A Student Engagement Evaluation Methodology Inspired from Usability Engineering for Extracting Course Design Requirements

Ms. Sitong Wang, University of Cincinnati

Sitong Wang is an undergraduate student from Chongqing University – University of Cincinnati Joint Coop Institution, major in Electrical Engineering. Her academic interest is in the Human-Computer Interaction (HCI) field. She had internship experience as a research assistant at Computer Science Department, University of Virginia in Fall 2018. She also took the role of being the research assistant in HCI Initiative group at The Hong Kong University of Science and Technology in Summer 2019. On the industry side, she was once a trainee user interface designer at a big data product development group of a technology company in Chongqing, China.

Dr. Panagiotis Apostolellis, University of Virginia

Panagiotis Apostolellis is a full-time Assistant Professor in the Computer Science Department at the University of Virginia. He received his PhD in Computer Science & Applications from Virginia Tech in 2017. He also holds a MSc in Computer Science from Virginia Tech and an MSc in Human-Centered Systems from the University of Sussex (UK). Additionally, Panagiotis has a Graduate Certificate in Human-Computer Interaction from Virginia Tech (2015). While a graduate student at Virginia Tech (2011-2017), he worked at the Center for Human-Computer Interaction under the guidance of Dr. Doug Bowman, researching the impact of audience interaction using serious games and VR on young student audiences visiting informal learning spaces. His teaching experience involves being an Adjunct faculty member and a Visiting Assistant Professor at Virginia Tech, serving as a Teaching Assistant in multiple CS courses, and teaching diverse audiences about IT and New Media Technologies. His dissertation received the Outstanding Research award for 2017-2018 from the CS department at Virginia Tech. Panagiotis has also extensive experience as a Senior Interactive Systems Designer and Developer at a cultural institution in his home country, Greece (2000-2011).
A Student Engagement Evaluation Methodology Inspired from Usability Engineering for Extracting Course Design Requirements

Abstract

Measuring student engagement inside the classroom and developing techniques for improving it has been traditionally very challenging for educators. This research paper describes a student engagement evaluation model that combined data from three sources: in-class observations using the Behavioral Engagement Related to Instruction (BERI) protocol, one-to-one student interviews, and anonymous online surveys. We tested this model on a higher-level elective Computer Science class with 134 students, focusing on user experience (UX) design. We used the exact same usability engineering process that students employed in the course to design software products on assessing the course’s effectiveness, with students acting as users and the class being the system under development. We present our exact methodology of data-driven analysis borrowed from UX design, which combines inductive reasoning (similar to narrative analysis or inductive coding) with deductive reasoning (similar to content analysis of qualitative data) used in social science research. Our paper includes a sample of the extracted requirements, similar to software non-functional requirements, which we used to redesign the course for the next semester. The results of this case study showed that the application of this mixed methods type of analysis informed by user-centered design of software systems was effective as a surrogate model of student-centered instructional design. Concluding, we extrapolate the lessons learned from this process and the significant implications we believe our industry-inspired methodology can have for engineering educators, in terms of evaluating student engagement in college classrooms.

1. Introduction

In the last few decades, educators have been looking for ways to increase learning outcomes by keeping students engaged during and outside of class. Popular pedagogical approaches include active learning [1], peer instruction [2], student-centered activities [3], gamification of instructions [4] etc. When evaluating the effectiveness of these designs, many instructors choose student engagement as one of the most significant methods of assessment of their interventions [3], [5], [6]. The level of engagement is usually measured by using student self-reported surveys or in-class observation instruments [5]. However, nontrivial efforts are needed for instructors to interpret and use the evaluation results in order to improve course design. In this paper, we present our work on combining such well-established methods of collecting student feedback with an engineering process for understanding and improving interactive software products. Using a third-year computer science course on human-computer interaction as a test case, we applied the exact same process taught in the course to improve the course itself, by extracting a list of course design requirements.

More specifically, inspired by the usability engineering process that students employed in the course to design software products [7], we applied the suggested methodology to move from engagement evaluation outcomes to course (re)design requirements by treating the course as
a product and the students as users of the product. Our main premise was applying a user-centered design approach to facilitate student-centered course design. As a tried engineering methodology, the usability engineering process includes an effective approach for consolidating and interpreting qualitative data (i.e., input from user research), combining inductive and deductive reasoning. Combining this methodology with an established observation protocol for in-class engagement, i.e., (BERI) [8], we anticipate we can add value to existing methods of interpreting student feedback for course evaluation and (re)design purposes.

The leading research question of this paper is the following:

- RQ: Can we evaluate the effectiveness of a college course in computer science by using the usability engineering process (taught in the case course)?

Adopting as an operational definition of engagement the combination of quality of student effort and involvement in productive learning activities [9], we present below the course used as a test case, our methodology for collecting and analyzing the data, and main findings from this process. We then discuss these findings in the light of existing efforts in student-centered course design and our future direction for evaluating the effectiveness of the interventions that derived from this work, in the form of course design requirements. Our value proposition lies in utilizing alternative—engineering-inspired—formative evaluation methods for the collection, analysis, and interpretation of qualitative data for course design requirements.

2. Background and Theoretical Foundations

2.1. Measuring student engagement

Student engagement is an important indicator of education quality. However, consensus exists among researchers that engagement is multifaceted and complex [6], [10]. In the context of education, engagement may contain a behavioral dimension (e.g., attendance and participation in class), a cognitive dimension (e.g., deep learning and self-regulation), and an affective dimension (e.g., belonging and identification with school). Though the latter two dimensions are certainly significant components of the student engagement framework, they are difficult to be defined and measured in practice [11]. In this paper, we mainly focus on the measuring the behavioral aspect of student engagement due to its operational feasibility [9], [11], [12] and close connection with the teaching practice, which is the improvement target of our research.

Behavioral engagement was usually measured by using student self-reported surveys or in class observation instruments [5]. However, due to subjective understandings of academic terms in the survey, reliability of student responses is sometimes questionable [13]. In addition, wide-angled self-reports miss the dynamic and situational snapshot of student engagement [14]. Others have used activity logs from a learning management system (LMS) to gauge student engagement [12], but although this can provide a more accurate measure of student involvement with the material, it misses the opportunity to accurately gauge in-class engagement. Direct observation can instead provide a more objective measure for capturing dynamic changes of in-class engagement. There exist several practical observation protocols for in-class engagement, such as the Reformed Teaching Observation Protocol (RTOP) [15],
the Classroom Observation Protocol for Undergraduate STEM (COPUS) [16], and the Behavioral Engagement Related to Instruction (BERI) [8] and so on. However, RTOP focuses more on the instructor’s course design instead of directly evaluating student engagement; COPUS provides behavioral patterns of what the students (and instructor) are doing, but it does not accurately capture engagement level. Among all protocols, BERI is the most suitable tool to quantifiably measure the level of student in-class engagement, while also capturing consistent patterns of engagement variation across class sessions.

2.2. Usability Engineering in course design

Usability engineering provides structured approaches to developing interactive systems with high usability and user friendliness [17]. In the context of software development, a variety of process models were developed by experts to provide guidance through various stages of development. For example, the waterfall model follows a linear progression of stages, moving from one phase to the next following a predetermined sequence [18]; the spiral model describes a progression of repetitions with each successive round expanding the scope a bit, like a spiral [19]. Based on existing models, the UX Wheel process model (thereafter called the UX process or lifecycle), proposed by Rex Hartson and Pardha Pyla [7], is an evaluation-centered iterative model that follows the progression of analyzing, designing, prototyping, and evaluating an interactive system or process (see Figure 1). Compared to the former methods, the UX process is more flexible, customizable, and complete [20].

![UX lifecycle](image)

Figure 1: The UX lifecycle [7] used as a tool for formative evaluation and course redesign.

By treating students as customers of courses, other researchers have started exploring the application of mature industrial principles in the academic field. For example, Pusca and Northwood [21] proposed applying lean principles—used by companies to make radical changes in their business model and understand customer needs—to develop and improve engineering courses. Adopting a user-centered design approach from the UX design profession, Vovoreanu and Connolly [22] got a better understanding of student needs and created a new curriculum based on the extracted design requirements. Similar to these
studies, our paper applies a tried engineering methodology from industry, combined with an established observation protocol, to explore engineering course design.

2.3. Qualitative data analysis and interpretation

In educational research, qualitative data analysis is essential to identifying important trends and interpreting findings. Certain types of qualitative data can also be used to investigate what type of intervention works [23]. Two very common and well-established techniques in qualitative data analysis, inductive and deductive reasoning, should arguably be combined to increase benefits in research [24]. Inductive reasoning moves from specific to general and involves a “bottom up” approach to build broader themes from the participants’ perspective [25]. Inductive coding is usually used when categories and themes are not predefined, and themes will be suggested by the data [26]. By contrast, deductive reasoning moves from general to specific and uses a “top down” method [25]. Qualitative data are coded with preset themes based on prior research and/or related assumptions and hypotheses.

For example, when studying student and faculty views on mathematics, Dewar and Bennett independently read the original responses and identified obvious repeated themes, using inductive coding to find emergent categories in describing math [27]. Similarly, deductive reasoning was utilized by Bruff to analyze qualitative responses about student difficulties using predetermined themes, derived from knowledge dimensions found in prior work [28]. In the context of assessing nursing practice, Fereday used a hybrid approach of inductive (emerging from participants’ discussions) and deductive coding (derived from an existing framework) using semantic analysis [29]. Similar to the latter work, the benefit of using the UX process on course design is that it includes a methodology combining inductive and deductive reasoning, which we used for formative evaluation of course effectiveness.

2.4. Student-centered instructional design

Course design has always been a significant topic in engineering education research. As higher education has shifted its focus from teaching to learning, student engagement and satisfaction have been the main goals for educators and course designers [30]. In recent years, traditional teacher-centered teaching models and their effectiveness have been questioned [31]. Student-centered learning, which changes the main character of class from instructors to students, has been increasingly adopted instead. By putting students in the center of class activities, student-centered learning creates more interactions and hands-on practices in the classroom, making students discover knowledge by themselves [32].

Adopting a similar philosophy, Traver proposed involving students in instructional design by applying the principle of user-centered design [33]. By making a mapping from elements of user interface design to instruction design, Traver claimed that there exists a nice analogy between user-centered design and student-centered instruction. Parallel to iterative user-interface design, where designers modify prototypes to accommodate users’ suggestions, student-centered instruction should adopt students’ feedback and voice. In this paper, we take this exploration a step further by adopting the systematic usability engineering process to enhance student-centered outcomes in the context of a higher-level computer science course.
3. Course Format

The Human-Centered Interaction (HCI) in Software Development course is a 3000-level computer science course at the University of Virginia and was designed to provide students a hands-on experience in identifying and designing a solution for a real-world problem. Students work in groups of 4-5 members and are matched based on skill level in terms of programming and design experience. There are four phases distributed throughout the semester, during which students need to identify a problem and an appropriate client doing a “rapid spin” around the UX lifecycle (phase 1), conduct contextual inquiry and analysis (phase 2), design a solution (phase 3), and finally prototype and evaluate their best design (phase 4). The major motivation for a project-based syllabus is to create a pedagogy of engagement, where students have the opportunity to engage with the disciplinary activities and practice the taught process like a UX professional [34].

The main objective of the course is for students to get excited about the UX design process and develop a deeper understanding about user-centered design and its value in software development. Considering the significance of appreciating the value of learning tasks in increasing motivation and consequently engagement in learning [30], we emphasize additional life-long skills like communication, collaboration, and creative thinking. Our ultimate learning goal is to create the type of significant learning [35] that produces creative problem solvers that can think in integrative ways for solving real-world challenges, working effectively in teams and applying HCI skills in a variety of contexts. An array of different active learning tasks, inside and outside of classroom, were designed to achieve the expected learning outcomes mentioned above.

To understand these tasks, let us describe a typical 2-day week: a) quiz on the reading of the week followed by lecture with added examples on that topic (Tuesday); b) in-class activities (ICAs) where students practice the part of the UX process taught that week (Thursday); c) project group meetings with the facilitation of the TAs (in lieu of office hours). Lecture time is augmented with complementary activities, such as ungraded polling questions (using mentimeter.com [36]), real-world examples with some brief activity, and mini individual presentations of good-bad-ugly UX examples (GBUX). In-class activities (ICAs) are complemented with sharing design artifacts to the whole class (using sharypic.com [37]) and mini group presentations of the day’s products in front of the class. Finally, two exams are included to test conceptual understanding, while two design studios are used as a form of peer assessment (formative evaluation of designs).

4. Course evaluation using the Usability Engineering process

Considering courses as analogous to interactive systems and students as the users of the designed systems, we applied the UX process taught in this course in improving the course itself. Following the analyze-design-prototype-evaluate lifecycle, we conducted the usability engineering process in the case course described above. Improved student engagement and satisfaction, in this context, were defined as the usability goals. During the Analyze phase, Contextual Inquiry and Analysis were adopted in the form of in-class observations,
interviews, and surveys. Second, in the Design phase, we considered feasible changes of the course design during the semester and beyond. Third, during Prototype, rapid prototyping techniques were used to make changes to the case course on the fly. Finally, during Evaluate, dynamic observation data were gathered to gauge student response to interventions as a means of formative evaluation. More changes informed the offering of the course in a future semester as well as follow-up research on the effectiveness of implemented changes—summative evaluation, not reported here. These phases are presented in more detail below.

4.1. Analyze

The goal of this first phase is to understand the status of an existing system (what works and what does not), understand its users (identify motivations, needs, and struggles), and come up with the main requirements of the envisioned system. Towards this goal, we followed the suggested process of Contextual Inquiry and conducted in-class observations, one-to-one interviews, and online anonymous surveys.

4.1.1 Contextual Inquiry as a formal evaluation methodology

In-class observations were used to measure students’ engagement level, our main usability goal set at the outset. We used the BERI observation protocol, developed and validated by Lane et al. in large university classes [8], as it allows quantifying student in-class engagement. Following instructions in this protocol, the first author acting as the observer would get the lecturer’s lesson plan before class and then randomly sit behind a group of 10 students in class, recording how many of them were engaged, disengaged, or of uncertain engagement. The observation is recorded for any major content change or every 2 minutes, and notes are written on every page of the lesson plan. Therefore, the instructor can relate the student engagement data to the lecture content after class. To refine the student behavior categories suggested by the BERI protocol, we did some pilot observations based on the suggested behavior set. We decided to add some course-specific behaviors such as “collaborative work,” “device use,” and “presenting.” We then used the expanded behavior pool in the formal observations, recording abbreviations of the observed students’ behavior on a spreadsheet simulating the classroom setup (e.g., ‘FB’ for social media use). Instead of directly marking the observed students as “engaged” or “disengaged,” we modified the protocol to record more details such as their exact behavior, their seat location on a diagram, questions from the audience, and specific instructor behaviors (e.g., asking/responding to questions).

To gather the completed engagement data for comparison and evaluation, we formally observed 19 lectures throughout the Fall 2018 semester. Before every class, the observer received a printed copy of lecture slides from the instructor and then randomly chose a group of ten students to observe. Classes were held in a large classroom with tiered seating and the observer sat behind the selected student group to get a clear view of their engagement. To cover as many students as possible, each time the location of observed students was recorded and next time the observer would choose a different area in the classroom. The observed students were coded anonymously as S1 to S10 and every student was observed for about 5 to 10 seconds until the behavior category was determined. The average time to finish one cycle
was approximately 2 minutes. The general information of the class, including the date, attendance, techniques used, and the environment was also noted for future reference.

For the interview, four qualitative questions were designed to motivate students to tell their stories regarding in-class engagement, including asking them to recall the last time they felt most engaged/disengaged and what part they liked/disliked the most in the class. Two questions asked about what parts of the class were most helpful for learning about HCI. Interviewees were recruited via email from the researcher and they provided verbal consent before the interview started. The interview was voluntary with no compensation. Seven students showed interest and consented to be interviewed. Each interview lasted around 45 minutes and all participants agreed to be audio recorded; recordings were then transcribed and used in the analysis described below. The whole interview process was done by an independent researcher and the interview data were not revealed to the instructor until after the final grades were released.

To increase our data pool, an online survey was designed for students to self-report their engagement in the HCI class. This anonymous survey included three questions specifically on engagement in different parts of the HCI course. Most questions were answered based on a five-point Likert scale; open-text questions were included to collect general student comments about the course. Students got one extra credit for completing the questionnaire. Around the middle of the term, the anonymous online survey was released using Qualtrics. In total, 139 responses were recorded on Qualtrics; after deleting the duplicate and unfinished responses, we ended up with 110 valid responses. This research and all data collection methods were approved by the University’s IRB-SBS, protocol #2018-0427-00.

4.1.2 Contextual Analysis as a method of analyzing qualitative data

As part of the Contextual Analysis stage, we used a work activity affinity diagram (WAAD) to help with interpreting the collected data. Hartson and Pyla claim in their book (used as the course textbook) that a WAAD can offer deep insights from all user (student, in our case) data in a visual way, through a collaborative data-driven consolidation and interpretation process [7]. The main purpose of building the affinity diagram was to extract the most significant HCI course requirements coming from the students, based on data from all three sources (although only a few comments were used from the observation data). Following the data-driven process described in the book (inductive reasoning), we compiled 228 work activity notes, numbered (to maintain connection with their source) and printed on post-it notes.

We invited three teaching assistants to join our team and help with the WAAD activity as domain experts. Creating the WAAD involves moving the post-it notes around and grouping them into clusters until they organically form into meaningful groups, effectively capturing the students’ thoughts and preferences about the HCI class in a hierarchical manner. The first level of the hierarchy included categories such as quiz, group meetings, lecture, textbook, in-class activities, project, and communication, covering almost all aspects of the course. The next levels gave a more detailed description of the students’ perceptions based on the work activity notes created from the interview and survey data. Figure 2 shows the team members while working on this process.
The second part of the process involves doing a walkthrough of the WAAD and extracting design requirements. These requirements come directly from the “user’s voice” and capture in essence what students see as significant course improvements/changes, a significant factor of student-centered instructional approaches. Based on the analysis of the affinity diagram, we extracted design requirements describing the system’s main supported tasks and qualities (similar to non-functional requirements in software engineering). The extraction process involved understanding which work activity data (post-it notes) under different categories could be best generalized and expressed as system requirements (deductive reasoning). We eventually ended up with 102 unique requirements for redesigning the course (our “system”), a representative sample of which is included in Table 1. These requirements were then prioritized based on their importance—potential impact on the course learning objectives (defined by the instructor) and their weight—the count of notes they derived from. Different changes needed different commitment level to be implemented and evaluated, so we only tried a few interventions during the semester (discussed next).

4.2. Design

Requirements are eventually feeding into design—the creative exploration of ideas about ways to implement the required supported tasks. Considering the risk of making changes during the semester, potentially impacting the flow of instructional activities, we initially noted the changes we deemed appropriate for the next semester (appear as “Planned” under the “Comments” in Table 1). We then chose the top priority requirements that we thought as more impactful when redesigning the course—based on student feedback—and came up with: (a) changing the quiz time administration; (b) enhancing in-class activities and adding other brief activities during lectures; (c) adding more rigorous assessment for the GBUX.
Table 1: A sample of our prioritized course redesign requirements, including the main categories and subcategories (themes), importance factor, and planned changes as comments.

<table>
<thead>
<tr>
<th>#</th>
<th>Category</th>
<th>Secondary</th>
<th>Third</th>
<th>WAAD ID</th>
<th>Count</th>
<th>Requirement</th>
<th>Importance</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Quiz</td>
<td>When</td>
<td></td>
<td>O12.1, 563.1,</td>
<td>8</td>
<td>The quiz should be set after the related lecture.</td>
<td>5</td>
<td>Planned; motivates students to attend</td>
</tr>
<tr>
<td>14</td>
<td>Group</td>
<td>Pros</td>
<td></td>
<td>I2.1, i3.13</td>
<td>2</td>
<td>The course should continue the form of group work.</td>
<td>5</td>
<td>Planned</td>
</tr>
<tr>
<td>15</td>
<td>Group</td>
<td>Pros</td>
<td></td>
<td>I1.3</td>
<td>1</td>
<td>The group should consist of students with diverse abilities.</td>
<td>5</td>
<td>Goal of HW1</td>
</tr>
<tr>
<td>23</td>
<td>Lecture</td>
<td>Content</td>
<td>GBUX</td>
<td>I1.13, i5.4,</td>
<td>3</td>
<td>GBUX should be designed to be more helpful.</td>
<td>5</td>
<td>Planned; &gt; rigorous assessment, &lt; points</td>
</tr>
<tr>
<td>24</td>
<td>Lecture</td>
<td>Content</td>
<td>GBUX</td>
<td>I2.3(3)</td>
<td>3</td>
<td>GBUX should be entertaining and relaxing.</td>
<td>5</td>
<td>This is the goal but need more structure</td>
</tr>
<tr>
<td>32</td>
<td>Lecture</td>
<td>Style</td>
<td></td>
<td>I2.12</td>
<td>1</td>
<td>Detailed slides should be provided for review reference</td>
<td>5</td>
<td>Slides are as a study guide, already Incorporate more real-world examples</td>
</tr>
<tr>
<td>45</td>
<td>Lecture</td>
<td>Engagement</td>
<td>Examples</td>
<td>I1.6, i1.14(16), 53.1, 566.1</td>
<td>18</td>
<td>More real-world examples demonstrated during lecture.</td>
<td>5</td>
<td>ICAs should continue being applied in the future courses.</td>
</tr>
<tr>
<td>66</td>
<td>ICAs</td>
<td>Positives</td>
<td></td>
<td>I1.4, 4.25</td>
<td>2</td>
<td>Design studio should be included in future course design.</td>
<td>5</td>
<td>Planned</td>
</tr>
</tbody>
</table>

First, when considering quiz administration, from the Quiz When subgroup of the diagram we can clearly see that students tend to feel tired right after having a quiz in class, since most of the activity notes included negative words such as “tiresome,” “pressured,” and “stressed.” Therefore, we considered changing the quiz time administration to improve student engagement. Second, students’ preference for in-class activities was strongly supported by pretty much all sources of data. There were 19 work activity notes emphasizing the Positive feelings (subgroup), under the main level In-class activities. Also, looking through the Lecture-Engagement-Other activities branch of the WAAD, we can find a post-it note saying, “The student enjoys the interactive parts of the class, such as the puzzle question in Design Thinking lecture.” Hence, we decided to enhance in-class activities and add other brief activities during lectures for the rest of the semester. Third, looking through the path of Lecture-Content-GBUX in the WAAD, work activity notes indicated that although students found GBUX fun, they did not think it was very helpful and sometimes they did not really pay attention to the content presented. It was interesting because in the anonymous survey, more than 80% of students reported GBUX as either “engaging” or “very engaging.” However, from the WAAD we found that such engagement was rather superficial.

4.3. Prototype

A common way to try out different design ideas and receive formative feedback in usability engineering (and user-centered design, in general) is called rapid prototyping. This involves low-cost, on-the-fly interventions for specific features of a system that is hypothesized can bring considerable value. Rapid prototyping was tried near the end of the semester for the first two design ideas presented above—(a) changing the quiz administration time; and (b) enhancing in-class activities and adding other brief activities during lectures. In order not to unfairly affect the grading of the GBUX assessment for some students, we deferred making any changes regarding design idea (c) to a future semester.
To test students’ mastery of reading materials, a 10-minute quiz was administered before every Tuesday lecture on that topic. This format was originally adapted from a flipped classroom approach, where students come prepared to class having read the material at home, allowing them to focus and participate with questions during lecture [38]. Following the rapid prototyping approach and based on specific student feedback, four different modes of quiz administration were attempted throughout the semester (in chronological order): (a) quiz at the beginning of the class (regular mode), (b) quiz at home, (c) quiz in the middle of class, and (d) quiz at the end of class. Following our second design idea, we injected more brief activities during lecture to capture student attention and help them (re)focus on the lecture. For example, we introduced the Systematic Inventive Thinking framework in IDEO’s Design Thinking lecture, including specific hands-on activities.

In the evaluation stage, we focused on formative assessments of our main usability goal—student engagement—to test the effectiveness of our prototyped interventions (design ideas). Engagement levels (eng) come from the BERI observations, while quiz scores are means from eight (8) points maximum per quiz; ANOVA was used when p values are reported.

4.4.1 Formative assessment

Regarding the first design idea about quiz administration time, engagement variations under different quiz modes were plotted following the order (a-d) mentioned above (see Figure 3). The vertical axis represents the average number of engaged students across the number of observed class sessions n that each mode was used for. Color-coded blocks indicate the specific activity taking place in class for specific intervals of the 75-min class time (used only when n=1). It is worth noting that the average engagement level for the four modes are (a)
4.01, (b) 5.40, (c) 4.89 and (d) 7.44 respectively, where the quiz at the beginning of class (the regular mode) ranks the lowest among all the conditions. In addition, we can see that the engagement number drops rapidly right after the quiz, when the quiz is either taken at (a) the beginning or (c) in the middle of the class. The explanation for this phenomenon was obtained from the interview and the survey. One student responded in the online survey: “I used so much energy on the quiz, and then having to focus on lecture felt exhausting.” Doing quizzes was analogue to checking something off a to-do list in some student responses. Too much energy and attention were consumed when having the quiz, leading to the dramatic decrease in engagement observed rightly after.

Meanwhile, having quizzes after class was given a lot of preference, through comments like “It was definitely more engaging in class without quizzes, because you don’t need to worry about the grades.” Plot (b) confirms student perception since on average these classes (n=3) had among the highest engagement level (M=5.40). Finally, having the quiz at the end of class (d) seemed to have the most positive impact on student engagement (M=7.44); we have to take into account, though, that this was the outcome of only one observed class. Students reported that this mode helped them fully understand the concepts before being tested and pushed them to sustain their focus when the professor was giving instructions.

In terms of performance, we did not find a significant effect of the quiz mode on student scores, when the quiz was administered in class. The only exception was when the quiz was taken at home (b), in which case the mean score (Mquiz=7.41/8) was significantly higher than all other modes (F=23.78, p<0.001). Although the quiz was open-notes, open-book in every case, we believe there is an apparent effect of stress when the quiz is taken in a classroom environment that immediately affects student performance. However, we do not claim that the higher score reflects greater learning gains and more investigation is needed for a safe conclusion. Nonetheless, the evaluation results indicated that the specific design of moving testing time right after lecture would improve student engagement and was adopted in the next iteration of the course.

Injecting various brief activities during lecture was a practice that was consciously adopted to capture student attention and help them (re)focus on the lecture. To test the effectiveness of this intervention, we selected two Tuesday lectures where such activities were used (Figure 4). As the legend indicates, the green areas cover the time intervals during brief activities. Referring to our observation notes, 28 minutes into the Design Thinking lecture (Figure 4a), the professor presented a puzzle question encouraging students to solve on paper. Similarly, in IDEO’s Design Thinking lecture (Figure 4b), the Systematic Inventive Thinking framework was introduced in the last twenty minutes of class, including specific hands-on activities. As apparently depicted in Figure 4, both activities rapidly sparked the curiosity and engagement of students, making them participate actively in the class.

We cannot overlook, however, that student engagement drops to a set level right after the activity, challenging the way lecture and brief activities in class are mixed. Also, worthwhile is the observation that the lecture with the lowest engagement level (Figure 4a) was the one where a quiz was administered at the beginning, confirming the choice to include a quiz later in the lecture, as a means to increase student attendance.
This observation evaluation outcome is supported by the survey statistics, where over 60% of students felt that they were either “engaged” or “very engaged” (M=3.48/5) during Thursday ICAs. Similarly, perceived learning value was rather high (M=3.44/5). Students’ preference for doing hands-on activities was demonstrated even more vividly when comparing the engagement level between Tuesday lectures and Thursday activities (Figure 5). Plotting the average engagement level of nine lectures and ten ICAs, we can confidently say that the design idea of enhancing within-lecture hands-on activities can efficiently sustain student engagement, echoing the substantial literature on active learning e.g., [39]–[41].

4.1.2 Summative assessment and future work

Due to the iterative nature of the UX process and instructional design, our work does not stop here; we keep changing (Prototype) and testing the outcomes (Evaluate) of our interventions. Summative evaluation of the planned changes in subsequent semesters include understanding the impact of note-taking during lecture on student performance (quizzes) and investigating the relationship between the personal relevance of a selected project idea with a student’s performance during groupwork. The former study is connected to our continuous effort on improving student performance through brief quizzes (#1 in Table 1), while the latter aims to help us understand effective ways to help with group projects (#14-15). A summative evaluation of the designs presented in this paper was not attempted, as our goal was more to present an alternative formative evaluation framework than prove that these were the best alternatives. Data from more course design interventions in future semesters are currently being evaluated and will be the focus of a subsequent paper.
5. Discussion

Revisiting our leading research question, we feel we managed to address it in a satisfactory manner. We successfully tested a viable way—UX process—for analyzing and redesigning a higher-level computer science course on HCI (RQ). This was a mixed methods approach, where we employed the power of a usability engineering methodology for gathering and interpreting course design requirements, while also using an established observation protocol (BERI) for evaluating our interventions in terms of student engagement—our value proposition. We further elaborate on the main learning points from applying this combined process from the perspective of our research question and main contribution.

5.1. Usability engineering process used for student-centered instruction

Following a user-centered design (UCD) approach, we treated our course as the product and students as the users. We conducted contextual inquiry through observations, semi-structured interviews and surveys, and then we did the qualitative data analysis through a work activity affinity diagram, which effectively combines inductive and deductive reasoning. Design requirements for course improvement naturally surfaced through this process and were well supported. Following this UCD philosophy, other techniques such as design thinking [42], focus groups [43], and personas [22] have been suggested to aid course design by several researchers. The BERI observation protocol was also used as a formative assessment tool for the effectiveness of our prototyped interventions and active learning strategies [44]–[46]. In a similar vein, student-centered learning gives priority to students’ interests, continuously assessing and acknowledging students’ voice as central to their learning experience [47].

Others have advocated about the potential power of UCD in student-centered education [33], but to our knowledge there has been no actual implementation of this specific methodology—the UX lifecycle—to course design. This analyze-design-prototype-evaluate iterative approach is very similar to what Denning and Dargan called action-centered design in software engineering [48]; paraphrasing their proposal (italics replacing original text), this is “a broader interpretation of design that is based on observing repetitive actions of students in a domain and connecting those action-processes to supportive instructional technologies.” Similarly, student-centered learning requires students to be active agents, responsible participants in their own learning, while the teacher becomes the facilitator aptly adapting instruction based on their gauged needs [49]. Our contribution lies in that we made a preliminary step in applying this usability engineering process—a mature and powerful tool in industry—to course evaluation and design for enhancing student-centered teaching. We propose that such a formative evaluation model expands beyond the field of computer science and can even be applied to non-engineering university disciplines.

5.2. The WAAD as an inductive and deductive analysis methodology

The WAAD, used during analysis, effectively combines inductive and deductive reasoning for qualitative data analysis. After extracting notes from the raw data gathered in contextual inquiry, inductive reasoning was used to generalize student comments into work activity notes. As a figurative example, an interview quote such as “I really liked the WAAD because
it helped me in getting an internship” would be transferred on a post-it note as “The WAAD is very helpful for students’ internship applications.” Since we had no predefined themes of how to redesign the course, this data-driven approach enabled the assignment of notes in affinity groups and letting themes emerge from the data. This is very similar to thematic analysis, a common technic of inductive coding in qualitative methods research [47].

Later, during the WAAD walkthrough, the designer is asked to switch mindsets from inductive to deductive reasoning [7]. Using the emergent themes, specific course redesign requirements are extracted from the coded, more general activity notes. The outcome of this process, i.e., the specific requirements, can be treated as the deductive or concept-driven coding done in qualitative methods research [50]. This deductive approach avoids any bias that can arise from having predefined codes [51] or, in our case, predetermined requirements about what changes need to be made to the course by the instructor. As such, the process helps surface the themes that are important to students, eliminating the risks involved in trying to push a specific redesign agenda from the instructor’s perspective; once more, serving the main goals of student-centered learning [52].

5.3. Limitations and threats to validity

Despite our fair attempts to have a representative sample of observed sets of ten students across the classroom, it was inevitable that for single sample observations (n=1, Figure 3c, d) we had a limited coverage of locations. Thus, the finding that administering the quiz at the end of class increases engagement (observation 6) might be an effect of student seating, since the observed group that triggered this finding happened to be seating in the first three rows of the room (see Figure 6). Similarly, the lecture topic and how interesting was perceived by students might have affected the observed engagement level.

<table>
<thead>
<tr>
<th>No.</th>
<th>Date</th>
<th>Weekday</th>
<th>Topic</th>
<th>Engagement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.11</td>
<td>Tuesday</td>
<td>Wheel &amp; CI (Start)</td>
<td>4.67</td>
</tr>
<tr>
<td>2</td>
<td>9.18</td>
<td>Tuesday</td>
<td>Contextual Analysis (Start)</td>
<td>5.30</td>
</tr>
<tr>
<td>3</td>
<td>9.25</td>
<td>Tuesday</td>
<td>DRVs &amp; OMs (Start)</td>
<td>3.25</td>
</tr>
<tr>
<td>4</td>
<td>10.2</td>
<td>Tuesday</td>
<td>Design Thinking (Start)</td>
<td>2.82</td>
</tr>
<tr>
<td>5</td>
<td>10.16</td>
<td>Tuesday</td>
<td>Conceptual Design (No)</td>
<td>6.43</td>
</tr>
<tr>
<td>6</td>
<td>10.30</td>
<td>Tuesday</td>
<td>UX Goals &amp; Prototyping (End)</td>
<td>7.44</td>
</tr>
<tr>
<td>7</td>
<td>11.5</td>
<td>Tuesday</td>
<td>Evaluation &amp; Rigorous Evaluation (Mid)</td>
<td>4.89</td>
</tr>
<tr>
<td>8</td>
<td>11.13</td>
<td>Tuesday</td>
<td>Design thinking by IDEO (No)</td>
<td>5.89</td>
</tr>
<tr>
<td>9</td>
<td>11.27</td>
<td>Tuesday</td>
<td>HCI Examples (No)</td>
<td>3.70</td>
</tr>
</tbody>
</table>

Figure 6: [top-left] The nine (9) lectures observed by topic and engagement level (out of ten students); [top-right] A line graph of the engagement level per lecture; [bottom] A visual representation of the classroom with the observer’s position for each observation.
Moreover, the room itself, being a large tiered room (see Figure 6-bottom) and the time of class (12:30 pm, a time commonly students have lunch) could have had their own impact on the observed engagement of students. As an example, some students were consistently captured eating in class, which unavoidably compromised their ability to concentrate on the lecture. Finally, observations bear the inherent subjectivity of the observer, although pilot sessions aimed at mitigating this factor. Additionally, there is always a risk of missing specific student behaviors, especially during class content transition times (e.g., switching from lecture to an active participation activity).

6. Conclusion

In this paper, we built a student-centered course evaluation model using lessons learned from UX design and combined with an objective observation protocol and multiple-source feedback from students. As the UX methodology (contextual inquiry, contextual analysis, and requirement extraction) work well in tandem with observational data, the requirements extracted are naturally grown from the students’ feedback rather than the professor’s personal experience and thoughts. The combination of inductive and deductive reasoning used during contextual analysis, a common UX methodology, and collected observational data gave the instructor a clean and systematic view of the whole course from the students’ perspective. This user-centered approach is well aligned with student-centered course design, which has been proven to be largely beneficial for increased student learning and engagement [30], [47], [49], [53]. Several recommended changes derived from this combined model in the form of design requirements (a sample presented here), which were used to guide course redesign for future semesters. By prototyping some of these changes and conducting formative evaluations (through in-class observations), we could see the impact of the attempted interventions on student engagement—a predetermined usability goal as an indicator of student engagement and involvement in learning activities [9]. Based on preliminary findings, we believe this model can act as a guiding post for future research on effective methods for course redesign, informed by one of industry’s practices on product design.

Acknowledgements

The authors would like to thank Rex Hartson and Doug Bowman for their influence in the original course design and guidance throughout. We also express our appreciation to the students who contributed to the development of this research study. Finally, many thanks to Lindsay Wheeler for her guidance and reviewing of this work.

References


D. Sawada et al., “Measuring Reform Practices in Science and Mathematics


[43] Z. E. H. Kahraman, “Using user-centered design approach in course design,” in


