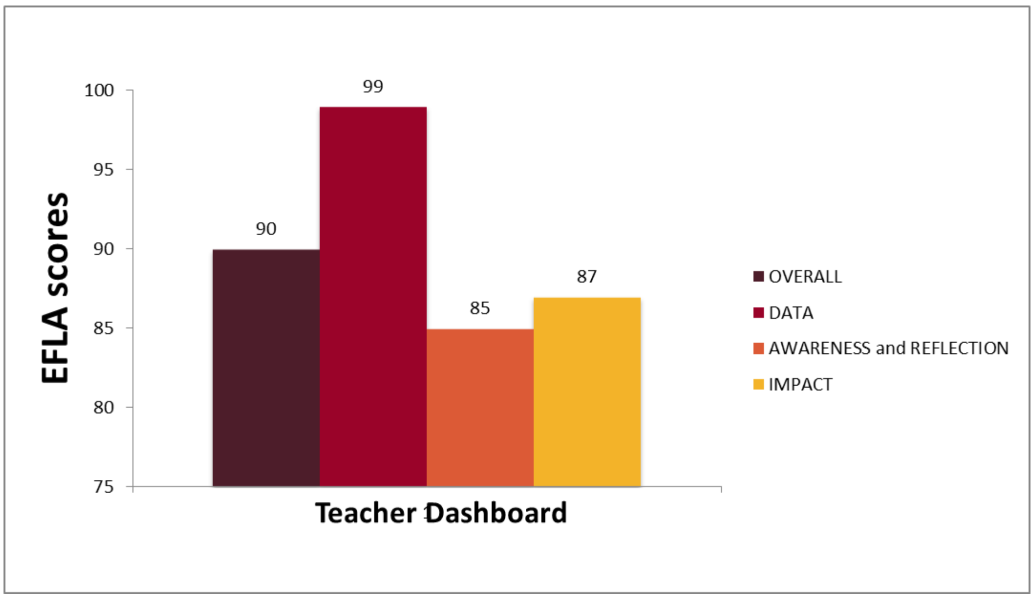


Figure 9. EFLA survey results

Out of 273 students, 242 completed the SUS; the mean SUS score for the student feedback interface was 73.6 (SD = 16.5; n = 242). Out of 242 students, 151 used mobile devices (SUS = 72.5; SD = 16.9) and 91 used a PC/laptop (SUS = 75.5; SD = 15.6) to provide feedback. The score reflects that the feedback interface is an acceptable system (Bangor et al., 2008). Along with the SUS, additional questions were asked to understand student perceptions of aspects such as distraction, anonymity, and value in providing voluntary feedback. For the question, *"Do you think with the feedback system you could provide valuable feedback on the lecture content?",* 94% of students said yes. The responses to other Likert-type items assessing these aspects showed general positive perceptions of the feedback interface. Figure 10 shows the distribution of responses.

5.1. Implications for Instructors: Real-Life Examples of Dashboard Use

In this section, we present the analyses of the video lectures performed by the researchers to demonstrate how instructors can use the dashboard. These could also act as guidelines for instructors.

5.1.1. Elements of Lecture Materials and Learning Experiences

The research on the following significant areas has been reported as relatively sparse (Kay, 2012): characteristics of video lectures that affect learning (e.g., quality of visuals used, pace, type of explanations offered, cognitive load, engagement) and its impact on instructors' pedagogical decisions (for example, design of in-class activities in the flipped classroom, revising and refining design of video lectures). Mirriahi & Vigentini (2017) argue that despite significant research evaluating the

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

efficacy of videos and multimedia for learning, there is a need to identify and investigate different ways to inform instructors of the changes required in the content of video lectures. We demonstrate how the dashboard could assist instructors in evaluating their lectures and lecture materials.

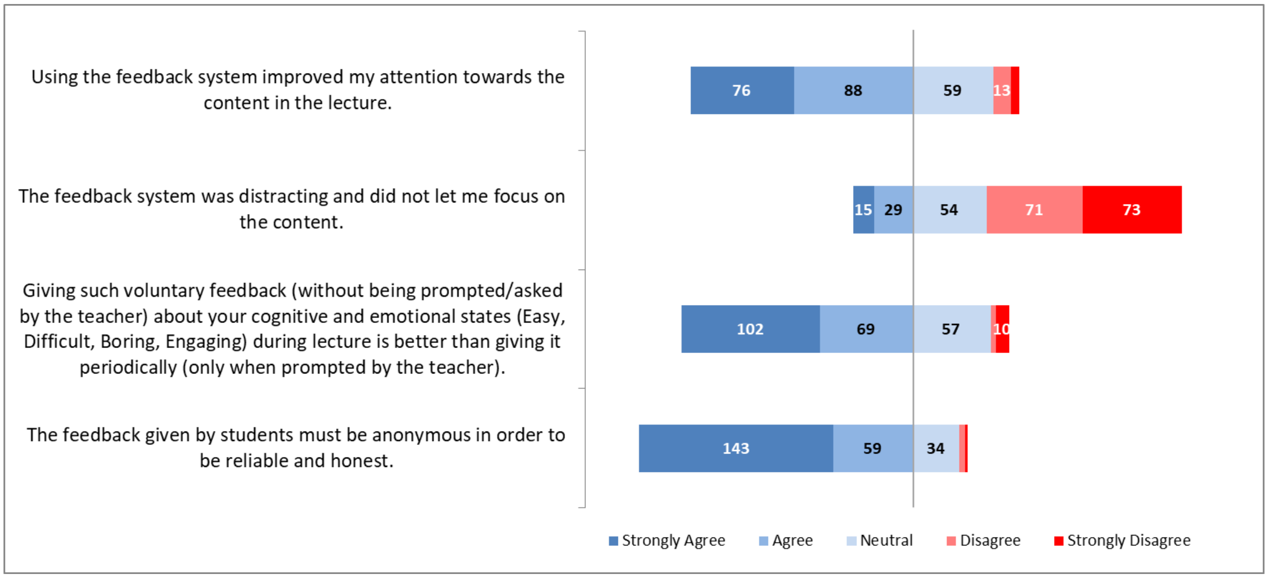


Figure 10. Distribution of responses to the survey items assessing student perceptions of distraction, anonymity, value in providing voluntary feedback, etc. Two questions measured the construct of distraction, one of which was reverse coded

Trigger question

In the video lecture "Segmented and Paged Memory" (Figure 11) from the Operating Systems course, at the start of the last minute (between 7 and 8 min), the instructor poses a question, "Which is best between segmented and paged memory?" and then summarizes the lecture by explaining the pros and cons of each memory management system using bullet points. This is the section where the net engagement rises suddenly (36 students clicking engaging and 14 students clicking easy between 7 and 8 min). This segment turned out to be optimal because of the trigger question bringing attention to the lecture summary, including a clear explanation given using structured bullet points in a sequence, supported by the DEBE Reasons showing "Good explanation" as the primary reason.

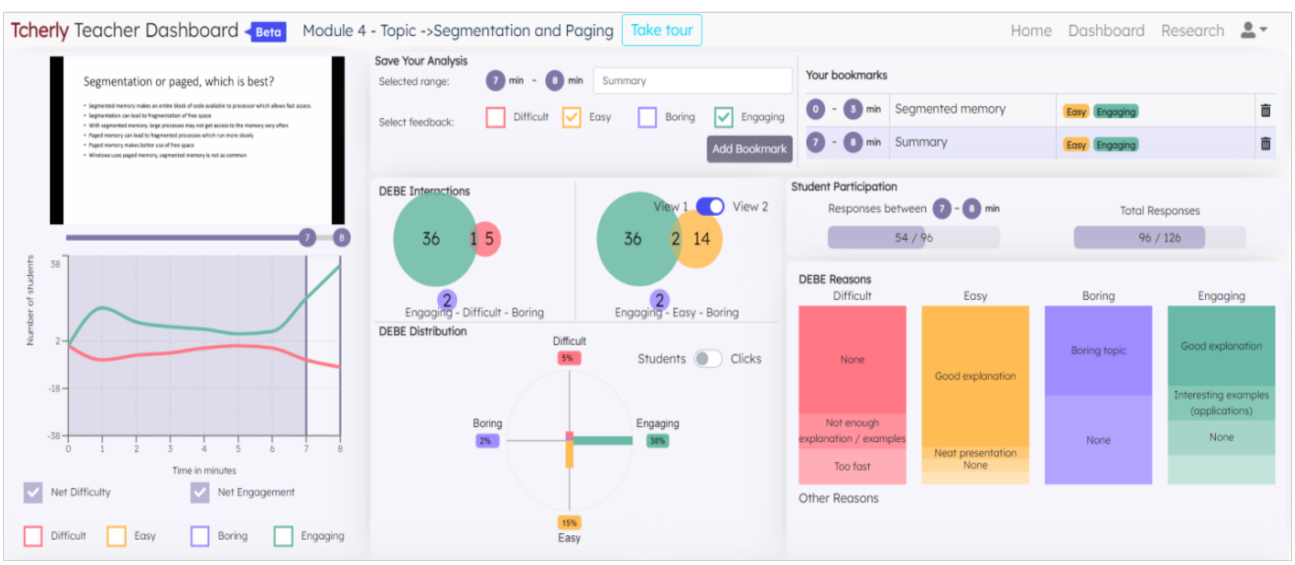


Figure 11. Screenshot of the dashboard from the Operating Systems course (Lecture: Segmented and Paged Memory)

Schematics and animations

In the first 2.5 minutes of the lecture "Segmented and Paged Memory" (Figure 12) from the Operating Systems course, the instructor explains the segmented memory management system using a schematic and animations, where a small peak of engagement appeared along with a small peak of easiness. The Student Participation section shows that approximately 59% of students (57/96) who provided feedback found this section easy and/or engaging. The DEBE Reasons section shows that "Good explanation" and "Presentation style" were cited by most students as the reasons for engagement and ease. Hence, this indicates that the schematics and animation supported information processing. More examples demonstrating how the dashboard could assist instructors in evaluating the elements of lecture materials are included in Appendix B.

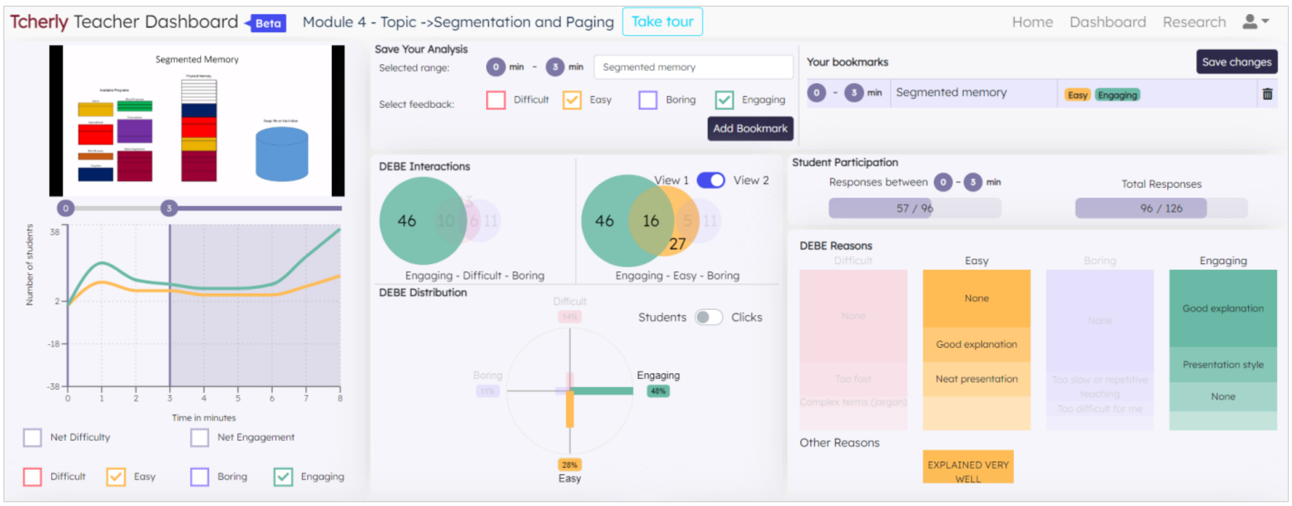


Figure 12. Screenshot of the dashboard from the Operating Systems course (Lecture: Segmented and Paged Memory)

5.1.2. Understanding Students' Learning Experiences

The research on event segmentation in cognitive science and cognitive neuroscience has demonstrated that a substantial level of agreement exists across individuals on the starts/ends of "meaningful events" when presented with a dynamic stimulus and asked to indicate such events. The CLAS system (Risko et al., 2012) extends event segmentation to the educational context, amalgamating individual event segmentation (important markers) to produce a consensus representation and providing it to students and instructors. The authors suggest that *"these 'importance markers' are akin to the points indicated in an event segmentation task"* (Newtson, 1973; Speer et al., 2003). Our system's philosophy aligns well with such an approach. Here, we demonstrate how the teacher dashboard could engage an instructor in meaning-making around a lecture by conducting a segment-by-segment analysis of aggregated meaningful events.

Cognitive-affective dynamics from line chart

Cognitive-affective dynamics play a vital role in student learning (Graesser & D'Mello, 2012; Tze et al., 2016), and understanding these dynamics could provide instructors valuable insights regarding student learning in relation to instructional materials. The video lecture "Demonstration of Java Programs" from the Java Programming course is 30 minutes long. It comprises four interlinked segments (Figure 13). The first segment (S1), between 0 and 6 minutes, discusses Java installation, the setting of environment variables, and how to edit, compile, and execute a Java program. The second segment (S2), between 6 and 16 min, demonstrates two simple Java programs (e.g., printing *"Hello, World!"*). In the third segment (S3), from 16 to 22 min, a program named *"TestArray"* (which includes a one-dimensional array) is demonstrated. This program is more complex than the previous two programs. The last segment (S4), between 22 and 30 min, demonstrates another program named *"a3DArray"* (including 3-dimensional arrays), which is even more complicated than the TestArray program. The difficulty level of the content increases with time, and the difficulty level of the different segments can be roughly ordered as follows: S4>S3>S2>S1. The difficulty plot on the line chart also mimics this progression. On the other hand, engagement remains almost steady until S2; however, it flips toward the middle of the lecture, around 16 minutes, in the transition from S2 to S3. After this flip, it never returns to the same level. Most importantly, this flip occurs around the time when difficulty breaches negative (easy) levels. Clearly, engagement is reduced with the onset of difficulty, and it flips between S2 and S3. However, between S3 and S4, engagement does not keep tanking and actually picks up again in S4, indicating that after the class reset its expectation between S1/S2 and S3 type of content, they re-engaged even when difficulty increased from S3 to S4. We hypothesized several such dynamics in Mitra and Chavan (2019). For example, had engagement dropped midway through S4 and not between S2 and S3, it could have indicated cognitive overload, where students started zoning out when the material became too difficult. If engagement had never peaked and dropped (or there was net boredom) from the start, then the rise in the difficulty level could have been attributed to the overall disengaged nature of the cohort and would have indicated the need for a complete overhaul of the lecture delivery and/or content.

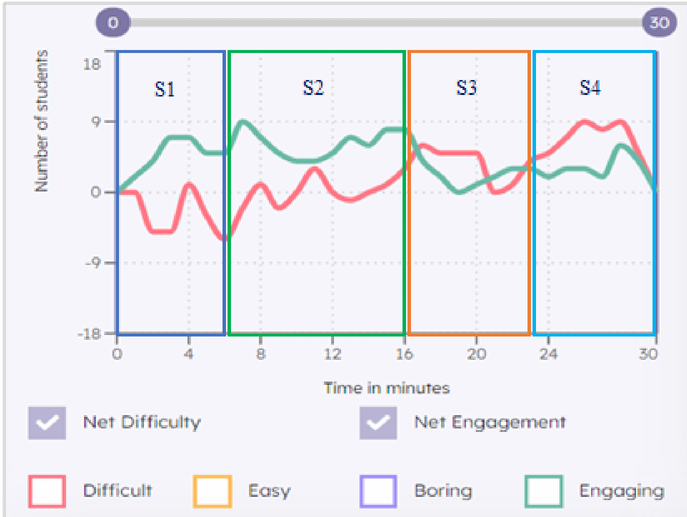


Figure 13. Cognitive-affective dynamics in the four interlinked lecture segments (Lecture: Demonstration of Java Programs)

Analysis of lecture segments with line chart and Venn diagrams

While the line chart was effective in suggesting the temporal interplay of cognition and affect, the addition of Venn diagrams could help in the optimization of lecture content. While S1 and S2 have similar engagement and low difficulty levels, as observable from the line chart, their fundamental difference is exhibited in the Venn diagram (Figures 14a and 14b). S1 being introductory in nature elicits students finding it either difficult, engaging, or boring (Figure 14a). There is less overlap, suggesting that students experienced this segment with either one of the cognitive-affective states. For S2, however, this clear demarcation is not evident, as the circles tend to overlap, suggesting that some students found the segment both difficult and engaging and less boring (Figure 14b). Thus, we observe that from S1 to S2, the difficulty increased (as evident from the line chart alone), and most students who found it difficult also found it engaging (as evident from the Venn diagram). Similar observations are possible if we analyze the complimentary Venn diagram of engaging-easy-boring. Between S3 and S4, we notice a distinct rise in difficulty. Additionally, notably, the two peaks are clearly due to the content difficulty of S3 and S4, as the 22nd minute distinctly separates the two peaks. Clearly, there is no reason for concern, as engagement seems not to have dropped further in S4, even with a further increase in difficulty. However, more importantly, the Venn diagram circles are still overlapping, indicating that students who found the content difficult also found it engaging (and boring to a lesser extent; Figure 14d). Had the circles shown separation as S1, it would have indicated a qualitative change in the learning experience of the class. If the Venn diagram had shown that the students found segment S4 either *difficult* (if difficult and boring did not overlap) or *difficult and boring* (if the two overlapped) and *those who found it engaging did not find it difficult,* then the interpretation would be different. In that case, the learning experience of the class in that segment, indicated by students finding the content either *difficult* or *difficult and boring*, would have been markedly different from the learning experience when a class found the content *difficult as well as engaging,* as it shows a class that is polarized into thinking of the content as either engaging or difficult. Such an outcome (the hypothesized outcome) is not desirable, but the observed outcome (Figure 14d) indicates an optimal learning experience.

5.2. Limitations

Although the present study has implications for instructors, LA researchers, and designers, it has certain limitations:

- This study was limited to a few video lectures and only examined instructors' perceptions of usability and usefulness. It provided no detailed insights into how instructors use the analytics, such as their interpretation of, pedagogical responses to, and actions based on analytics (Wise & Jung, 2019), or the effectiveness of the teacher dashboard to support awareness and reflection. Our future work will focus on understanding instructors' use of analytics and the actions that instructors take based on feedback.
- Instructors for the design prototyping but not the evaluation and implementation studies were all male, which is not uncommon in the country. Even then, some of the study results may not be adequately generalizable because of

inadequate gender representation. Furthermore, all participants were from an engineering background. Such samples may have certain characteristics, such as tolerance for low-usability designs, which may have influenced study results.

(a) (b) (c) (d)

Figure 14. Analysis of lecture segments using different features of the dashboard. In the Venn diagram, red = difficult, yellow = easy, purple = boring, green = engaging

6. Conclusion

While prior work in video learning analytics has explored the use of rich video interaction data to model learner engagement, investigate learning patterns, and measure their effects on learning achievement, the implications of such work for instructors employing blended or online learning in universities are minimal because much of these data are largely used for research, with very little focus on developing a suite of analytics that can be used by instructors. We characterized this gap into two subtypes and tried to address both in our research. First, the data informing such dashboards must be more indicative of students' reasons behind clicks (or actions), and second, the teacher-facing dashboard must be more responsive to teachers' visualization and interaction needs. We addressed the first gap through DEBE (pronounced as Debbie) methodology (section 3.2), which samples the experiences of students and their reasons for choosing those states throughout the duration of a lecture through a student feedback interface. Next, to address the second gap, we aggregated individual student data into a class-level time series of cognitive-affective states, which was then used to form the base-level information for the teacher-facing dashboard whose design underwent an iterative prototyping process with instructors and designers.

In this paper, we described the design and development of the teacher-facing dashboard followed by a preliminary evaluation of the high-fidelity dashboard prototype and its potential for instructors through a set of real-life examples of the dashboard use that instructors can use as guidelines. The design study (section 4) demonstrates how two different stakeholders, i.e., instructors and designers, can meaningfully participate in the prototyping phase to design visualizations, dashboard format (i.e., positioning of different elements), features, and supports that will assist instructors in using the analytics effectively. The instructors' perceptions of the usability and usefulness of the dashboard were positive. The students' perceptions of the usability of the feedback interface were also positive. Real data gathered and displayed with the Tcherly Dashboard offered potential ways instructors can use the data (section 5.1), including 1) evaluate the elements of the lecture materials (e.g., the effectiveness of asking a trigger question and the schematics and animations used), 2) understand cognitive-affective dynamics in the lecture (e.g., a concomitant decrease in engagement with an increase in difficulty in a specific segment of the lecture or sustained engagement in a lecture segment even with an increase in difficulty), and 3) analyze and compare student experiences across different lecture segments to identify optimal/suboptimal segments of the video lecture. Further investigation of instructors' use of analytics will improve the understanding of how instructors' activities and actions are informed by dashboard analytics.

Declaration of Conflicting Interest

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Acknowledgments

The authors would like to thank two anonymous reviewers and the editor whose comments greatly improved the manuscript. The authors would also like to thank students from IDP in Educational Technology and all participating teachers and designers for their help with this study. This work was supported by an IIT Bombay internal research grant, RD/0517-IRCCSH0-001, to Ritayan Mitra and a sponsored research grant, RD/0117-NEXT000-002, to the Educational Technology Programme at IIT Bombay.

References

Aldridge, J., Bell, L., & Fraser, B. (2010). Using students' perceptions of the learning environment to guide teacher action research to improve senior secondary classrooms. Paper presented at the AARE Annual Conference, Canberra 2009. Australian Association for Research in Education. https://www.aare.edu.au/data/publications/2009/ald091340.pdf

Aldridge, J. M., & Fraser, B. J. (2008). Outcomes-focused learning environments: Determinants and effects. Brill Sense. https://brill.com/view/title/37331

Atapattu, T., & Falkner, K. (2018). Impact of lecturer's discourse for students' video engagement: Video learning analytics case study of MOOCs. *Journal of Learning Analytics, 5*(3), 182–197. https://doi.org/10.18608/jla.2018.53.12

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human–Computer Interaction, 24*(6), 574–594. https://doi.org/10.1080/10447310802205776

Bargeron, D., Grudin, J., Gupta, A., Sanocki, E., Li, F., & Leetiernan, S. (2002). Asynchronous collaboration around multimedia applied to on-demand education. *Journal of Management Information Systems, 18*(4), 117–145. https://doi.org/10.1080/07421222.2002.11045706

Bell, L. M., & Aldridge, J. M. (2014). Investigating the use of student perception data for teacher reflection and classroom improvement. *Learning Environments Research, 17*(3), 371–388. https://doi.org/10.1007/s10984-014-9164-z

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society, 15*(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878

Brooke, J. (1996). SUS: A 'quick and dirty' usability scale. *Usability Evaluation in Industry, 189*(194), 4–7. https://doi.org/10.1201/9781498710411-35

Brooks, C., Amundson, K., & Greer, J. (2009). Detecting significant events in lecture video using supervised machine learning. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. Graesser (Eds.), *Frontiers in Artificial Intelligence and Applications* (pp. 483–490). Artificial Intelligence in Education, vol. 200. IOS Press. http://dx.doi.org/10.3233/978-1-60750-028-5-483

Brooks, C., Demmans Epp, C., Logan, G., & Greer, J. (2011). The who, what, when, and why of lecture capture. In P. Long, G. Siemens, G. Conole, & D. Gašević (Eds.), *Proceedings of the 1st International Conference on Learning Analytics and Knowledge: Connecting the technical, pedagogical, and social dimensions of learning analytics* (LAK '11), 27 February–1 March 2011, Banff, AB, Canada (pp. 86–92). ACM Press. https://doi.org/10.1145/2090116.2090128

Brooks, C., Greer, J., & Gutwin, C. (2014). The data-assisted approach to building intelligent technology-enhanced learning environments. In J. A. Larusson & B. White (Eds.), *Learning analytics: From research to practice* (pp. 123–156). Springer. https://doi.org/10.1007/978-1-4614-3305-7_7

Brooks, C., Thompson, C., & Greer, J. (2013). Visualizing lecture capture usage: A learning analytics case study. WAVe 2013 Workshop at LAK '13, 8 April 2013, Leuven, Belgium, (pp. 9–14). http://ceur-ws.org/Vol-983/paper3.pdf

Bustingorry, S. O. (2008). Towards teachers' professional autonomy through action research. *Educational Action Research, 16*(3), 407–420. https://doi.org/10.1080/09650790802260398

Chavan, P., Gupta, S., & Mitra, R. (2018). A novel feedback system for pedagogy refinement in large lecture classrooms. In J. C. Yang, M. Chang, L.-H. Wong, & M. T. Rodrigo (Eds.), *Proceedings of the 26th International Conference on Computers in Education* (ICCE 2018), 26–30 November 2018, Manila, Philippines (pp. 464–469). Asia-Pacific Society for Computers in Education. https://apsce.net/icce/icce2018/wp-content/uploads/2018/12/C4-14.pdf

Chavan, P., & Mitra, R. (2019). Developing a student feedback system using a design-based research approach. *Proceedings of the 10th International Conference on Technology for Education* (T4E), 9–11 December 2019, Goa, India (pp. 1–8). IEEE Computer Society. https://doi.org/10.1109/T4E.2019.00-59

Chen, Y., & Hoshower, L. B. (2003). Student evaluation of teaching effectiveness: An assessment of student perception and motivation. *Assessment & Evaluation in Higher Education, 28*(1), 71–88. https://doi.org/10.1080/02602930301683

Corrin, L., & De Barba, P. (2015). How do students interpret feedback delivered via dashboards? *Proceedings of the 5th International Conference on Learning Analytics and Knowledge* (LAK '15), 16–20 March 2015, Poughkeepsie, NY, USA (pp. 430–431). ACM Press. https://doi.org/10.1145/2723576.2723662

Cross, A., Bayyapunedi, M., Cutrell, E., Agarwal, A., & Thies, W. (2013). TypeRighting: Combining the benefits of handwriting and typeface in online educational videos. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '13), 27 April–2 May 2013, Paris, France (pp. 793–796). ACM Press. https://doi.org/10.1145/2470654.2470766

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience.* HarperPerennial.

de Barba, P. G., Kennedy, G. E., & Ainley, M. D. (2016). The role of students' motivation and participation in predicting performance in a MOOC. *Journal of Computer Assisted Learning, 32*(3), 218–231. https://doi.org/10.1111/jcal.12130

Demmans Epp, C., Perez, R., Phirangee, K., Hewitt, J., & Toope, K. (2019). User-centered dashboard design: Iterative design to support teacher informational needs in online learning contexts. Paper presented at the American Educational Research Association (AERA) Annual Meeting, Toronto, Canada, April, 5–9. https://doi.org/10.3102/1432967

Ellis, R. A., Han, F., & Pardo, A. (2017). Improving learning analytics: Combining observational and self-report data on student learning. *Journal of Educational Technology & Society, 20*(3), 158–169. https://www.jstor.org/stable/10.2307/26196127

Fiorini, S., Sewell, A., Bumbalough, M., Chauhan, P., Shepard, L., Rehrey, G., & Groth, D. (2018). An application of participatory action research in advising-focused learning analytics. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (LAK '18), 5–9 March 2018, Sydney, NSW, Australia (pp. 89–96). ACM Press. https://doi.org/10.1145/3170358.3170387

Fraser, B. J. (2012). Classroom learning environments: Retrospect, context and prospect. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 1191–1239). Springer. https://doi.org/10.1007/978-1-4020-9041-7_79

Gajos, K. Z., Kim, J., Li, S.-W., Cai, C. J., & Miller, R. C. (2014). Leveraging video interaction data and content analysis to improve video learning. In Proceedings of the CHI 2014 Learning Innovation at Scale workshop, 27 April 2014, Toronto, Canada. https://dash.harvard.edu/handle/1/22719144

Gašević, D., Mirriahi, N., Dawson, S., & Joksimović, S. (2017). Effects of instructional conditions and experience on the adoption of a learning tool. *Computers in Human Behavior, 67*, 207–220. https://doi.org/10.1016/j.chb.2016.10.026

Graesser, A. C., & D'Mello, S. (2012). Emotions during the learning of difficult material. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (pp. 183–225). https://doi.org/10.1016/B978-0-12-394293-7.00005-4

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102%2F003465430298487

Hoban, G., & Hastings, G. (2006). Developing different forms of student feedback to promote teacher reflection: A 10-year collaboration. *Teaching and Teacher Education, 22*(8), 1006–1019. http://dx.doi.org/10.1016/j.tate.2006.04.006

Holstein, K., Hong, G., Tegene, M., McLaren, B. M., & Aleven, V. (2018). The classroom as a dashboard: Co-designing wearable cognitive augmentation for K–12 teachers. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (LAK '18), 5–9 March 2018, Sydney, NSW, Australia (pp. 79–88). ACM Press. https://doi.org/10.1145/3170358.3170377

Jacobs, L. C. (1987). University faculty and students' opinions of student ratings. *Indiana Studies in Higher Education* No. 55.

Kay, R. H. (2012). Exploring the use of video podcasts in education: A comprehensive review of the literature. *Computers in Human Behavior, 28*(3), 820–831. https://doi.org/10.1016/j.chb.2012.01.011

Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014). Understanding in-video dropouts and interaction peaks in online lecture videos. *Proceedings of the 1st ACM Conference on Learning @ Scale* (L@S 2014), 4–5 March 2014, Atlanta, GA, USA (pp. 31–40). ACM Press. https://doi.org/10.1145/2556325.2566237

Kizilcec, R. F., Bailenson, J. N., & Gomez, C. J. (2015). The instructor's face in video instruction: Evidence from two large-scale field studies. *Journal of Educational Psychology, 107*(3), 724. https://psycnet.apa.org/doi/10.1037/edu0000013

Kizilcec, R. F., Papadopoulos, K., & Sritanyaratana, L. (2014). Showing face in video instruction: Effects on information retention, visual attention, and affect. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14), 26 April–1 May 2014, Toronto, ON, Canada (pp. 2095–2102). ACM Press. https://doi.org/10.1145/2556288.2557207

Larson, R., & Csikszentmihalyi, M. (2014). The experience sampling method. In *Flow and the foundations of positive psychology: The collected works of Mihaly Csikszentmihalyi* (pp. 21–34). Springer. https://doi.org/10.1007/978-94-017-9088-8_2

Li, N., Kidziński, Ł., Jermann, P., & Dillenbourg, P. (2015). MOOC video interaction patterns: What do they tell us? Proceedings of the 10th European Conference on Technology Enhanced Learning: Design for Teaching and Learning in a Networked World (EC-TEL 2015), 15–17 September 2015, Toledo, Spain (pp. 197–210). Lecture Notes in Computer Science, vol. 9307. Springer. https://doi.org/10.1007/978-3-319-24258-3_15

Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research, 11*(3), 253–388. https://doi.org/10.1016/0883-0355(87)90001-2

Marsh, H. W., & Roche, L. (1993). The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness. *American Educational Research Journal, 30*(1), 217–251. https://doi.org/10.2307/1163195

Michos, K., Lang, C., Hernández-Leo, D., & Price-Dennis, D. (2020). Involving teachers in learning analytics design: Lessons learned from two case studies. *Proceedings of the 10th International Conference on Learning Analytics and Knowledge* (LAK '20), 23–27 March 2020, Frankfurt, Germany (pp. 94–99). ACM Press. https://doi.org/10.1145/3375462.3375507

Mirriahi, N., & Vigentini, L. (2017). Analytics of learner video use. C. Lang, G. Siemens, A. Wise, & D. Gašević (Eds.). (2017). *The handbook of learning analytics*, 1st ed. (pp. 251–267). Society for Learning Analytics Research (SoLAR). https://doi.org/10.18608/hla17.022

Mitra, R., & Chavan, P. (2019). DEBE feedback for large lecture classroom analytics. *Proceedings of the 9th International Conference on Learning Analytics and Knowledge* (LAK '19), 4–8 March 2019, Tempe, AZ, USA (pp. 426–430). ACM Press. ACM. https://doi.org/10.1145/3303772.3303821

Montrieux, H., Vangestel, S., Raes, A., Matthys, P., & Schellens, T. (2015). Blending face-to-face higher education with web-based lectures: Comparing different didactical application scenarios. *Educational Technology and Society, 18*(1), 170–182. https://www.jstor.org/stable/jeductechsoci.18.1.170

Morris, L. V. (2009). Little lectures? *Innovative Higher Education, 34*(2), 67–68. https://doi.org/10.1007/s10755-009-9108-1

Namuddu, J., & Watts, P. N. (2020). Choice of device to view video lectures: An analysis of two independent cohorts of first-year university students. *Research in Learning Technology, 28*. https://doi.org/10.25304/rlt.v28.2324

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology*, 28(1), 28. https://doi.org/10.1037/h0035584

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pham, P., & Wang, J. (2015). AttentiveLearner: Improving mobile MOOC learning via implicit heart rate tracking. In C. Conati, N. Heffernan, A. Mitrovic, & M. Verdejo (Eds.), *Proceedings of the 17th International Conference on Artificial Intelligence in Education* (AIED '15), 22–26 June 2015, Madrid, Spain (pp. 367–376). Springer. https://doi.org/10.1007/978-3-319-19773-9_37

Pham, P., & Wang, J. (2017). AttentiveLearner^2: A multimodal approach for improving MOOC learning on mobile devices. In E. André, R. S. Baker, X. Hu, M. M. T. Rodrigo, & B. du Boulay (Eds.), *Proceedings of the 18th International Conference on Artificial Intelligence in Education* (AIED 2017), 28 June–1 July 2017, Wuhan, China (pp. 561–564). Springer. https://doi.org/10.1007/978-3-319-61425-0_64

Pham, P., & Wang, J. (2018a). Adaptive review for mobile MOOC learning via multimodal physiological signal sensing: A longitudinal study. *Proceedings of the 20th ACM International Conference on Multimodal Interaction* (ICMI '18) 14–18 October 2019, Bolder, CO, USA (pp. 63–72). ACM Press. https://doi.org/10.1145/3242969.3243002

Pham, P., & Wang, J. (2018b). Predicting learners' emotions in mobile MOOC learning via a multimodal intelligent tutor. In R. Nkambou, R. Azevedo, & J. Vassileva (Eds.), *Proceedings of the 14th International Conference on Intelligent Tutoring Systems* (ITS 2018), 11–15 June 2018, Montreal, QC, Canada (Lecture Notes in Computer Science, vol. 10858, pp. 150–159). Springer. https://doi.org/10.1007/978-3-319-91464-0_15

Preston, G., Phillips, R., Gosper, M., McNeill, M., Woo, K., & Green, D. (2010). Web-based lecture technologies: Highlighting the changing nature of teaching and learning. *Australasian Journal of Educational Technology, 26*(6). https://doi.org/10.14742/ajet.1038

Rhine, S. (1998). The role of research and teachers' knowledge base in professional development. *Educational Researcher, 27*(5), 27–31. https://doi.org/10.2307/1176739

Richardson, J. T. E. (2005). Instruments for obtaining student feedback: A review of the literature. *Assessment & Evaluation in Higher Education, 30*(4), 387–415. https://doi.org/10.1080/02602930500099193

Risko, E. F., Foulsham, T., Dawson, S., & Kingstone, A. (2012). The collaborative lecture annotation system (CLAS): A new TOOL for distributed learning. *IEEE Transactions on Learning Technologies, 6*(1), 4–13. https://doi.org/10.1109/TLT.2012.15

Rivera-Pelayo, V., Munk, J., Zacharias, V., & Braun, S. (2013). Live interest meter: Learning from quantified feedback in mass lectures. *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge* (LAK '13), 8–12 April 2013, Leuven, Belgium (pp. 23–27). ACM Press. https://doi.org/10.1145/2460296.2460302

Roche, L. A., & Marsh, H. W. (2002). Teaching self-concept in higher education. In N. Hativa & P. Goodyear (Eds.), *Teacher thinking, beliefs and knowledge in higher education* (pp. 179–218). Springer. https://doi.org/10.1007/978-94-010-0593-7_9

Santagata, R., Zannoni, C., & Stigler, J. W. (2007). The role of lesson analysis in pre-service teacher education: An empirical investigation of teacher learning from a virtual video-based field experience. *Journal of Mathematics Teacher Education, 10*(2), 123–140. https://doi.org/10.1007/s10857-007-9029-9

Scagnoli, N. I., Choo, J., & Tian, J. (2017). Students' insights on the use of video lectures in online classes. *British Journal of Educational Technology, 50*(1), 399–414. https://doi.org/10.1111/bjet.12572

Scheffel, M. (2017). The evaluation framework for learning analytics. Open Universiteit in the Netherlands, Welten Institute, Research Centre for Learning, Teaching and Technology. https://core.ac.uk/download/pdf/95759389.pdf

Scheffel, M., Drachsler, H., Stoyanov, S., & Specht, M. (2014). Quality indicators for learning analytics. *Journal of Educational Technology & Society, 17*(4), 117–132. https://www.jstor.org/stable/jeductechsoci.17.4.117

Schön, D. A. (1987). Educating the reflective practitioner: Toward a new design for teaching and learning in the professions. Jossey-Bass.

Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., Gillet, D., & Dillenbourg, P. (2016). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies, 10*(1), 30–41. https://doi.org/10.1109/TLT.2016.2599522

Sharma, K., Alavi, H. S., Jermann, P., & Dillenbourg, P. (2016). A gaze-based learning analytics model: In-video visual feedback to improve learner's attention in MOOCs. *Proceedings of the 6th International Conference on Learning Analytics and Knowledge* (LAK '16), 25–29 April 2016, Edinburgh, UK (pp. 417–421). ACM Press. https://doi.org/10.1145/2883851.2883902

Speer, N. K., Swallow, K. M., & Zacks, J. M. (2003). Activation of human motion processing areas during event perception. *Cognitive, Affective, & Behavioral Neuroscience, 3*(4), 335–345. https://doi.org/10.3758/CABN.3.4.335

Srivastava, N., Nawaz, S., Newn, J., Lodge, J., Velloso, E., M. Erfani, S., Gašević, D., & Bailey, J. (2021). Are you with me? Measurement of learners' video-watching attention with eye tracking. *Proceedings of the 11th International Conference on Learning Analytics and Knowledge* (LAK '21), 12–16 April 2021, Irvine, CA, USA (pp. 88–98). ACM Press. https://doi.org/10.1145/3448139.3448148

Srivastava, N., Velloso, E., Lodge, J. M., Erfani, S., & Bailey, J. (2019). Continuous evaluation of video lectures from real-time difficulty self-report. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (CHI '19), 4–9 May 2019, Glasgow, Scotland, UK (Paper No. 586). ACM Press. https://doi.org/10.1145/3290605.3300816

Stöhr, C., Demazière, C., & Adawi, T. (2020). The polarizing effect of the online flipped classroom. *Computers & Education, 147*, 103789. https://doi.org/10.1016/j.compedu.2019.103789

Tze, V. M., Daniels, L. M., & Klassen, R. M. (2016). Evaluating the relationship between boredom and academic outcomes: A meta-analysis. *Educational Psychology Review, 28*(1), 119–144. http://dx.doi.org/10.1007/s10648-015-9301-y

van Es, E. A., Stockero, S. L., Sherin, M. G., Van Zoest, L. R., & Dyer, E. (2015). Making the most of teacher self-captured video. *Mathematics Teacher Educator, 4*(1), 6–19. https://doi.org/10.5951/mathteaceduc.4.1.0006

Vatrapu, R. K. (2012). Towards semiology of teaching analytics. Workshop Towards Theory and Practice of Teaching Analytics at the European Conference on Technology Enhanced Learning (TaPTA-2012) 18 September 2012, Saarbrücken, Germany. CEUR Workshop Proceedings. http://ceur-ws.org/Vol-894/paper1.pdf

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186–204. http://dx.doi.org/10.1287/mnsc.46.2.186.11926

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist, 57*(10), 1500–1509. https://doi.org/10.1177%2F0002764213479363

Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: An overview and future research opportunities. *Personal and Ubiquitous Computing, 18*(6), 1499–1514. https://doi.org/10.1007/s00779-013-0751-2

Wise, A. F., & Jung, Y. (2019). Teaching with analytics: Towards a situated model of instructional decision-making. *Journal of Learning Analytics, 6*(2), 53–69. https://doi.org/10.18608/jla.2019.62.4

Xhakaj, F., Aleven, V., & McLaren, B. M. (2016). How teachers use data to help students learn: Contextual inquiry for the design of a dashboard. *Proceedings of the 11th European Conference on Technology Enhanced Learning* (EC-TEL 2016), 13–16 September 2016, Lyon, France (pp. 340–354). Lecture Notes in Computer Science, vol. 9891. Springer. https://doi.org/10.1007/978-3-319-45153-4_26

Xiao, X., Pham, P., & Wang, J. (2017). Dynamics of affective states during MOOC learning. In E. André, R. S. Baker, X. Hu, M. M. T. Rodrigo, & B. du Boulay (Eds.), *Proceedings of the 18th International Conference on Artificial Intelligence in Education* (AIED 2017), 28 June–1 July 2017, Wuhan, China (pp. 586–589). Springer. http://dx.doi.org/10.1007/978-3-319-61425-0_70

Xiao, X., & Wang, J. (2015). Towards attentive, bi-directional MOOC learning on mobile devices. *Proceedings of the 17th ACM International Conference on Multimodal Interaction* (ICMI '15), 9–13 November 2015, Seattle, WA, USA (pp. 163–170). ACM Press. https://doi.org/10.1145/2818346.2820754

Xiao, X., & Wang, J. (2016). Context and cognitive state triggered interventions for mobile MOOC learning. *Proceedings of the 18th ACM International Conference on Multimodal Interaction* (ICMI '16), 12–16 November 2016, Tokyo, Japan (pp. 378–385). ACM Press.. https://doi.org/10.1145/2993148.2993177

Yoon, M., Lee, J., & Jo, I.-H. (2021). Video learning analytics: Investigating behavioral patterns and learner clusters in video-based online learning. *The Internet and Higher Education, 50*, 100806. https://doi.org/10.1016/j.iheduc.2021.100806

Appendix A

Details of different sections and features of the dashboard: As shown in Figure A, the Tcherly Dashboard is composed of six panels with coordinated views: Overview (A), DEBE Distribution and DEBE Interaction (B), Student Participation (C), DEBE Reasons (D), Question Generator and Action Tracker (E), and Save Your Analysis (F).

A. Overview: After choosing a lecture from a course to analyze, instructors can obtain an overview of aggregate lecture feedback on a line chart (A1) illustrating student feedback variation over time, synced with the lecture video (A2). With sliders above the line chart (A1), the instructors can select a specific part of the lecture for analysis. This panel was regarded as an important entry point for analysis by the instructors during the design study. The line chart helps instructors identify areas of interest (e.g., the peak of difficulty or peak of engagement), and the interactive nature of the line chart allows instructors to select variables of their choice to plot on the chart (e.g., difficult vs. easy, difficult vs. boring). The lecture video synced with the feedback assists the instructor in revisiting specific parts of the lecture based on the feedback (the video plays based on the slider positions; e.g., if slider positions are 3 and 5 min, then a video between 3 and 5 min plays). A change in the slider position and the selection of a particular variable on the line chart (e.g., difficult) reflects changes in the following panels: DEBE Distribution and DEBE Interaction (B), Student Participation (C), and DEBE Reasons (D). For example, with slider positions of 3 and 5 min, the relevant data between 3 and 5 min are shown in those panels. If difficult is selected on the line chart, then the data related to the “difficult” variable are highlighted, fading out the data for other variables.

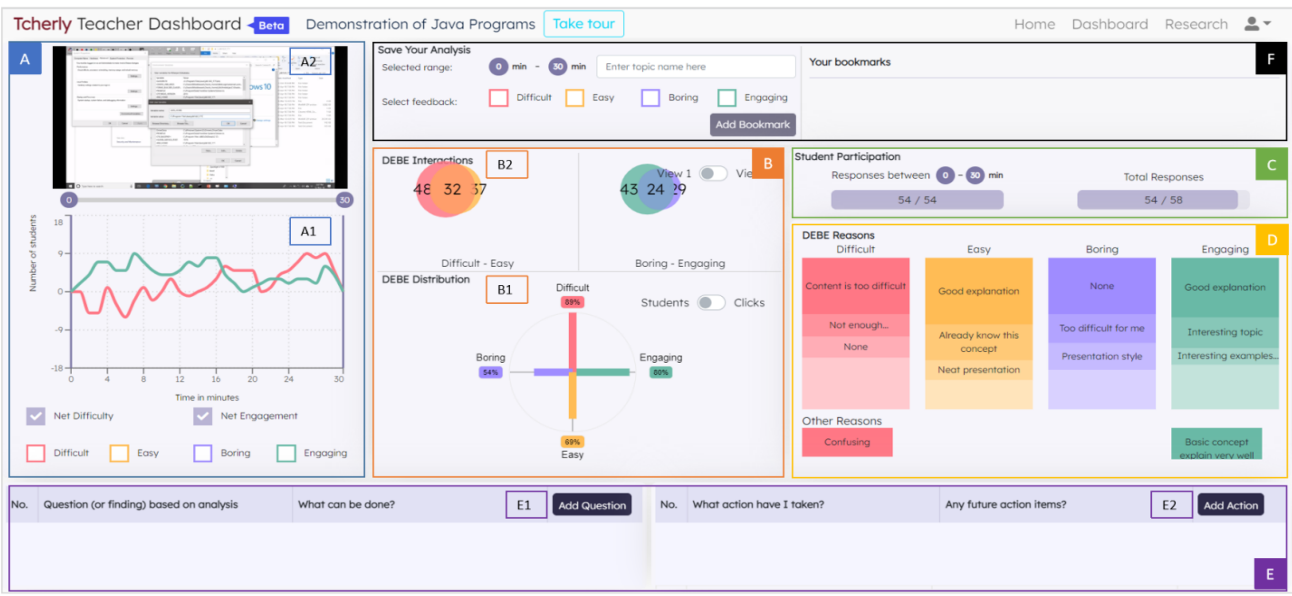


Figure A. Tcherly dashboard

B. DEBE Distribution and DEBE Interaction: The DEBE Distribution (B1) gives instructors a clear idea about the percentage of students giving particular feedback (e.g., difficult) in the selected part of the lecture. The radial column chart in this section provides the percentage distribution of student and click data (on toggle) for each feedback option. In the default setting, users see the percentage distribution of students for the four feedback types. Using the toggle option, the user can switch to the percentage click distribution. The DEBE Interaction (B2) allows instructors to understand student feedback behaviour in selected parts of the lecture. The Venn diagram helps the instructor know how the content in a specific segment of the lecture or across different segments is being received by students by comparing the degree of overlap between different variables. See section 5.1.2 for more details.

C. Student Participation: This section provides an indicator of student participation in the feedback-giving activity. This becomes important when the instructor wants to take actions (e.g., make changes in instruction or revise instructional materials) based on the student feedback but wants to know how many student responses will inform those changes. In an online context (e.g., flipped classroom), it also helps the instructor obtain information on how many students accessed the video lecture as an out-of-class activity. This section has two metrics: Total responses and Responses between x–y min. Total responses indicate the number of students who provided feedback in the lecture out of the total number of students who watched the lecture. Responses between x–y min indicate the number of students who provided feedback in a specific

ISSN 1929-7750 (online).

part of the lecture (x–y min) out of the total number of students who provided feedback in the entire lecture.

D. DEBE Reasons: This section (D) allows instructors to understand why students gave particular feedback. The heatmap shows the distribution of student reasons for specific feedback in the selected part of the lecture, along with a subsection "Other Reasons" reported by students apart from the options given on the feedback interface. *For example*, the heatmap of "Difficult" shows the percentage distribution of students' reasons for perceived difficulty in the lecture or a specific part of the lecture selected using sliders.

E. Question Generator and Action Tracker: Using Question Generator (E1), the instructor can record the important questions based on her analysis of feedback or annotate findings. Using Action Tracker (E2), the instructor can record the actions she took or future actions based on the questions or findings recorded in the Question Generator. For example, the instructor may pose the following questions to herself during analysis: According to students, what did not work in concept ABC/segment between x and y min in the lecture? In my view, what did not work? What should be done to address the problem in the face-to-face session or the next course offering?

a. Example: Question generator (E1)
   i. Questions or findings based on analysis — The slide design is problematic. Students felt that the graph had too much content (plots and equations).
   ii. What can be done? — Sequence the graph from simple to complex and revisit this part in the next class (or face-to-face lecture). Revise the design of the slides for the next course.

b. Example: Action tracker (E2)
   i. What action have I taken? — Revisited this part in the class (or during face-to-face lecture) with revised slides (used stepwise slides).
   ii. Any future action items? — Look for alternative simple figures/graphs and revise the slide design further.

F. Save your analysis: Save your Analysis (F) helps the instructor bookmark or record the specific analysis she has done (e.g., the instructor has analyzed a difficulty peak that appeared in a particular part of the lecture). "Add Bookmark" allows the instructor to select for which feedback the bookmark will be made and write the name of the related topic/subtopic/concept. Using "Your bookmarks," the instructor can navigate to the different analyses she has done on the lecture feedback. The bookmarks appear on the right side of this section. When a specific bookmark is clicked on, it gives a view of the detailed feedback analysis for the bookmarked section (DEBE Distribution and Interaction, DEBE Reasons, the actions you have taken, etc.).

Appendix B

Demonstration of practical applications/examples accompanied by a good explanation

In a video lecture on the *"Graph coloring problem"* in the Analysis of Algorithms course, the instructor explains the concept with brief examples for a significant part of the video. However, toward the end (between 13 and 16 min), the instructor gives a practical application — a concrete example — of graph coloring in newspaper printing, which students indicated as easy and/or engaging (Figure B1).

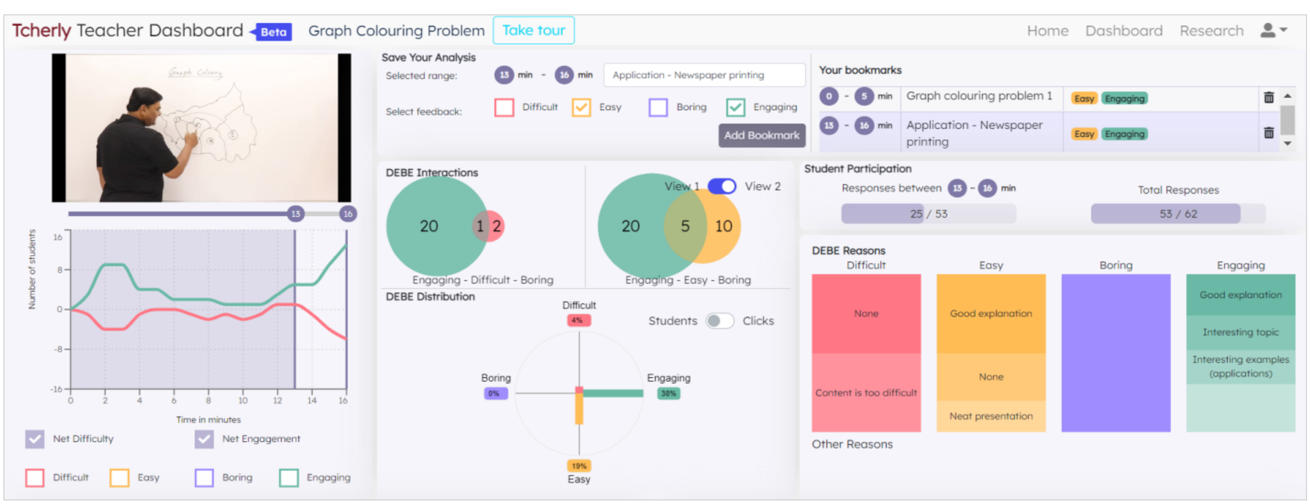


Figure B1. Screenshot of the dashboard from the Analysis of Algorithms course (Lecture: Graph Coloring Problem)

In another video on the "Insertion sort problem" from the Analysis of Algorithms course, to introduce and explain the concept of "insertion sort," the instructor started the lecture with a real-life example with the teaching assistant (TA) of the course sorting answer paper in descending order based on the marks (Figure B2). The feedback (easy and/or engaging) and the reasons cited by students (DEBE Reasons) indicate that this section worked well due to the practical example and the associated explanation by the instructor.

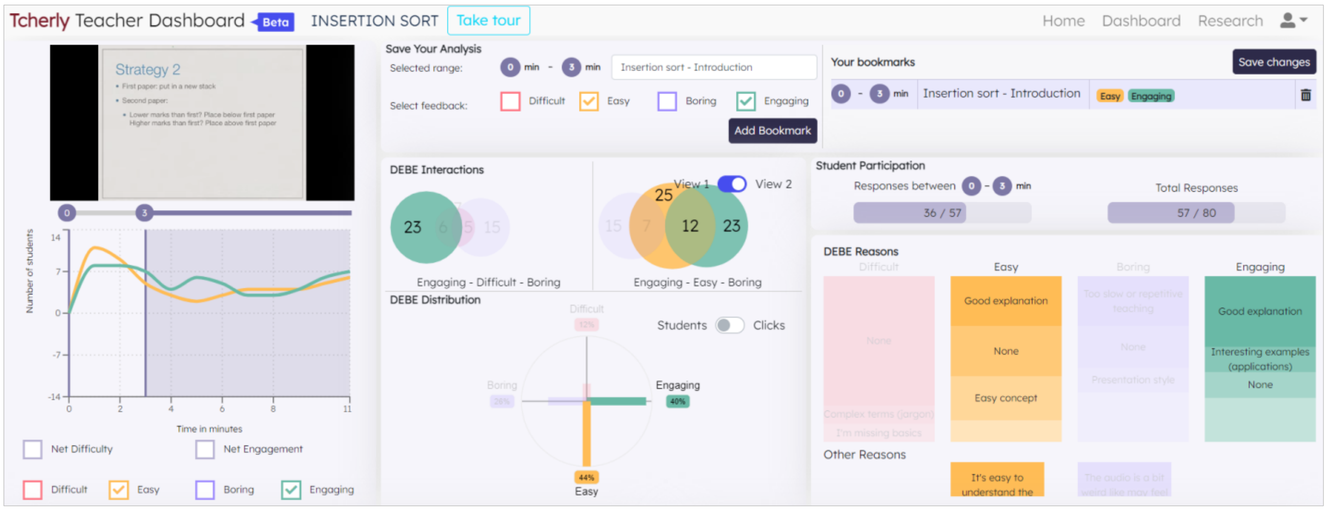


Figure B2. Screenshot of a dashboard from the Analysis of Algorithms course (Lecture: Insertion Sort)

Quality of audio

While watching a video lecture (a classroom lecture capture) presenting *"a Pentium 4 case study"* in the Microprocessor course, some students reported audio quality issues using the "*Other"* option on the student feedback interface. The student responses included the following: *Voice is too low, cannot hear properly, sound quality is low, voice is not clearly audible, voice is not clear*. This shows that the DEBE data and dashboard could also help identify audio and video quality issues in video lectures.

Summarizing the lecture using bullet points

Similar to the video lecture "Segmented and Paged Memory" discussed in section 5.1.1 (Figure 11), in a video lecture named "Programmable Interrupt Controller 8259A" of the Microprocessor course, toward the end of the lecture (between 11 and 14 min), the instructor first uses numbered bullet points to explain the operation of the 8259A and then to summarize the lecture by highlighting key points in the lecture (Figure B3).

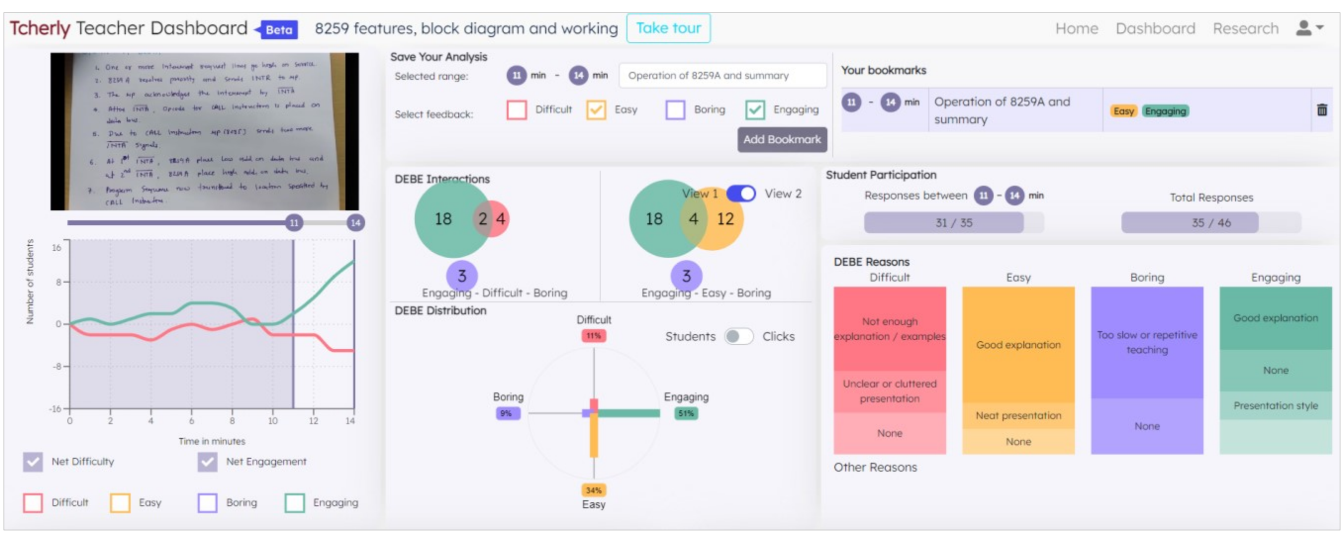


Figure B3. Screenshot of the dashboard from the Microprocessor course (Lecture: 8259A)