Measuring student behaviour dynamics in a large interactive classroom setting

Vasileios Triglianos⋆, Sambit Praharaj†, Cesare Pautasso⋆, Alessandro Bozzon†, Claudia Hauff†
⋆University of Lugano, Faculty of Informatics, Lugano, Switzerland
{name.surname}@usi.ch
†Delft University of Technology, Web Information Systems, Delft, the Netherlands
s.praharaj@student.tudelft.nl
{a.bozzon,c.hauff}@tudelft.nl

ABSTRACT

Digital devices (most often laptops and smartphones), though desired tools by students in a higher education classroom, have in the past been shown to serve more as distractors than supporters of learning. One of the reasons is the often undirected nature of the devices' usage. With our work we aim to turn students' digital devices into teaching and communication tools by seamlessly interleaving lecture material and complex questions in the students' browser through ASQ, a Web application for broadcasting and tracking interactive presentations. ASQ's fine-grained logging abilities allow us to track second by second to what extent students are engaging with ASQ which in turn enables insights into student behaviour dynamics. This setup enables us to conduct "in situ" experiments. Based on the logs collected in a longitudinal study over a ten week period across 14 lectures with more than 300 students, we investigate (i) to what extent ASQ can be reliably employed to assess attention and learning in the classroom, and (ii) whether different in-class question spacing strategies impact student learning and engagement.

ACM Reference format:
Vasileios Triglianos⋆, Sambit Praharaj†, Cesare Pautasso⋆, Alessandro Bozzon†, Claudia Hauff†. 2016. Measuring student behaviour dynamics in a large interactive classroom setting. In *Proceedings of ACM Conference, Washington, DC, USA, July 2017 (Conference'17),* 9 pages.
DOI: 10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In post-secondary classroom-based learning, teaching units of forty-five or ninety minute units are common. Past studies have shown that students' attention rates during such sessions vary significantly with regular episodes of inattention [4, 10, 14, 18]. A second complicating factor in the modern classroom is the distraction that modern technology — such as personal laptops, smartphones and tablets — affords to students [6, 8, 17, 21], a factor that has been shown to decrease several metrics associated with student learning including retention, attention and exam grades.

One potential solution to this issue is to ban laptops completely from the classroom or at least create "laptop-free zones" within large classroom areas [1]. Such a drastic measure though is not realistic to transfer to all higher education classrooms. Instead of resisting this new wave of classroom technologies, we aim to gain a deeper understanding into how it fits into students' learning habits.

In this work we report on a *case study* in a large higher education classroom that attempts to tackle student inattention and distractions through modern technology in tandem. One proactive approach to reducing inattention during such classes is active learning [5] — students not only passively listen to the lecturer, but also actively participate in the learning process. In order to facilitate active learning components in a classroom with hundreds of students, we adapted and deployed ASQ [19], an open-source Web-mediated teaching tool that incorporates interactive teaching elements (by offering a variety of practice questions) and keeps detailed logs of students' ASQ-related Web browser activities. We selected a 10-week undergraduate course (*Web and Database Technology*) for more than 300 Computer Science students as our target due to the diverse range of possible question types (from multiple-choice, text-based to programming and database queries) and conducted a longitudinal study of ASQ's usage.

Encouraging positive use of digital technologies in the classroom is not new. Previous work addressed the effect of personal response systems (colloquially known as "clickers") on student attention and engagement [7, 11]. Other studies have explored the impact of laptops in the classroom for note-taking [17, 21] or undirected use [8]. While providing insights on the relationship between digital technologies and students behaviour, many prior works suffer from one main technical limitation: they assume the students' devices to be complementary to, and not integrated with, the lecture experience. In our work, we take the next step and turn the students' devices from potential distractors into a teaching and communication tool by *seamlessly* interleaving lecture material and *complex* questions such as programming questions in each student's browser.

Previous work is also largely affected by a methodological limitation: experiments took place in a controlled setting, where the students' identities were known, and they were explicitly assigned to experimental conditions that could have harmed their learning experience.

These issues of privacy and fairness are not compatible with the requirements of a real-world course, where students must be guaranteed equal treatment, and privacy must be preserved.

These requirements were at the center of our use case, and defined our experimental methodology: participation was optional, students' identity concealed, and the learning set-up equal for all participants. Intuitively, changing the experimental conditions

Conference'17, Washington, DC, USA
2016. 978-x-xxxx-xxxx-x/YY/MM…$15.00
DOI: 10.1145/nnnnnnn.nnnnnnn

might make our experiments not comparable with previous work. To this end, we first focused on the following research question:

RQ1 *To what extent can a Web-based, privacy-preserving teaching tool be reliably used to assess the attention and learning outcomes of students?*

To answer this question, we analyse the ASQ user logs collected throughout our target course, and compare the obtained results with the findings collected from a systematic analysis of the literature. For instance, it is well-known that the activity of answering practice questions is benefiting students by increasing their engagement in the classroom [16]. Confirming these results, provides evidence of the suitability of ASQ as a platform for longitudinal "in-situ" experiments, thus enabling the investigation of additional research questions.

Next to exploring the impact of ASQ on several student behaviour metrics, we also investigated one dimension of practice questions that so far has not received a lot of attention: their spacing in time as they are interleaved with slides during the entire class duration.

RQ2 *Does the practice question strategy — questions are deployed uniformly across the entire lecture or with bursts of several questions at the same time – have any impact on student engagement and student learning?*

The remainder of the paper is organized as follows: §2 discusses related work; §3 introduces the ASQ platform, while §4 describes the experimental setting; §5 discusses the findings of our analysis and §6 summarizes the main lessons learned.

2 BACKGROUND

In this section we will first describe the impact of laptops and other digital devices (such as personal response systems) on the higher education classroom experience and then move on to discuss studies of question spacing in classrooms.

2.1 Digital technology in the classroom

Wood et al. [21] examined the impact of multi-tasking in lectures in particular through digital means: 145 undergraduate students took three classes of 20 minutes, each followed by a 15-item multiple choice quiz. Students were randomly assigned to one of seven experimental conditions. The learning performance (the percentage of correctly answered quiz questions) was found to be significantly and negatively impacted by multi-tasking (also confirmed in [15]), in particular when it involves highly-engaging social networks such as Facebook.

Closer to our own experimental setup, in one of the seven conditions the participants were free to use or not use their laptops (i.e. no multi-tasking was forced on them); it was found that those students opting not to use digital technologies achieved a higher learning performance than those that did. A similarly designed study and result was reported in [17], where not only the participants' multi-tasking was investigated but also the impact it had on participants in direct view of the multi-tasker — the distraction to the non-multi-tasking peers was significant and they in turn reached lower test scores than non-distracted peers. A much larger study by Aguilar-Roca et al. [1] across 800 students and 15 lectures did not find detrimental effects to students in the proximity of laptop users as measured in exam grades. As [17, 21], Fried [6] relied on an undergraduate psychology student class to reach the same conclusion through weekly self-reports and the students' test performance. In [8] a binary setup was employed: students were either allowed to use their laptop in class for any activity of their choice or disallowed to use their laptop at all. Students in the "open laptop" condition were found to remember less of the lecture content than students in the "closed laptop" condition. Finally, Ravizza et al. [13] performed the most natural experiment by routing students' Internet traffic in class through a proxy that logged all online activities during class time which was subsequently classified as either class-related or class-unrelated. They found (not surprisingly given past research) that class-unrelated Internet usage (e.g. the use of social networks or emailing) was common among students that chose to use their laptop in class and was negatively related to class performance. More surprisingly, class-related Internet usage did not benefit students, their class performance (usually measured through an after-lecture quiz or final course grades) did not increase over students that did not use their laptops in class.

Despite the varying experimental setups (cf. Table 1), the multitude of studies converge on the same conclusion: the use of digital devices in the classroom is not advantageous for students' learning performance due to students' multi-tasking behaviour and the available distractions. Nevertheless, students themselves perceive technology in the classroom as mostly useful instead of distracting [3, 9]. Since the complete ban of technology in the classroom is often not feasible (though less radical ideas such as laptop-free zones within a large classroom have shown promise [1]), we aim in our work to take advantage of technology to establish an additional communication channel between students and lecturers.

Previously, the use of personal response systems[1] ("clickers") has been explored as one potential positive use case for interactive technologies (besides laptops) in large classrooms. Mayer et. al al [11] found students engaged in in-class multiple-choice question answering through clickers to have higher learning gains than students engaged through the same questions without clickers. Notably, this latter group did also not fare any better than the control group of students who did not receive those in-class questions. Gauci et. al [7] made clicker usage in their classroom voluntary and found students who participated in answering in-class questions this way (in contrast to [11] it was not possible to answer questions without a clicker) to achieve higher exam results than those who did not. Importantly, low-performing students (those with low marks in a prerequisite course) were found to benefit more than mid- and high-achieving students. These results show that a guided and restricted usage of technology can benefit students' learning.

2.2 In-class questions

It is a well-established fact that active learning, in all its various forms, is effective in increasing students' learning performance compared to the traditional lecture setup where students are passive recipients of information [12]. In large classes, active learning is most often associated with questions that allow immediate automated feedback, such as multiple-choice questions or (relevant in our use case) programming questions evaluated through unit

[1] Such personal response systems can either be dedicated pieces of hardware or software installed on mobile phones and laptops.

testing. Recently, Weinstein et al. [20] explored the benefit of quiz spacing in the classroom: in a within-subject design (45 students) two spacing strategies — interspersed and at-the-end — were compared with each other. Similarly to our own study, the questions (five per lecture) in the interspersed condition followed directly the slides containing the necessary information, while in the at-the-end condition all questions were placed at the end of the lecture. The achieved quiz scores in the interspersed condition were found to be significantly higher than in the at-the-end condition. This difference in learning performance though vanished when the students were tested one last time nearly three weeks after the last lecture.

To summarize, previous works (often conducted in simulated classrooms with assigned conditions) have shown that undirected use of digital devices in the classroom leads to distractions and ultimately degrades the students' learning performance. At the same time, the directed use of technology has shown promise. With respect to in-class quizzes, there is little doubt in the literature that interactive classes improve students' learning performance, however, there is very little work discussing and exploring the benefits of question spacing. Our work adds additional knowledge to this issue and explores to what extent a platform such as ASQ can enable students laptop to function as directed devices in the classroom.

3 OVERVIEW OF THE ASQ PLATFORM

ASQ [19] is a Web-based tool that allows lecturers to deliver interactive lectures directly on students' Web browsers. ASQ has been designed and developed to increase the lecturer's awareness of the level of understanding in the classroom, and to turn student devices from potential distractors to a novel communication channel. Lecturers control the progression of slides from their own device, and changes to the current slide are automatically propagated to all connected student browsers.

In ASQ lecture slides are encoded in HTML5. Slides might contain interactive exercises of various types, including – but not limited to: multiple-choice, multiple answers, open questions, text highlight, SQL queries, and Javascript functions. Figures 1–2 show two examples of question types: *SQL*, which comprises a text editor for the writing of SQL queries performed on an in-browser database instance, and a results pane to visualise the query results; and *text highlighting*, where students are asked to highlight the type (synchronous and asynchronous) of Javascript methods.

The answers submitted by students are available to the lecturer for review and discussion in real-time.

3.1 Capturing students' interactions

Figure 3 provides a high-level overview of ASQ's distributed event collection and analysis architecture. The student's version of the presentation slides is a Web application which establishes a low latency bi-directional communication with the application server via the WebSockets protocol. Student interactions with the client-side Web application generate *events* that are captured in the Web browser, emitted to the server and stored in the event log database. ASQ does not require students to log in, and events are captured as soon as the browser connects to a running ASQ presentation session. Closing the browser tab that renders the ASQ presentation will disconnect the student. Each student is given a unique identification token for each presentation session and active connection. The token expires when the browser session expires, i.e. when the Web browser that established the initial connection to ASQ closes. This allows us to associate students with the events they generate, while preserving students' privacy across multiple lectures.

ASQ tracks different types of events (Table 2) that are generated by the browser during a presentation session; occurrences of such events are immediately sent to the ASQ server. Examples of events include: (i) connecting to the ASQ presentation; (ii) submitting an answer to a question; (iii) switching to another browser tab, or to another application, which may make the ASQ window invisible.

Figure 4 provides an example of the sequence of events emitted by three students during the first 15 minutes of a lecture. The first student has a low level of engagement as he immediately hides the ASQ window and then disconnects after 3 minutes without answering any questions. The second student shows a high level of engagement with the slides (which are never hidden after the initial connection) and also submits one answer to a question. The third student presents several context switches where ASQ is repeatedly hidden and shortly thereafter becomes visible again. This behavior continues also after the student submits an answer.

4 METHODOLOGY & USE CASE

In order to answer our research questions, we first need to define engagement metrics based on the events emitted by each student's browser (§4.1). In §4.2 we then introduce the question spacing strategies we explored in our use case, followed by a brief description of our annotation of question difficulty in §4.3. Finally, in §4.4 we introduce our target undergraduate course in more detail.

4.1 Modeling slide and question engagement

For each student, events are aggregated in order to compute *slide* and *question* engagement metrics. The former refers to engagement during the non-interactive parts of the presentation session, i.e. the content slides presented by the lecturer. The latter refers to engagement during the interactive (question-containing) slides of the session. We consider a student engaged with the slides if they are visible to the student in his or her browser. Conversely, we consider a student not to be engaged if the lecture slides are not visible to him or her (e.g. due to activities such as browsing the Web, emailing and so on).

We use the `tabvisible` and `tabhidden` events to detect whether the ASQ Web browser tab of the student is visible or not as well as the `answersubmit` event to detect if a student has submitted an answer to a question. We assume that a session s (full lecture from the time a presentation starts and ends in ASQ) of length $T(s)$ starts at second 1 and ends at second $T(s)$. For every student v, for every second t of session s we create an indicator variable $visible(v, s, t)$ which is 1 if the ASQ tab is visible and 0 otherwise. The mean slide engagement $MSE(v, s)$ of a student across s is the number of seconds the ASQ tab is visible, normalised by the session length:

$$MSE(v, s) = \frac{1}{T(s)} \sum_{t=1}^{T(s)} visible(v, s, t) \quad (1)$$

Table 1: Related work overview: column 2 reports the number of lectures and the length of each lecture (in minutes). Column "Simulated" indicates whether the experiment was conducted in a simulated classroom (✓) or a natural classroom (✗) setting. Students' behaviours can be determined as follows: assigned condition (behaviour determined by experimental condition assigned), self-reports (students report on their behaviour/distraction), human observers (observers sit behind students and observe them) or online activity logging (through a proxy server or ASQ in our work). The "Learning performance measurement" reports how students were evaluated on their learning performance.

	#Lectures (time)	#Students	Simulated	#Exp. conditions	Logging type(s)	Learning performance measurement	Class	Incentive
[21]	3 (20 minutes)	145	✓	7: Facebook – Texting – Natural Technology Use – Word Processing – pen and paper – MSN – email	Assigned condition	15 multiple-choice (MC) questions immediately after the lecture	Research methods, Statistics	$15 or course credit
[17]	1 (45 minutes)	44	✓	2: Multitasking – Non-multitasking	Assigned condition	20 MC questions immediately after the lecture	Introductory Psychology	Course credit
[6]	20 (75 minutes)	137	✗	2: Open laptop – Closed laptop	Weekly self-reports	4 exams and 10 homework assignments	General Psychology	None
[8]	1 (N/A)	44	✗	2: Open laptop – Closed laptop	Assigned condition and voluntary online activity logging	20 MC and open questions immediately after the lecture	Communications	None
[15]	1 (60 minutes)	64	✓	2: Open laptop – Closed laptop	Assigned condition	10 MC questions immediately after the lecture	N/A	$15 or course credit
[13]	15 (100 minutes)	84	✗	3: Class-related Internet usage – Nonacademic Internet usage – No Internet usage	Online activity logging	Final exam	Introductory Psychology	Course credit
[7]	36 (50 minutes)	175	✗	2: Personal response system (PRS) usage – No PRS	PRS logging	Midterm, final exam	Psychology: Control of Body Function	None
[1]	13 (50 minutes)	800	✗	2: Zoned laptop use – Uncontrolled laptop use	Assigned condition	Final exam	Bio 93: DNA to Organisms	None
Our work	14 (90 minutes)	89-319	✗	2: High engagement with ASQ – Low engagement with ASQ	ASQ activity logging	123 questions (MC, highlight, fill-in-the-blank) interspersed in lectures	Web & Database Technology	None

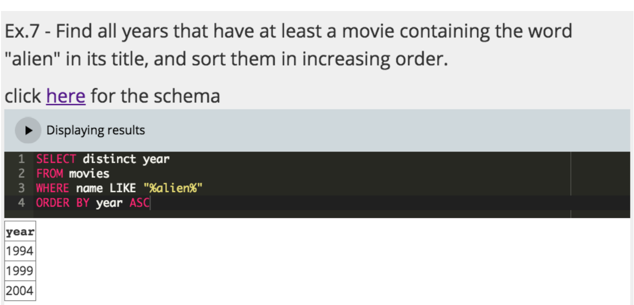

Figure 1: SQL question in ASQ.

Figure 2: Text-highlighting question

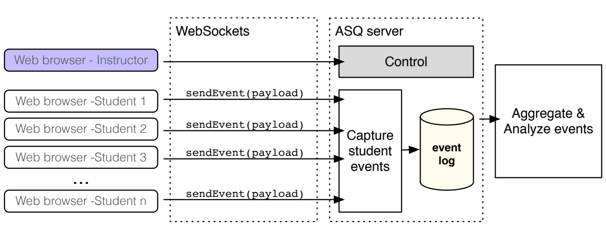

Figure 3: ASQ's architecture.

A student v's question engagement is exclusively defined over the interactive (question) slides in s; it is the fraction of questions $q \in Q(s)$ (with $Q(s)$ being the set of all questions in session s) that v submitted an answer for:

$$\mathrm{MQE}(v, s) = \frac{1}{|Q(s)|} \sum_{q \in Q(s)} submitted(v, q) \qquad (2)$$

Here, $submitted(v, q)$ is 1 if v submitted an answer to q and 0 otherwise.

4.2 In-class question strategies

As noted in §2 there is little research (apart from [20]) exploring the advantages or disadvantages of certain question spacing strategies in the classroom. To fill this gap and inspired by [20] the course

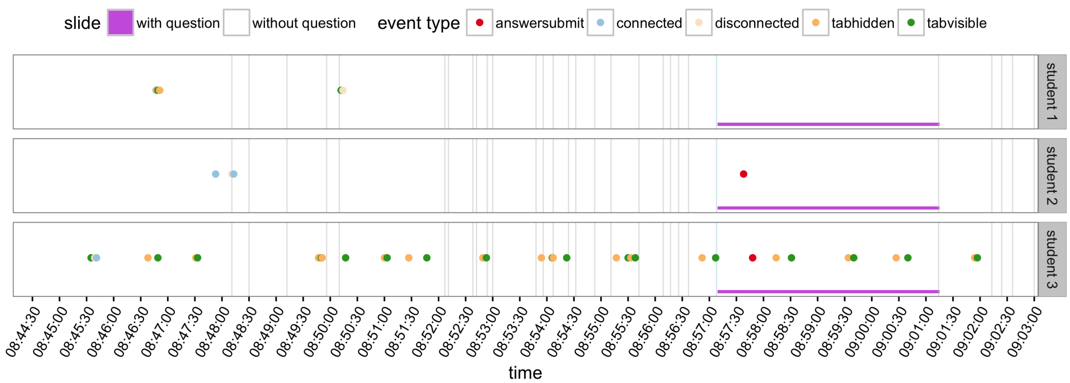

Figure 4: Sequence of events for three example viewers during the first 15 minutes of Lecture 2, *HTML*. Each vertical line represents a slide change — there are 29 slides shown in total, one of which is a question slide.

Table 2: Events monitored by ASQ relevant for this paper.

Event Name	Description
tabhidden	0 pixels from the browser tab that displays the ASQ web app becomes visible on the user's screen.
tabvisible	At least 1 pixel from the browser tab that displays the ASQ web app becomes visible on the user's screen.
answersubmit	A student submits an answer for an ASQ question (an exercise can have multiple questions).
connected	A student connects to the ASQ server.
disconnected	A student disconnects from the ASQ server.

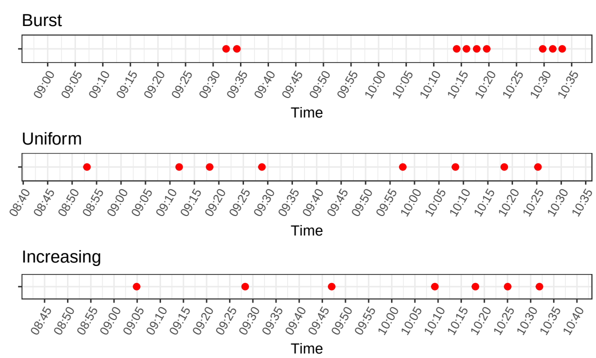

Figure 5: A temporal view of question strategies - each dot represents a question.

instructors initially designed the set of questions in each lecture according to one of the following three question strategies (with the strategies being randomly assigned to each lecture):

Burst (b): questions appear in bursts, each burst contains at least two questions; bursts can be randomly spread across a 90 minute lecture.

Uniform (u) questions are distributed uniformly in time across a 90 minute lecture;

Increasing (i) questions are placed with increasing frequency towards the end of each 45 minute period (two such periods exist per lecture with a 15 minute break in-between).

Figure 5 shows three example lectures (i.e. Cookies & session, SQL continued, Database introduction) and their question distribution across time.

After both class instructors had trialed all three question strategies in the first six lectures, it was decided to drop the increasing questions strategy from further consideration as it turned out to be very challenging to align the lecture material with this strategy. Subsequently, in the remaining 8 lectures only the burst and uniform question strategies were implemented.

4.3 Question difficulty

In our analysis, we also consider the difficulty of the questions posed to the students. All questions were created by the instructors of the course and manually annotated in accordance with the revised Bloom's taxonomy [2] — questions on the *remembering* and *understanding* level were considered to be *easy* and questions on the *applying* level were considered to be *difficult*. None of the questions belonged to higher Bloom levels.

4.4 Course overview

The data for this study was collected during the 2016/17 edition of a first year Bachelor course teaching Web technology and database concepts to Computer Science students at the Delft University of Technology. The course took place between November 14, 2016 and January 20, 2017. Table 3 presents the topic overview of the 14 lectures[2]; each lecture lasted 90 minutes, after 45 minutes a fifteen minute break occurred. The course topics were split across two instructors (I_1 and I_2), the in-class question strategy was randomly assigned. Overall, six lectures implemented the bursts

[2]In total, the course contained 15 lectures; one had to be excluded from our analysis due to a faulty logging mechanism.

question strategy and six lectures implemented the uniform question strategy, equally distributed across both instructors. Every lecture contained between seven and 13 question of various types: multiple-choice, multiple-answer, highlighting, SQL-query creation and fill-in-the-blank questions; in all but one lecture (CSS) the easy exercises outnumbered the difficult ones.

5 RESULTS

In this section we report on the analysis of the usage logs created by ASQ during the 10 weeks and 14 logged lectures of the course. First, in §5.1 we analyse the *slide* and *question* engagement of students across lectures, and compare the findings with previous work to assess the suitability of ASQ as a privacy-preserving platform for "in-situ" experiments. In §5.2 we study whether question strategies affect students' engagement with the lecture material.

5.1 RQ1: ASQ as an experimental platform

Table 3 reports the course attendance statistics, and data about the percentage of students using ASQ to visualise slides and to interact with questions. In each lecture, the students were counted by one of the authors ten minutes after the official start of the lecture.

Engagement across lectures. Over the course of the class, student attendance dropped from initially 319 to approximately 100 students — an attendance drop also seen in comparable courses at the same institution.

The usage of ASQ also fluctuated: while in the first three lectures more than 90% of the students used ASQ at least for some time during the lecture, novelty wore off and in later lectures the usage varied with between 65% and 80% of students connected to ASQ. Of those students that used ASQ, between 63% and 85% answered at least one question, while at most 27% of students submitted answers to all of a lecture's questions; in three lectures (all on database topics) fewer than 3% of students were such all-answer-submitting students. Across the 14 lectures we only observe 4 lectures (Conceptual design, CSS, HTML and HTTP) where at least half of the ASQ connected students submitted answers to more than half of the posed questions. These results are lower than expected, but not completely surprising. Not all students are equally compelled to experience lectures through their own devices (some students prefer pen and paper when being given the choice [1]), and their preference could vary according to topic. Also, questions varied in complexity; for topics like *SQL* (where SQL queries needed to be produced by students during the lecture), answering all questions was objectively harder than for lectures with mostly multiple-choice or multiple-answer questions. These results show that ASQ has been well-received by the course's students, and that a significant number of such students has been actively engaged with the lecture material and with questions.

Engagement within lectures. The fine-grained logging abilities of ASQ allow us to also investigate the students' engagement with ASQ beyond submitting responses to questions. Let us consider the students' *slide* and *question* engagement, as defined in §4.1.

The distribution of students engaged with slides for each lecture is plotted in Figure 6. Focusing on the first lecture (HTTP), we observe that half of the students are engaged with the slides

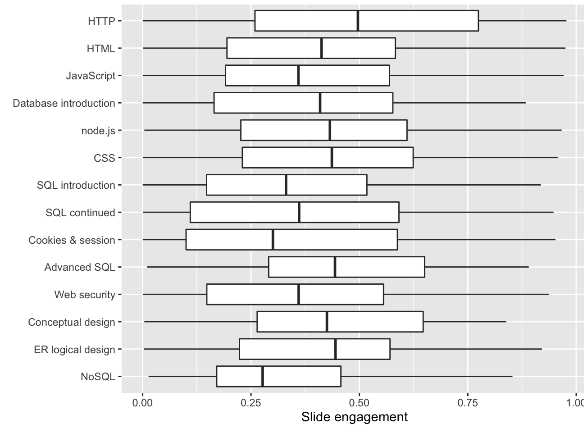

Figure 6: Distribution of slide engagement (computed per student) across all lectures. A value of 1 (0) indicates that ASQ was always (never) visible during class time.

between 25% and 76% of the lecture time. In subsequent lectures, the slide engagement drops, reaching its lowest median w.r.t. the median slide engagement in the final lecture (NoSQL) where half the students using ASQ have the non-question slides visible or in focus less than 30% of the time. Again, we confirm prior works, e.g. [8], that have shown learners to be distracted by the availability of digital devices (laptops in particular). Even though ASQ offers a directed, continuous and structured laptop use during the course of a lecture, students are still distracted (which we infer from the fact that the ASQ window is hidden most of the time).

For each lecture, we compute the Spearman rank correlation coefficient between students' question and slide engagement percentages. A correlation close to 1.0 would indicate that students with high slide engagement are also interacting with almost all questions. Table 3 shows the results of this analysis in the $\rho(S, Q)$ column. In the first lecture we observe a correlation of $\rho_{HTTP} = 0.72$ which indicates a strong relationship between slide and question engagement. In later lectures this relationship decreases in strength, reaching its lowest point in the final lecture with $\rho_{NoSQL} = 0.39$. This trend could be explained by the increasing complexity of the addressed topics, that might have discouraged students from participating in some or all of the question activities.

In contrast to previous studies, e.g. [6, 8, 21], we cannot compare our ASQ-using student population to the non-using population, as we only measure learning performance through questions posed within ASQ. Due to our privacy-aware setup, students attending lectures did not login to ASQ, nor did they provide their university IDs. As all question activities are formative assessments and students do not identify themselves when connecting to ASQ it is unclear why so many chose to not *attempt* to answer some of these questions. The lack of identification prohibits us to use the final exam score (as done in previous works) as learning performance measurement.

Although we *cannot* gather insights about students that chose to not consume lecture material through ASQ, we observed clear trends for students that did use the platform: slide and question engagement are strongly and positively correlated. We find this

Table 3: Lecture overview, reported in temporal order of delivery. Legenda. INS: instructor ($I1$ or $I2$). QS: the question strategy (b=burst, i=increasing and u=uniform). #EQ and #HQ: number of easy and hard questions in the slides. #STU: number of students in the classroom. %ASQ: the percentage of students in the classroom connected to ASQ. %SQ: the percentage of students connected to ASQ and answering at least one question (≥1 question submitted). %AQ: percentage of students connected to ASQ and submitting an answer to all questions. $\rho(S,Q)$: Spearman rank order correlation between *slide* and *question* engagement scores. %CA: percentage of correct answers. Correlations significant at the $p < 0.01$ level are marked with a †.

Lecture	INS	QS	#EQ	#HQ	#STU	%ASQ	%SQ	%AQ	$\rho(S,Q)$	%CA
HTTP	I1	b	6	4	319	99.7%	74.5%	26.1%	0.72†	62.16
HTML	I1	u	4	4	238	94.5%	79.1%	27.6%	0.58†	47.75
JavaScript	I1	i	5	5	192	91.1%	81.1%	8.0%	0.69†	48.14
Database introduction	I2	i	6	1	204	83.8%	78.9%	10.5%	0.60†	49.66
node.js	I1	b	5	3	208	64.9%	74.1%	15.6%	0.66†	65.53
CSS	I1	u	2	5	163	82.8%	79.3%	21.5%	0.72†	64.88
SQL introduction	I2	b	4	5	196	75.0%	79.6%	8.2%	0.57†	52.05
SQL continued	I2	u	6	2	157	87.9%	76.8%	2.9%	0.44†	54.41
Cookies & session	I1	b	5	4	133	97.0%	78.3%	11.6%	0.51†	49.76
Advanced SQL	I2	u	7	3	89	87.6%	62.8%	0.0%	0.64†	65.52
Web security	I1	u	4	4	151	70.9%	74.8%	14.0%	0.50†	48.46
Conceptual design	I2	b	11	2	83	65.1%	83.3%	13.0%	0.42†	60.26
ER logical design	I2	u	4	4	125	68.0%	84.7%	2.4%	0.40†	54.49
NoSQL	I2	b	8	0	89	76.4%	69.1%	11.8%	0.39†	61.61

Table 4: Overview of correlations between *slide* and *question* engagement, computed separately for each instructor, each question strategy and the four combinations of instructor and strategy. N is the number of items to correlate. %CA is the percentage of correct answers. Correlations significant at the $p < 0.01$ level are marked with a †

Condition	N	$\rho(S,Q)$	%CA
Question strat. *uniform*	768	0.50†	51.92
Question strat. *bursts*	851	0.61†	59.65
Instructor $I1$	1224	0.63†	55.47
Instructor $I2$	741	0.45†	55.02
$I1$ +*uniform*	467	0.58†	50.25
$I1$ + *bursts*	582	0.67†	60.85
$I2$ + *uniform*	301	0.39†	55.58
$I2$ + *bursts*	269	0.47†	56.75

alignment of results an indication of the suitability of a platform like ASQ for teaching and experimental purposes.

5.2 RQ2: Impact of question strategies

Are questions strategies influencing the engagement levels of the students? To address **RQ2**, we analyse slide and question engagement for lectures adopting *uniform* and *burst* question strategies. The results are reported in Table 4 (rows 2-3).

The correlation between slide and question engagement scores is higher for the *burst* than the *uniform* condition. This means that in the *uniform* condition with one question on the topic of the previous 5-10 slides, students are more likely to engage with the question than in the *burst* condition where at least some of the questions are likely to be about material more than 10 slides in the past. Drilling down further, we also consider in the analysis the role that the instructors, with their individual lecture style (and question styles) as well as addressed topics, might have played on the engagement level of students. Table 4 (rows 3-8) show that instructors clearly played a role. Despite the absolute differences in correlations, we observe that for both instructors the *burst* strategy leads to a higher correlation between slide and question engagement than the *uniform* strategy.

6 LESSONS LEARNED

This section summarises what the authors learned from the usage of ASQ in the classroom. The fourth and fifth authors are the instructors and designers of the *Web and Database Technology* course.

On the effectiveness of online questions to reduce distractions. The results presented in the previous sections are mixed: while initially many students adopt ASQ, over the course of the class the adoption rate decreases. Although ASQ offers a directed use of a digital device to follow and engage with a large classroom lecture, a significant fraction of students are engaged elsewhere through browser activities (as also indicated in prior works). A qualitative inspection of the collected statistics show that the engagement of students with ASQ varies across lectures, and also within the same lecture. Take for instance the graphs in Figure 7–9; they show for each lecture second the number of students connected to ASQ (those are all students with ASQ either visible or hidden) and the number of students having ASQ visible on their screen. `Javascript` (Figure 9) is an example of lecture where it is possible to observe a growing engagement trend, both before and after the lecture's break, that is correlated with the progression of questions. During the CSS lecture (Figure 8), questions provided only local increases in the number of engaged students, with no observable trend. Finally, lecture HTTP shows how the effect of questions changes during the

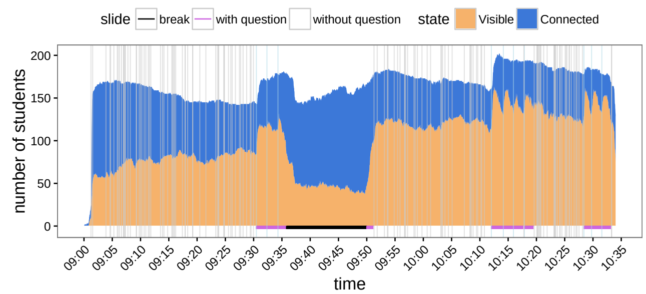

Figure 7: Temporal evolution of connected and slide engaged students during the *HTTP* lecture. Evidence of distracted behaviour in the second half of the lecture.

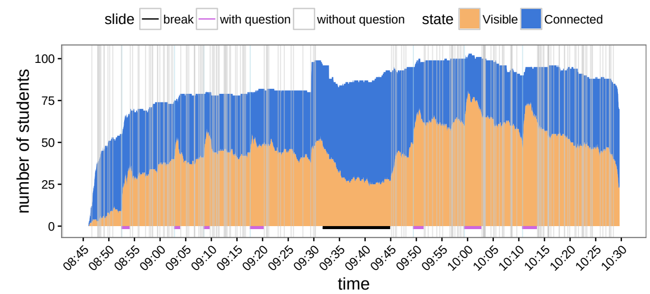

Figure 8: Temporal evolution of connected and slide engaged students during the *CSS* lecture. No clear slide engagement trend can be observed.

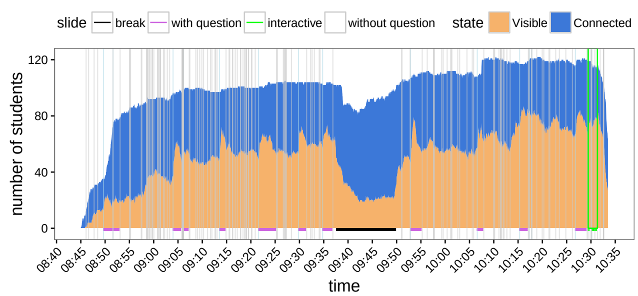

Figure 9: Temporal evolution of connected and slide engaged students during the *Javascript* lecture. Evidence of increasing slide engagement trend. Three interactive non-question slides are marked in green.

lecture. Before the break, the burst of questions helped in increasing students' engagement. After the break, each peak due to a new question is immediately followed by a sudden drop, which indicates that students were prone to distract themselves after answering the question.

Interactive slides beyond questions. In the JavaScript lecture (Figure 9), apart from the question slides a small number of additional slides contained interactive features to showcase JavaScript's abilities in the browser (such as a typing game and a text selection tool). However, when we explore whether those slides led to additional engagement (i.e. bringing students back from the other activities) we do not find this to be the case — for very few connected students the ASQ status changes from invisible to visible. This result is not supporting one of the authors' assumptions that more attractive/engaging ASQ slides lead to more direct student engagement.

Lecturer experience. The design of lecture material for a system like ASQ can be challenging for the instructor. In addition to obvious technical complexities (such as the creation of slides in HTML5), the design of a lecture experience enhanced with several, pertinent, and rich questions is not trivial. While this is an intrinsic difficulty in the design of lectures, we felt the problem to be exacerbated by the need for pre-defined question strategies, which sometimes "forced" the redistribution of content to account for the temporal allocation of questions.

7 CONCLUSIONS

Large student cohorts are becoming more common in post-secondary education. Bachelor-level courses are lectured in large classrooms that can easily include hundreds of students. Under such conditions, students are easily distracted — especially through the ubiquitous availability of digital devices.

To gain a better understanding on the *positive* role that technology can play in large higher-education classrooms, in this work we reported on a case study where a Web-based, privacy-preserving teaching tool (ASQ) has been put into use to support lectures with interactive teaching elements.

We conducted a longitudinal study over 10 weeks of lectures, operating under strict *privacy* and *fairness* requirements, both obvious ethical concerns in real-world courses. We analysed the resulting usage logs, and established the suitability of ASQ as platform for "in-situ" privacy-preserving research in education. We investigated how different practice question strategies affected students' engagement with lecture material. We show that question strategies can be of influence, but also observe the important role played by lecturers and lecture topics.

These results provide plenty of inspiration for future work. ASQ will allow us to easily perform new use cases in other education settings (e.g. different courses, or education level). We plan to investigate (i) how different instructors influence student learning and student engagement and what are the causes for the observed differences, (ii) how more complex question types (e.g. to implement an entire program) can be broken down and incorporated into the lecture, and (iii) how the real-time insights we have about students' engagement can be reflected back to them in the live classroom to raise awareness and self-reflection.

Acknowledgements. This research was partially supported by the Extension School of the Delft University of Technology.

REFERENCES

[1] Nancy M Aguilar-Roca, Adrienne E Williams, and Diane K O'Dowd. 2012. The impact of laptop-free zones on student performance and attitudes in large lectures. *Computers & Education* 59, 4 (2012), 1300–1308.

[2] Lorin W Anderson, David R Krathwohl, and Benjamin Samuel Bloom. 2001. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* Allyn & Bacon.

[3] William M Baker, Edward J Lusk, and Karyn L Neuhauser. 2012. On the use of cell phones and other electronic devices in the classroom: Evidence from a survey of faculty and students. *Journal of Education for Business* 87, 5 (2012), 275–289.

[4] Diane M Bunce, Elizabeth A Flens, and Kelly Y Neiles. 2010. How long can students pay attention in class? A study of student attention decline using clickers. *Journal of Chemical Education* 87, 12 (2010), 1438–1443.

[5] Scott Freeman, Sarah L Eddy, Miles McDonough, Michelle K Smith, Nnadozie Okoroafor, Hannah Jordt, and Mary Pat Wenderoth. 2014. Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences* 111, 23 (2014), 8410–8415.

[6] Carrie B Fried. 2008. In-class laptop use and its effects on student learning. *Computers & Education* 50, 3 (2008), 906–914.

[7] Sally A Gauci, Arianne M Dantas, David A Williams, and Robert E Kemm. 2009. Promoting student-centered active learning in lectures with a personal response system. *Advances in Physiology Education* 33, 1 (2009), 60–71.

[8] Helene Hembrooke and Geri Gay. 2003. The laptop and the lecture: The effects of multitasking in learning environments. *Journal of computing in higher education* 15, 1 (2003), 46–64.

[9] Robin H Kay and Sharon Lauricella. 2011. Exploring the benefits and challenges of using laptop computers in higher education classrooms: A formative analysis. *Canadian Journal of Learning and Technology/La revue canadienne de lfiapprentissage et de la technologie* 37, 1 (2011).

[10] Sophie I Lindquist and John P McLean. 2011. Daydreaming and its correlates in an educational environment. *Learning and Individual Differences* 21, 2 (2011), 158–167.

[11] Richard E Mayer, Andrew Stull, Krista DeLeeuw, Kevin Almeroth, Bruce Bimber, Dorothy Chun, Monica Bulger, Julie Campbell, Allan Knight, and Hangjin Zhang. 2009. Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary educational psychology* 34, 1 (2009), 51–57.

[12] Michael Prince. 2004. Does active learning work? A review of the research. *Journal of engineering education* 93, 3 (2004), 223–231.

[13] Susan M Ravizza, Mitchell G Uitvlugt, and Kimberly M Fenn. 2016. Logged In and Zoned Out: How Laptop Internet Use Relates to Classroom Learning. *Psychological Science* (2016), 0956797616677314.

[14] Evan F Risko, Nicola Anderson, Amara Sarwal, Megan Engelhardt, and Alan Kingstone. 2012. Everyday attention: variation in mind wandering and memory in a lecture. *Applied Cognitive Psychology* 26, 2 (2012), 234–242.

[15] Evan F Risko, Dawn Buchanan, Srdan Medimorec, and Alan Kingstone. 2013. Everyday attention: Mind wandering and computer use during lectures. *Computers & Education* 68 (2013), 275–283.

[16] Henry L Roediger and Andrew C Butler. 2011. The critical role of retrieval practice in long-term retention. *Trends in cognitive sciences* 15, 1 (2011), 20–27.

[17] Faria Sana, Tina Weston, and Nicholas J Cepeda. 2013. Laptop multitasking hinders classroom learning for both users and nearby peers. *Computers & Education* 62 (2013), 24–31.

[18] John Stuart and RJD Rutherford. 1978. Medical student concentration during lectures. *The lancet* 312, 8088 (1978), 514–516.

[19] Vasileios Triglianos and Cesare Pautasso. 2013. ASQ: Interactive Web Presentations for Hybrid MOOCs. In *Proc. of WWW.* 209–210. DOI:http://dx.doi.org/10.1145/2487788.2487894

[20] Yana Weinstein, Ludmila D Nunes, and Jeffrey D Karpicke. 2016. On the placement of practice questions during study. *Journal of Experimental Psychology: Applied* 22, 1 (2016), 72.

[21] Eileen Wood, Lucia Zivcakova, Petrice Gentile, Karin Archer, Domenica De Pasquale, and Amanda Nosko. 2012. Examining the impact of off-task multitasking with technology on real-time classroom learning. *Computers & Education* 58, 1 (2012), 365–374.