Increasing student engagement in higher education using a context-aware Q&A teaching framework

Jan Knobloch
Technical University of Munich
Munich, Germany
knobloch@in.tum.de

Jonas Kaltenbach
Technical University of Munich
Munich, Germany
kaltenbj@in.tum.de

Bernd Bruegge
Technical University of Munich
Munich, Germany
bruegge@in.tum.de

ABSTRACT

Modeling in Software Engineering is a complex task which includes interaction and discussion. Often multiple iterations are needed to end up with a satisfying design to handle given problems like complexity or change. This behavior can also be seen in education in software engineering, where learners tend to iterate over their models after having multiple discussions with their instructors or peers about possible solutions. This is an important part of the learning process, however, it is increasingly harder to implement for large scale on campus courses due to high student to lecturer ratios. One of the biggest problems is that lecturers are not able to allocate their time during class to support all students equally. This issue leads to new teaching methodologies and automated or semi-automated tools to support in-class interaction of students. However, these tools are lacking an automated mapping between questions asked and the teaching context provided. This context has to be inserted manually or is implicitly available during the lecture, however is lost after the lecture is over. In this paper we describe the adaption of a lecture-style instruction and the introduction of a context-aware Q&A teaching framework to increase student interaction by parallelizing it with content delivery. We achieve this while also lowering barriers for students to participate, even in multi-classroom setups. The stated approach also allows the creation of a knowledge repository which persists student interaction including its teaching context. This repository can be used by students to prepare for upcoming exams as well as by instructors to optimize their teaching content. Using our approach, we are able to show that there is an increase in student participation, leading to increased student examination performance for active students.

CCS CONCEPTS

• **Applied computing → Interactive learning environments**; Education;

KEYWORDS

question & answer system, context, teaching, education, interaction

ICSE-SEET'18, May 27-June 3, 2018, Gothenburg, Sweden

ACM ISBN 978-1-4503-5660-2/18/05.
https://doi.org/10.1145/3183377.3183389

ACM Reference Format:
Jan Knobloch, Jonas Kaltenbach, and Bernd Bruegge. 2018. Increasing student engagement in higher education using a context-aware Q&A teaching framework. In *ICSE-SEET'18: 40th International Conference on Software Engineering: Software Engineering Education and Training Track, May 27-June 3, 2018, Gothenburg, Sweden.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3183377.3183389

1 INTRODUCTION

Typically, models in software engineering are used to abstract and communicate with other developers [5]. To properly teach software engineering modeling concepts and its implications it is important that students actively [12] participate in exercises and follow in-class content while also being able to express their thoughts and ideas about certain concepts to strengthen their knowledge on the related material. In certain course environments, new teaching methodologies such as flipped classroom [22], team teaching [9] and others have been invented to tackle the problem of apathetic students during class, by increasing student engagement. Still, many of those teaching concepts do not scale for large class sizes with over 1000 students [16][14] or lack the needed tool support. This is especially problematic for Massive On Campus Courses (MOCCs) which we define as follows: A MOCC is a course held on-campus using multiple classrooms for a single lecture by applying live-streaming technologies to deliver teaching content to all participants at the same time but in different locations. In addition to the introduction of new teaching methodologies new teaching tools have been introduced in teaching environments focusing on larger class sizes. These tools can be categorized into two major groups. There are *e-learning platforms* like *Moodle* [24] or *Sakai* [28] which are mainly used for lecture content storage and lecture organization. Other platforms like *OnlineTED* [26], *Pingo* [27], and *Tweedback* [33] focusing on *in-class interaction*, mainly serving as audience response systems. The first systems of their kind where expensive and came with bundled hardware like clickers to achieve quiz show like feels and other functionality. Yet, setup cost and configuration efforts reduced its impact on teaching environments and lead to a redesign of many tools in the Bring Your Own Device (BYOD) [1] direction. However, two major aspects have been left aside while implementing and using the different teaching tools offered: teaching context and scalability.

Applying the stated interaction tools to MOCCs especially in regard to multi-room lecture setups introduces new challenges of student interaction. Students often have a hard time asking questions to their instructors while participating using a live-stream for content delivery. In this particular case, the appropriate teaching context is even more important to phrase and answer student questions in an accurate manner. In this paper we apply the term *context awareness* [29] to teaching as follows: *Context aware teaching information* includes all processes and information provided at a specific time and location based on goals of the instructor and the needs of the learner during a lecture or an exercise. Context aware teaching information is a discrete unit of knowledge including the mapping to its relevant lecture material. This unit can be in the form of clarifications of lecture content, references towards other existing teaching material or additions to teaching content.

This paper provides a context-aware Q&A teaching framework to scale beyond single class room setups and describes the needed adaptations to a traditional teaching methodology. Its contribution is *threefold*:

(1) First, it provides a framework to collect and supply the needed context for peers or teaching staff to answer questions accurately. This also allows the creation of a knowledge repository for students in the form of Q&A reports for each lecture, which contains the teaching context to keep questions traceable towards lecture material and therefore reusable over time. Furthermore, the teaching staff can use the system as an audience response system to query students about their state of knowledge on specific topics.
(2) Second, it enables instructors to graphically analyze interactions related to specific teaching content. The data collected by the system allows the statistical analysis of student interaction.
(3) Third, it provides reproducible modifications to an existing teaching methodology which can be applied to MOCCs as well as traditional lecture setups. Our research shows an increase of student participation as well as an improvement in examination results for students participating in class using the framework.

2 RELATED WORK

Our work is inspired by existing tools that are able to improve the classroom learning experience. It incorporates proven concepts while extending the functionality to integrate teaching material context. This section introduces each of the tools in more detail, highlighting pros and cons during their use in public universities and according to our personal experience.

2.1 OnlineTED

OnlineTED [26] is a web based audience response system for higher education that uses the students own internet-enabled devices to participate in in class quizzes. It's main objective is the assessment of students knowledge during a lecture and serve as a starting point for a discussion. The quizzes are supposed to increase the students' attention and promote interaction even with large classes. Early results show that web based audience response system seem to perform better than traditional systems using clickers but are highly dependent on a available Wifi or mobile data connection.[32] While it can't be used as question and answer system for students, having the ability to probe students knowledge during class provides instructors with a powerful tool to adjust lectures on the fly to improve the learning outcome.

2.2 Tweedback

Tweedback [33] is a system similar to OnlineTED to gather feedback during a lecture. It consists of three independent feedback channels: pre-built quizzes to be asked by the teacher, anonymous questions from students on a chat wall, and so called problem buttons to communicate concrete problematic situations such as "voice to quiet" or "please give an example". The results from these channels are shown to the instructor in a web interface and a summary of the lecture can be received via email after the lecture. The questions by students are asked anonymously and students can up-vote questions of interest as well as provide answers to them. When a question gets enough up-votes the instructor is notified and can decide to address the question. This approach has several problems. Questions that concern only a small number of students get very few up-votes and therefore are never addressed. The inherit problem to this approach is that question and answer get split up into two different communication channels using a software platform for asking questions and spoken word for answers provided by lecturers. This makes it difficult to persist the generated knowledge due to distribution of artifacts. Question, answer, and context are spread over different channels and technologies. This system claims to indicate critical situations to the professor, but fails at addressing individual needs of students as all data collected is focused on the crowd rather than the individual student. Furthermore, the idea of providing feedback to the instructor is a good approach but instructors need to divide their attention which can have a negative impact on the lecture quality. Another downside of Tweedback is that generated knowledge can't be made available for students after the lecture as knowledge repository because the data collected is only captured per lecture. This also prohibits complex data analysis for teaching staff without doing a large amount of additional work.

2.3 Slack

Early 2017, *Slack* [18] rolled out a feature called *Threads*. This feature allows users to reply to a message directly and have message and replies organized in a small sub-conversation. The motivation behind this feature was to group messages that talk about the same topic to avoid cluttering the channel with messages. Without threads, a *Slack* channel can quickly get very hard to follow because multiple sub-conversations overlap. The message that starts a thread can be seen as a sort of context that all replies relate to. The implementation of this feature shows that the developers of *Slack* have realized how important context can be for a textual conversation. The framework to be introduced will make use of this new feature by opening new threads for each question stated by students.

2.4 AMATI

AMATI (Another Massive Audience Teaching Instrument) [19] tried to unite many of the features that tools such as Tweedback and OnlineTED offer by implementing live quizzes, a question wall, and a mood chart. Live quizzes function very similar to other tools by serving multiple choice question to students and showing a graphical representation of results. The question wall allowed students to ask questions anonymously and had a up-vote functionality. Teaching assistants and tutors were able to answer questions while students could "like" answers to show they understood the answer. To add context to a question a hashtag based system was created. Students could apply a tag to their question or select from a predefined set of hashtags. This allowed grouping of questions with similar context but did not allow a direct connection to the teaching content delivered by the instructor. AMATI also does not have the ability to generate meaningful question reports for students but this was highlighted as future work in [19].

3 CHALLENGES IN HIGHER EDUCATION

While several studies [4][10][15] provide different results on the impact of high student numbers per class on student performance, statistics definitely show that student enrollment numbers are increasing year by year in the last decade. The federal statistics office in Germany published the following numbers in 2017[13]: 230,670 students have been enrolled in Germany in the winter-term 1998/1999, whereas there have been 432,589 enrolled students in Germany in the winter-term 2015/2016. This is a 87,53% increase over the last 17 years. This evidently leads to larger class sizes and in many cases also to higher student to lecturer ratios.

3.1 Synchronous in-class questions

This situation can have a severe impact on knowledge transfer in traditional lectures since asking questions is usually done by the raise of hand method. The lecturer then has the option to accept or dismiss the question and continues with the lecture content afterwards. Halting the lecture leads to lost lecture time which is considered to be the time that could be used to deliver new content to students but instead is used to elaborate on previous delivered content due to a question. The idle time for students not concerned with the question is the same as the lost lecture time but multiplied with the number of students not concerned with the question. The lost lecture time and student idle time can be modeled with mathematical formulas to show the influencing factors. The lost lecture time shall be called t_{lost} and the student idle time t_{idle}. Given the number of students n_s, numbers of questions asked n_q and average time to answer a question $\bar{t}_q$, the lost and idle times can be calculated with the following formulas.

$$t_{lost} = n_q \bar{t}_q \quad (1)$$

$$t_{idle} = (n_s - 1) n_q \bar{t}_q \quad (2)$$

Lets assume a class that allows all students to actively participate with a size of $n_s = 1000$ where only 5 percent of the in-class participants intend to ask a question $n_q = 50$ with an average answer time of one minute $\bar{t}_q = 1\ minute$. Following this assumption a lecturer would have to spend up to $t_{lost} = 50\ minutes$ of in-class lecture time to address each question accordingly. This calculation does not even consider the idle times of students who know the answer to the given question already. This highlights the need of new means of interaction which also scale with high in-class student numbers.

3.2 Classroom Setups

The issue highlighted in the previous section gets even worse considering the fact that due to high student numbers a single classroom is often not sufficient anymore to accommodate all registered students. Students have the option to participate watching a livestream of the lecture in a separate classrooms or use the possibility to stream lecture content from home, if available. Student participation however is impeded using separate rooms or video streaming since it is not supported by traditional means of interaction.

3.3 Barriers to Interaction

Due to the distinct nature of lecture participation highlighted above, new barriers of interaction can occur which lead to additional technological and methodological challenges for the teaching staff. These barriers can be categorized in physical, psychological, and technological barriers which will be covered in the following subsections.

3.3.1 Physical Barriers. For students not participating in the main hall, it is impossible to ask question via the traditional method of raise of hand. In addition to the physical barrier from the students perspective the teacher also has a hard time interacting with students at home or in different rooms. Therefore new means of interactions in both directions teacher-to-student and student-to teacher have to be found and established.

3.3.2 Psychological Barriers. In addition to physical barriers there are psychological barriers, as students are intimidated by phrasing questions to a large audience especially when they do not know how to phrase their question properly[2].

3.3.3 Technological Barriers. Physical and psychological barriers aside, new means of interaction between multiple lecture rooms have to be established. This also leads to the issue of context-awareness as students asking questions during class tend to refer to certain lecture material while phrasing their question. Given an example from a software foundation course, namely Introduction to Software Engineering (ITSE) a student phrased an in-class question in the following way:

> "Is nickname not an attribute of any class?"

The provided answer of the teaching staff was:

> "You can put it in league and have a hash map mapping nickname to player"

The question was probably about a model of a software that handles some sort of leagues and players but this question-answer tuple doesn't carry any value for students not participating in the same lecture hall as the context or reference towards the material discussed is missing. This relationship of questions and their context is especially important when there is a delay using a streaming service and students would like to phrase a question regarding lecture content which may have already has been passed in the main lecture hall.

4 TEACHING METHODOLOGIES IN HIGHER EDUCATION

Looking back at the different challenges in higher education, new student centered approaches have been defined to enhance student participation in classrooms. Many of different approaches which can be found nowadays are based on the concept of increasing interaction during class hours. However, the problem of having a single instructor for many students in MOCCs is still persistent.

4.1 Content Delivery vs. Deepening Understanding

Approaches like flipped classroom [22] amongst others tend to delegate content delivery out of the classroom and therefore allow more time for discussion and exercises inside the classroom. A very similar separation can be seen in traditional teaching setups, where the separation is being done by splitting content delivery during class, from exercises being held in tutorial sessions. This concept however is often realized by the use of student tutors, as instructors can not deal with the high amount of students on their own. Other approaches are combining exercises and content delivery inside the classroom [11] [6] to deepen the understanding of students.

4.2 Automated Corrections vs. Instructor Explanations

Online teaching approaches like Massive Open Online Courses (MOOCs) deal with this problem by offering automated corrections for exercises. However, automatic verification can be difficult to achieve if the existing solution space is large. For complex answers MOOCs often rely on peer reviews to achieve reasonable results as the automated correction and verification for complex questions especially in regards software engineering is not trivial. Furthermore, verification of correctness does not necessarily explain the reasoning behind a given correction. Often, peers are also not able to provide the additional knowledge needed for in-depth corrections and insights. Fred G. Martin describes his experiences towards personal communication using small discussions each week in a MOOC as follows [21]:

> "Most of my students got a lot out of the fall Stanford course - and our weekly discussion sections made a difference."

Especially in software engineering and modeling, learning is not only about correcting student solutions but actually to guide students, to achieve knowledge transfer. Yet, the traditional concept of knowledge transfer by asking questions during class is getting less and less popular due to the fact students prefer anonymity [7] in large classes and students often have a hard time to properly formulate their questions [2].

4.3 On Demand Delivery

Often, instructors have a hard time to split their attention between different students in a short amount of time. This is especially the case when instructors have to deal with heterogeneous student groups and questions are only of concern for small groups of the audience. Therefore having smaller classes in general or having more time to devote to each student can be seen as beneficial[3][25]. This leads back to subsection 3.1 as synchronous questions only allow to deal with a certain amount of questions per lecture unit and also stalling the content delivery thread. Therefore we propose a new teaching concept which we call a *Virtual One-to-One Teaching*, which increases the personal communication between single students and their instructors.

4.4 The Virtual One-to-One Teaching Concept

Before introducing our software framework, certain modifications of a traditional teaching methodology need to be applied in order to use the framework properly. The following subsections will cover each individual transformation step towards our *Virtual One-to-One Teaching* approach.

4.4.1 Question Moderator. Traditional lectures are a teacher-centered approach, focusing on the the dialogue between the lecturer and the class as a synchronous process. To circumvent the issues described in subsection 3.1 we propose a new moderator role to be included into the teaching setup. The moderator person will take the synchronous load away from the instructor and will focus on answering questions from students using the provided framework. This reduces idle times for participating students waiting for new content and allows the instructor to keep transferring new lecture content. This approach is inspired by the methods used in team teaching, however the second teacher, in our case the moderator, has to moderate all incoming questions, deliver answers to students and is also able to accept given answers from student peers to according questions. The moderator should not replace or interfere with the lecturer, yet provide a means for students to ask questions without stalling the lecture. A moderator can be any person which is qualified to answer student questions properly, by having the required knowledge of a certain subject. In our case study we used teaching assistants as well as tutors to provide answers to incoming questions. The role of the moderator can also be used to flag answers given by student peers as valid answers.

4.4.2 Additional content projector. Traditional lectures usually use a single projector setup to transfer lecture content. We propose the use of an additional content projector. This second projector takes the role of automatically providing moderated question-answer-context sets, the moment the a moderator decides to share them. This allows students to receive knowledge generated from questions by other peers, while keeping focused on the main lecture thread.

4.4.3 Review breaks. Lecture breaks should be used efficiently for reviews by the moderator and the lecturer. The lecturer should be provided with useful questions and answers to adjust the focus of lecture content, to ensure key aspects are properly transferred and understood by students.

4.4.4 Internet access. To ask questions, students should be able to access the internet to use the provided framework. Students have the minimal requirement to have a device with internet access at hand in order to access the framework in order to ask questions. If students are not participating in the lecture hall, they can also use their hand-held devices to view the information shown on the additional content projector as described in 4.4.2.

4.4.5 Optional: Lecture livestream. When dealing with more students than a single lecture hall can hold, other means of knowledge delivery are needed. We suggest the lecture should be made available online for enrolled students allowing them to participate from remote locations, compared to providing a live stream only towards additional lecture halls.

5 INTRODUCING CAQAS

First, this chapter introduces the software to support the previously described methodology in order to increase in-class interaction. The software bundle is called CAQAS (Context Aware Question & Answer System) and consists of seven components that are described in the following sections. Second, the process on how to ask questions and how the system handles the answering process will be described.

5.1 CAQAS Software Components

5.1.1 Context Aggregation. To accurately capture context during a presentation a small software called *PresentersClient* is deployed on the computer running the presentation in either Microsoft PowerPoint of Apple's Keynote. This software automatically detects the current active slide information. This includes *slide number* and *presentation name*. When connecting to the server screenshots of the individual slides are generated and transferred to the server. This allows tracking of the ongoing presentation.

5.1.2 Context Providing. A server is responsible for providing context information via a RESTful API to messaging client integrations. It serves as a central data hub for all other components and ensures data integrity. It handles data flow between the other components.

5.1.3 Chatroom integration. We used *Slack* as chat room to serve as a central point for question asking and answering. A slack bot integration connects to the chat rooms via a real time messaging API provided by *Slack* (*Slack RTM API*[30]). The bot integration receives all messages sent in the channels and automatically detects questions by looking for question marks in messages. This component runs on a server and therefore is available during and after lectures. It tracks all questions from start to finish to prevent missing questions. We make use of *Slacks* new feature called *threads* to keep the structure of question channels clear and allow for discussions to be grouped together. We encourage the use of these threads by immediately opening a new thread when a question is posted. An example of such a thread can be seen in Figure 1.

Figure 1: A question thread in Slack.

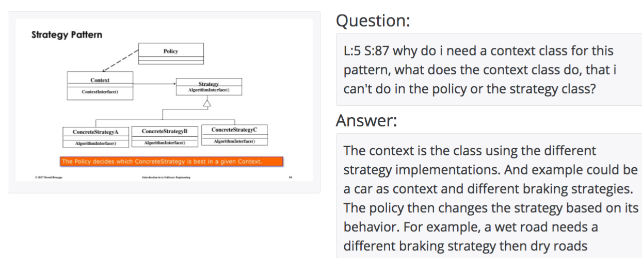

Figure 2: Question-Answer wall detailing a Question-Answer set including lecture context

5.1.4 Question Wall. During lectures answered questions are shown on a secondary projector in the lecture hall as soon as a satisfying answer has been reached. This allows students to follow up on questions while still focusing on the lecture content. The *Question Wall* updates itself automatically the moment the moderator decides to share the new question-answer-context set. It also provides options to filter stated questions by certain categories namely: general questions for lecture organization and content-related questions for in-class content. This gives the opportunity to have students focus only on content relevant information during lecture participation time. In addition to reviewing the question-answer-context sets provided, it is possible for students to review the full discussion of the answering process. This allows students to trace references and explanations if the pure answer is not sufficient for clarification. A screenshot of the question wall highlighting the lecture context, and the associated Q&A set can be seen in Figure 2.

Students can choose to open the question wall on their own devices, as it is publicly available. This allows students to set a personal filter for question-answer-context sets to be displayed, in addition students can use the question wall while not participating live in the lecture but via live stream.

5.1.5 Knowledge Repository. Since the *Question Wall* provides students only with the last 10 questions asked, new means of storing and providing a knowledge repository for studying purposes has to be introduced. The CAQAS framework allow students to access generated Question-Answer reports in the form of portable document format (PDF) files for each session. The Question-Answer sets provided contain their relevant teaching context and are sorted in the order questions occurred during class. In addition to the Question-Answer reports generated for content related questions, an additional report for general questions can be compiled containing organizational questions that have been asked during the course.

5.1.6 Teacher Analytics. In conjunction to the features mentioned above, CAQAS offers an analytics dashboard for teachers to highlight important statistics in various areas. The dashboard provides the teacher with comprehensive information such as tool participation rate, questions and answers given including their context or the average time to answer questions needed by the moderator. Figure 3 shows a plot of questions asked per slide and lecture, detailing a specific outlier in lecture 3, Slide 29 as there has been a disproportionately high number of questions asked for this single slide in particular. These insights can provide the lecturer with new means of lecture analytics and content revision. To provide a dashboard in CAQAS an open source framework named *Metabase* [23] has been used, to easily configure interesting data plots for different needs. Therefore it is possible to enhance and modify existing dashboards to retrieve even more detailed information if necessary.

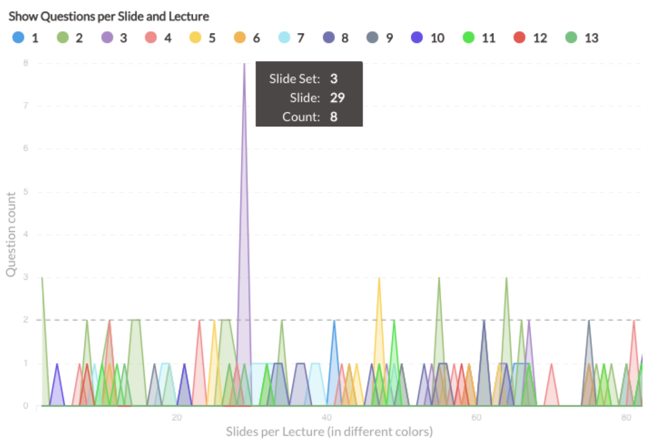

Figure 3: Number of Questions asked per Slide and Lecture

5.1.7 Teacher Audience Response System. CAQAS also provides the possibility to the teacher to query the student corpus with specific content-related questions to receive an impression of the knowledge state of the student audience. This can be done by sending a question via Slack in conjunction with providing different answer possibilities. Those answer possibilities can than be send to the *CAQAS bot integration*, which takes care of the summarization. Results will be presented on the *Question Wall* plotting the student votes graphically.

5.2 Question Asking Process

This section describes the process of asking and answering questions and explains the roles of the software components introduced in the previous section. When a student asks a question it is submitted to the corresponding *Slack* question channel. The *CAQAS bot integration* then automatically detects the question and asks the *CAQAS Server* for the current context. If a lecture is running, the bot responds to the question with a screenshot of the current slide. This opens a new thread in the *Slack* channel to signal to students that a question has been recognized by the system and guides the discussion. Students can override this behavior by supplying context information in the question itself or in a reply. For example, the text "s:32" will change the context to slide 32 of the running presentation.

If no presentation is currently running students have to supply the context for their question, if applicable. The *CAQAS bot integration* records all responses to a given question. The moderator can than decide to mark a given response as a correct answer. This triggers several actions. The *CAQAS bot integration* notifies the original author of the question that a final and moderated answer has been submitted by sending the question, answer, and context in a personal message. This allows students to keep focusing on the lecture as they don't have to monitor their question unless they want to participate in the discussion. Additionally, the question-answer-context sets are send to the *Question Wall* where they are displayed for all students. If a student is not satisfied with the answer provided, it can still be changed by a moderator and a new notification will be send when a reviewed answer is available.

6 EVALUATION

This section describes our evaluation applied during the ITSE'16 and ITSE'17 course. First, it describes the study design, followed by our descriptive analysis and quantitative analysis.

6.1 Study Design

In the previous iteration of the course, ITSE'16, the same chat room application *Slack* was used for student-teacher interaction as in ITSE'17 during and after the lecture. The addition in ITSE'17 was the integration of the context aware question and answer system described in section 5. This allows a comparison of collected data to evaluate the impact of context information on student interaction. In ITSE'16 all question and answer pairs were recorded and categorized into general and content related questions however the implicit lecture context was not saved. This can be directly compared to the data collected in ITSE'17 as shown in Table 1.

Additionally, students were given a survey to evaluate their experience with the framework. The survey was made available to students for a duration of 30 days at the end of the semester in 2017. 301 participants filled out the survey, out of which 224 were complete responses. The survey consisted of 29 questions in four categories. In the first question group "General" some basic data about the participant like field of study and current semester were gathered as well as information about participation and tool use. The group consists of seven questions, with two of them directly relating to the use of CAQAS. The results from these questions are used to filter the remaining groups and to only show relevant questions to the participants. A student who did not use CAQAS to ask questions should not have to evaluate the different aspects of the tool. The second group "Methodology" was also mandatory for all participants and is supposed to provide comparable data to evaluate the change in exercise methodology, compared to last years iteration of the course. The group consists of four questions total. Two dichotomous questions are used to evaluate the participants opinion of question asking process in CAQAS:

- Do you prefer questions being **asked** via Slack & CAQAS compared to **asking** the professors directly?
- Do you prefer questions being **answered** via Slack & CAQAS compared to the professor **answering** questions directly?

This section also contains an open ended feedback question to allow students to give general feedback regarding to tools and methodologies used in the ITSE lecture. In the third group "Tool Evaluation" the tools used in ITSE are evaluated in more detail, in terms of general impression, distraction, and ease of use. The group consists of 10 questions total. For CAQAS, the distraction levels by using a second projector for the Question-Answer wall are measured in two ordinal-polytomous questions. The questions are answered using a five point Likert scale.

- In the beginning of the semester: Did using multiple projectors showing different content distract you from following the lecture?
- After a few lectures using multiple projectors: Did using multiple projectors showing different content distract you from following the lecture?

The fourth section "CAQAS specific" consists of seven mixed type questions. This group is only presented to participants who have used CAQAS at least once. This group focuses on the individual features of CAQAS.

6.2 Descriptive Analysis

6.2.1 Data Description. In ITSE'16 students asked a total of 67 questions. None of these questions had a reference to context and consisted only of a question and answer pair. The categorization of questions can be seen in Table 1. With the introduction of context the number of questions asked increased by 431% to a total of 356 questions. Student numbers increased by 25% from 1142 to 1432, however this cannot account for such an increase in the total number of questions asked. The categorization of questions in ITSE'17 can be seen in Table 1. Furthermore, only 6 out of 356 questions were asked anonymously. Unsurprisingly, the majority of general questions (62%) were asked in the first week of the lecture. This might be due to the semi-anonymous nature of the chat room we used. Students needed to sign up with their university identifier but were allowed to choose a username themselves. This allowed us to match our collected data to exam grades while giving students the freedom to choose their own user names. The anonymous question feature allowed students to hide their chosen username when asking questions.

6.2.2 Survey results. Interestingly, more people attended the lecture via live stream(48%) than participating on campus in person (40%). Attendance via lecture recording was the least popular option (12%). This may indicate that the location aspect has a bigger impact on attendance than time constraints. The percentage of participants who used CAQAS is 31.75% according to our survey. This is significantly higher than the percentage reported by our own collected data, which was 17.1%. This disparity is due to the fact that not all students who signed up for the course actually participated. Since there is no way for us to exactly determine the number of participating students the percentage of students using CAQAS is believed to be close to the result of the survey.

Question Type		# Questions ITSE'16	# Questions ITSE'17
General		14	127
Content	Context	0	166
	No Context	53	63
Total		67	356

Table 1: Number of Questions asked in ITSE'16 and ITSE'17 by category

Category	Responses
Appreciation	19
Improvement suggestion	5
Tutorial request	5
Bug report	3
Overwhelmed by tools	2
Communication too Cluttered	2
Felt distracted	2
Tools are impractical	2
Dislike of the tools	1
Not satisfied with quality	1

Table 2: Classification of open question results from the survey.

A overwhelming majority (68.54%) of students reported that they liked the fact, that questions were asked via *Slack* and not directly to professor. A smaller majority (55.61%) reported that they liked questions being answered via *Slack* more. 64.28% of participants reported that they asked more questions via *Slack* than they would have by raise of hand.

The open feedback question received 69 results with 42 of them concerning CAQAS. The responses were classified into categories to analyze trends. The classification resulted in 10 categories. Each response was matched to exactly one category. Responses that would fit into multiple categories where split up into several unique responses. The results can be seen in Table 2.

In terms of distraction students showed that the use multiple projectors takes time to get used to but does not impact their ability to follow the lecture. Table 3 shows the results for the survey regarding the use of multiple projectors.

6.3 Quantitative Analysis

First, the number of questions asked by each student was mapped to the number of points scored in the final exam. To have consistent data the students who chose not to attend the exam were removed from the dataset as their score of zero points is not representative of their knowledge state. Furthermore, students who failed the exam intentionally to retry at a later point in their study were excluded as well. This was done by removing all entries with an exam score of less than 20 points. The 20 point threshold was determined by manually examining the dataset and exams. The points scored in

Distraction	First lecture	During the semester
Very Much	4.18	2.09
Much	5.86	4.60
Moderately	17.99	10.88
Slightly	24.69	17.57
Not at all	46.86	64.44
No Answer	0.42	0.42

Table 3: Distraction by the use of multiple projectors. All numbers in %

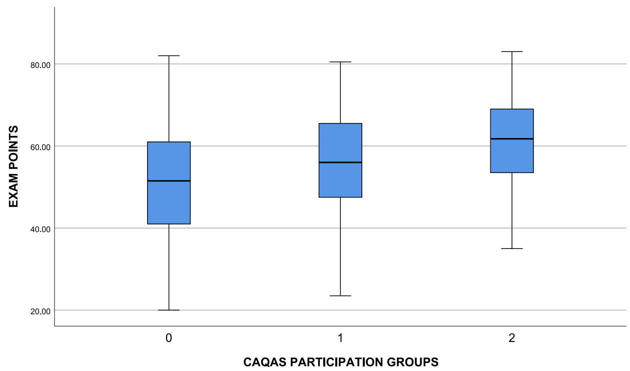

Figure 4: Boxplot of scores achieved by groups

the final exam were chosen as the dependent variable and were coded as a scale variable ranging from 0 to 90 points. The number of questions asked were mapped to the independent variable.

To perform a one-way analysis of variance test (ANOVA) [31] the data was classified into three student groups:

(0) Group 0 - No questions asked via CAQAS
(1) Group 1 - A single question asked via CAQAS
(2) Group 2 - More than one question asked via CAQAS

To ensure validity of the ANOVA result, the data was inspected beforehand. First, the groups were checked for possible outliers. The analysis of the generated box plot and descriptive statistics of the individual groups showed no outliers as seen in Figure 4. Second, the groups were checked for homogeneity of variances using the Levene's test provided by IBM SPSS Statistics Software [17]. The Levene's test showed no violation of the homogeneity of variances, therefore a one-way ANOVA was executed.

The ANOVA showed a statistically significant difference between the three groups described above ($F(2, 827) = 15.789, p \leq 0.001$). A Tukey post hoc test revealed that students who asked more than one question ($60.750 \pm 11.988, p \leq 0.001$) and students who asked exactly one question ($55.823 \pm 12.462, p = 0.012$) scored significantly higher in exam than students who asked no questions at all (50.858 ± 13.600). There was no statistically significant difference between the students who asked one question and students who asked two or more questions ($p = 0.125$). The results of the ANOVA are shown in Table 4 and the results of the post hoc Tukey test are shown in Table 5. To allow better interpretation of the results the effect size was calculated using the η^2 method shown in Equation 3 [20].

$$\eta^2 = \frac{SS_{between}}{SS_{total}} \tag{3}$$

The effect size for the ANOVA resulted in $\eta^2 = 0.037$. Using Cohen's scale [8] this can be interpreted as a small effect of questions asked on the points scored in the final exam.

ANOVA
POINTS

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5690,739	2	2845,370	15,789	0,000
Within Groups	149034,925	827	180,212		
Total	154725,664	829			

Table 4: Results of the ANOVA test

Multiple Comparisons
Dependent Variable: POINTS
Tukey HSD

(I) GROUP2		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
0	1	-4.96552*	1,73912	0,012	-9,0489	-0,8822
	2	-9.89245*	1,96374	0,000	-14,5032	-5,2817
1	0	4.96552*	1,73912	0,012	0,8822	9,0489
	2	-4,92692	2,52522	0,125	-10,8560	1,0021
2	0	9.89245*	1,96374	0,000	5,2817	14,5032
	1	4,92692	2,52522	0,125	-1,0021	10,8560

*. The mean difference is significant at the 0.05 level.

Table 5: Results of the post hoc Tukey test

7 THREATS TO VALIDITY

It was not possible for the survey to check participants status of enrollment in the course, therefore it is possible that people not taking the course took part in the survey. While we do not believe this was the case it is important to mention it for future research. Additionally, the survey software used has no way of prohibiting multiple entries in anonymous surveys. We used cookies to stop students filling out the survey multiple times but clearing the browser cache or a different device could easily circumvent this measure. Furthermore, due to time constraints, the survey had to be executed before students took the final exam. It is unclear if this had an impact on the result but should be considered when trying to replicate the results.

8 CONCLUSION & FUTURE WORK

We were able to show that student interaction in question and answer systems can be increased by providing the teaching context as seen in Table 1. We also showed that students participating in class scored significantly higher scores in the final exam. The CAQAS framework can achieve increased interaction with adaptations to traditional teaching methods. While our results needs to be validated for courses outside the software engineering domain, we believe that it is applicable to all courses where an iterative process of question-answer-context sets is valuable. The CAQAS framework was introduced in the beginning of the ITSE'17 class, however there was no proper tutorial, on how to use certain given features. Lacking a proper tutorial, many students where able to use basic functions of the CAQAS framework for example asking questions, however applying context-changes to questions asked or using the possibility to lookup slide references was mainly done by a moderator. This shows that CAQAS was easy to use for basic features, however it was more complex to modify a given lecture context or adding missing references to a particular question.

Students seem to have used the provided Question-Answer reports to a similar extend as the lecture material itself, according to download numbers, this indicates lecture reports have been a valuable asset for lecture and exam preparation.

The peer instruction mechanism students could use to answer questions, which than have been reviewed by a moderator worked nicely, however most of the time students used this feature outside of class hours and for exam preparation where students helped each other to prepare.

The introduced moderator role during class worked well to select and answer questions allowing the lecturer to minimize lost lecture time, however additional moderators have been proven useful depending on the number of students and the usage of the framework.

The Up-voting feature for interesting questions did not work out as intended, as it turned out to be a high cognitive load for students to follow the complete chat discussion while following the lecture content.

Answering questions in a timely manner after lecture hours and especially during peak exam preparation times was hard to do for a single moderator, suggesting to introduce multiple moderators for the upcoming ITSE course.

8.1 Future Work

During lectures, questions can arise with higher frequency than outside of lectures often overwhelming teaching assistants and moderators. When multiple questions appear in a short period of time it is possible for the first question to be pushed off-screen by the following questions. This can lead to the question being overseen and the student not receiving an answer. Currently, there is no way for students to "push" their question back to the front. The student has to ask the question again to get an answer. This problem can be fixed by tracking unanswered question with CAQAS and notifying moderators when a question has been unanswered for a certain period of time.

The generated question reports for each lecture are a very static and therefore are a non interactive form of knowledge repository. Students can only manually see through the reports for each lecture. This leads to many question being asked multiple times because the task of looking through reports to check if a question has been asked already is cumbersome. This problem is most noticeable in the channel for general questions, where students ask about certain dates, regulations, or other organizational matters multiple times. Ideally a feature like this would integrate with the question asking process, suggesting similar questions to students as they type theirs. This could reduce the workload for moderators and remove clutter from the chat room. Reducing duplicate questions also lowers the distraction caused by the tools because students have to check the question wall or chat room less often.

The usability of CAQAS is largely determined by the seamless integration into the classroom. Therefore, keeping the amount of cognitive attention needed to use the system needs to be as low as possible. The current implementation that uses *Slack* has some limitations in terms of clutter and distraction simply by being locked in to the *Slack* platform. Further improvements can be made by switching to a different a platform or building a standalone version of CAQAS. It is important to note that the introduction of new tools will increase the number of tools students have to use, which was another criticism brought forward by students.

REFERENCES

[1] Rahat Afreen. 2014. Bring your own device (BYOD) in higher education: opportunities and challenges. *International Journal of Emerging Trends & Technology in Computer Science* 3, 1 (2014), 233–236.

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

[3] Joshua D Angrist and Victor Lavy. 1997. *Using Maimonides' rule to estimate the effect of class size on student achievement*. Technical Report. National Bureau of Economic Research.

[4] Joshua D Angrist and Victor Lavy. 1999. Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics* 114, 2 (1999), 533–575.

[5] Bernd Bruegge and Allen H Dutoit. 2004. *Object-Oriented Software Engineering Using UML, Patterns and Java-(Required)*. Prentice Hall.

[6] Adam Butler, Kayah-Bah Phillmann, and Lona Smart. 2001. Active learning within a lecture: Assessing the impact of short, in-class writing exercises. *Teaching of Psychology* 28, 4 (2001), 257–259.

[7] Jane E Caldwell. 2007. Clickers in the large classroom: Current research and best-practice tips. *CBE-Life sciences education* 6, 1 (2007), 9–20.

[8] Jacob Cohen. 1988. Statistical power analysis for the behavioral sciences . Hilsdale. *NJ: Lawrence Earlbaum Associates* 2 (1988).

[9] James R Davis. 1995. *Interdisciplinary courses and team teaching: New arrangements for learning.* Greenwood Publishing Group.

[10] Ronald G Ehrenberg, Dominic J Brewer, Adam Gamoran, and J Douglas Willms. 2001. Class size and student achievement. *Psychological Science in the Public Interest* 2, 1 (2001), 1–30.

[11] Richard M Felder and Rebecca Brent. 2003. Learning by doing. *Chemical engineering education* 37, 4 (2003), 282–309.

[12] 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.

[13] Federal Statistics Office Germany. 2017. First-year students: Germany, semester, nationality, sex. (2017). Retrieved August 29, 2017 from https://www-genesis.destatis.de/genesis/online/logon?sequenz=tabelleErgebnis&selectionname=21311-0010

[14] Folker Hanusch, Levi Obijiofor, and Zala Volcic. 2009. Theoretical and Practical Issues in Team-Teaching a Large Undergraduate Class. *International Journal of Teaching and Learning in Higher Education* 21, 1 (2009), 66–74.

[15] Caroline M Hoxby. 2000. The effects of class size on student achievement: New evidence from population variation. *The Quarterly Journal of Economics* 115, 4 (2000), 1239–1285.

[16] Gwo-Jen Hwang, Chiu-Lin Lai, and Siang-Yi Wang. 2015. Seamless flipped learning: a mobile technology-enhanced flipped classroom with effective learning strategies. *Journal of Computers in Education* 2, 4 (2015), 449–473.

[17] IBM. 2017. IBM SPSS - IBM Analytics. (2017). Retrieved October 20, 2017 from https://www.ibm.com/analytics/de/de/technology/spss/

[18] Slack Inc. 2017. Slack: Be less busy. (2017). Retrieved August 29, 2017 from https://slack.com/

[19] Jan Knobloch and Enrico Gigantiello. 2017. AMATI: Another Massive Audience Teaching Instrument. In *SEUH*.

[20] Timothy R. Levine and Craig R. Hullett. 2002. Eta Squared, Partial Eta Squared, and Misreporting of Effect Size in Communication Research. *Human Communication Research* 28, 4 (2002), 612–625. https://doi.org/10.1111/j.1468-2958.2002.tb00828.x

[21] Fred G. Martin. 2012. Will Massive Open Online Courses Change How We Teach? *Commun. ACM* 55, 8 (Aug. 2012), 26–28. https://doi.org/10.1145/2240236.2240246

[22] E. Mazur. 2013. *Peer Instruction: Pearson New International Edition: A User's Manual.* Pearson Education Limited. https://books.google.de/books?id=2UipBwAAQBAJ

[23] Metabase. 2017. Metabase. (2017). Retrieved October 23, 2017 from https://metabase.com/

[24] Moodle. 2017. Moodle - Open-source learning platform | Moodle.org. (2017). Retrieved August 29, 2017 from https://moodle.org/

[25] Frederick Mosteller. 1995. The Tennessee study of class size in the early school grades. *The future of children* (1995), 113–127.

[26] OnlineTED. 2017. OnlineTED - Voting made simple. (2017). Retrieved August 29, 2017 from https://www.onlineted.de/

[27] PINGO. 2017. PINGO - Peer Instruction for very large Groups, Classroom-Response-System, Audience-Feedback. (2017). Retrieved August 29, 2017 from http://trypingo.com/

[28] Sakai. 2017. Introducing Sakai 11. (2017). Retrieved October 20, 2017 from https://www.sakaiproject.org/

[29] Bill Schilit, Norman Adams, and Roy Want. 1994. Context-aware computing applications. In *Mobile Computing Systems and Applications, 1994. WMCSA 1994. First Workshop on*. IEEE, 85–90.

[30] Slack. 2017. Real Time Messaging API. (October 2017). Retrieved October 20, 2017 from https://api.slack.com/rtm

[31] Barbara G Tabachnick, Linda S Fidell, and Steven J Osterlind. 2001. Using multivariate statistics. (2001).

[32] National Institutes of Health US National Library of Medicine. 2014. OnlineTED.com - a novel web-based audience response system. (2014). Retrieved October 20, 2017 from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3935166/

[33] Jonas Vetterick, Martin Garbe, and Clemens H Cap. 2013. Tweedback: A Live Feedback System for Large Audiences. (2013), 194–198 pages.