

Edvard Mattias Bø

# Investigating the effects of visualizations in teacher dashboards for adaptive self-assessment systems

Master's thesis in Informatics: Software Engineering

Supervisor: Michail Giannakos

Co-supervisor: Zacharoula Papamitsiou

June 2022



Edvard Mattias Bø

# **Investigating the effects of visualizations in teacher dashboards for adaptive self-assessment systems**

Master's thesis in Informatics: Software Engineering  
Supervisor: Michail Giannakos  
Co-supervisor: Zacharoula Papamitsiou  
June 2022

Norwegian University of Science and Technology  
Faculty of Information Technology and Electrical Engineering  
Department of Computer Science



# Abstract

SmartU is an adaptive self-assessment system built to help students learn while fostering motivation and engagement. This thesis describes the design of a teacher dashboard for use within this system. The dashboard will support the teacher to get an overview of how the quizzes and questions are used and whether the students understand the educational material. Thus, the teacher can improve the quality of the questions by utilizing the dashboard, and can further help students in their learning.

The thesis investigates how visual analytics can help teacher monitor the progress of students who use adaptive assessment. To address this, the dashboard was created, and tested on participants with teaching experiences. Data was collected through questionnaires, interviews and observational data. Results from the data analysis indicate that the teachers found the dashboard to be useful, usable and motivational to use.



# Sammendrag

SmartU er et adaptivt selvevalueringssystem som er laget for å hjelpe studenter å lære samtidig som det er motiverende og engasjerende. Denne oppgaven beskriver hvordan et dashboard for lærere ble utviklet og testet i dette systemet. Ved å lage et dashboard, kan lærere få oversikt over hvordan oppgavene gjør det, og om studentene forstår pensum. På denne måten kan de forbedre kvaliteten på spørsmålene ved hjelp av dashboardet, og dermed hjelpe studentene med å lære mer og forstå bedre.

Denne forskningen undersøker hvordan visuell analyse kan hjelpe lærere å følge med på fremgangen til studenter som bruker adaptiv evaluering. For å utforske dette ble et dashboard utviklet og testet på deltakere med undervisningserfaring. Data ble anskaffet gjennom en spørreundersøkelse, intervju og observasjonsdata fra brukertester. Resultatene fra dataanalysen indikerer at lærerne synes dashboardet var nyttig, brukervennlig og motiverende å bruke.





# Acknowledgements

This thesis was written for the Department of Computer Science (IDI) at the Norwegian University of Science and Technology (NTNU) during the spring semester of 2022.

I would like to specially thank my supervisors, Michail Giannakos and Zacharoula Papamitsiou for their brilliant guidance, feedback and support throughout the work of this thesis. Their thoughts and experience has been of great value, and I appreciate all the work they have done in their guidance.

Last but not least, I want to thank family and friends, for their moral support throughout the semester. Without their motivational words, the work with this thesis would not have been the same.

*Edvard Mattias Bø  
Trondheim, June 13, 2022*



# Contents

<b>Abstract</b> . . . . .	<b>iii</b>
<b>Sammendrag</b> . . . . .	<b>v</b>
<b>Acknowledgements</b> . . . . .	<b>vii</b>
<b>Contents</b> . . . . .	<b>ix</b>
<b>Figures</b> . . . . .	<b>xiii</b>
<b>Tables</b> . . . . .	<b>xv</b>
<b>Acronyms</b> . . . . .	<b>xvii</b>
<b>1 Introduction</b> . . . . .	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Research objectives and Research questions . . . . .	2
1.3 Research Design & Research Methods . . . . .	3
1.4 Contributions . . . . .	4
1.5 Thesis Structure . . . . .	5
<b>2 Background Theory and Related Work</b> . . . . .	<b>7</b>
2.1 Literature Review . . . . .	7
2.2 Previous studies . . . . .	8
2.2.1 The role of a dashboard . . . . .	8
2.2.2 Mirroring and advising dashboards . . . . .	8
2.2.3 Teacher dashboard for use with intelligent tutoring systems . . . . .	9
2.2.4 Effects of a teacher dashboard for an intelligent tutoring system . . . . .	9
2.2.5 Temporal learning . . . . .	10
2.2.6 Visualization and interactivity in a Teacher Decision Support System . . . . .	11
2.2.7 Design principles of a dashboard . . . . .	11
2.2.8 Summary of most important findings . . . . .	11
<b>3 Design and Implementation</b> . . . . .	<b>13</b>
3.1 SmartU . . . . .	13
3.2 Functional Requirements . . . . .	13
3.3 Development Tools . . . . .	15
3.3.1 Figma . . . . .	15
3.3.2 Google Drive . . . . .	15
3.3.3 SharePoint . . . . .	15
3.3.4 Github . . . . .	15

3.3.5	React . . . . .	16
3.3.6	React-Vis . . . . .	16
3.3.7	Material UI . . . . .	16
3.4	Iterative Design Process . . . . .	16
3.4.1	Iteration 1 . . . . .	16
3.4.2	Iteration 2 . . . . .	19
3.4.3	Iteration 3 . . . . .	21
3.4.4	Iteration 4 . . . . .	21
3.4.5	Iteration 5 . . . . .	22
3.5	Final User Interface . . . . .	23
<b>4</b>	<b>Methods . . . . .</b>	<b>31</b>
4.1	Study Design . . . . .	31
4.1.1	Participants . . . . .	31
4.1.2	Physical test-setup . . . . .	31
4.1.3	Procedure . . . . .	32
4.2	Data Collection . . . . .	33
4.2.1	User test & observation . . . . .	33
4.2.2	Questionnaire . . . . .	34
4.2.3	Interviews . . . . .	35
4.3	Data Analysis . . . . .	36
4.3.1	Qualitative analysis . . . . .	36
4.3.2	Quantitative analysis . . . . .	37
<b>5</b>	<b>Results . . . . .</b>	<b>39</b>
5.1	Participants & Demographics . . . . .	39
5.2	Descriptives . . . . .	40
5.3	System Usability Score . . . . .	40
5.4	Pearson's correlation . . . . .	41
5.5	Results from Interviews . . . . .	42
5.5.1	In what way does visual analytics help teachers to follow the progress of students who use adaptive assessment . . . .	42
5.5.2	Visualizations and their usefulness . . . . .	44
5.5.3	Participants' reflections on motivation . . . . .	47
<b>6</b>	<b>Discussion . . . . .</b>	<b>51</b>
6.1	Effect of visualizations . . . . .	51
6.2	Useful features . . . . .	52
6.3	Motivation . . . . .	54
6.4	Implications for design . . . . .	55
6.5	Limitations . . . . .	56
6.5.1	Mock data . . . . .	56
6.5.2	Diversity and sample size . . . . .	57
6.5.3	Inexperienced researcher . . . . .	57
6.5.4	Limited interaction with the dashboard . . . . .	57
<b>7</b>	<b>Conclusion and Future Work . . . . .</b>	<b>59</b>
	<b>Bibliography . . . . .</b>	<b>61</b>

<b>A Study</b> . . . . .	<b>65</b>
A.1 Form of Consent . . . . .	66
A.2 Introduction to SmartU . . . . .	69
A.3 SmartU Post-Experiment Questionnaire . . . . .	70
A.4 Observation guide . . . . .	81
A.5 Interview Guide . . . . .	83
A.6 Interviews . . . . .	84
A.6.1 Interview 1 . . . . .	85
A.6.2 Interview 2 . . . . .	87
A.6.3 Interview 3 . . . . .	89
A.6.4 Interview 4 . . . . .	91
A.6.5 Interview 5 . . . . .	93
A.6.6 Interview 6 . . . . .	96
A.6.7 Interview 7 . . . . .	99
A.6.8 Interview 8 . . . . .	101
A.6.9 Interview 9 . . . . .	103
A.6.10 Interview 10 . . . . .	106
<b>B Design, Implementation &amp; Product</b> . . . . .	<b>109</b>
B.1 User tests . . . . .	110
B.1.1 User test guide . . . . .	111
B.1.2 Observations . . . . .	112
B.2 Final User Interface . . . . .	113
<b>C Results</b> . . . . .	<b>117</b>
C.1 Questionnaire & Descriptives . . . . .	117
C.2 SUS-Score . . . . .	118



# Figures

1.1	Oates' research model. The chosen strategies for this study have been highlighted . . . . .	3
3.1	A photograph of the first low fidelity design of the index page . . .	17
3.2	A photograph of the first low fidelity design of the topic page . . . .	18
3.3	A photograph of the first low fidelity design of the detailed overview page . . . . .	19
3.4	A photograph of the first low fidelity design of the question page . .	19
3.5	Displaying the digital prototypes of the four main screens of the dashboard. . . . .	20
3.6	First coded implementation of Course and Question page . . . . .	21
3.7	Displaying the final design of the four main screens of the dashboard.	23
3.8	Color Palette . . . . .	25
3.9	Landing page . . . . .	26
3.10	Course page . . . . .	27
3.11	Detailed overview page . . . . .	28
3.12	Question page . . . . .	29
4.1	Test setup for user test . . . . .	32
5.1	Gender and age of participants . . . . .	40
B.1	Home screen . . . . .	113
B.2	Course screen . . . . .	113
B.3	Detailed overview screen . . . . .	114
B.4	Question screen . . . . .	114
B.5	Edit question screen . . . . .	115





# Tables

4.1	Examples of events/actions to note during observation . . . . .	34
4.3	Categories in the questionnaire . . . . .	35
5.1	Participants' age . . . . .	39
5.3	Questionnaire mean variables of categories . . . . .	40
5.5	Average SUS-score per question. . . . .	41
5.7	Results of Pearson's product-moment correlation test . . . . .	42
5.9	Useful features . . . . .	44
5.11	Participants thoughts on motivation . . . . .	48
C.1	Overall evaluation of the usability of the Dashboard (OEUD) . . . .	117
C.3	Usefulness of the Dashboard (UOD) . . . . .	118
C.5	Attitude towards the Dashboard (ATD) . . . . .	119
C.7	System Usability Score (SUS) . . . . .	120



# Acronyms

- ATD** Attitude towards the Dashboard. xv, 35, 40–42, 51, 54, 119
- IDI** Department of Computer Science. vii
- ITS** Intelligent Tutoring System. 9
- LA** Learning Analytics. 8
- LAAM** Learning Analytics Acceptance Model. 10
- MVP** Minimum Viable Product. 14
- NSD** Norwegian Centre for Research Data. 34
- NTNU** Norwegian University of Science and Technology. vii, 1, 56
- OEUD** Overall evaluation of the usability of the Dashboard. xv, 35, 40–42, 51, 117
- SUS** System Usability Score. xv, 35, 40–42, 51, 54, 59, 120
- TA** Teaching Analytics. 8
- TAD** Teaching Analytics Dashboard. 2, 5, 11, 56
- TLAV** Temporal Learning Analytics Visualization. 10
- UOD** Usefulness of the Dashboard. xv, 35, 40–42, 51, 54, 118



# Chapter 1

## Introduction

This chapter provides an overview of the thesis, starting with the motivation behind conducting the research, the research objectives, and the research questions. Next, it presents the research design and research methods for addressing the defined goals and concludes with the contributions of the research and the outline and structure of the thesis.

The thesis begins with a review of the results from previous studies reported in the relevant literature on learning analytics dashboards. Next, it presents wireframes sought to explain the implementation of a teacher dashboard onto SmartU, a system developed by Papamitsiou [1] and further utilized and adapted by Westermoen and Lunde [2]. Then, it describes the design and implementation of the teacher's dashboard developed for SmartU and the evaluation study conducted to assess the application, along with the results and conclusions.

### 1.1 Motivation

SmartU is an abbreviation for *Self-assessment Measured with Analytics on Run-Time for YOU*. Papamitsiou created it during her Ph.D. research, [1] and it was further utilized and adapted as a master thesis at NTNU [2] that showed promise as a Learning analytics dashboard [3]. It focuses on the needs of the students and what they need to learn as much as possible with an adaptive and personalized learning experience. What this system is missing is an interface for teachers.

For a system like this to function properly, it is equally important to provide good functionality for the teachers as well as the learners. Therefore, this thesis will focus on how a dashboard can facilitate teachers to help them provide quizzes and questions of good quality. Furthermore, by providing data and insight into how the questions are used, the teachers are able to monitor their questions and make alterations when needed.

## 1.2 Research objectives and Research questions

This thesis will describe the planning, development, and testing of a Teaching Analytics Dashboard (TAD) for the learning analytics application SmartU. The system builds upon an older version, called LAERS [4], which has been further developed by Westermoen and Lunde [2]. They looked at the effect of the use of learning analytics on learners' motivation, attitudes, and performance in a learning activity.

For the present research, a visual analytics dashboard has been designed, developed, and evaluated. The goal of this dashboard is to provide insight to the teacher with regards to how the students, quizzes, and questions in the system are doing. It will be important to give the teacher information about how students are interacting with the system, in terms of how much time they use, their performance, their effort, and how much time they spend watching meta-data. Additionally, the dashboard will provide tips and feedback to the teacher to help them notice things that require the teachers' attention. For instance, if questions are too easy or too difficult. It could also be that students finish the question too quickly or spend too long time on it. Either of these factors would indicate that the question may need some revision. The goal of the dashboard is to aid the teacher in finding these factors.

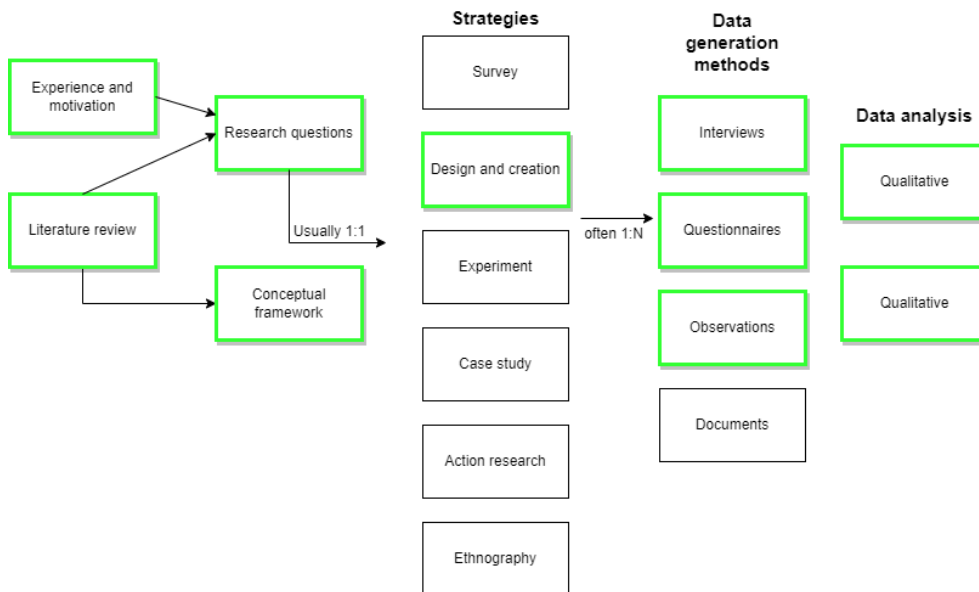
The SmartU system is built to help students learn and improve within courses. It is meant to help students understand course material and works as a hub where they can monitor their progression and how well they have understood the concepts of the selected course. This way, it becomes a system students can use as a benchmark to monitor their skill level. Therefore, it makes more sense for this dashboard to focus on how the *questions and quizzes* are doing rather than if the students are learning and if they have understood the material. Although the performance of the questions and the students is tightly connected, and there will be a focus on how the students are learning as well, it is an important distinction. For the teacher using this system, it is equally important to provide quizzes and questions that help the students monitor their progress, as it is to follow up on how they are learning.

This research aims to look at how visualizations can help teachers follow students' progression, the effect the visualizations have on teachers' motivation for assessing, and to detect how visualizations provide the most valuable insight. To verify the hypothesis, this study attempts to answer the following three questions:

1. How can visual analytics help teachers to follow the progress of students who use adaptive assessment?
2. How do visualizations provide useful insight?
3. What are teachers' motivations towards systems that provide visualizations?

### 1.3 Research Design & Research Methods

Before the study could be conducted, the research had to be planned, and methods had to be decided. The planning was conducted in accordance with the research process described by Oates [5]. The selected research methodology can be found in Figure 1.1, where the research model of Oates has been used to outline the methods used for this research. First, a thorough literature review was conducted to form a solid foundation of knowledge on the relevant topic. The literature review provided an insight that created a reliable basis for formulating the research questions.



**Figure 1.1:** Oates' research model. The chosen strategies for this study have been highlighted

#### Research Strategy

For this study, the design and creation strategy was chosen as the research strategy. It focuses on developing a new IT product, such as a teacher dashboard and can offer various elements as a contribution to knowledge. One of them is instantiations, a working system that demonstrates constructs, models, methods, ideas, genres, or theories that can be implemented in a computer-based system [5, p. 108]. The design and creation strategy is a problem-solving approach. It involves an iterative process involving five steps: awareness, suggestion, development, evaluation, and conclusion [5, p. 111]. These steps were kept in mind while creating and developing the dashboard, explained in further detail in Chapter 3.

Other research strategies were also considered but ultimately not deemed fit-

ting for this study. Neither ethnography nor case study seemed relevant for this type of research. Action research was considered, but it was considered more important to conduct tests in a controlled environment to focus on the research questions defined. A survey would have required a larger pool of participants and regarded out of this project's scope. Finally, experiment was evaluated as a possible research strategy for this study. However, there was no experimental and control group to measure any effect before and after introducing a new element. Therefore, it was deemed unrelated to this thesis.

### **Data generation methods**

This study collected data through observation, questionnaires, and interviews. Therefore a triangulation of data generation methods was used [5]. Thus, both qualitative and quantitative data were collected for analysis, allowing for a broader basis for evaluation. Additionally, it provided the opportunity to cross-check the findings from one method to another, to validate the findings, as emphasized by Oates [5]. Observations were used during the iterative design phase and the final user testing to help understand if the users found the system intuitive and comprehensible. The questionnaire was given to the participants immediately after completing the user test. It represented the quantitative data, and helped capture the participants' attitudes and motivation toward the dashboard. The interviews were conducted immediately after the questionnaires. They provided qualitative data and gave an understanding of what elements provided the most useful insight and why.

### **Data Analysis**

As previously mentioned, the data generation provided both qualitative and quantitative data. The interviews followed a semi-structured approach, and were recorded and transcribed into text. Subsequently, they were categorized by being color-coded, according to which research question the quote addressed. The results from the questionnaire were analyzed using R Studio to calculate mean values and standard deviation. Additionally, R Studio was used to test whether there was any correlation between the categories of the questionnaire, using Pearson's correlation test. The qualitative and quantitative analysis is presented in Chapter 4.3

## **1.4 Contributions**

This thesis contributes to the academic community by providing a mixed-methods study that looks at the usability of a teacher dashboard in a smart self-assessment system. It combines qualitative and quantitative empirical results to highlight the strengths and weaknesses of the system, to present what kind of visual analytics



teachers find most insightful, and what effect these visualizations have on their motivation. Combined, these elements can provide insight into what information teachers find most useful in a TAD, which can improve their teaching skills and the learning outcome of students. Furthermore, this could open up new directions and discussions for future work.

## 1.5 Thesis Structure

This thesis consists of seven chapters and is structured as follows:

- **Chapter 1: Introduction** - introduces the thesis by explaining the problem description, the research questions, the contributions, and the methods planned to be used to address the research questions.
- **Chapter 2: Background Theory and Related Work** - Presents the process and findings of the literature review.
- **Chapter 3: Design and Implementation** - Explains the process of creating the Teaching Analytics Dashboard (TAD), presenting the various iterations and the thoughts behind the decisions.
- **Chapter 4: Methods** - Explains the methods used in this study by elaborating on the study design, data collection, and data analysis.
- **Chapter 5: Results** - Presents the findings of the quantitative and qualitative analysis.
- **Chapter 6: Discussion** - Discusses the findings against each other and looks at how it compares to related studies.
- **Chapter 7: Conclusion & Future Work** - Presents the conclusion of the thesis and main points of the discussion while looking forward to highlight future possibilities.



## Chapter 2

# Background Theory and Related Work

Prior to conducting any research in an academic setting, an understanding of the current body of knowledge should be established [5]. This understanding ensures that the research will contribute positively to the academic community by providing new and relevant knowledge concerning State-of-the-Art. On that basis, this chapter presents a review of related work in the area of learning analytics dashboards for teachers and teaching analytics.

### 2.1 Literature Review

Before any research questions or wireframes could be implemented onto SmartU, a literature review of the State-of-the-Art had to be conducted to find out what had been done previously, what had been deemed helpful and valuable, and what had not. To find papers to review, the scientific paper databases of Google Scholar, Scopus and ERIC were searched. Subsequently, the snowballing method was applied. Snowballing refers to using the reference list of a paper to identify additional papers [6].

**Search terms used:** *Teacher Dashboard, Learning analytics, Teaching analytics*

The search terms were chosen to cover a wide range of studies surrounding teacher dashboards. These searches yielded many results, so the first fifteen for each query were scanned for relevance to reduce the number of papers needed to review. Next, these papers were filtered out based on relevance to the research topic. After filtering out, 19 articles were considered more central. These were reviewed and classified in a spreadsheet, categorizing what they were about and their main findings and contributions. Finally, some new articles were added out of the snowballing methodology, where some of the articles led to other interesting articles. Ultimately, the method raised the total amount of reviewed articles to 25.

## 2.2 Previous studies

This section describes the relevant previous studies within the field of LA, TA, and dashboards. It focuses on which visual analytics teachers have found useful earlier and how to display the information most effectively.

### 2.2.1 The role of a dashboard

Several literature reviews have been conducted on learning analytics [7–10]. These studies have looked at how studies have conducted their research and their strengths and weaknesses.

The literature review conducted by Schwendimann et al. [9] observed that research on the effects of learning dashboards is still in its early stages. They highlight the importance of providing insight also when developing a dashboard for future research. Their request is that researchers explain what technology has been used, provide insight details about the educational context, describe what constructs were the target of the evaluation, and give insight on the learner. By providing these details, future research better understands what has been done, helping further develop the field and supporting comparison between studies.

Durall and Gros [11] argue that one of the main focuses of learning analytics research is to empower teachers and learners to make informed decisions about the learning process, mainly by visualizing the collected learner data through dashboards. Yoo et al. [8] claim that an educational dashboard allows teachers to know students' learning status in real-time and in a scalable way. Furthermore, they argue that the dashboard may play a role in motivating students, improving their self-directed learning ability, and helping them achieve their learning goals effectively.

These arguments emphasize the importance of displaying data understandably, so the teachers are able to extract necessary information as effortlessly as possible and help them make informed decisions. Yoo et al. [8] also emphasize the importance of data visualization, explaining that it converts the abstract and complex to the concrete and visible by amplifying human cognition. This point indicates that the primary focus should be to use visualizations to present the data.

Verbert et al. [10] pointed to how not all teachers feel the need for the same information. It is nearly impossible to create a dashboard that fits all, pointing out how one size does not fit all. This argument emphasizes the importance of making information available in different formats.

### 2.2.2 Mirroring and advising dashboards

In a study from 2020, van Leeuwen et al. compared teachers' use of mirroring and advising dashboards [12]. They explain that a mirroring dashboard provides information about learners but leaves all subsequent detection and interpretation of relevant information to the teacher. An advising dashboard is described as a

dashboard that provides alerts about learners that may require support as well as advice about what problem a learner could be facing. Their research aimed to measure how well teachers could detect problematic groups with the two different dashboards.

The experiment was done by dividing 35 primary school teachers into two groups, where one group used the mirroring dashboard, and the other used the advising dashboard. They observed eight situations, which simulated groups in 4th grade working with fractions, and were asked to identify groups who struggled or were problematic.

The study showed that the teachers who used the mirroring dashboard on average identified the problematic group 6.65 times out of 8, compared to 7.50 in the advising condition. Additionally, the mirroring group reported a significantly higher cognitive load and a lower confidence level in their decision-taking compared to the advising group. This research shows that simple visual cues can have a considerable effect, giving better results and providing a less effortful and more confident decision-making process.

If the teacher can rely on the system to perform the initial assessment, it relieves the teacher of processing all information, thus saving mental effort.

### **2.2.3 Teacher dashboard for use with intelligent tutoring systems**

Aleven et al. created a dashboard for teachers who use Intelligent Tutoring System (ITS) in their classrooms [13]. Their experience was that ITSs often are not designed to involve teachers. They argue that ITSs have led to improved learning for students but could be more effective if they were designed to provide data to teachers so they could help their students as well. What they found in their study was that teachers like to have power over the dashboards and their decisions. Thus, they would not prefer having the dashboard have complete control over the students.

The study found that the best solution was to provide a high-level summary of actions that require the teachers' attention the most. At the same time, they emphasized the importance of providing insight into how these conclusions have been reached. This is also supported by van Leeuwen et al. [12]. They look at what information a computer-supported collaborative learning teacher dashboard should provide. In their research, they point out that teachers based their actions on the suggestions they got from their system but that they were also very interested in seeing how all the groups they were observing were doing. This research emphasizes the importance of both suggestions and provides a deeper insight into statistics.

### **2.2.4 Effects of a teacher dashboard for an intelligent tutoring system**

Khakaj et al. [14] built upon the work of Aleven et al. [13] and looked at what the effects were of a dashboard in an ITS. They focused on a scenario where a teacher

uses a dashboard when preparing for a class session to see if it helped in focusing the class discussion on the topics most in need of discussion. They also investigated the influence this dashboard had on student learning. To investigate their hypothesis, they used a program called Luna, a high-fidelity dashboard prototype.

In their study, they conducted an investigation where five 7th grade teachers prepared for classes. One time they used Luna in their preparation (experimental condition), and one time they did not use it (control condition). There were nine classes in the control condition and 8 in the experimental condition.

They found that the teachers referenced the information gained from Luna on average 12.6 times per class. The results showed that the information affected the teacher's knowledge about the class, both overall and on an individual level. They also found that Luna prompted changes in the teachers' lesson plan, especially knowledge about where students were struggling. Teachers incorporated 44.6% of the statements they learned from the dashboard in their lesson plan. The study also showed that teachers focused more on what the students did not master and the students who did not do so well rather than what the students did master and the students that did well.

They concluded that the dashboard's information was helpful to the teachers as they prepared for a class session. The information provided by the dashboard affected the teacher's knowledge, lesson plans, and what they covered in the class session.

### **2.2.5 Temporal learning**

In 2015 Papamitsiou and Economides [15] looked at how temporal learning analytics visualizations could be helpful in gaining insight into students' perceptions and discover how they behave when dealing with assessment tasks. They wished to shed light on how students learn, reveal learning habits and learner-strategy patterns, and provide a deeper understanding of learning mechanisms in general. They emphasized how teachers' awareness of what and how learners are doing is essential to assess learner progress.

To perform this research, they created a Temporal Learning Analytics Visualization (TLAV) tool, which obtained necessary temporal information from another assessment environment which consists of a testing mechanism and a tracker that logs the data. They asked 32 secondary education teachers from different domains to evaluate the TLAVs. The Learning Analytics Acceptance Model (LAAM) was used for the teachers to assess the system. They had the teachers answer a questionnaire, where they rated the TLAVs effectiveness, perceived usefulness, and ease of use. They found that temporal data is perceived as easy to use and valuable for teachers.

### **2.2.6 Visualization and interactivity in a Teacher Decision Support System**

Mottus et al. [16] wrote an article in 2013 where they pointed to important factors to help teachers support students in an online learning environment. They wrote about what a teacher dashboard should include to empower teachers to make well-informed educational decisions.

The article points out the significant challenge of displaying complex information in a visual and accessible format and how it is particularly difficult to display for people who do not work with analytics on a day-to-day basis, like teachers or professors. The article also warns about how easy it is to get caught up in selecting visually stimulating data that does not necessarily provide as much practical information.

Finally, the article emphasizes the importance for the professor to be able to “see” the students by means of a page or data visualization that represents them. This point is especially helpful for teachers who may never be directly in contact with their students.

### **2.2.7 Design principles of a dashboard**

In 2013, Few [17] developed a set of design principles for dashboard design. He emphasized the importance of displaying important information so it can be monitored at a glance. Regarding visual perception, three considerations are highlighted in his design principles. First of all, humans have limited working memory. Thus, only three of four chunks of visual information can be stored at a time. Therefore, to design a good graphical pattern, graphs and visualizations are better for efficient perception and memory retention rather than numbers. Second, visual aids such as color, spatial position, and form should be utilized appropriately to ensure rapid perception. Finally, he points out the importance of following Gestalt’s design principles regarding similarity, continuity, closure, proximity, perception, organization, and symmetry.

These are useful principles to keep in mind during the dashboard development. As pointed out by Mottus et al. [16], it is difficult to present complex data understandably. However, these suggestions are meant to help tackle these difficulties.

### **2.2.8 Summary of most important findings**

The literature review provided helpful information by highlighting valuable aspects to keep in mind and be wary of while developing a dashboard. Schwendimann et al. [9] emphasize how important it is to report what has been done properly. By properly reporting, the academic community within TAD will easier be able to benefit from this research.

As emphasized by Papamitsiou and Economides [15], it will be important to include temporal data for the teachers to use in their evaluation. Additionally,

a high-level summary of actions that require the teachers' attention the most, and providing insight into how these conclusions have been reached, will both be important to include, as pointed out by Aleven et al. [13] and van Leeuwen et al. [12]. Finally, it will be important to create a system that can deal with the initial assessment, to save mental effort for the teacher.

By following the principles laid down by Few [17], of designing a dashboard which is understandable and easy to use, it should be possible to create a dashboard that empowers teachers and learners to make informed decisions about the learning process, mainly by visualising the collected learner data through dashboards, as described by [11]. At the same time, the design principles should help avoid creating a complex dashboard, which is difficult to understand, as Mottus et al. [16] and Yoo et al. [8] warn about.



## Chapter 3

# Design and Implementation

This chapter gives an overview of the process of designing and developing the interface for the teacher dashboard of SmartU, according to the Design and Creation strategy. First, the design of the SmartU interface version developed by Westermoen and Lunde [2] will be presented, in Section 3.1 SmartU. Next, describing the requirements for designing the new interface in Section 3.2 Functional Requirements Functional Requirements. Section 3.3 Development Tools describes which tools were used in the project, while Section 3.4 Iterative Design Process explains the iterative process of designing the user interface. Finally, Section 3.5 Final User Interface gives an introduction to the current teacher dashboard by showing the various screens and explaining the thoughts behind the elements.

### 3.1 SmartU

As mentioned in the introduction, SmartU is an abbreviation for *Self-assessment Measured with Analytics on Run-Time for YOU*. It was initially developed by Papamitsiou [1] during her Ph.D. research on learning analytics and later modified and re-configured by Westermoen and Lunde [2]. The teacher dashboard developed during this master thesis is created independently from the existing SmartU learner application. The present application only utilizes some of the data generated previously during the studies conducted for the evaluation of the SmartU learner interfaces. The data was utilized here as the content to be used in the application's backend and displayed on the frontend as visual analytics for the teachers. In creating the designs and layouts for this dashboard, the designs of SmartU were used as inspiration to attempt to give a consistent feeling so that the two projects could one day merge seamlessly.

### 3.2 Functional Requirements

*Functional requirements* have been defined in many articles, and there is a broad consensus about how to define it. Glinz [18] splits the definition into two threads.

The first thread emphasizes on functions. A functional requirement describes "a function that a system (...) must be able to perform" [19], and "what the product must do" [20], "what the product should do" [21]. The second thread focuses on behavior, where the requirements describe the inputs to a system, the outputs, and the relationship between them [22]. This thread focuses more on what a system should do given a specific input.

Non-functional requirements focus on the performance characteristics of the system, making it a different type of requirement [18]. However, because this thesis aims to be used in a controlled environment, these types of requirements are not included.

The functional requirements were split into three categories, in accordance with the article to Bradner [23]. Those three are requirements that must, should, or may be implemented in a complete prototype. This way, the features needed for a Minimum Viable Product (MVP) are separated from the features that would expand the MVP, or the features that would have been intriguing to implement.

The requirements are based on the findings from the literature review. As mentioned in Section 2.2.8, it was important to give the teacher a high-level overview, a system that can handle the initial assessment while also providing deep insight into how these suggestions have been made. In addition to providing insight, it is important to follow the design principles of Few [17] to ensure that the dashboard is user-friendly and understandable. The requirements are the following:

*Requirements that **must** be implemented:*

- Teachers must be able to get statistics of students' performance
- Teachers must be able to see visualized data regarding a specific question
- Teachers must be able to see visualized data regarding an overview of the questions in a course.
- Teachers must be able to see visualized data of all questions in a course.
- Teachers must be able to edit a question if needed.
- The dashboard must provide suggestions to the teacher.
- The dashboard must be aesthetically pleasing and user-friendly on a standard desktop setup (e.g., using Google Chrome on a 1920x1080 display)

*Requirements that **should** be implemented:*

- Teachers should be able to reset statistics for an edited question if necessary.
- Teachers should be able to apply filters and view all questions based on what they find interesting or useful.
- Teachers should be able to either accept or decline the suggestions provided by the system.

- The dashboard should support multiple available courses.

*Requirements that **may** be implemented:*

- Allow for user creation and log in
- Allow teachers to create new question in a course
- Allow teachers to create a new course.

### 3.3 Development Tools

To create and design the teacher dashboard, it was important to find and use proper tools to make the process as good and as effortless as possible. This section describes which tools have been used in the design and implementation phases.

#### 3.3.1 Figma

Figma [24] is a design and collaboration tool to help teams design and create wireframes for a project. It allows for real-time collaboration and the creation of functioning prototypes. Figma enabled an easy and efficient progression in the design phase, making it effortless to create visually satisfying prototypes.

#### 3.3.2 Google Drive

Google Drive [25] is a cloud storage service that lets the user store, collaborate and share files within a project. By having files in the cloud, collaboration with the supervisors have been much easier, allowing them to access material and give feedback without having to meet physically or send files back and forth. This cloud storage was solely used to store plans and ideas, not to store personal data.

#### 3.3.3 SharePoint

Microsoft SharePoint [26] is a part of Microsoft Office 365, which provides safe and secure cloud storage of test results from the study. Due to a data agreement between Microsoft and NTNU, safe, accessible cloud storage was ensured. This storage allowed the data to be stored safely while also being efficiently accessible.

#### 3.3.4 Github

Github [27] is a software development platform, meant for version control. It allows collaboration on software development projects, by managing the source code, and allowing several developers to collaborate on the same files. Due to this project solely being done by one student, it was mainly used to store the code in the cloud, in case something happened to the local computer.

### 3.3.5 React

React [28] is an open-source front-end JavaScript library, for building user interfaces. This library allowed for relatively easy implementation of the design prototypes made in Figma, to be coded into an actual product. It has a lot of built-in functions and libraries, which helps the developer, so they do not have to make everything from scratch.

### 3.3.6 React-Vis

React-Vis [29] is a collection of React components to render data visualization charts. It has been developed by Uber, and allows the developer to create common charts easily. This visualization tool was particularly useful in this project, as it contained so many graphs and visualizations.

### 3.3.7 Material UI

Material UI [30] is a library that lets the developer import and use different components to create a user interface in React applications. This way, the developer does not have to create everything by themselves. Not only does this make the development process faster, but it also ensures that the end product looks good.

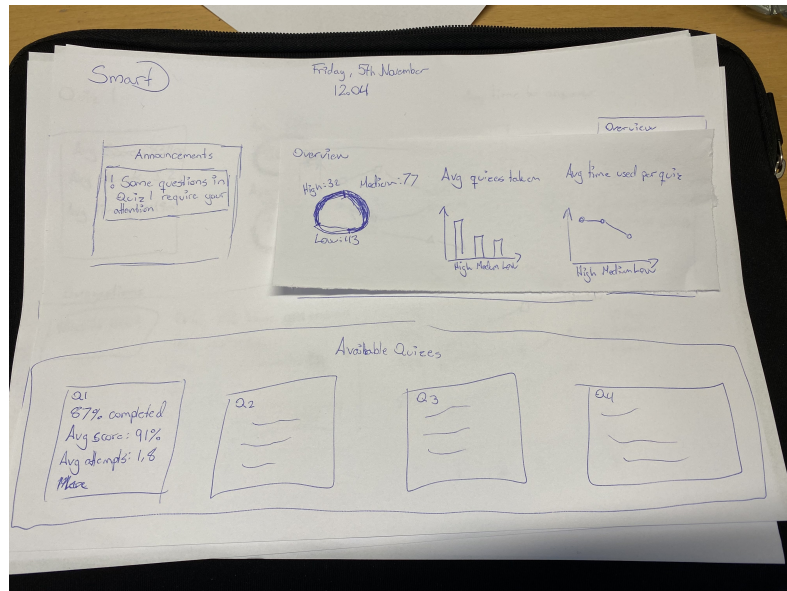
## 3.4 Iterative Design Process

After performing the literature review, it was time to implement what had been read about into low-fidelity designs. As a starting point, the first low-fidelity layouts were created using only pen and paper. These were used to conceptualize and visualize the ideas gained from reading the literature. The flow between the screens was not a significant concern in this phase, but more what kind of information should be accessible, and where it should be accessible.

### 3.4.1 Iteration 1

The index page of the system is shown in Figure 3.1. Here one can see an Announcements box, highlighting certain aspects that could be of interest to the teacher. For example, it could be that a question is too easy or too hard. This element was implemented to comply with the recommendations of Aleven et al. [13] and van Leeuwen et al. [12], who emphasized the importance of providing help to the teachers by showing them where to look. The teacher is taken to the given question to view more detailed information by clicking on the suggestions. To the right, one can see three statistics showing the teacher a brief overview of how the students are doing. Here the teachers can see how many students are in each mastery level, how many quizzes they have taken, and how much time each mastery level group, on average, spends solving one quiz. Below these components, one can see an overview of all the available topics. This overview shows how many

students have completed a quiz within the topic, what the average score is, and more many times students of average have taken a quiz on the given topic. If the user clicks on one of the topics, they are taken to the topic page of the system.



**Figure 3.1:** A photograph of the first low fidelity design of the index page

The topic page shows detailed information about how the students are performing within the given topic and can be seen in Figure 3.2. In the top left corner, the user gets a brief overview of how the students are doing in their quizzes within the topic. It provides information about the average score, the average number of attempts, and the average time spent per quiz. In the top middle, the teacher can see the average effort and performance in the topic, and in the top right, one can see a bar chart displaying more detailed temporal data. The frequent use of temporal data included in the system is based on the proven effects it has had in the research conducted by Papamitsiou and Economides [15].

The graph shows how long time students in each mastery level, on average, spent on each question. It is divided into three, showing how much time they spent answering right, wrong, and total. In the bottom left, the user receives suggestions, like on the index page. These are meant as a heads up to give the teacher help as to know where to look for rooms for improvement, as previously discussed. Finally, the bottom right displays the average score of the different mastery groups per attempt, showing how they improve as they try to take more quizzes within the topic.

Figure 3.3 shows a rough idea of how the detailed overview page is intended to look. The idea is that the top half will be dedicated to displaying visual analytics about all the questions in a course, while the bottom half will be a spreadsheet containing more data about the questions one can see in the top half. The main idea behind this screen is to provide complete insight for the teachers into how the questions are doing. As Alevén et al. [13] pointed out, teachers want to have

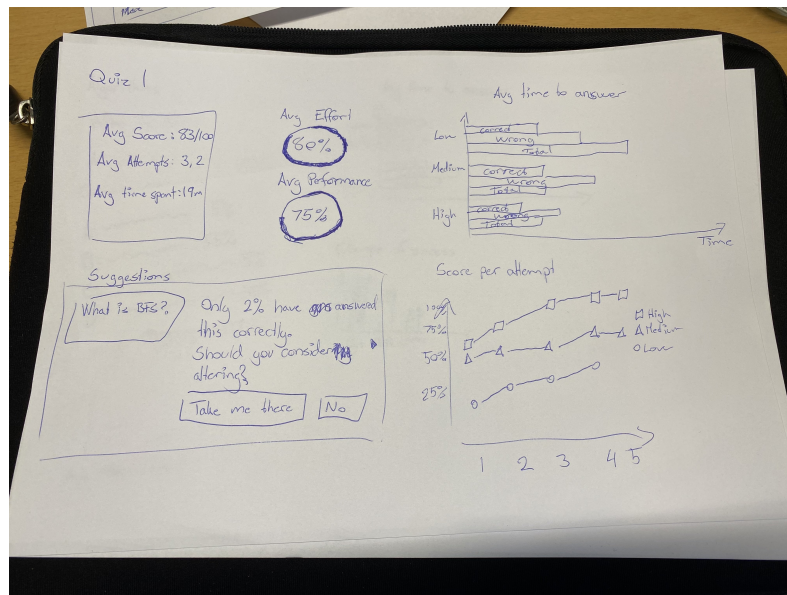


Figure 3.2: A photograph of the first low fidelity design of the topic page

access to the data, not only the suggestions, to give them a sense of control. This screen will also empower the teacher to make informed decisions, as pointed out by Durall and Gros[11]. If the user clicks on one of the questions, they are taken to a page that shows more detailed information about that specific question. The questions are displayed both as a bar chart and a spreadsheet, so the teacher can get an overview in the bar chart and see more details in the spreadsheet. Furthermore, as pointed out by Verbert et al. [10], not all teachers have the same preferences. Some may prefer the spreadsheet, others the bar chart. This way, the data is displayed to please both preferences.

Figure 3.4 shows the first iteration of how the user will see data about a specific question. Here they will be able to see the question, the alternatives, and how many have selected the various answers. In the top right, they will see the bar chart familiar to the one on the topic page, showing how much time each mastery level has, on average, spent answering this question right, wrong and total. In the bottom left, the user can see how much time each mastery level has spent viewing the meta-data. Finally, they can see the calculated chance of each mastery level answering this question correctly in the bottom right.

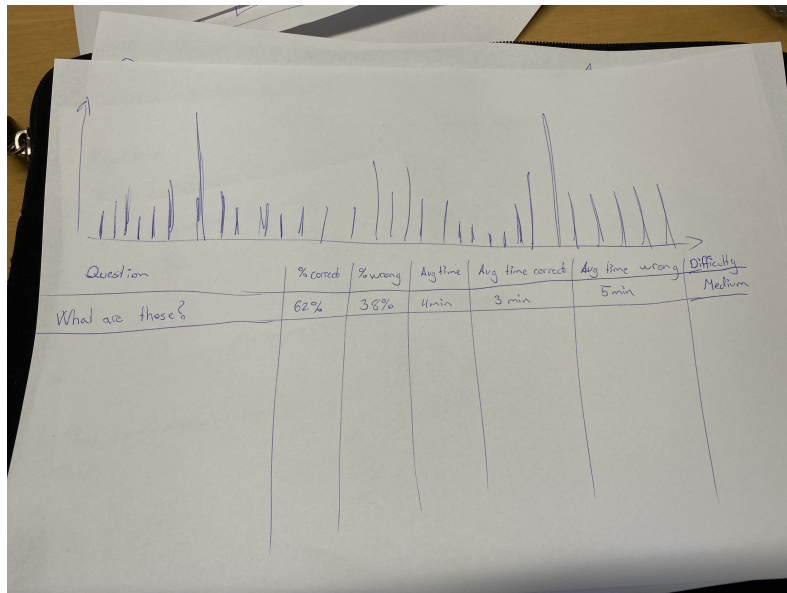


Figure 3.3: A photograph of the first low fidelity design of the detailed overview page

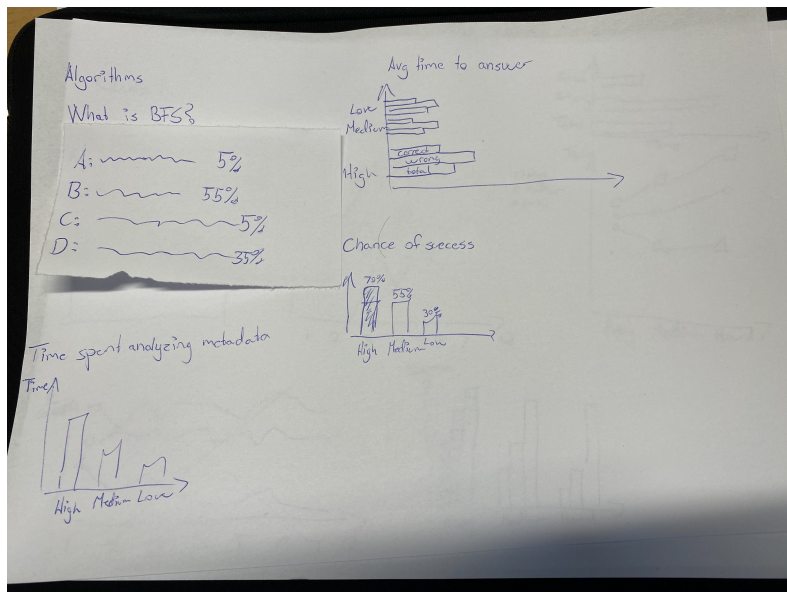


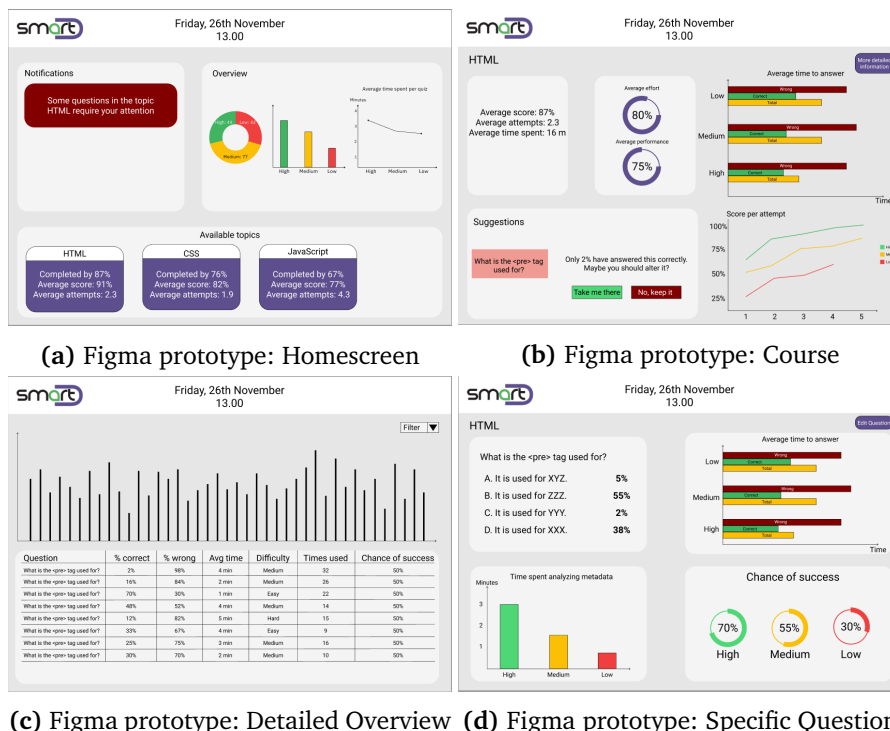
Figure 3.4: A photograph of the first low fidelity design of the question page

### 3.4.2 Iteration 2

After creating the paper prototypes, it was decided to develop the draft further digitally. This was created using Figma [24]. When creating the screens in Figma, inspiration was taken from the designs of Westermoen and Lunde [2]. As the existing SmartU learner interfaces and the new teacher interface are meant to work as an integrated system, it was deemed useful to have a design that was

consistent. Four screens were created and can be seen in Figure 3.5. These were created with the previously mentioned paper prototypes as templates. Although a lot is similar, some changes were made after discussions with professors and peers. While creating these, the design principles of Few [17] were taken into account to ensure that the visualizations were presented as comprehensible as possible.

On the index page, the announcements were changed to be called "Notifications", as this was a more descriptive term. In the far top right of the topic page in Figure 3.5b there was added a button saying "Take me to detailed info". This button takes the user to the screen shown in Figure 3.5c, which shows a more detailed view of all the questions within the topic. On that screen, a filter functionality was added. The questions available will be based on filters, where one can, for example, choose to see the questions based on different metrics, like times used or success chance. The spreadsheet will show the question, and necessary information like how many got it right, average time spent, difficulty, and times used in quizzes. The screen viewing a single question (shown in Figure 3.5d) has also been reworked to fit the screen better, so each component takes up roughly a quarter of the screen. A button in the top right corner has also been added, which will allow the teacher to alter the given question, which was designed later on.



**Figure 3.5:** Displaying the digital prototypes of the four main screens of the dashboard.



### 3.4.3 Iteration 3

Once the digital prototypes were in place, it was time to implement the prototypes in software. To implement the prototypes, ReactJS [28] was used, which is elaborated in further detail Section 3.5. Because visualizations were a substantial part of the dashboard, a thorough process was made to find a suiting third-party tool that could handle all visualizations. Furthermore, the design contained many different visualization types, so it was important to find a tool that supported this, while also being well documented and easy to use. On that basis, React-Vis [29] was the best fit for this project (more details can be found in Section 3.3.6). Once starting to implement the dashboard, some limitations with React-Vis were encountered. The most notable was that it did not support a horizontal view of grouped bars in a bar chart, as Figure 3.5b and Figure 3.5d display with *Average time to answer*. Therefore, these were altered to display the data vertically. Also, the radial chart did not have a good way of displaying the rest of the percentage with a thin line, as shown in Figure 3.5b and Figure 3.5d. Therefore it was decided to remove this detail, as shown in Figure 3.7b and Figure 3.7d.

Additionally, it was made clear that the digital prototypes had no clear way of maneuvering from the detailed overview (Figure 3.5c) to the specific question (Figure 3.5d). The idea was that it should be possible to go to all of the specific questions from there to view more info. Therefore, a button was added next to each question in the table, so the user could go to each question from there, as shown in the final design in Figure 3.7c. Figure 3.6 shows the first implementation of the course and question page.



Figure 3.6: First coded implementation of Course and Question page

### 3.4.4 Iteration 4

After successfully implementing a version of the dashboard, parts of it were tested (pre-evaluated) with real users to get feedback and thoughts. These users were friends, three in total, who all study informatics. The tests were conducted by showing them screens of the four pages, asking them where they would look to solve specific tasks, and asking them to provide initial thoughts on the design.

Responses from the user tests were that the design was intuitive and clear. The users liked the structured layout and thought it was mostly easy to find informa-

tion. However, all participants thought the detailed overview was a bit confusing. They did not understand that the lines represented a bar chart (Figure 3.5c) and wished there was a hover effect on it to make things easier. In general, there was too much information there, and it was not presented clearly. Two of them also commented on the question page (Figure 3.6b), where they thought it looked too empty. One of the participants said: *"It felt like there was a bit too much air there, maybe you could show some other statistics there?"*. Finally, all of the participants also commented on the summary box on the course screen (Figure 3.6a). They thought it was a bit boring to have that statistic represented as text and suggested finding something more in style with the rest of the dashboard to display there. Additionally, the users did not understand what Effort and Performance meant, so this did not provide any valuable information.

### 3.4.5 Iteration 5

The feedback provided valuable insight, and adjustments were made to the dashboard based on the suggestions. The bar chart was revised to display actual bars, as can be seen in the final version in Figure 3.7c. Furthermore, a hover effect was added to each bar, so when the user hovers over a bar, they will see which question it refers to and information about how many answered correctly and wrong. This implementation provided more coherence between the bar chart and the spreadsheet.

Additionally, revisions were made to the page, which showed detailed information about a single question (Figure 3.7d), as was remarked in the user test. Therefore, another element was added in the top row, between the question and the average time to answer. This element displayed information about how many times the question was used, difficulty, performance, and effort.

Finally, some adjustments were made to the summary box on the course screen (Figure 3.7b). To comply with the feedback from the user test, the element in the top right now showed info as a text instead of just listing it up. Additionally, the information here was changed, because it became clear that the original information was available for the users in other places on the screen (i.e., performance is shown in the pie chart to the right, and average attempts is shown on the landing page). Therefore, a pie chart was added, showing the course's distribution of hard, medium, and easy questions. Furthermore, because the user testers did not understand Performance and Effort, small question marks were added next to them. This implementation gave the user the opportunity to hover over the question marks, to get information about what they mean.

### 3.5 Final User Interface

After successfully iterating through the design process, the design was good enough for testing for this study. In the end, the finalized teacher dashboard interface consisted of five frames which can be seen in Figure 3.7. Detailed images of all the screens can be found in Appendix B. In the following subsection, design choices and thoughts for the final user interface are discussed.

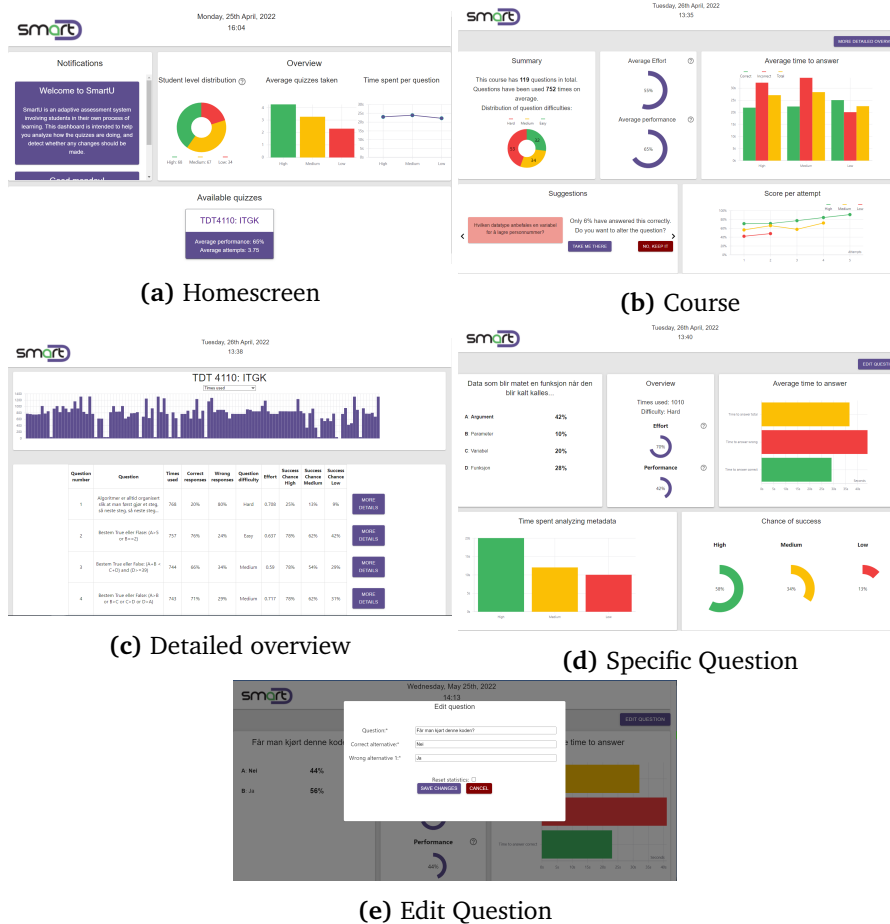


Figure 3.7: Displaying the final design of the four main screens of the dashboard.

#### Technical details

Due to this project's scope, a functioning backend was not set up. This decision had to do with the fact that it was not strictly necessary for the hypothesis to be tested. By having access to a JSON file containing necessary data about student activity in a course from an earlier study, it was possible to create a dashboard and load data from the JSON file locally. However, it was aimed to develop everything

as dynamic as possible so that if the dashboard was to be merged with a backend one day and more courses were added, it would be as seamless as possible.

The data from the JSON file was loaded upon entering the site and stored in local storage. The application checks if there are any values in the local storage. If there is not, it loads in the required data. This way, the user would only alter the data in the local storage, not the JSON file itself. This decision was made to make testing as effortless as possible. The user is asked to edit some questions and reset statistics in the user test. By loading in the same JSON file for all participants, one ensures that all the users get the same experience and alter what they want without the researcher being concerned about how this affects the next participant.

To create a visually stimulating and easily maintainable user interface, ReactJS [28] was used, as previously mentioned. Using React, it was possible to create smaller components within each page, which were highly customizable and could be updated individually. Additionally, it made it easy to edit components without worrying about affecting the rest of the page.

## Colors

As previously mentioned, this dashboard was designed with the design of Westermoen and Lunde's [2] SmartU system in mind. Therefore, the colors chosen in this dashboard are mostly the same as they used to ensure consistency. The following is an explanation of the colors and what they represent in the interface:

- Red (#F03F3F) was primarily used to represent elements of "negative" character, such as time to answer wrong, low mastery level, and hard questions.
- Yellow (#FBBF05) was used as a "medium achieving" color in terms of graphs and performance and also used to display the total time spent answering questions.
- Green (#40B461) was used to promote positive feelings, such as excellent performance in terms of graphs and statistics, time to answer correctly, and easy questions.
- Dark red (#820000) was used to visualize canceling a process or not doing an action. It was used to close the "Edit question" screen as a means of canceling and not changing and to indicate not to change a question that was suggested (Figure 3.7b).
- Purple (#5D4E8E) was the primary color of the system used throughout the application in either navigation bars or other elements needing contrasts.
- Pink (#EF9A93) was used as a secondary color, to display the question card in the suggestions (Figure 3.7b).

- White (#FFF) was used as the background color on all boxes on all screens.
- Grey (#E6E6E6) was used as the background color of all screens.

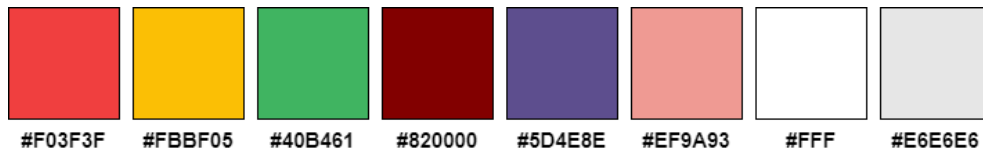


Figure 3.8: Color Palette

## Landing page

Figure 3.9 shows the landing page of the dashboard. This page provides teachers with a general overview of how students are doing. It shows them how students are distributed among the three mastery levels (i.e., novice, intermediate, and advanced), how many quizzes on average the students in the three mastery levels have taken, and how long time on average they spend. Next to the student mastery level distribution, there is a question mark icon, which the user can hover over. This functionality gives the teacher information about the student's mastery level and what it means. The landing page also shows notifications for the teachers to get notified. Although, for now, there are not any notifications, however, the idea is that administrators can send notifications to teachers, or students can send messages to teachers. Then they get notified via this functionality. Also, if the teachers have quizzes in several courses, this functionality can notify them about available suggestions in each course. Finally, on the bottom section of the page, the teacher can see all courses they are affiliated with. The courses are clickable and take the user to the respective course page. For now, there is only one course. If there were more courses, they would appear next to the current course. If it overflows, a slider will appear so the teacher can scroll sideways.

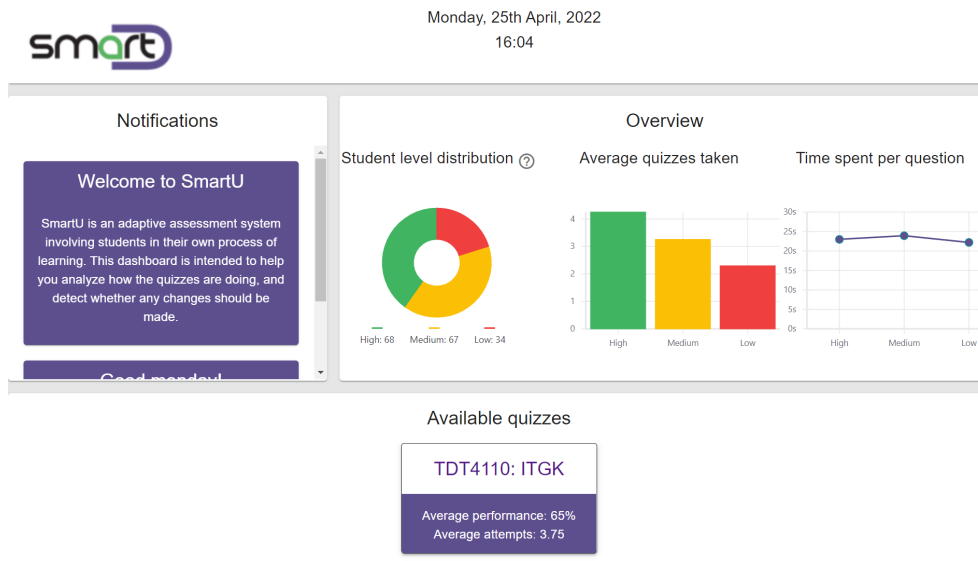


Figure 3.9: Landing page

### Course page

If teachers click on a course on the landing page, they are taken to the affiliated course page (Figure 3.10). Here, the teacher is presented with detailed information about how the quizzes generated for this course are doing. In the top left, they are presented with a summary. It shows teachers how many questions the course has, how many times they have been used, and how the questions are distributed based on their difficulty (i.e., easy, medium, hard). In the top middle, the teacher can see the average effort and performance of the students. Because some users (during the pre-evaluation testing) were unsure what effort and performance meant, small question mark icons have been added next to those features as well, similarly to the student mastery level on the landing page. If teachers hover over the question mark icons, they will get a textual explanation of what the displayed information corresponds to. A bar chart in the top right shows how much time, on average, students spend answering a question. The use of temporal data visualization is based on the research of Papamitsiou and Economides [15]. It is split into the mastery levels and shows how long they spend answering correctly, wrongly, and on average.

In the bottom left, the teacher receives suggestions from the system. These work as simple visual cues to help the teacher know where to look, as suggested by van Leeuwen et al. [12]. Here they are shown the question and a small text explaining why they advise the teacher to look into it. The teacher can then choose to look into it or keep the question as it is. If they click "Take me there" they are taken to the question screen for that question. If they click "No, keep it", the suggestion will be removed. If they have clicked through all suggestions, the box will say: "Things are looking good, there are no suggestions in this course". The

bottom right shows statistics of how the students score per attempt, showing their development over time.

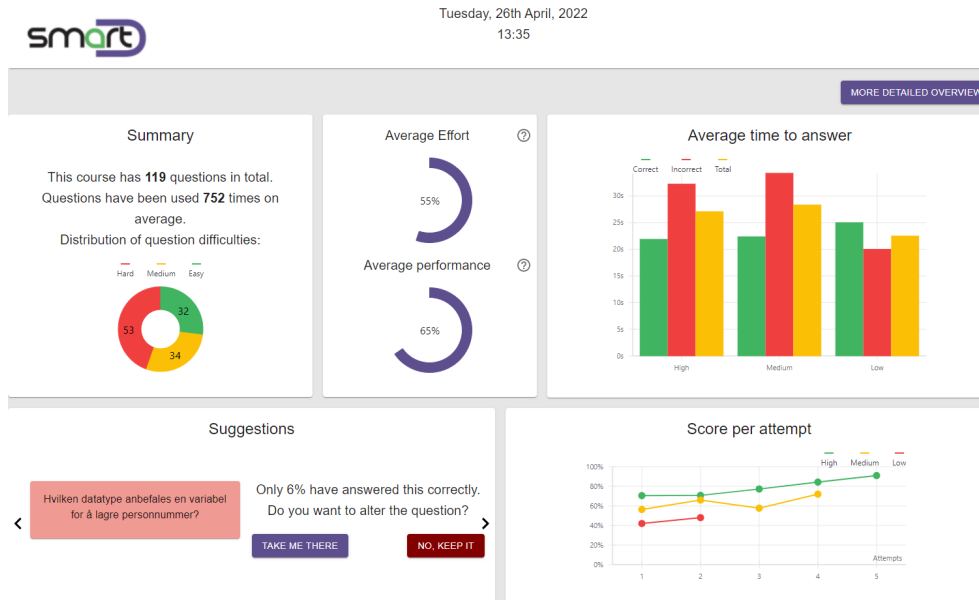


Figure 3.10: Course page

### Detailed information about all questions

In the top right corner of the course page, there is a button labeled "More detailed overview". This button will take the teacher to a detailed overview of all questions in the selected course, as seen in Figure 3.11. The top section of the page shows the teacher which course it is displaying questions from and includes a bar chart, where each bar represents a question in the course. If the teacher hovers over a bar, a modal will show what the question is and how many have answered correctly and incorrectly. In this section, the teacher can choose different information to view in the bar chart. This information includes the number of times questions have been used, the number of correct responses and the effort put in.

Below the bar chart is a spreadsheet showing all questions in the course. This spreadsheet contains the question number, question, times used, correct and incorrect responses in percent, difficulty, effort, and the success chance for high, medium, and low master levels. Next to each question is a button labeled "More details". If the teacher clicks this button, it will lead to the question page for that specific question.

The detailed overview is presented to give the teacher relevant information about how the questions in the form of various metrics. This functionality was implemented based on the research from Alevén et al. [13] and van Leeuwen et al. [12], who emphasized the importance of providing insight. It is also meant to empower the teacher to make more informed decisions, as stressed by Dural and

Gros [11].

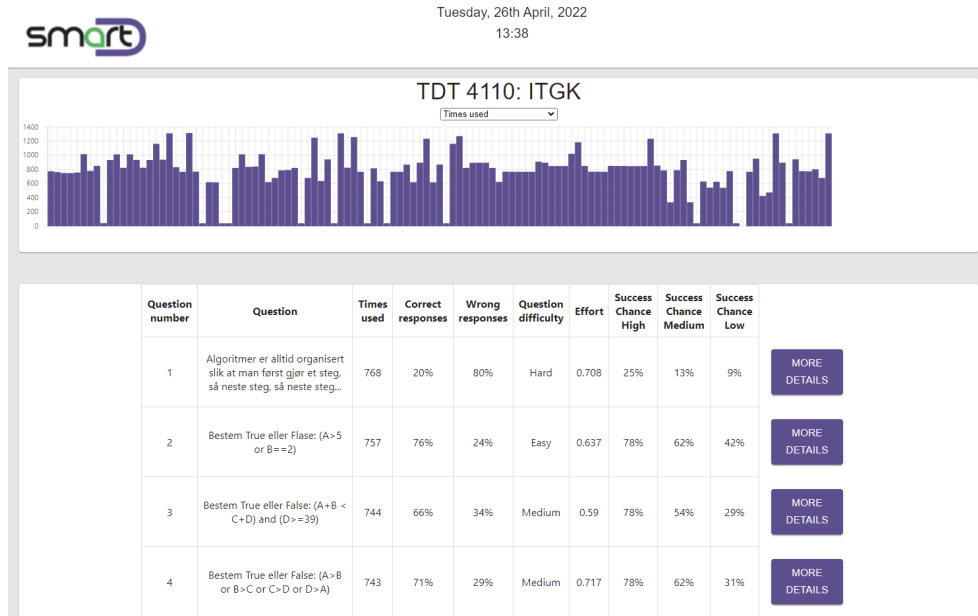


Figure 3.11: Detailed overview page

### Question page and edit question

The question page (Figure 3.12) shows detailed information about the specific question. In the top left, one can see the question and the corresponding alternatives. Next to the responses, the user can see how many percent answered the various alternatives. The correct one is marked with bold font. Next, an overview is displayed in the center of the screen. It shows how many times the question has been used, its difficulty, the effort, and the performance. Finally, the average time to answer is displayed in the top right, showing the time spent answering wrong, correct and total.

In the bottom left, a bar chart displays how long time the students spent analyzing the metadata available for them in the different mastery levels. To the right, the chance of success is displayed. This chart is also split into showing the different mastery levels.

A button labeled "Edit question" is placed in the far top right. This button opens a modal for the user, allowing them to edit the question they are currently on. Here they can edit the question itself and the alternatives. Alternative 3 and 4 can be blank, but the other fields have to contain text. The user can also choose to reset the statistics for the question when they change it to ensure the data represented is correct. However, they can choose not to reset it and keep it as it is.



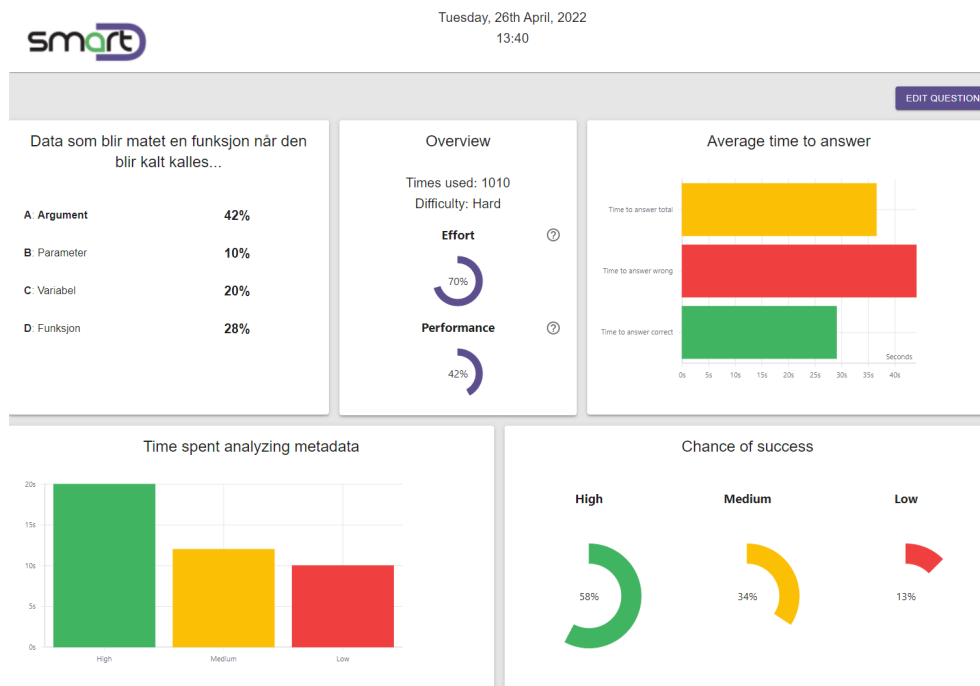


Figure 3.12: Question page



## Chapter 4

# Methods

This chapter describes how the study was conducted. The study has been conducted in compliance with what has been described in chapter 1.3. Section 4.1 Study Design describes how user tests were conducted and how the participants were observed during the user test. 4.2 Data Collection describes the data collection process. This process includes the observation notes, questionnaires, and interviews. Ultimately, Section 4.3 Data Analysis describes how the collected data was analyzed by using qualitative and quantitative methods.

### 4.1 Study Design

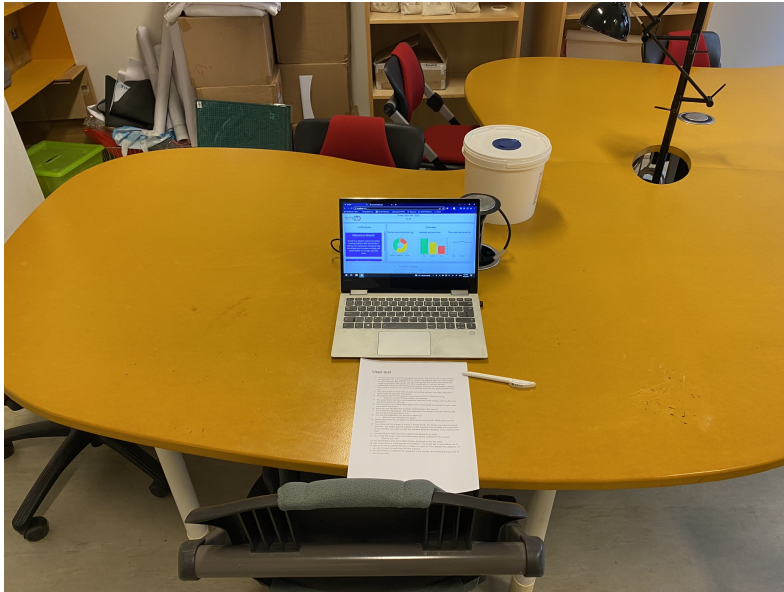
This section describes how the study was conducted by introducing the participant group, explaining the physical test setup, and explaining the steps of the experiment itself.

#### 4.1.1 Participants

In total, 20 participants were partaking in the study. For recruiting participants, it was a requirement that they would have taken the "*TDT 4110 Information Technology, Introduction*" course, as the only data set available in the dashboard was from this course. It was also necessary that they had had some teaching experience, either from previous jobs as a substitute teacher, as a teaching assistant at the university, or were studying to become teachers. The sample of participants is further described in Section 5.1

#### 4.1.2 Physical test-setup

A room was booked to conduct the user tests to ensure a controlled test environment. To create a relaxed atmosphere, drinks and snacks were made available to the participants. The researcher used his own laptop, where the user test had been prepared and set up before the arrival of every participant, as can be seen in Figure 4.1.



**Figure 4.1:** Test setup for user test

### 4.1.3 Procedure

Due to there only being one researcher, only one participant could take the test at a time. The researcher observed the participants as they were interacting with the dashboard. None of the participants had any previous experience with the SmartU dashboard. Each session lasted for a maximum of 1 hour.

To ensure that all participants got the same experience and environment, the study consisted of three steps, namely a) briefing, b) user test and observation, and c) debriefing. By following these steps, it was easy to keep track of the process of collecting data and consent during testing while also ensuring the participants got the necessary information.

#### 1. Briefing

Before the study, all participants were given an informed consent form (available in Appendix A) that contained information about the study, its purpose, and the rights of the participants. They were asked to sign this form to satisfy the requirements for handling the personal data collected. Next, participants were given a 1-page document explaining what SmartU is and how it works, to help them understand the statistics and other relevant information that they were about to process. Finally, participants were informed that an observer would be observing their interaction with the system and environment during testing.

#### 2. User test and observation

Participants were handed a document containing 16 user stories for the user test,

which they had in front of them during the entire test. They were asked to imagine that they were a teaching assistant in the course "*Information Technology, Introduction*", sitting at home or school, using the dashboard to monitor how students were performing in the course. They were asked to imagine these environments to relieve them from any external pressure and make the process as natural as possible.

In the user test, which can be seen in Appendix B, participants were asked to maneuver the dashboard and solve tasks, assessing how the quizzes and students were doing, and fixing specific errors in questions. While performing these tasks on the dashboard, the observer took notes following the observation guide. The observation as a data collection is elaborated in Section 4.2.1.

### 3. Debriefing

After the participants had conducted the user test, they were given a questionnaire (available in Appendix A) to fill in. They used the same computer they had used for the user test to answer the questionnaire. The questionnaire was used for collecting quantitative data and is explained in detail in Section 4.2.2. Immediately after completing the questionnaire, 10 of the participants assisted by giving input in a follow-up interview. All interviews were conducted in Norwegian, as all interviewees were more comfortable speaking Norwegian than English. The interviews were recorded and stored locally before being transcribed. After transcription, the recordings were deleted from the local device. The interviews are described in further detail in Section 4.2.3.

After completing the questionnaire and the interview, for those who had that, all participants were given a gift card to the local cinema worth 150 NOK.

## 4.2 Data Collection

This section describes the data collection process. It explains the methods used in the user test and observations, questionnaire, and the interviews, and why these methods have been chosen.

### 4.2.1 User test & observation

It was essential to observe and take notes of the users' interaction with the dashboard during the user test. By following the observation guide (available in Appendix A), the observer was able to make meaningful observations about the test environment, participants' emotional state, utilization of aids, and other unforeseen events during the test [5].

Observations of participants can offer insight into a study, as discussed by Mustante and DeWalt [31]. Therefore, it was essential to take notes of the aspects mentioned above of the testing environment. It could explain unexpected results and provide further insight into how the participants experienced the dashboard.

The observation guide assisted the observer by explaining how to behave and what to look for before, during, and after the observation. It also provided a matrix with categories and examples of what should be noted while the user test was ongoing, as shown in Table 4.1.

**Table 4.1:** Examples of events/actions to note during observation

Category	Includes	Researchers should note
Test environment	Ambiance, atmosphere, alterations to environment	Any differences in test environment between tests should be noted
Subjects emotional state	The users' state of mind	Is the user stressed or nervous? Does the state of mind change during the tests? Is the test subject focused on the task?
Unexpected events	Failing to navigate, bugs in the system	Did anything unexpected happen during testing? Note when, why, and how it happened
Help	Frequency, topic of questions	What does the user ask you to help out with and how often/ many times are you asked to help?
Disturbances	People entering/leaving the room, phone notifications, noises, etc	Any unwanted noises or disturbances hindering the focus on the testing should be noted.

### 4.2.2 Questionnaire

As mentioned previously, the questionnaire was a part of the debrief of the study, along with the interview. In the questionnaire, the participants answered about their opinion and attitude toward the dashboard. The questionnaire was made in Microsoft Forms, in accordance with the approved application submitted to the Norwegian Centre for Research Data (NSD), regarding the handling of personal data.

The questionnaire was split into different categories, summarized in Table 4.3. The decision to categorize the questionnaire was based on Oppenheim's [32] article. The article claims that by dividing the questionnaire into categories containing questions relating to the same perception, one achieves a more realistic image

of the participants' opinions. To capture these opinions, the majority of the questions requested answers on a 5-point Likert-like Scale, ranging from 1 to 5, where one represented "Strongly Disagree", three represented "Neutral" and five represented "Strongly Agree" [32].

**Table 4.3:** Categories in the questionnaire

Category	Acronym	Constructs
System Usability Score	SUS	Usability score
Overall evaluation of the usability of the dashboard	OEUD	General Usability
Usefulness of the dashboard	UOD	Usefulness, understanding, intuitiveness
Attitude towards the dashboard	ATD	Usage, motivation, positive and negative emotions

The first two categories, *System Usability Score* (SUS) and *Overall evaluation of the usability of the Dashboard* (OEUD), comprised of questions regarding the usability of the dashboard. These categories helped indicate whether the participants thought the dashboard met usability requirements. The SUS schema is a standardized method of measuring usability [33], and the OEUD category was created to see if the dashboard met the definition of usability as described in ISO 9241-210:2019: *the extent to which a system, product, or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.*

The other two categories, *Usefulness of the Dashboard* (UOD) and *Attitude towards the Dashboard* (ATD) contained questions regarding the participants' attitudes and perceived usefulness of the dashboard. This category included questions about motivation, feelings, and intuitiveness. These questions were constructed from a previous relevant study [2] and altered to fit the focus area of this study.

### 4.2.3 Interviews

The interviews made up the qualitative part of the data collection and were held immediately after the participant had tested the dashboard and answered the questionnaire. The reason for this decision was both to be as efficient as possible and to get the participants' reflections while they still had them fresh in mind. The interviews followed a semi-structured approach, where the interviewee followed an interview guide, available in Appendix A. The interview guide contained some predefined questions that the interviewer would build the interview around, which were created based on the research questions of this study. The predefined questions were used as conversation starters to let the participants talk about their experience with the dashboard. It also allowed the interviewer to come up

with follow-up questions for the participant to make them elaborate on topics of interest. Some of the predefined questions used in the interviews were the following:

- Did you find the dashboard interesting? If so, how come?
- Could you mention some features that you found useful in the dashboard?
- Do you think that the dashboard would improve your motivation to analyze questions and quiz material? If you do, why would it do that?
- Were there any parts you found difficult to understand?

### 4.3 Data Analysis

This section describes how the analysis of the collected data was performed. Given the mixed methods approach for data collection, both qualitative and quantitative data analysis was necessary.

#### 4.3.1 Qualitative analysis

To analyze the interviews described in Section 4.2.3, the interviews had to be transcribed. The transcribed interviews can be found in Appendix A. After the transcription, all interviews were read through, and the broad, recurring themes were noted for later use in the analysis.

For the analysis, sections in the interview were color-coded to be categorized based on the research questions. Each research question then worked as a category, and statements were given a color according to which research question they addressed. When working through the interviews, certain recurring themes started to form within each research question. These were then color-coded and categorized to create more structure. Thus, a tree-like structure was established, providing a detailed separation of the content of the interviews.

In the last iteration, the data was prepared for presentation by being cleansed. It was important that the actual quotes of the interviewees were intact and not altered. Thus, when an answer to a question or a quote was ambiguous, the correct context was added, followed by "red.anm.". This is an abbreviation for *redaksjonell anmerkning*, meaning *editorial remarks* in English. These remarks were clearly marked with parentheses.

The qualitative results provide insight into all of the research questions. It provides insight into how visual analytics can help teachers follow up on students' progress. However, it also provides insight into which visualizations provided the most useful insight and motivation. By asking them what they liked best and about their motivation, one receives a deep understanding of how they perceived the dashboard.



### 4.3.2 Quantitative analysis

For the quantitative analysis, the data had to be cleaned before the answers could be analyzed. All of the answers were in textual form ranging from "Strongly disagree" to "Strongly agree". These were coded into values, so Strongly disagree equaled 1, and strongly agree equaled 5. Subsequently, the data was imported into R Studio, to calculate mean values and standard deviation, as mentioned in Section 1.3.

The quantitative results provide insight into how the participants experienced the dashboard. It tells how useful they found it, their attitude towards it and how it affected their motivation. Therefore, it answers both the research question regarding motivation and how visual analytics can help teachers follow up on students' progress.

Additionally, Pearson's correlation was used to check if there was any correlation between the answers in the questionnaire categories.



# Chapter 5

## Results

This chapter presents the results from the conducted study. It begins with introducing the demographics of the participants (Section 5.1) before presenting the descriptive statistics from the analysis of responses to the questionnaire (Section 5.2). Next, Section 5.3 describes the results from the System Usability Score, and Section 5.4 presents the results from the conducted correlation tests. Finally, Section 5.5 describes the main findings from the conducted interviews.

### 5.1 Participants & Demographics

As Figure 5.1a shows, the study consisted of more males (65%) than females (35%). Table 5.1 and Figure 5.1b show the distribution of the participants' age.

Most of the participants (14, 70%) had experience as learning assistants (l ring-sassistent), while 6 (30%) had experience as teaching assistants (undervisning-sassistent). Furthermore, 4 (20%) had experience with teaching from primary education, and 2 (10%) were studying to become teachers in science (lecturer in science). Keep in mind that the participants could have experience in multiple areas.

**Table 5.1:** Participants' age

	<b>N</b>	<b>Minimum</b>	<b>Maximum</b>	<b>Mean</b>	<b>Std. Deviation</b>
Age	20	22	26	23.65	1.03999

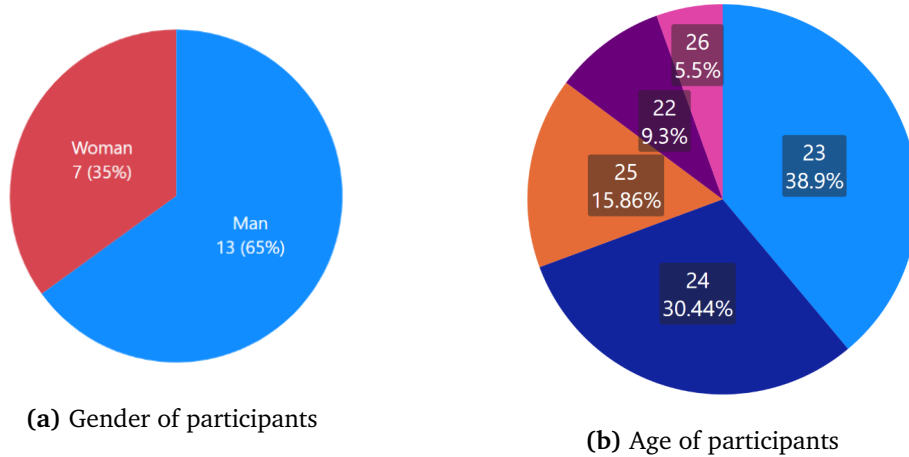


Figure 5.1: Gender and age of participants

## 5.2 Descriptives

This section presents the descriptive statistics (i.e., min, max, mean, and std. dev) of the result from the analysis of the answers from the questionnaire in Section 4.2.2, calculated using R Studio. Table 5.3 summarizes those results. These results indicate that there is a high perceived Usefulness of the Dashboard (UOD) and that the Attitude towards the Dashboard (ATD) was also high. The Overall evaluation of the usability of the Dashboard (OEUD) was slightly lower but had a somewhat lower standard deviation compared to UOD and ATD. The complete set of descriptives, including descriptives of each individual question of the questionnaire, can be found in Appendix C.

Table 5.3: Questionnaire mean variables of categories

	N	Min	Max	Mean	Std Dev
Overall evaluation of the usability of the dashboard	20	3.5	4.4	4.005	.2762
Usefulness of the dashboard	20	3.6	4.9	4.442	.317
Attitude towards the dashboard	20	3.4	4.9	4.243	.3915

## 5.3 System Usability Score

Table 5.5 shows the average score given on each question of the SUS section of the questionnaire. The average SUS-score of the 20 participants was 82.875. Furthermore, from all the SUS-scores, the lowest score was 67,5, the highest score was 92.5, and the median score was 85 (SD = 7.17). According to Sauro [34],

the score indicates that the dashboard has good usability. The score is further elaborated on in Section 6.1, and the full list for the SUS-scores can be found in Appendix C.

**Table 5.5:** Average SUS-score per question.

Question	Score
Q1 - Frequent Use	4.25
Q2 - Unnecessarily complex	1.95
Q3 - Easy to use	4.45
Q4 - Would need support	1.50
Q5 - Well integrated functions	4.25
Q6 - Too much inconsistency	1.55
Q7 - Learn it quickly	4.55
Q8 - Slow or complicated	1.40
Q9 - Felt confident	3.85
Q10 - Required training	1.80
Avg SUS-score	82.875

## 5.4 Pearson's correlation

Pearson's correlation tests were conducted between the mean scores of the four categories; SUS, OEUD, UOD and ATD, to see if there was any correlation between the mean scores. The results of these can be seen in Table 5.7.

The Pearson's correlation test between SUS and OEUD showed,  $r(18) = .48, p = .033$ . This coefficient indicates that there is a low correlation between these two values. While the correlation test between OEUD and UOD showed,  $r(18) = .27, p = .260$ . Both these results indicate that there is a low correlation between these values.

The Pearson's correlation test between SUS and ATD showed,  $r(18) = .55, p = .0127$ . Furthermore, the correlation test between ATD and UOD showed,  $r(18) = .54, p = .014$ . Both results indicate that the values can be considered moderately correlated.

Finally, the Pearson's correlation test between SUS and UOD showed,  $r(18) = .71, p = .001$ . Additionally, the correlation test between OEUD and ATD showed,  $r(18) = .75, p = .001$ . These results indicate that the values can be considered correlated.

**Table 5.7:** Results of Pearson's product-moment correlation test

	SUS	OEUD	UOD	ATD
SUS	X			
OEUD	.48 *	X		
UOD	.71 **	.27	X	
ATD	.55 *	.75 **	.54 *	X
Notes: * $p < 0.05$ (2-tailed); ** $p < 0.01$ (2-tailed); Pearson correlation coefficients (N = 18)				

## 5.5 Results from Interviews

This section presents the findings from the qualitative analysis described in Section 4.3.1. The results presented are extracted from the ten interviews that were conducted for this research. The interviews were conducted in Norwegian. However, the quotes presented in this section are translated into English by the researcher. The complete interviews in Norwegian can be found in Appendix A.

As previously mentioned, statements were categorized into three different categories based on the three research questions of this thesis. This categorization led to certain sub-categories forming these three categories, thus creating a tree-like structure.

### 5.5.1 In what way does visual analytics help teachers to follow the progress of students who use adaptive assessment

For the first research question, it was interesting to investigate how the participants experienced the dashboard and how they thought it would help them in the assessment. The general opinion of visual analytics used in the dashboard was that it made the experience easier and more understandable. When asked about how they perceived the dashboard, all of the participants said they found it interesting.

One participant appreciated how the dashboard gave such a good overview on different levels:

**Participant:**

*“One gets a good overview of how things are going, and one can also get an overall overview, and a more detailed overview if one wants that. So I felt like it covered the needs one might need in a situation like this.”*

When asked if they preferred visual analytics over textual representation, they all stated that they preferred visualized data. One participant replied:

**Participant:**

*“Definitely (better with graphs than textual representation, red. anm.). It’s worse to open a page with a lot of numbers, where I have to find out things myself. This makes it easier to find out where the problems lie.”*

Other participants agreed that visualizations were better than textual representation. Another participant added by pointing to how visual analytics saved them of mental effort. They explained:

**Participant:**

*“I think the visualizations help a lot more, and then you do not have to ... Not that comparison of numbers is so hard. But one could say it takes a bit more brain-power than quickly looking at a graph. So then it is easier to see instantly ‘Oh, here something is off’, without having to look at a lot of numbers and compare them to each other.”*

One participant exemplified by pointing to their experience with working with multiple-choice quizzes in Blackboard and had some interesting insight as to how these two dashboards compared:

**Participant:**

*“I am used to using Blackboard, from when I was a learning assistant, and that is a much harder tool, so it was easier to gain insight in how students did in the quizzes here. It is much more modern, and a lot less complex than what Blackboard has. So definitely, I would have used this, at least to make multiple-choice quizzes. It seems very easy to edit things and a lot easier to find information if one needs the information. So yeah, I would definitely have used this over Blackboard, because Blackboard is very complicated.”*

When asked if they had statistics and if they visualized them, the participant replied that they had minimal statistics and that the system was challenging to use. They appreciated how much easier this was and how easy it was to edit questions. They explained:

**Participant:**

*“They have very little statistics, I think. I do not think I can remember them having it. If I remember correctly, they only have some numbers there, I cannot recall there being any graphs. It is quite a while since I made a quiz in Blackboard, but I am pretty sure I have done it before. And it was very difficult for students to see which alternative was correct and so on. So I think this is a lot better than that. I also think it was difficult to edit questions, if I am not mistaking. So yeah, Blackboard is just a lot more complicated, because there are many unnecessary steps. This (SmartU dashboard, red. anm.) is less clicks away from doing what you want.”*

From these answers, it becomes clear that the participants appreciated the various overviews the dashboard presents and the fact that the data is visualized, not textual. The visualization saves mental effort and helps the teacher see things quicker.

**5.5.2 Visualizations and their usefulness**

All the participants were asked to mention some features they found extra useful. As displayed in Table 5.9, five features in the system were explicitly mentioned.

**Table 5.9:** Useful features

Name	Files referenced	Statements in total
Suggestions	9	9
Bar chart	8	12
Temporal data	9	11
Data about specific question	8	10
Summary about all questions	2	2

Of the mentioned features, four stand out as being mentioned most often. Temporal data was one feature that was often mentioned. The ability to see how long time students spent to answer the questions and analyze metadata was something almost all of the participants mentioned as useful. One of the participants explained that they found it useful because:

**Participant:**

*“And I liked to see how long time students spent. That is pretty nice, then one can see that ‘okei, this is one they struggle with. Maybe I should say something more about it, because they spend a lot of time on it’.”*



However, not everyone found the temporal data useful. One participant explained that they did not quite understand how to use this data, as it varies how long time a question takes to answer:

**Participant:**

*“Temporal (statistics) was also useful, but I did not feel like it was... It was more a bonus. I do not know how much I would have used it. It is hard to know how long time a question takes. Some use a lot of time, some do not. It is hard to find a correlation, but a small bonus.”*

Another aspect that many participants mentioned as useful was the detailed insight into each question. One participant explained that it helped them see where the problems are easily:

**Participant:**

*“I really liked getting insight into each question. I think that is a good way to catch questions quickly. There you can see instantly ‘Oh, this is a hard question. This was poorly formulated’, things like that. So I think that could be an important tool to find out where the problems are.”*

One of the most frequently mentioned feature was the suggestions. This feature referred to the section in the application where the system highlighted some potentially bad questions, and asked if the user wanted to look into them. When asked to explain why they found it so helpful, one participant explained how it made it quicker to achieve the goal, which is to identify bad questions. Additionally, they thought it would make it easier for teachers because it demanded less of them. They explained:

**Participant:**

*“This again makes it easier to not have to analyze which questions have performed badly, and work through each question, when it comes up immediately. And that is very nice. There are fewer steps to do what one wants to do. And then one can fix the question or the problem a lot quicker than what one would have done. Another thing is that I do not know if professors and teachers, who do this on a regular basis, bother to check all the questions. Then it becomes much easier to just get some suggestions, rather than to start to analyze by oneself. So it demands less effort for them, and then it help so one solves problems quicker, or bothers to do something about it rather than having to go through each question, and have to analyze.”*

One participant pointed out that the suggestions could help identify errors in the questions, which could prevent students from learning something wrong because of a technicality:

**Participant:**

*“One of the worst things there is, is if a question has a technical error that affects the knowledge of students. Where one teaches something that is wrong. So if you have a system that can detect that, then it is very good.”*

Not all participants were as comfortable with the suggestions. Some pointed out it was a good feature but wanted to have insight as to how it made recommendations:

**Participant:**

*“But it looked like there were quite a lot (suggestions, red. anm.), I probably would have preferred a list of why it is a suggestion. There is a bit of text there, but then I cannot compare it to the other questions, how good or how bad they are doing compared to the other questions. That is why I preferred to see it compared on one line, where one could clearly see why it is suggested. (...) Yes (would prefer to make the decision oneself, red. anm.), because then I can see it myself. Give me the statistics, then I can look at them myself.”*

Another participant also liked the suggestions, but wanted to be able to influence the suggestion system by telling the system what kind of statistics to look at to determine when to notify them:

**Participant:**

*“And I thought the suggestions were good, so you do not have to look for it yourself. So you only get ‘Maybe you should look at this?’. And maybe if one uses the system over time, one might want to affect what suggestions appear. So you could say ‘Generally I want to see all questions where over 70% have answered wrong’.”*

One of the features that was mentioned the most was the bar chart containing all of the questions, where one could see statistics for all the questions in the course. When asked to explain why they perceived it as useful, one participant explained:

**Participant:**

*“I thought it (the bar chart, red. anm.) was spot on. One could see all the questions, and the hover effect showed how many had answered wrong and correct, and so on. (...) but it worked very well.”*

Another participant pointed out how the bar chart made it very easy to spot patterns:

**Participant:**

*“I thought the long bar chart was very interesting. I liked that a lot. When one could see all of the 119 questions in the course. Then one could see the average of the various variables that one could use as a filter. This way one could see patterns very easily.”*

However, not all participants shared the same enthusiasm for the bar chart. One participant preferred to see the statistics in the spreadsheet rather than the bar chart. They explained:

**Participant:**

*“Mostly I like visual representations. But I felt that when you have a table below that is so clean, I did not feel that I needed the (visual, red. anm.) statistics. But usually, I am more fond of the visual (representation, red. anm.)”*

The responses from the participants show that some elements stood out as more useful than others. Participants appreciated having access to detailed statistics of all the questions and seeing all of the questions compared in one bar chart. However, one participant thought it was better to see this data in the spreadsheet. The participants also liked the suggestions, as it helped them know where to look. Finally, participants found the temporal data interesting, as it could tell them a lot about how the questions were doing.

### 5.5.3 Participants’ reflections on motivation

The last research question revolved around if the teachers found the dashboard motivating. Table 5.11 shows the distribution of the answers. Most of the participants found it motivating to some extent. Furthermore, from the table, one can read that none answered that they did not find it motivating.

As seen in Table 5.11, not all students were as sure they would be too motiv-

**Table 5.11:** Participants thoughts on motivation

Motivation	Amount
Definitely	5
Possibly	3
Unsure	2
No	0

ated by the visual analytics presented in the dashboard. Three (3) of the students said they most likely would find it motivating. They responded that it was hard to tell, given the short amount of time they had with the system. Two (2) of the students were more unsure. They pointed to their lack of experience with teaching in this type of course. As one of the participants said:

**Participant:**

*“Maybe (would find the dashboard motivating, red. anm.). I have not been teaching assistant in so many courses like this, I have had more project based courses, so it’s hard to tell.”*

Even though some participants were uncertain, five (5) of the participants were in no doubt that a system like this would make it more motivating to follow up on students’ progress and see how questions are doing. However, they pointed to different factors as to why it would motivate them. For example, many participants pointed to ease of use being a contributing factor to their motivation:

**Participant:**

*“This (the dashboard, red. anm.) makes it easier to find out where the problem lies, and that makes it more motivating to get started with it as well.”*

Another participant pointed out how it was motivating to know that the system does work for them by picking out suggestions. This way, the teacher saves mental effort. They explained:

**Participant:**

*“It (suggestions, red. anm.) makes it even more motivating because in many ways it does the job for you. (...) All you need to do is click in and see.”*

Other participants pointed out how this tool could help them see what students do not understand to help them prepare for teaching assistant meetings. One participant explained:

**Participant:**

*“Yes. I think it would be nice to see how long time they had used, and if they have spent a lot of time but gotten the wrong answer, I would spend more time explaining more thoroughly. While if a person has answered quickly, and gotten right. Or answered quickly and answered wrong, you could think it is a sloppy mistake. So you do not need to explain things in too much detail. So you save quite a lot of time on the feedback. (...) Yes (it would be nice to use to prepare for meetings with students, red. anm.). I think it would be nice to get an overview to see what people struggle with.”*

Additionally, to help identify what material students struggle with, one participant pointed out how this dashboard could motivate teachers to make more frequent changes to the questions:

**Participant:**

*“This (the dashboard, red. anm.) could motivate for more frequent changes. Change the quizzes more frequent. If one notices on a question that people answer very quickly and everyone gets it correct one can assume they are sharing answers with each other. Then one can assume one has to change things up. So definitely very useful.”*

Finally, one student pointed out that having the statistics presented in a comprehensible way was motivating:

**Participant:**

*“I really enjoy studying statistics. And this is very nice, and a very comprehensible way to see how things are going. So that is the best thing there is, to not have to count things yourself and ... So when things happen automatically it is very nice (...) So yeah, I had been motivated.”*

The responses show that, in general, the participants found the dashboard motivating. The participants pointed to ease of use, and the feeling of the system doing some of the work, as reasons to why they found it motivating. They also

suggested it could motivate to more frequent changes.

## Chapter 6

# Discussion

This Chapter covers the discussion of the findings presented in Chapter 5. The results provided interesting insight regarding the main hypothesis of this study, which was that visualized statistics could help increase teachers' motivation and help them detect problems in a learning system. The hypothesis was split into three research questions regarding the effect of visual analytics, motivation, and useful features.

### 6.1 Effect of visualizations

The first research question was: *How can visual analytics help teachers follow students' progress who use adaptive assessment?*

From the quantitative data presented in Section 5.2, we can see that the general opinion of the dashboard is good. The average scores of OEUD, UOD and ATD, as presented in Section 5.2 mean that the average answer for each claim in the questionnaire was between "Agree" and "Strongly agree". These results indicate that the participants found the dashboard understandable, useful, and motivational.

Additionally, the dashboard received an average SUS score of 82.875 (SD=7.17). According to Sauro [34], scores above 68 would be considered above average, and above 80.3 is within the top 10th percentile of scores. This score indicates that the dashboard is perceived as very usable and learnable [34].

The correlation tests showed a correlation between SUS and UOD, and OEUD and ATD. The correlation between SUS and UOD could show that the users who found the dashboard user-friendly also saw the system's usefulness. The correlation between OEUD and ATD could imply that the users who found the overall usability to be best also were the ones who were most motivated by the dashboard. Both of these correlations could demonstrate the importance of creating a user-friendly dashboard, as it is hard to become motivated or see the usefulness of a product if one does not find it user-friendly and usable.

From the interviews, it was clear that the participants found the use of visual analytics in a dashboard to be interesting, engaging, and understandable. For ex-

ample, one participant stated they got a good overview overall and detailed overview of each question, which felt like it covered the needs one might need in a teaching situation. The participants also appreciated the ease of use regarding how well the statistics were laid out, and how easy it was to edit questions and get recommendations. These statements are in high agreement with the results revealed from the quantitative data and confirm that the participants found the dashboard to be useful.

The participants appreciated having visualizations and statistics to present the data, compared to having to read the numbers themselves, as it was less demanding to grasp the insights from the visualizations. One participant said that by having visual analytics, it became much easier to see where the problems were, compared to having to compare numbers on a spreadsheet. The quantitative data also confirm this statement, where 65% of the participants answered "*Strongly agree*", 25% "*Agree*", and 10% "*Neutral*" when prompted with the statement "*I believe that a graphical representation is more fitting than a textual representation of the data displayed in the dashboard*".

One of the participants compared the dashboard to their experience with teaching within Blackboard. Their opinion was that this dashboard made acquiring insight and managing quizzes much more effortless. One of the factors the participant mentioned as preferring this dashboard over Blackboard was that it had none or very little visualization of data. By presenting the data visually and in a comprehensible matter, it became easier to extract the necessary information. This point could show that the presentation of visual analytics made the information easier to process for the teacher.

All of the positive feedback received from both the qualitative and quantitative data shows that the dashboard managed to use visual analytics to provide information to the teacher in a comprehensible and structured manner. This feedback shows that by following the design principles of Few [17], the dashboard managed to present the data comprehensibly. Furthermore, these results demonstrate how the dashboard manages to reach the goal set by Durall and Gros [11]; that the main focus of a dashboard should be to empower the teacher to make informed decisions about the learning process.

From the literature review, Mottus et al. [16] warned about the significant challenge of displaying complex information in a visual and accessible format and how it is challenging to display for people who do not work with analytics on a day-to-day basis, like teachers. They also warned about getting caught up in presenting visually stimulating data and not necessarily valuable data. The fact that the participants of this study gave such good feedback proves that the dashboard managed to stay clear of these potential pitfalls.

## 6.2 Useful features

The second research question revolved around how visualizations provide useful insight, asking: *How do visualizations provide useful insight?*



Based on the literature from the background work, some elements were considered interesting to add, as they had previously proved to provide valuable insight. One of these features was temporal data, as Papamitsiou and Economides [15] showed that it could provide insight into students' perceptions and discover how they behave when dealing with assessment tasks. Looking at Table 5.9, one can see that temporal data was mentioned as a useful feature 11 times. The participants said that it provided valuable insight, showing which questions students struggled with or that it could indicate if the questions were too easy because students spent so little time. This result supports the findings of Papamitsiou and Economides [15], who found that temporal data is perceived as easy to use and useful for the teachers. The participants had limited time with the dashboard, so the fact that so many pointed to temporal data as an interesting factor proves the point of its ease of use. However, not all participants found it useful. One participant called it "a bonus", as they felt they did not get too helpful information from the temporal data.

Another feature from the literature review, shown by Alevén et al. [13], was the importance of providing a high-level summary to the teacher of actions that require the teacher's attention the most. They also pointed out the importance of providing insight to the teacher by giving them access to statistics on which the recommendations are based. This point was also supported by van Leeuwen et al. [12], who pointed out the importance of both suggestions and providing a deeper understanding of the statistics.

These points were taken into account during the dashboard development, and they were among the features that received the most praise. The suggestions were mentioned in 9 interviews and held up as a valuable feature for many participants. The participants pointed out the ease of use, not having to look for the problems oneself. This feedback supports the study conducted by van Leeuwen et al. [12]. They emphasized how simple visual cues could considerably affect the result and make it less effortful and more confident decision-making for the teacher.

However, some participants requested the opportunity to see why the system suggested these questions and have the chance to influence the recommendation system. These statements support the findings of Alevén et al. [13] and van Leeuwen et al. [12], showing that teachers appreciate getting suggestions and not having to look for the problems themselves, but also want to have control and knowledge about why the system suggests the particular recommendations. This also emphasizes the importance of the screen that displayed all of the questions in the course. The bar chart in this screen was mentioned 12 times within the interviews, making it the most mentioned feature. The participants found it very useful because they could filter on different values, and it provided a straightforward overview of how all the questions in the course were doing. Many participants expressed that viewing the data in the bar chart was much more effortless than viewing it in the spreadsheet below. However, one participant preferred to view the data in the spreadsheet. This feedback shows how it is essential to display the data differently to satisfy the needs of various teachers.

Additionally to the features mentioned above, the participants also appreciated the ability to get an overview of each course and see detailed information about each question. Some participants found it most helpful to view the course as a whole, while others found it most beneficial to view specific questions. However, most participants found it helpful to vary, viewing the whole course and then viewing particular questions if there was anything they found interesting.

To summarize, the participants appreciated getting a high-level overview of the course, but also having the opportunity to get detailed insight into statistics if they felt like they needed to. The participants also appreciated the use of visualizations, stating it made it easier to analyze and spot outliers, and thought the suggestions were useful, as they relieved them of mental effort. This indicates that visualizations help the teacher by saving mental effort, and presenting data more comprehensibly, making it easier to analyze.

### 6.3 Motivation

The third and last research question was about the motivation of the teachers towards the visualization in the dashboard: *What are the teachers' motivations towards systems that provide visualizations?*

From the quantitative analysis presented in Section 4.3.2, the participants found the dashboard motivating. When presented with a claim stating the statistics motivated them to monitor the course question, 55% answered "Strongly agree" and 45% answered, "Agree".

The correlations tests showed a moderate correlation between SUS and ATD, and ATD and UOD. These results could imply that the perceived usefulness and ease of use played a part in the attitude of the participants, thus implying that usability makes it more motivating.

The qualitative findings substantiate this implication. In the interviews, participants said a dashboard like this would make it more motivating for them to follow up on students' progress. Some participants pointed to how the visual analytics dashboard made it easy to spot flaws, which made it more motivating to get started assessing questions. This finding supports the claims of Yoo et al. [8], who emphasized the importance of visualization, as it helped convert abstract and complex information into something concrete and visual. Other participants implied the suggestions made it even easier, thus more motivating, as they pointed out the errors for the teacher. Simple visual cues help the teacher make more confident decisions while also being less effortful than if they had to find these themselves, as pointed out by van Leeuwen et al. [12]. The boosted confidence of the teacher is also supported by the quantitative results, where 60% answered "Agree", 35% "Strongly agree" and 5% "Neutral", when asked whether they felt more informed and confident conducting an assessment when utilizing the dashboard.

Another participant also confirmed this finding, as they thought a dashboard like this would motivate for frequent changes in the questions because the system suggests changes, and it was so easy to make changes to the questions. As the

participant pointed out, not all courses are good at changing their course material. This lack of updates leads to students having the chance to share answers with each other and learning the answers, not necessarily the course material. As the participant pointed out, teachers can notice when students do not put in an effort by looking at the statistics. Additionally, they can be assisted in making necessary changes with the suggestions, so they do not have to do it all by themselves. Both of these factors can motivate them to change questions and make more confident decisions to ensure students keep learning and do not take shortcuts.

Some participants also pointed to how the dashboard could motivate them by helping them prepare for teaching assistant meetings to know what students had struggled with and prepare for which type of questions the students might have. This finding supports the findings of Xhakaj et al. [14], who found their dashboard to be helpful for teachers to prepare for class sessions, as they got an indication of what the students understood and did not. This shows how the visual statistics in this dashboard can not only help teachers make as good questions as possible but also help them understand what students are struggling with and give an indication as to what they need to focus on in teaching.

## 6.4 Implications for design

Although the dashboard received good feedback from the participants, it is by no means perfect. The user tests and interviews unfolded some weaknesses and rooms for improvement regarding the dashboard's design. It was unsurprising that the users uncovered some weaknesses, as developing a user interface requires many iterations. The dashboard had only been briefly user-tested once, so it was expected that it would not be perfect.

One of the most significant sources of confusion, mentioned in both the interviews and from the observations (available in Appendix B), was the use of colors in the dashboard. Green, yellow and red were used to represent multiple purposes throughout the system. For example, it was used to show the distribution of mastery levels (High, Medium, and Low), the question difficulty (Hard, Medium, and Easy), and to showcase right and wrong answers in the visualization of temporal data (Correct, Incorrect, and Total). This inconsistency led to many participants becoming confused or misinterpreting what the data they were looking at was actually representing.

Another detail that would have been fixed would have been to mark the correct alternative of a question more clearly, on the question page (Figure 3.7d). Not all students noticed the bold font indicating the correct alternative, as a lot of the surrounding text was bold as well.

Finally, some wording changes would have been made, as some participants misinterpreted what was intended to be communicated. For example, eight (8) participants misinterpreted the statistic "*times used*". This statistic showed how many times a question had been used in quizzes, but many thought it represented time spent on a question, or simply did not understand what it meant.

This demonstrates the importance of an iterative design process when designing a TAD. The use of terms and colors can easily be misinterpreted, leading to confusion and errors. By iterating more frequently with user-tests, one uncovers misunderstandings as the aforementioned, thus eliminating flaws and creating a better TAD.

## 6.5 Limitations

Even though this thesis has gone mostly according to plan, some parts were not planned thoroughly enough. This has, in some ways, limited the study. These aspects will be covered in this section.

### 6.5.1 Mock data

As previously mentioned, the data used in this dashboard was a JSON file containing information about how users had performed in a previous study. It consisted of 119 questions and contained detailed information about the questions given to students in the course "*TDT 4110 Information Technology, Introduction*" at NTNU. However, it did not contain all of the necessary data for this study. Therefore, some mock data had to be added to the data set, so all the desired data could be visualized and tested. Examples of data that were added are:

- Users were created. There was only one user in the data set. More were added for the system to seem in operation.
- User interactions with the system. Attempts on quizzes, scores, and average time spent were generated.
- Time to answer correctly and wrong was added.
- Time spent viewing metadata was added.
- The number of students who answered the various alternatives was added. The number of correct and wrong answers are from the original dataset, but how many answered the various wrong alternatives was made up.
- Some of the question statistics were altered to create tasks for the user test.

The mock data implies that the data viewed by the teachers was not entirely representative of how it would have been if it contained student interaction. Though it is not ideal, the data was attempted to be as realistic as possible and does not seem to have affected the results, as no participant pointed out that they noticed any inconsistencies.

### 6.5.2 Diversity and sample size

By only having access to a data set regarding one course, the study became limited regarding eligible participants to recruit. This factor limited the pool to consist of people with teaching experience who also had had the course "*TDT 4110 Information Technology, Introduction*". This group is generally experienced with both maneuvering a dashboard and viewing statistics. Additionally, the limitations lead to a very homogenous group. Therefore, one can question whether the results apply to all types of teachers. It would have been interesting to test the dashboard on a broader group of teachers, for example, within non-technical courses.

Finally, the sample size should have been bigger to give a more representative view of how teachers, in general, perceive the use of visual analytics in a teacher dashboard.

### 6.5.3 Inexperienced researcher

The researcher of this thesis has minimal experience conducting a study of this magnitude. Although it seems to have gone well, and the results seem promising, the study definitely would have benefited from a bit more experience. Especially in the data collection phase, formulating a questionnaire and conducting interviews, some more experience could have helped. For example, the questions in the questionnaire could have been better developed to make sure they were more precise and more targeted to the research questions. Also, the interviews could have been performed in a better way to prevent bias. Additionally, it became apparent during the transcription that the interviewer should have formulated some questions better to ensure more open-ended questions.

Finally, the user test had some poorly formulated tasks, making the users uncertain. In many cases of uncertainty in the user tests, the user would have solved the task smoothly had it not been that the user stories were worded poorly.

### 6.5.4 Limited interaction with the dashboard

Another limitation of the study is the fact that the participants were given a limited opportunity to interact with the dashboard. First of all, they were given a limited time with the dashboard. Had the users had more time to use the dashboard, they would have gained a more thorough opinion of the dashboard. Furthermore, they were presented with data from a course that many of them had not taught. This way, it was not easy for them to know exactly which data was useful for them in the specific course.



## Chapter 7

# Conclusion and Future Work

This thesis has looked at how the usage of visual analytics in a dashboard affects teachers. To investigate this aspect, the thesis was broken down into three research questions, investigating how it can help teachers, how visualizations provide insight, and the effect on the teachers' motivation. In order to answer these questions, a teacher dashboard for SmartU was designed and developed, and then it was tested in a controlled environment on 20 participants with teaching experience.

The results from this study, as discussed in Chapter 6, show that the participants found the dashboard to be user-friendly, useful, and motivational. The average SUS-score indicates that the dashboard had a good overall perceived usability. The teachers found it interesting to analyze the visualized data in the dashboard. Among the most useful features, the suggestions, temporal data, and the bar chart containing all of the questions were pointed out the most by the users. However, the diversity in the results regarding what was perceived as useful proves that it is important to provide different views to satisfy the needs.

Both the qualitative and the quantitative data suggest that the participants found the dashboard motivating for various reasons. They appreciated that the suggestions did much work for them and also found it motivating to have so much data available and presented in a visually stimulating manner.

### Future work

This thesis has opened up other opportunities for further research and improvement of the dashboard. As seen in Section 6.5, this research has its limitations. These provide some interesting points for further research and development. By investigating these paths, it would provide meaningful and interesting future work directions:

- Conduct experiments with a larger and more diverse group of participants. By going broader in the investigation, one can make more general assumptions about teachers.

- Test the dashboard in a longitudinal study during their course. A longer study would provide insights on how the dashboard can be used over time and potential comments for functionalities that will come up after intensive use.
- Add more functionality to the dashboard and test the effects of this. The functionality could, for example, be a customizable user interface, as Verbert et al. pointed out [10], not all teachers make use of the same information. It could also be to allow the teacher to specify what kind of suggestions they want to get, to give them more insight into the decision making. Mottus et al. [16] pointed to the importance of being able to "see" the student, even though they are in a digital environment. It could be interesting to let the teacher gain more insight into each student to see if this had any effect.
- Improve the design of the current system. As pointed out in Section 6.4, the design is not perfect. By performing more design iterations, one could create a better experience for the users.



# Bibliography

- [1] Z. Papamitsiou, 'Supporting learners' capacity for autonomous decisions using learning analytics,' 2018.
- [2] J. Westermoen and M. Lunde, 'Smartu: Investigating the effects of visualizations in adaptive self-assessment systems,' M.S. thesis, NTNU, 2020.
- [3] Z. Papamitsiou, M. Lunde, J. Westermoen and M. N. Giannakos, 'Supporting learners in a crisis context with smart self-assessment,' in *Radical Solutions for Education in a Crisis Context*, Springer, 2021, pp. 207–224.
- [4] Z. K. Papamitsiou and A. A. Economides, 'Towards the alignment of computer-based assessment outcome with learning goals: The laers architecture,' in *2013 IEEE Conference on e-Learning, e-Management and e-Services*, IEEE, 2013, pp. 13–17.
- [5] B. J. Oates, *Researching information systems and computing*. Sage, 2005.
- [6] C. Wohlin, 'Guidelines for snowballing in systematic literature studies and a replication in software engineering,' in *Proceedings of the 18th international conference on evaluation and assessment in software engineering*, 2014, pp. 1–10.
- [7] I. Jivet, M. Scheffel, H. Drachsler and M. Specht, 'Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice,' in *European conference on technology enhanced learning*, Springer, 2017, pp. 82–96.
- [8] Y. Yoo, H. Lee, I.-H. Jo and Y. Park, 'Educational dashboards for smart learning: Review of case studies,' *Emerging issues in smart learning*, pp. 145–155, 2015.
- [9] B. A. Schwendimann, M. J. Rodriguez-Triana, A. Vozniuk, L. P. Prieto, M. S. Boroujeni, A. Holzer, D. Gillet and P. Dillenbourg, 'Perceiving learning at a glance: A systematic literature review of learning dashboard research,' *IEEE Transactions on Learning Technologies*, vol. 10, no. 1, pp. 30–41, 2016.
- [10] K. Verbert, X. Ochoa, R. De Croon, R. A. Dourado and T. De Laet, 'Learning analytics dashboards: The past, the present and the future,' in *Proceedings of the tenth international conference on learning analytics & knowledge*, 2020, pp. 35–40.

- [11] E. Durall and B. Gros, 'Learning analytics as a metacognitive tool.,' in *CSEDU (1)*, 2014, pp. 380–384.
- [12] A. van Leeuwen and N. Rummel, 'Comparing teachers' use of mirroring and advising dashboards,' in *Proceedings of the tenth international conference on learning analytics & knowledge*, 2020, pp. 26–34.
- [13] V. Alevan, F. Xhakaj, K. Holstein and B. M. McLaren, 'Developing a teacher dashboard for use with intelligent tutoring systems,' *technology*, vol. 34, pp. 44–50, 2010.
- [14] F. Xhakaj, V. Alevan and B. M. McLaren, 'Effects of a teacher dashboard for an intelligent tutoring system on teacher knowledge, lesson planning, lessons and student learning,' in *European conference on technology enhanced learning*, Springer, 2017, pp. 315–329.
- [15] Z. Papamitsiou and A. A. Economides, 'Temporal learning analytics visualizations for increasing awareness during assessment,' *International Journal of Educational Technology in Higher Education*, vol. 12, no. 3, pp. 129–147, 2015.
- [16] A. Mottus, S. Graf, N.-S. Chen *et al.*, 'Visualization and interactivity in the teacher decision support system,' in *2013 IEEE 13th International Conference on Advanced Learning Technologies*, IEEE, 2013, pp. 502–503.
- [17] S. Few, *Information Dashboard Design: Displaying data for at-a-glance monitoring*. Analytics Press Burlingame, 2013, vol. 5.
- [18] M. Glinz, 'On non-functional requirements,' in *15th IEEE international requirements engineering conference (RE 2007)*, IEEE, 2007, pp. 21–26.
- [19] 'International standard - systems and software engineering–vocabulary,' Institute of Electrical and Electronics Engineers, Tech. Rep., 2017.
- [20] S. Robertson and J. Robertson, *Mastering the requirements process: Getting requirements right*. Addison-wesley, 2012.
- [21] I. Sommerville, *Software engineering*. Pearson Education India, 2004.
- [22] A. M. Davis, *Software requirements: objects, functions, and states*. Prentice-Hall, Inc., 1993.
- [23] S. Bradner, 'Key words for use in rfcs to indicate requirement levels,' Tech. Rep., 1997.
- [24] Figma. 'Figma, the collaborative interface design tool.' (), [Online]. Available: <https://www.figma.com/> (visited on 26/04/2022).
- [25] Google. 'Google drive, a safe place for all your files.' (), [Online]. Available: <https://www.google.com/drive> (visited on 26/04/2022).
- [26] Sharepoint. 'Sharepoint, team collaboration software tools.' (), [Online]. Available: <https://www.microsoft.com/nb-no/microsoft-365/sharepoint/collaboration/> (visited on 26/04/2022).

- [27] Github. 'Github, the worlds leading software development platform.' (), [Online]. Available: <https://www.github.com/> (visited on 26/04/2022).
- [28] React. 'React, a javascript library for building user interfaces.' (), [Online]. Available: <https://www.react.com/> (visited on 26/04/2022).
- [29] react-vis. 'React-vis, data visualization components.' (), [Online]. Available: <https://uber.github.io/react-vis/> (visited on 26/04/2022).
- [30] M. UI. 'Material ui, the react component library you always wanted.' (), [Online]. Available: <https://mui.com/> (visited on 26/04/2022).
- [31] K. Musante and B. R. DeWalt, *Participant observation: A guide for fieldworkers*. Rowman Altamira, 2010.
- [32] A. N. Oppenheim, *Questionnaire design, interviewing and attitude measurement*. Bloomsbury Publishing, 2000.
- [33] J. Brooke, P. Jordan, B. Thomas, B. Weerdmeester and I. McClelland, 'Usability evaluation in industry,' 1996.
- [34] J. Sauro, *A practical guide to the system usability scale: Background, benchmarks & best practices*. Measuring Usability LLC, 2011.



## **Appendix A**

### **Study**

## A.1 Form of Consent

# Request for participation in research project SmartU

"User-testing of teaching analytics dashboard"

## Background and Purpose

This research project aims to investigate the effects visualizations have on teachers when assessing how their quizzes are performing in a course at university. The project builds upon the SmartU system, which is an adaptive self-assessment system. In relation with the research project, a teaching analytics dashboard (TAD) have been developed as a mean to test the visualizations in a real life context. The user-test will contain hands-on testing of the TAD at NTNU Campus Gløshaugen in Trondheim, Norway, where the teachers will get familiar with the dashboard and test themselves in the available self-assessments within the system.

The participants of our project will be professors and teachers at NTNU Campus Gløshaugen in Trondheim, Norway, volunteering to participate in the user-testing.

The responsible for the project will be an Associate Professor at the department of Computer and Information Science (IDI) at NTNU, Trondheim, Norway. (see general information section).

## What does participation in the project imply?

For the purpose of the research project, data will be collected using observations, web-based questionnaires. Questions that will be used for the questionnaire, will concern participants' attitudes (e.g. sensemaking, satisfaction, easiness, difficult/easy/challenging parts of the activity) toward the use of visualizations in the TAD. Data regarding participants' gender, age and field of teaching will be collected using a questionnaire on a computer belonging.

The duration of the user-testing will be approximately half an hour, consisting of the mentioned hands-on user-testing of the system and a follow-up questionnaire prior to the user-testing.

Participants can request to see the questionnaire and ask for any additional information regarding any other data collection instrument before giving consent.

## What will happen to the information about you?

All personal data will be treated confidentially. Only the project group (see general information section below) will have access to the personal data. The list of names of the participating students will be stored in NTNU Sharepoint according to the data processing agreement between NTNU and Microsoft. Only the researchers and data controller will have access to the data.

We state that the participants will not be recognizable in the publication. The project is scheduled for completion by June 2022, then all data be anonymized.

## Voluntary participation

It is voluntary to participate in the project, and you can at any time choose to withdraw your consent without stating any reason. If you decide to withdraw, all your personal data will be made anonymous.

## Participants' rights

Participants have the right to request access to/deletion/correction/limitation of personal data, the right to data portability, and the right to send a complaint to the Data Protection Officer at NTNU or The Norwegian Data Protection Authority about the processing of personal data.

## General information-project group:

The leader of the project is Michail Giannakos, Associate Professor at Department of Computer and Information Science at NTNU, e-mail: [michailg@ntnu.no](mailto:michailg@ntnu.no), address: Sem Sælands vei 9, IT-bygget \* 103, phone number: +47 73593469.

If you would like to participate or have any questions concerning the project, please contact:

Edvard Mattias Bø, e-mail: [edvardmb@stud.ntnu.no](mailto:edvardmb@stud.ntnu.no), mobile number: +47 94489480 Master student at the Department of Computer and Information Science at NTNU.

Data Protection Officer (Personvernombud) at NTNU (Thomas Helgesen, [thomas.helgesen@ntnu.no](mailto:thomas.helgesen@ntnu.no))

The study has been notified to the 'NSD – The Norwegian Centre for Research Data AS ([personverntjenester@nsd.no](mailto:personverntjenester@nsd.no), 55 58 21 17) has assessed that the processing of personal data in this project is in accordance with data protection legislation.

## Consent for participation in the study

I have received information about the project and I am willing to give my consent for my participation.

Participant's name: \_\_\_\_\_

---

(Signed by participant, date)



## A.2 Introduction to SmartU

### SmartU

SmartU is the result of a previous master's thesis. They developed an adaptive self-assessment system, which allowed students to take quizzes that were designed to best fit the individual's expertise based on the previously provided responses and to successfully detect and determine the individual's mastery level.

Specifically, the activities involve multiple-choice questions with 2-4 possible answers per question - where only one is correct. The questions are delivered to the students one-by-one (one at a time). Each time they answer a question, SmartU considers the correctness of the response and the difficulty of the question and decides which is the next most appropriate question for their mastery level. The goal is to detect their mastery level using the least number of needed questions. Due to the adaptivity, the number of questions that will be assigned varies between 10 to 20 questions. There is a question bank for each course, and the system chooses the questions. Each question has an assigned difficulty; hard, medium, or easy.

### Task-related metadata

When students work through questions, they are given the opportunity to view metadata about the question they are currently asked. This metadata includes performance (percentage of how many people got the question correct), effort (average peers putting effort into the question in percent. This is calculated based on how much time they spend answering, based on how much time it is estimated that they use. They can also view how much time on average their peers spent on answering the question, both when they got it wrong and when they got it right.

### The dashboard

The dashboard you are about to test will give you an overview of the questions and students' performance in a given course. It will allow the teacher to monitor, study and edit the questions in the course, to make sure that students are learning as well as possible.

### A.3 SmartU Post-Experiment Questionnaire

## SmartU Dashboard

\* Obligatorisk

### Introduction

#### 1. Gender \*

- Woman
- Man
- Non-binary
- Prefer not to say

#### 2. Age \*

3. Teaching experience? \*

- Learning assistant
- Teaching assistant
- Teacher education (Lektor i Realfag)
- From primary education
- Other

## Usability of SmartU

The System Usability Scale (SUS) is a reliable tool for measuring the usability of a user-interface. It consists of a 10 item questionnaire with five response options for respondents; from Strongly agree

### 4. System Usability Scale \*

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
As a teacher, I think I would use the SmartU dashboard frequently.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found the teacher dashboard unnecessarily complex.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I thought the teacher dashboard was easy to use.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I think that I would need the support of a technical person to be able to use the teacher dashboard.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found the various functions in the teacher dashboard were well integrated.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree

I thought there was too much inconsistency in the dashboard.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would imagine that most people would learn to use SmartU very quickly.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found SmartU very slow or complicated to use.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I felt confident using the dashboard.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I needed to learn a lot of things before I could get going with the dashboard.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5. Overall evaluation of the usability of SmartU \*

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I think that the navigation in the dashboard was nearly	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

effortless.

I think that when I needed help to learn how to use the dashboard, the system provided me with sufficient information.

I think that a user who has never seen the dashboard before can learn how to accomplish basic tasks fast.

I think that experienced users should be able to complete tasks using the dashboard quickly.

I think that it is not frequent that users make errors while using the dashboard.

I think that the interaction with the dashboard is clear and understandable.

Strongly disagree      Disagree      Neutral      Agree      Strongly agree

Using the

using the dashboard makes me happy to accomplish my assessment tasks.

Using the dashboard gives me enjoyment for my continuous teaching.

Using the dashboard leads me to exploration in my continuous assessment.

## Graphs and visualizations used in the dashboard

This section consists of questions regarding the usage of graphs and visualizations in the graphical user interface of the SmartU system and NOT the available statistics during an assessment. The answers will be on a scale from 1 to 5, press the option that is the closest match to your

### 6. Usage of the dashboard \*

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Not relevant
It was easy to understand what the graphs reflected.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I was able to make sense of the graphs and extract information.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 7. Usefulness of the dashboard \*

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Not relevant
I think the dashboard helped my understanding of how the students were doing.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I believe that a graphical representation is more fitting than a	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Not relevant



making them a textual representation of the data displayed in the dashboard.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found it useful to view stats of the students' performance over time	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I understood that the mastery-level reflected level of knowledge as a result of previous performances	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found it easy to understand that the color green was used to represent a positive association.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found it easy to understand that the color yellow was used to represent a mediocre/neutral association.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found it easy to understand that the color red was used to represent a negative association.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	<input type="radio"/> Strongly disagree	<input type="radio"/> Disagree	<input type="radio"/> Neutral	<input type="radio"/> Agree	<input type="radio"/> Strongly agree	<input type="radio"/> Not relevant

association.

I think that the graphical representations in the dashboard provided all the information I needed.

#### 8. Attitude towards the dashboard \*

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Not relevant
I believe that the provided visualizations in the dashboard made information easy to extract.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I was motivated to further use the dashboard to see progression unfold in the visualizations.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I felt more informed and confident conducting an assessment when utilizing the dashboard.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Not relevant

The design and colors in the dashboard made me feel confident and comfortable while using it.

The design and colors in the dashboard made me feel happy.

I would like to continue to use this dashboard to follow up on students' progress.

All the available statistics and suggestions made me more motivated to monitor the course questions.

9. Other comments regarding the use of dashboard.



## A.4 Observation guide

# Observation guide

**Introduction:** Before the experimental study and observation begins, the observer should introduce itself and the role it will take during the experimental study. A connection should be established between the observer and the test-subject, creating a relaxed and comfortable setting for the study to be conducted in. The perfect environment for the study to be conducted in is if the study could be conducted in an environment-setting familiar to the test-subjects so they would interact with the system as they normally would.

**Interaction with test-subject:** In the role of an observer, interaction with the test-subject should be kept to a minimum, except for when the test-subject asks for help. It is important that when providing help, the help does not take too long, or become very frequent.

**Capturing contextual information:** Capture the ambience, atmosphere and environment in the room during the test. Also try to capture the test-subjects' emotional state, are they stressed/nervous/relaxed? Does the emotional state change during the test?

The aim of this point is to develop a set of notes that describe the scene as much as possible.

**Capture crucial information:** When observing, you should note what the test-subject is doing while testing. You do not have to take notes of every single action or mouse-click, but your notes should summarize the events that happened and unexpected actions the user made.

**Examples of events/actions to note:**

Category	Includes	Researchers should note
Test environment	Ambiance, atmosphere, alterations to environment	Any differences in test environment between tests should be noted
Subjects emotional state	The user's state of mind	Is the user stressed or nervous? Does the state of mind change during the tests? Is the test subject focused on the task?
Unexpected events	Failing to navigate, bugs in the system	Did anything unexpected happen during testing? Note when, why and how it happened
Help	Frequency, topic of questions	What does the user ask you to help out with and how often/ many times are you asked to help?
Disturbances	People entering/leaving the room, phone notifications, noises, etc	Any unwanted noises or disturbances hindering the focus on the testing should be noted.

**End of testing and observation:** It is important to thank the test-subjects for their time and to provide them with the interviewer's contact details. Depending on circumstances, it may also be worth letting respondents know how they can obtain the project reports because this provides them with a sense of ownership of the material that they have shared.

## A.5 Interview Guide

### Semistructured Interview Guide

**Capturing answers:** Recording of answers will be done through taking notes and audio recording. This procedure allows the interviewer to highlight key points to probe further on relevant topics, while making sure no information is lost due to the transcripts of the recordings.

**Develop a report with the respondent:** Obtaining meaningful information from respondents will be easier if they are comfortable opening up to the interviewer. This can be done by asking non probing questions related to their hobbies, their spare time, and so on.

**Ask questions that lead to detailed answers:** It is important that you phrase questions in a way that gets respondents to provide detailed answers, rather than simple “Yes” or “No” answers.

**Examples of questions:**

- Did you find the teacher dashboard interesting? If you did, how so?
- How did you feel about the homepage and presented visualizations?
- Do you think that the system would improve your understanding of how well the quizzes were working? If you do, why would it do that?
- Could you mention some features that you found useful in the system?
- How did you feel about the statistics provided about the quizzes?
- Could you mention some features which were easy or hard to understand?
- Did you feel like the statistics helped you understand what students understood and did not? How did the statistics help you?
- Would you use the system again? If so, what encourages you to do so?
- What do you think of the mastery element to visualize your knowledge?

It is good to have a set of questions on hand, but the interviewer must also be prepared to expand on, or probe, the predetermined questions as the need arises. This is the essence of qualitative interviews.

**End the interview:** Deciding when to end an interview may depend on a number of factors. E.g. interviewers may feel that they have exhausted their questions, and that they are no longer getting new information or if the respondent seems tired or has other commitments to attend to. It is good practice for interviewers to summarize the key points that they feel the respondent has provided, because this gives the respondent a final chance to expand or clarify any points. Finally, it is important to thank the respondent for their time and to provide them with the interviewer's contact details. Depending on circumstances, it may also be worth letting respondents know how they can obtain the project reports because this provides them with a sense of ownership of the material that they have shared.

## **A.6 Interviews**

The following section presents all the interviews conducted during the study. These were conducted in Norwegian, as that was the native language of all the participants.



## A.6.1 Interview 1

# Interview 1

I: Syns du noen av de integrerte tingene i dashbordet var interessante eller nyttige?

P: Hovedforsiden var nyttig. Da fikk man visualisert det viktigste med en gang. Jeg føler de viktigste er average performance, for å se hvordan studentene ligger an generelt. For hvis avg perf er høy, vet man at man gjort mye riktig. Likte også den store visualiseringen med stolpene, hvor man så alle spørsmålene i et fag. Når man kunne filtrere og se gjennomsnittlig tid brukt. Der kan man se med én gang hvor problemet ligger, hvilket spørsmål som sliter. Tabellen med alle spørsmålene var ikke like nyttig, var mange parameter, og var ikke visuell. Tror jeg ville brukt stolpediagrammet mye.

I: Var det noen av filtrene du syns var mer nyttig enn andre?

P: Ja for eksempel gjennomsnittlig tid brukt. Det sier litt om hvor vanskelig det er. Og antall ganger brukt, for å se hvilke som blir brukt. De som ikke blir brukt indikerer jo at noe er problematisk, så det er interessant. Og performance er spennende å se. Greit å se hvilke som er vanskeligst.

I: Hva følte du rundt all statistikken? Syns du det ble litt mye data? Litt overveldende?

P: Det var litt mye, men nå brukte jeg det ikke så lenge da. Det var forvirrende med grafer hvor man hadde fargekodet ulike ting med de samme fargene. Det var forvirrende, men om man leser så skjønner man hva det er. Grafer er alltid nice, så jeg syns ikke det var for mye av det. Men det var noen grafer som var litt like. Dashbordet over faget var bra, for der hadde man ulike representasjoner. Man hadde det lilla sirkel diagrammet, og så hadde man de stolpediagrammene, og så hadde man den brøk-greia. Det er bra, for da er det lettere å se forskjellen.

I: Syns du det var bedre at det var grafer, kontra tekstlig representasjon?

P: Ja helt klart.

I: Tror du et dashbord som dette ville gjort det mer motiverende for deg å gå gjennom hvordan elever har gjort det?

P: Ja det gjør det jo, fordi det er lett å se hvor jeg skal sette inn innsatsen, og det gjør det mer motiverende. Det er verre å åpne en side med masse tall, hvor jeg må finne ut av ting selv. Dette gjør det lettere å finne ut hvor problemene ligger, og da er det mer motiverende å komme i gang med også.

I: Var det noen features du syns var spesielt gode?

P: som jeg nevnte, likte jeg veldig godt den store oversikten i stolpediagrammet, og performance-sirkelen. De ga meg mest ved første øyekast.

I: Det var en forslagsfunksjon, hvor systemet prøver å hjelpe deg hvor du skal se. Syns du den fungerte?

P: Ja den hadde jeg glemt! Den var veldig bra! Det gjør det desto mer motiverende, for den gjør jo jobben for deg på mange måter. Den viser deg hva som kan være lurt å gjøre. Alt man trenger er å bare klikke seg inn og gå og se. Den var lur!

I: Hva syns du om posisjoneringen av den?

P: Det tenkte jeg ikke så mye over. Det er jo veldig mye informasjon på en side. Notification var bra, for den må du ha synlig tidlig tenker jeg. Og den generelle oversikten lå også bra. Det lå egentlig bra der det var.

I: Følte du at all statistikken hjalp deg forstå hvordan studentene gjorde det, og hvordan spørsmålene gjorde det? Følte du at du fikk den informasjonen du trengte?

P: Nå husker jeg egentlig ikke så mye hvordan det så ut. Men jeg føler forslagene sa veldig tydelig fra om hvilke spørsmål som var "outcasts", så det kom tydelig frem.

I: Var det noen ting du syns var trøblete eller vanskelig å forstå?

P: De stolpegrafene hvor du hadde tre stolper under hver kategori (high, medium low).

I: Hvor lang tid man hadde brukt på å svare riktig, galt og totalt, per nivå man var i?

P: Ja, for der er det mange parametere å ta innover seg. Der er det tid i forhold til riktig svar, og hvilket nivå. Der var det litt mye å ta innover seg. Også på grunn av fargebruken, at de samme fargene ble brukt til flere ting. Og den store tabellen når du trykket på "Get more details". Det er veldig mye informasjon. Selv om det er forståelig at det er sånn, er det litt overveldende, og man må vite hva man ser etter når man ser på den. Det er vanskelig å gjøre så mye informasjon forståelig. Men i alle fall det stolpediagrammet med tre stolper per kategori, det var litt kronglete.

## A.6.2 Interview 2

### Interview 2

I: Syn du dette dashbordet virket interessant for deg som studass? I såfall, på hvilken måte.

P: Jeg vil absolutt si det. Nå skal jo systemet og oppgavene være smarte i seg selv, men her kan man endre på oppgavene hvis det er feil i en eller annen retning. Og det med at det var superenkelt å endre alt mulig, i hvert fall av det jeg så, gjør det kjempeenkelt å fortsette å gjøre det bedre for studentene.

I: Ja. Du nevner dette med å oppdage om det er noe feil med spørsmålene. Følte du at du fikk den nødvendige informasjonen du trengte for å kunne se slike ting?

P: Om det er feil i oppgavene?

I: Ja, man får jo presentert en del statistikk her. Føler du den statistikken er til hjelp? Gir den deg relevant informasjon, eller føler du det bare er støy?

P: Til å begynne med føles det ut som støy, for det er mange måter å se ting på. Men jeg regner med at om man bruker dette en liten stund så får man sikkert en form for favoritt av disse tingene, for å se om det er helt feil for studentene. For eksempel om man ser der. \*peker på chance of success\* og så kanskje man vil få mer informasjon om hvorfor det er sånn, ved å se på de andre (statistikkene). Så ja, som jeg sa: Støy til å begynne med, men når man først kom inn i det tror jeg det funker.

I: Du nevner favoritt. Var det noen av disse elementene som pekte seg ut som spesielt nyttige?

P: Da jo holdt på med dette så jeg først på de enkle tingene, som om folk har gjort det riktig. Og så går det an å sammenligne med hvor lang tid man bruker på å svare. Jeg prøvde egentlig å se etter det når jeg testa. Sånn, "er det en sammenheng mellom hvorfor folk har svart helt feil og bruken av å analysere metadata?". Og den sammenhengen tror jeg det er viktig å se på.

I: Ja. Så å se på sammenhengen mellom de ulike tingene?

P: Ja. Si et spørsmål har veldig lav performance. Da kan man prøve å få folk til å se på metadata for eksempel, hvis studentene ikke har gjort det. Sånne ting.

I: så for deg var det nyttig å se de ulike statistikkene i sammenheng med hverandre?

P: Ja, det vil jeg si

I: Var det noen ting i dashbordet du synes var vanskelig å skjønne eller uklart?

P: Ja, Effort var litt vanskelig å forstå. Jeg forsto det etter en liten stund. Det gir mening, men den er litt mer komplisert enn de andre.

I: Hva syns du om den oversikten med alle spørsmålene? Hvor man kunne se

P: Den syns jeg var spot on egentlig. Man kunne se alle spørsmålene, og hover effekten viste hvor mange som har svart feil og riktig og så videre. Hvis man ikke hadde fått informasjon når man hovret over, hadde jeg blitt forvirra, men det funka veldig bra. Og når jeg skulle se på spm 2 på testen; super easy. Veldig lett å gå inn, og veldig lett å skjønne.

I: Så bra. Du fikk jo også litt forslag underveis når du jobbet gjennom. Hva syns du om det?

P: Veldig nice. Er det automatisk? Det går på prosent osv?

I: Tanken er det ja. Men det gjør ikke det i dette eksempelet.

P: Syns det var dritbra.

I: Føler du at det er et element som ville gjort det mer motiverende for deg som studass?

P: Absolutt. Spesielt hvis den klarer å få med seg alt som er viktig. For jeg tenker sånn; hvis det bare kommer sånn innimellom eller kommer på noen, så hadde jeg sikkert bare gjort suggestionsene egentlig, og ikke gjort noe mer sikkert. Men det får deg til å handle. Hvis du er en studass som gjør lite, så er det der perfekt. Og ikke bare fordi det er der, men hvis du er en studass som går inn og ser men ikke endrer så mye, så får du kanskje endret noe.

I: Ja mm. og som en helhet da, føler du at et dashboard som dette ville gjort det mer motiverende å skulle rette studenters arbeid?

P: Absolutt, det tror jeg ikke det er noe spørsmål om egentlig. Dette er livlig, enkelt og noe helt annet enn å bare få inn oppgaver. Og veldig enkelt å holde oversikt, i stedet for å ta én og én.

I: Ja, sant. Nei men bra, har du noe mer du vil tilføye? Noen tanker?

P: Mmm, nei jeg tror jeg har fått ut alt jeg mente.

I: Nei men bra, supert! Takk for hjelpa!

P: Bare hyggelig.

### A.6.3 Interview 3

## Interview 3

I: Nå har du jo prøvd det dashboardet. Syns du det virket interessant?

P: mmm ja

I: Hvorfor det?

P: Det ga innsikt i hva studenter driver med, så det er nice. Og så liker jeg å kunne jeg se hvor lang tid studentene bruker. Det er ganske fint, da kan man se "okei det her er en de grubler på. Kanskje jeg burde si noe mer om det, siden de bruker lang tid på den".

I: Mmm, ja. Hvilke aspekter ved dashboardet likte du best, eller syns var mest interessante.

P: Som sagt, tid likte jeg godt, det var veldig bra. Og at man kunne se individuelle spørsmål, da fikk man en dyp innsikt i spørsmålet. Likte også den oversikten hvor jeg kunne se alle spørsmålene. Da kunne man spesifikt se akkurat hvilke spørsmål folk sliter med. Det var nyttig.

I: Hvorfor likte du den delen best?

P: Fordi da kan du hjelpe flest fokusere på det problemet. Du får fokusert på de tingene flest folk sliter med.

I: Mm. Var det noen andre aspekter du likte?

P: Ehm... Fargeskjemaet var nice. Gode klare farger. Rød, grønn, gul er enkelt. Intuitivt og forståelig.

I: Du fikk jo litt forslag der. Hva syns du om de?

P: Det er bra. De er veldig nice. Da må du ikke gå inn og se spesifikt, du kan bare raskt få et estimat. Men det så ut som det var ganske mange, jeg ville kanskje fått en liste over hvorfor det er et forslag. Det står jo litt, men da kunne jeg ikke se det sammenlignet med de andre spm (som i den store oversikten), hvor dårlig eller hvor bra de gjør det sammenlignet med de andre. Det er derfor jeg likte bedre å se det sammenlignet på en linje, hvor man kunne klart se hvorfor den er foreslått.

I: Så du kunne egentlig tenkt deg å ta de avgjørelsene selv?

P: Ja, for da ser jeg det selv. Gi meg statistikken, så kan jeg se på det selv.

I: Ja, bra. Syns du et dashboard som dette ville gjort det mer motiverende for deg som studass?

P: Ja det tror jeg ville gjort det lettere. For da kan jeg vite at dette er noe folk skjønner, så kan jeg heller lese meg opp på det før studasstimen, så vet jeg litt hva som kommer.

I: Så all statistikken du fikk presentert her gjorde det både mer motiverende men også mer forståelig for å skjønne hva studentene sliter med?

P: Ja veldig.

I: Så bra. Var det noen deler du synes var vanskelig å forstå?

P: Ja, men det var fordi jeg glemte at studentene kunne se metadata. Kunne kanskje stått "times students spent analyzing metadata". Vært formulert litt bedre. Og så liker jeg at notifications er en bjelle og en bar man kan se. Så slipper jeg å ha den der hele tiden.

I: Ja sant, nei men bra. Var det noen andre ting?

P: Nei, tror det var det. Hadde vært digg om man kunne sortere den store tabellen med alle spørsmålene, sånn sortere by success chance osv. Jeg synes times used og success chance var nyttige i den store grafen.

I: Ja. Er det noen andre ting?

P: Kanskje flere vanskelighetsgrader enn 3. Det blir hardt å kun dele inn spørsmålene etter tre vanskelighetsgrader.

I: Ja, supert, takk for hjelpen!

#### A.6.4 Interview 4

## Interview 4

I: Nå har du jo fått testet dashbordet til SmartU. Det jeg lurer på da er; syns du det dashbordet virket interessant?

P: Ja, hvis jeg hadde hatt den dataen tilgjengelig på spørsmålene, hadde jeg definitivt brukt det. Det jeg er mest bekymret for er hvordan man skal få tak i dataen. For man må jo ha litt oppsett for å samle det. Men selve dashbordet syns jeg fungerte bra, og hvis jeg hadde hatt tilgang på den dataen hadde det vært kult.

I: Ja. Var det noen deler du likte spesielt godt, eller fant spesielt nyttige?

P: Mmm.. Syns spesielt hva folk har svart på spørsmålene og hvor lang tid de har brukt er spennende. For å se om noen spørsmål er dårlige eller gode.

I: Ja, å se distribusjon, hvilke alternativ de har svart, og hvor lange tid de har brukt?

P: Ja, på hvert spørsmål. Jeg syns kanskje det mest nyttige var å kunne gå inn på hvert spørsmål.

I: Ja.

P: og så syns jeg det virket lett å se hvor mange som svarer feil på spørsmålet, så man kan velge hvilke spørsmål man har lyst til å se. Og så syns jeg de forslagene var ganske bra, så man slipper å lete etter det selv. Så du bare får "Ah dette bør du kanskje se på?". Og det er kanskje sånn at hvis man skal bruke det over lengre tid, så har man kanskje lyst til å kunne påvirke hva som kommer i forslagene. Så kan man si "Generelt vil jeg se alle spørsmålene hvor over 70% har svart feil"

I: Ja! At du kan si litt hva du selv ser etter?

P: Ja. Jeg vil i alle fall kontroll over hva den foreslår. I hvert fall vite hva det er. Hvis det skal være en snarvei til å finne ut hva som er galt, vil jeg ha muligheten til å påvirke hva den skal se etter og gi meg.

I: Ja sant. Var det noen deler du syns var vanskelig?

P: Jeg syns det var vanskelig å forstå effort til å begynne med, men jeg tror når man blir vant med det at det er nyttig å ha det der det er.

I: Ja den er litt vrien å skjønne i starten. Man blir jo presentert for en del statistikk, syns du det blir presentert på en forståelig måte?

P: Jeg syns det var veldig forståelig. Det jeg er mer bekymret for er når man har brukt det over lenger tid, vil man kanskje få mer innsyn i detaljene. For hver av grafene, så man kan

se rådataen. Men for meg som ikke har sett den dataen før var det veldig lett å få et overblikk.

I: Så du kunne tenkt deg et enda dypere innblikk?

P: Ja, det kan hende det bare er innimellom da. Og jeg vet ikke hva det skulle vært. Men kanskje hvis jeg ser "det er én student som har brukt veldig lang tid på alle spm", som trekker det opp veldig. Jeg kunne tenkt meg å ha muligheten til å sjekke ut detaljene, og kanskje redigere bort hvis det er noen outliers.

I: Ja skjønner. Tror du at du ville blitt mer motivert for arbeidsoppgavene dine som studass om du hadde hatt et sånn her dashbord tilgjengelig?

P: Tja, jeg vet ikke hvor relevant dette ville vært som studass. Jeg tror det hadde passet bedre et hakk opp. Som studass har man ikke så mye å si på hvilke spm som blir stilt og sånt. Et nivå opp må man passe på hvor bra øvingsopplegget er og sånne type ting. Men definitivt nyttig til når man skal velge hvilke spørsmål man skal stille. Ofte har undassen noe å si, andre ganger er det professoren. Jeg har vært undass, og da har vi hatt ansvaret for spørsmålene som blir stilt. Så der er det veldig aktuelt å kunne ha en sånn oversikt. Hvor man kan se hvilke spm som kanskje burde endres på.

I: Og da tror du noe sånt som dette ville gjort det mer motiverende og lettere å jobbe med?

P: Ja. Veldig mye raskere å se hva som funker og hva som ikke funker. Det kunne motivert mer til hyppigere endring. Bytte øvingsopplegget oftere. Hvis man ser at "Her bruker folk veldig kort tid og får riktig hele tiden" så kan man anta det er kok på gang. Da kan man anta man må endre noe sånt. Så, definitivt kjempenyttig.

I: Så bra! Er det noen andre inntrykk du vil dele?

P: Nei, syns det var nice.

I: så bra, supert. Tusen takk for hjelpa!



## A.6.5 Interview 5

# Interview 5

I: Ja, da har du fått prøvd dashbordet litt. Hva syns du?

P: Jeg syns det var et ganske kult konsept. Jeg er vant til å bruke Blackboard fra da jeg var studass, og det er et mye vanskeligere verktøy, så det var mye lettere å få innsikt i hvordan studentene holdt på med quizene her. Ikke at jeg har hatt så mange quizer som studass, men jeg tenker det var mye bedre enn hva blackboard har styrt med.

Jeg likte veldig godt innsyn i hvert spørsmål. Det syns jeg er en god måte å få ta tak i spørsmål fort. Der du kan se med en gang "åja dette er et vanskelig spm. Dette var dårlig formulert" sånne ting. Så jeg tror dette kan være et viktig verktøy for å ta tak i de problemene der. Hvis for eksempel multiple choice er en viktig del av vurderingen. Ja...

I: Så bra. Du nevner du likte godt å få innsyn i hvert spørsmål. Var det noen andre deler du likte godt?

P: Ja! Skal vi se. Jeg tror effort og performance syns jeg var kult, men det var ikke helt.. Det indikerer på en måte hvordan quizene ligger an. Hvis man hadde flere quizer kan man fort se gjennom, og se på effort og perf, for å se hvordan ting ligger an, og da kan du ta tak i problemer hvis du ser at performance er veldig lavt med en gang. Så det syns jeg er bra. Tidsmessig (statistikk) var også nyttig, men jeg følte ikke det var... Det var mer en bonus. Jeg vet ikke hvor mye jeg ville brukt det. Det er vanskelig å vite hvor lang tid spørsmål tar. Noen bruker lang tid, noen ikke. Det er vanskelig å finne noen korrelasjon, men en liten bonus. Score attempt er også en fin ting å se. Jeg tror på blackboard ser man bare antall. Så det er fint å se hva success raten (score) per attempt er. Tja.. Tror ikke jeg har noe mer enn det...

I: Nei, men det var bra! Du sammenligner jo litt med Blackboard, tror du det hadde vært mer motiverende å rette det studenter har gjort i et dashbord som dette? Føler du motivasjon øker med noe sånt?

P: Ja, definitivt. Det er mye mer moderne, og mye mindre komplisert enn det Blackboard har. Så definitivt ville jeg brukt det her, i alle fall til å lage multiple choice quiz. Det virker veldig enkelt å endre ting, og mye lettere å finne informasjon hvis man trenger den informasjonen. Så ja, jeg ville definitivt brukt det her ovenfor Blackboard, fordi Blackboard er veldig komplisert.

I: På hvilken måte?

P: De har veldig lite statistikk mener jeg. Jeg tror ikke jeg kan huske at de har det. Hvis jeg husker riktig har de bare noen tall som sto der, jeg kan ikke huske det var noen grafer på noe vis. Det er veldig lenge siden jeg lagde en quiz og sånt i BB, men jeg mener jeg har gjort det før. Og det var veldig vanskelig for student å se hva som var riktig svar og sånt. Så jeg tror det her er myemye bedre enn det. Jeg tror også det var vanskelig å endre spørsmål, hvis jeg ikke tar helt feil. Så ja, Blackboard er bare mye mer komplisert, fordi det er mange unødvendige steg. Det her er færre clicks unna å gjøre det du vil gjøre.

I: Ja... men forstår jeg deg riktig da, med at du likte dette bedre fordi det var mer statistikk, og fordi den ble fremstilt med grafer, og ikke som tall.

P: Ja. Jeg syns det... Ja. Det syns jeg. Jeg syns visualiseringen hjelper mye mer, og da slipper man å ... ikke at sammenligningen av tall er så vanskelig. Men man kan si det tar litt ekstra hjernekapasitet, enn å raskt se over grafer. Så da er det lettere å se med en gang at "åja, her er det noe feil", uten å måtte se på tall og sammenlikne det med hverandre.

I: så bra. Du fikk jo litt forslag, fra systemet. Hva syns du om det?

P: Suggestions, mener du?

I: mm

P: eehm... jeg syns det var kult. Ehm.. skal vi se nå må jeg bare huske. Jeg gikk ikke så mye gjennom det. Sånn jeg ser forslag vil jeg tro den viser alle spørsmål man kanskje bør endre på, siden folk har svart feil og sånt. Da blir det igjen litt enklere å slippe å analysere hvilke spørsmål som har gjort det dårlig, og gå gjennom hver, når det bare kommer opp med en gang. Og det er jo veldig digg da. Det er færre steg til det man vil gjøre. Og da kan man ta tak i spørsmålet eller problemet mye fortere enn hva man ville gjort. En annen ting er at jeg vet ikke om så mange professorer og lærere, som kanskje gjør det her til vanlig, gidder å gå gjennom alle spørsmålene. Da blir det mye enklere å bare få noen forslag enn å selv begynne å analysere. Så det krever mindre jobb for dem, og da vil det kanskje være sånn at man tar tak i problemer fortere, eller gidder å gjøre noe med det enn at man må gå gjennom hvert spørsmål, og må analysere. Og som kanskje ikke gir noe. Og så skjønner man kanskje ikke helt hvorfor heller, mens her får man en tekst som forklarer.

I: mmm så bra. Var det noen deler du syns var vanskelig å forstå?

P: Ja jeg nevnte såvidt i spørreundersøkelsen. Disse tre fargene som indikerte tre forskjellige ting: correct, incorrect og neutral. Det kan komme litt feil ut fra high, medium og low. For det er samme farge som er brukt, som viser ulike ting. Så det hadde vært fint å separere de fargene der. Og kanskje ikke indikere at low er dårlig. \*ler\* selv om det er dårlig. Jeg skulle ønske det var kanskje mer navn på aksene. Jeg er veldig vandt til å ha, og det er litt enklere å se med en gang hva grafene viser. Så slipper man å se andre steder, for man ser det veldig naturlig. Den refresh greia syns jeg var litt kronglete. Jeg skjønte i etterkant at man ikke alltid vil resete statistikken. Det skjønte jeg ikke i første omgang. Da følte det veldig naturlig at det var reset selv

Litt vanskelig å skjønne hva som var riktig svar, siden det bare var bold tekst blant mye annen bold tekst. Hadde vært fint å ha noen farger der. Litt mer spacing hadde vært fint, hvis jeg får lov til å si det. Det hadde vært fint å kanskje ha tall på det sirkelbarene

Det står jo under der, men det hadde vært fint å ha det på selve greia. Nå vet jeg at det er en prototype, men kanskje ha tekst ved siden av fargene i her.

\*peker på kakediagrammet i Overview\*

Ehm ja. Det er det jeg kommer på i hvert fall.

I: Veldig bra! Veldig mange gode tilbakemeldinger. supert!

### A.6.6 Interview 6

## Interview 6

I: Ja, da har du jo fått prøvd dashbordet. Syns du det virket interessant?

P: Ja.

I: Ja. Har du lyst til å fortelle litt om hvorfor?

P: Ja, fordi som læringsassistent så er det spennende å se hvordan folk ligger an. Og dette var et enkelt og oversiktlig system på det her. Sånn, det var gøy at det var litt smart, ved å regne ut effort gi forslag. Og så virket det veldig oversiktlig og greit. Du får detaljert, fin informasjon om alle det forskjellige spørsmålene. Ja, og notifications hvis du trenger det. Og det var ganske greit oppsatt. Det var bare "Yey det er rett der, og det er ikke vanskelig å finne frem"

I: Så bra. Var det noen deler du likte spesielt godt?

P: Jeg syns suggestions greia var veldig bra. Fordi, det var sånn "oi ja, kanskje den gir deg forslag du ikke har tenkt på fra før"

I: Ja ikke sant...

P: Mmm ja.. Hva var det du spurte om?

I: Om det var noen deler du likte spesielt godt?

P: Ehm.. ja. Alt egentlig. Jeg syns det er gøy å se performance, og hvor lang tid har de brukt på å svare på det spørsmålet, i forhold til om det har svar riktig eller feil. Og så den andre tingen... Hva var det jeg så på da? Ja! Hvor mye de bruker den meta-greien, om det hjelper på en måte. For da vet du om det er nyttig eller ikke. Og så likte jeg veldig godt alle fargene! Jeg syns det var veldig fint med farger. Og det her  
\*Peker på barchart\*  
Gjorde det veldig lett å se. Sånn kontrastene mellom oppgavene.

I: Ja. Det gjorde det enkelt å ...?

P: Å se hvilke spørsmål var veldig bra.. Eller liksom spørsmål blir brukt mest.

I: Ja hvilke som skiller seg ut?

P: Ja. mmm. Det var lett å se ekstremalpunktene, men de i midten blir litt borte i mengden.

I: Ja.. mmm skjønner det. Tror du et system som dette ville gjort det mer motiverende for deg som studass?

P: Ja. Jeg tror det hadde vært nice å se hvor lang tid de hadde brukt, og om de har brukt veldig lang tid og fått feil svar, så hadde jeg brukt mer tid på å forklare ting nøyere. Mens hvis en person som hadde brukt veldig lite tid, og svar riktig. Eller brukt lite tid og svart feil, så kan man liksom tenke at det er en slurvefeil. Så trenger du ikke forklare ting i detalj så mye. Så sparer man en del tid på tilbakemeldingene. Og så er det også nice å se hvor mye effort de har putta inn vs performance. Kanskje øvingsopplegget er 100% bra hvis det er 100% effort, og 30% performance.

I: ja. Så du tror dette systemet ville gjort det mer motiverende for deg?

P: Ja, da blir det mer motiverende å skrive lange tilbakemeldinger til folk som har strevd litt.

I: Ja sant. Nei men bra. Ehm.. Følte du at du fikk den informasjonen du trengte fra dette dashbordet?

P: Ja. spesielt det med tid, og om de svarer riktig eller feil. Du ga meg jo masse gøy informasjon her. Og så hvert spørsmål også så går det jo an å se hvor mange det var som svarte feil. Og så var det veldig nice å ha med easy, medium, hard. Det er nice å få oversikt, og å skille mellom de også.

I: Var det noen deler av dashbordet du synes var vanskelig?

P: Ja jeg skjønte først ikke spørsmålene har blitt brukt 752 ganger on average. Skjønte ikke helt det. Men ellers synes jeg ting egentlig er ganske selvforklarende. Noen ganger går jeg bare for fort inn i en graf, og da skjønner jeg ikke hva den var. Fordi her sånn:

\*Peker på avgTTA\*

Så tenkte jeg først at det var hvor mange har svart rett på et spm med high, og hvor mange har svart feil. Men det er litt brukerfeil, for jeg kunne bare lest overskriften.

I: Ja så da misforsto du hva grafene representerte. Men du forsto det selv da du leste overskriften.

P: Ja. For jeg så på siden, så sto det sekunder. Og da var ting litt feil.

I: Ja

P: Men ja, jeg synes egentlig ting var veldig lett å forstå.

I: Ja, så bra.

P: Ja.. mmm. Ja fint å ha en average (total time spent to answer). Litt helhetsinntrykk, det er nice.

I: Bra. Noen andre kommentarer eller tanker?

P: Mmmm. Godt spørsmål. Ting var veldig å forstå egentlig. Det var veldig oversiktlig.

I: Det er veldig bra. Det er godt å høre

\*spørsmål om hvordan quizen fungerer\*

I: Nei men bra, takk for hjelpa!

P: Bare hyggelig!

### A.6.7 Interview 7

## Interview 7

I: Ja, da har du fått prøvd dette dashbordet. Syns det virket interessant ?

P: Ja

I: Ja, kan du fortelle om hvorfor?

P: Det er jo interessant å se hvordan studentene gjør det på quizene. Sånn med tidsbruk og kanskje spesielt også om mange har svart feil på et spørsmål eller brukt lang tid. Så kan man se hva folk sliter med. Og litt gøy å se om man gjør det bedre når man har gjort flere forsøk.

I: Ja sant. Så du syns det er interessant at du får et innblikk i hvordan de gjør det, med all den dataen?

P: Ja

I: Var det noen elementer du syns var spesielt nyttige eller likte spesielt godt?

P: Hvor lang tid de har brukt, og hvordan de har gjort det, med avg performance. Og denne var gøy.

I: Ja, de forslagene du får opp.

P: Ja den var gøy. For da får du opp "Denne bør du kanskje se litt på", uten at du trenger å gjøre det selv. Den gjør det for deg, så den likte jeg godt.

I: Ja, den syns du var nyttig.

P: Ja

I: Så bra. Det er jo en del ulike screens her, som gir oversikt på litt ulike måter. Var det noen du følte ga bedre innsikt enn andre?

P: Det spørs litt. Hvis man skal se generelt i faget, hvor folk sliter så kan det være greit å se på denne, hvor man ser alle spørsmål. Men hvis du da har et tema eller noe mer spesifikt du vil se om folk har forstått er det veldig greit å se spesifikt på spørsmålene. Jeg syns egentlig begge deler er nice.

I: Så bra. Tror du et dashbord som dette ville gjort det mer motiverende for deg som studass å gå inn og rette?

P: Kanskje. Nå har ikke jeg vært studass i så mange fag som dette, jeg har vært i mer prosjektbaserte fag, så det er vanskelig å si. Men hadde det vært sånne quizer, hadde det vært nice å se "okei du ser disse sliter med dette", så hvis de kommer og spør om hjelp, så

kan man være litt "okei skjønnte du det? Det var mange som slet med akkurat dette. Jeg så at du brukte litt tid på den". Man får litt oversikt. Det blir lettere å se hvor folk sliter, sånn at man vet hvor man skal legge lista når man snakker med folk.

I: Så det ville vært fint å bruke dette til å forberede seg, til møte med studenter?

P: Ja. Jeg tror hadde vært fint å få en oversikt for å se hva folk sliter med.

I: Så bra. Følte du at du fikk den informasjonen du trengte av statistikken?

P: Jeg synes det er nyttig, det man får presentert. Effort er kanskje det jeg ville brukt minst, tror jeg. Det er den jeg i alle fall nå følte jeg fikk minst ut av. Men ellers, synes jeg det er nyttig

I: Så bra. Var det noen aspekter eller elementer du synes var vanskelig å skjønne?

P: Jeg stussa på effort, så det var bra det var et spørsmålsteget ved den, som ga med informasjon om hva det betydde. Og så stussa jeg litt på "Chance of success", men jeg tror jeg skjønnte det riktig. Men jeg måtte bare tenke litt ekstra. Utenom det, nice, men... En liten kommentar.

I: Bring it

P: Fordi både den, den og den er tid. Men disse to er stående, og den er liggende. Så da må man endre hvordan man ser på grafen. Så det hadde kanskje vært nice om de var konsistente.

I: Ja, så det følger de samme reglene

P: Ja, så slipper man å tenke... Så kan man tenke litt mindre når man leser det.

I: Ja, det er noe med det. Noen andre kommentarer?

P: nei, ikke som jeg har tenkt på. Ellers så ser det bra ut!

I: Så bra! fint.



### A.6.8 Interview 8

## Interview 8

I: Okei, nå har jo du fått prøve dette dashbordet? Syns du det virket som et fornuftig tillegg til SmartU?

P: Ja det syns jeg. Man får en god oversikt over hvordan det står til, og man kan også få en overordnet oversikt, samt en mer detaljert oversikt hvis man ønsker det. Så jeg følte den dekket de behovene man måtte ha i en sånn situasjon.

I: Så bra! Var det noen aspekter du syns var mer nyttig enn andre?

P: Jeg likte at man kunne gå inn på hvert enkelt spørsmål, og få mye informasjon på det. Og så syns jeg også den store oversikten med alle spørsmålene var nyttig. Syns mye var nyttig egentlig.

I: Så bra! Var det noen andre elementer du vil trekke frem?

P: Ehm.. jeg likte at samtidig som man kan få veldig detaljert informasjon, så kan man også få et overordnet blikk på hvordan det står til også. Og så følte jeg at selv om det er veldig mye data, spesielt på enkeltspørsmålene, så er det likevel veldig greit å sette seg inn i. Ofte kan sånn være vanskelig å forstå, for det blir veldig mye på en gang. Det hjalp veldig med de små spørsmålstejnene, som forklarte hva ting var. Average effort og sånt.

I: Så du syns at selv om det var veldig mye statistikk, så ble det presentert på en forståelig måte?

P: Ja, det syns jeg.

I: Så bra. Tror du et dashboard som dette ville gjort det mer motiverende for deg som studass å skulle se hvordan studentene gjør det?

P: Mmm, ja det vil jeg tro. Nå har jeg ikke erfaring som studass i slike type fag, med denne typen oppgaver. Men jeg vil tro at det er nyttig. Som i det ene faget så sliter jeg med å vite hva studentene forstår og ikke forstår, for det kan være vanskelig å se basert på en enkelt øving. Så jeg tror noe som det her kunne vært fint, for å se hvordan det står til akkurat på et enkelt område.

I: Så bra. Var det noen elementer du syns var vanskelig å forstå?

P: Det var nok litt med å se hvilket alternativ som markert riktig, i spørsmålet. Det strevde jeg litt med. Og så tenker jeg at siden mye annet er fargekodet, med rødt og grønt, at noe lignende kanskje hadde gitt mening å ha på spørsmålene også.

I: Ja skjønner, bra. Du har kanskje nevnt det litt før, men følte du at du fikk den informasjonen du trengte i all den statistikken?

P: Ja, jeg følte det var ganske dekkende. Det gjorde jeg. Jeg synes det var bra. Nå har ikke jeg erfaring med å se på sånn her type statistikk, så det er vanskelig å si hva man bør vektlegge. Men jeg følte at det man fikk var veldig nyttig og bra.

I: Så bra. Du fikk jo litt forslag fra systemet. Hva synes du om det?

P: Kan du gjenta?

I: Du fikk noen forslag hvor systemet sier sånn...

P: åja de ja!

I: Har du noen tanker rundt det?

P: Nei jeg synes det virka bra å ha, for det kan være vanskelig å oppdage sånne ting selv, så da kan det være fint at systemet hjelper og gjør det litt for deg. Og oppklarer det. Så det likte jeg, og så likte jeg at man kan velge å gjøre noe med det, eller bare la det være.

I: Nei men bra. Har du noe mer du vil legge til, eller du bet deg merke i?

P: Ehm... nei jeg synes det var bra jeg. Jeg føler jeg har fått sagt det jeg ville si.

I: Så bra! Takk for hjelpa!

P: Bare hyggelig!

### A.6.9 Interview 9

## Interview 9

I: Okei, nå har du fått prøvd dashbordet litt. Syns du det virket som et interessant tillegg til SmartU?

P: Ja, jeg syns dette komplementerer SmartU systemet, slik jeg har forstått det veldig bra.

I: Ja, kan du fortelle litt om hvorfor?

P: Ja jeg tenker siden det fokuserer på hvordan studentene gjør det i forhold til quizene, og utviklingen da. Og det er jo det det handlet om, med tanke på self-assessment for students.

I: Ja

P: så du kan i alle fall få mønstrene fra den statistikken her. Det er i alle fall mulig å få det, for å se om det er fremgang eller ikke. Og se hvilke studenter som gjør det bedre og ikke. Det er veldig bra visualisert med tanke på hvilke type spørsmål som gjør det bra og ikke og man kan få ut et mønster fra det meste av dette.

I: Ja, så du syns det var enkelt å se hvilke spørsmål som gjorde det bra og ikke?

P: Ja, det vil jeg si

I: Så bra. Var det noen av elementene du likte spesielt godt, eller du følte ga deg litt mer?

P: Jeg syns det var veldig interessant med det lange søylediagrammet. Det likte jeg veldig godt. Når man så alle 119 spørsmålene i faget. Da kunne man se snittet på de forskjellige variablene man kunne filtrere på. Så kunne man se et mønster veldig lett. Så kunne man se om det var en periode hvor folk dabbet av, eller om noen type spørsmål som tydelig ikke er gjort det bra på. Og så ser man tydelig hvilke spørsmål folk brukte veldig liten tid på. Da er det interessant å se hvorfor

I: Ja, det ga en god oversikt?

P: Ja. Og de man bruker lang tid på kan være fordi de er vanskelig formulert, og sånne ting.

I: Ja ikke sant.

P: Så jeg likte den veldig godt. Men jeg likte også veldig godt det faktum at man redigere spørsmålene i ettertid, hvis man oppdager at ting er feil. Det er veldig nice. For lærere, å kunne fikse det.

I: Ja, sant. Synd du det var en enkel prosess, å skulle endre spørsmål?

P: Ja, det var veldig enkelt. Jeg har jo ingen forkunnskaper, men jeg syns det var veldig selvforklarende.

I: Ja. Er det noen andre elementer du vil trekke frem.

P: Ja... ikke negativt, men jeg var litt usikker på det der med high/low/medium, og mye farger. Men når jeg først leste om det så skjønte jeg det. Samme med effort. Sånn som den average time to answer, hvor gul er total. Det assosierte ikke jeg med total, men medium. Da trodde jeg det var hard/medium/easy. Så akkurat den statistikken var forvirrende, men det var også fordi jeg trodde det var hard/medium/easy, så det var derfor jeg ble forvirra. Men det står jo at gul er total, så hvis jeg hadde brukt mer tid hadde jeg skjønt det.

I: Ja, så du tror at hvis du hadde brukt systemet over tid så hadde du kommet greit inn i det?

P: Ja, definitivt. Det var bare akkurat den statistikken som var litt forvirrende første gang. Og så liker jeg også suggestions, det er ikke feil det at den kommer frem. Hva er det som velger hva som skal komme som suggestion?

I: Tanken er at den skal se på lav performance.

P: Ja sant, da kommer det som et forslag der?

I: Da blir det foreslått.

P: Ja så da klarer systemet å plukke ut ting som har feil svar, som den 77% greia (se brukerhistorie)?

I: Ja, det er liksom tanken. At forslagsgreiene skal kunne gjøre det litt lettere for lærere

P: Ja, for da vil jeg trekke frem den også som en veldig bra feature.

I: Ja, så bra. Du tror det ville hjulpet deg som lærer?

P: Ja definitivt. For det kjedeligste som finnes er jo spørsmål som har tekniske feil som går utover kunnskapen til elevene. Hvor man lærer bort feil. Så hvis man har et system som kan plukke opp det, så er det veldig bra.

I: Så bra. Følte du at du fikk den informasjonen du trengte av all statistikken? Var det noe du savnet?

P: Tvert imot. Jeg syns jeg fikk mer enn jeg så for meg. Av det jeg selv hadde klart å tenke meg frem til. Så jeg ble egentlig positivt overrasket over hvilke typer statistikk det var mulig å få inn. Så jeg syns denne viste mer enn nok, hvert fall. Sånn som tid brukt på hvert spørsmål, og alle de snitt prosentene på hvordan folk har gjort det, og også det med vanskelighetsgrad, hvor lang tid man har brukt på de ulike vanskelighetsgradene. Er sånne ting jeg ikke hadde kommet på selv.

I: så bra. Tror du at er dashboard som dette ville gjort det mer motiverende for deg å skulle følge opp hvordan studentene gjør det?

P: Ja. Det vil jeg si. Ikke sånn sinnsykt mye, men definitivt hadde det motivert meg. Det er vanskelig å si når vi har hatt så kort tid med systemet. Jeg er veldig glad i å se på statistikk. Og det er veldig fint, og en veldig oversiktlig måte å se hvordan ting går. Så det er det beste som finnes, å slippe å telle ting selv og... Så når sånn ting går automatisk er det veldig nice.

I: Ja, så bra.

P: Så ja, jeg hadde blitt motivert.

I: Det er bra. Ehm.... Eh.. Er det noe annet du har lyst til å trekke frem? Noe du synes var vanskelig eller tråkleite å forstå?

P: Neeei... Egentlig ikke. Det eneste jeg tenkte på var "More detailed overview" knappen, i forhold til "detailed overview" (vet ikke om den finnes). Jeg synes det var litt likt forklart.

I: Okei ja

P: Så når jeg gikk inn på "More detailed overview", så skjønnte jeg det var statistikk over alle spørsmålene, siden det var på hovedsiden, men det var ikke noe mer enn det. Og siden det var en ganske lik knapp oppe i høyre hjørne føltes det veldig likt. Men jeg skjønnte det var forskjellig. Og siden det var over grupperingene, som en slags header, så lurte jeg på om det var samme knapp (på alle sidene). Men jeg husker det var forskjellige navn, for det står jo her... Det står jo "Edit question", ja. Det gjør det jo. Men jeg mente det sto detailed... Kanskje det var lenger nede, faktisk.

I: Ja det kan hende. Men ja, du synes det var litt lite forklarende?

P: Ja eller ja. Jeg bare tenkte på det. Jeg ble bare et millisekund usikker på om jeg ville komme til alle de spørsmålene, eller om jeg ville komme til noe annet. Men det var ikke noe big deal.

I: Nei men det er bra. Nei men bra. Da har jeg fått svar på det jeg lurte på

P: Oki, nice

I: Takk for hjelpa!

P: Null stress.

### A.6.10 Interview 10

## Interview 10

I: Ja, da har du jo fått prøvd dette dashbordet litt.

P: Stemmer det

I: Syns du det virket som et interessant tillegg til SmartU?

P: Når du sier SmartU, hva mener du da?

I: Det systemet hvor studentene kunne ta quizeene.

P: Ja, absolutt! Jeg syns det var veldig interessant. Og det at du kunne se statistikk på hvordan de gjorde det og sånt. Det var spennende. Og se hvor god performance de har og sånne ting. Som gjør at du kan få litt bedre oversikt enn hva du ellers ville fått. Det blir en tydelig tilbakemelding på "dette får de til, dette får de ikke til"

I: Mhm.. ja. Var det noen elementer du likte ekstra godt?

P: Jeg syns performance og average effect (effort) var kul. I tillegg til at du bruker grønn, rød og gul på alt. Det gjør det veldig lett å skjønne. Det blir veldig tydelig hva som er rett og hva som er galt.

I: Ja skjønner. Så det å se på performance ga deg mye?

P: Mhm

I: så bra. Var det noen andre ting?

P: Det er litt interessant å se på average time da. Å se hvor lang tid folk bruker på å svare. Hvor nyttig det er vet jeg ikke. Men da vet man kanskje om folk er usikre eller ikke. Eller tenker at "okei, er det sammenheng mellom om de har svart riktig og brukt lang tid?", fordi kanskje da har de tenkt skikkelig nøye, eller kanskje har de bare stått mellom to og gjetta riktig.

I: Mmm ja det er interessant det med tid. Har du noen formening om du likte best å få en overordnet oversikt over faget, eller om du likte å kunne gå inn på hvert spørsmål.

P: I utgangspunktet syns jeg det var kult å se oversikt over faget i starten. Ble fort veldig mye og detaljert når jeg gikk inn på spørsmål i starten. Så det var fint at du fikk den overordnede først, og så hvis du vil kan du dykke ned i det. For først tenkte jeg "oj her var det mye å sette seg inn i" når det er sykt mye som dukker opp samtidig. Men jeg syns egentlig det ble gjort på en ryddig måte. Jeg syns det var bra.

I: Så bra. Du fikk jo litt forslag der. Hva syns du om det?

P: Det er et kult konsept som jeg ikke kan huske å ha sett noe annet sted. Men det gir jo kanskje en pekepinn på at "okei mange har svart feil på et spørsmål, kanskje det er formuleringen som er dårlig. Og ikke spørsmålet som er feil" Så det er veldig kult. Og det gjør det lettere for lærere å kunne se enkelt hvilke spørsmål som gjør det dårlig. Og så kanskje det er formulert riktig men at de kan ikke dette godt nok. Så det er en kombinasjon av disse to tingene. Enten dårlig formulering eller dårlig kunnskap. Men det var absolutt kult å få det highlighta.

I: Følte du at du fikk infoen du trengte?

P: Ja, det syns jeg absolutt. Nå har jeg ikke vært så mye borti et sånt system før. Men til å kunne sette seg i den rollen hvor man skal evaluere studenter, så syns jeg at jeg fikk det jeg trengte.

I: Så bra. Var det noen deler du syns var litt vanskelig å forstå?

P: Eneste var det at jeg gjerne skulle hatt en tilbakeknapp. Man bruker jo den mer enn man tenker. Men så var det jo da, den siste her. Som går på hvert spørsmål. Det var ikke vanskelig å forstå, men det var litt mye å sette seg inn i. Den overordnede oversikt her, med alle de blå strekene. Her er det mye data som må prosesseres på en gang. Mens de andre sidene hadde mye tydeligere skiller med grå linjer mellom hver rute, som gjør at du enkelt kan se at dette er én ting og dette er en annen, mens her er det én svær rute med masse data. Og jeg vet ikke om det er noen bedre måte å gjøre det på, og det er ikke så tungvint. Men det er bare at når det er mye der, så må du sette deg inn i det. Men sånn kommer det til å være uansett.

I: Ja det er ikke så lett å presentere så mye data oversiktlig. Men det presentert på to litt ulike måter på det screenet. Øverst har du søylediagram, mens under har du tekstlig presentasjon. Har du noen formening om hva du likte best?

P: I utgangspunktet så er jeg fan av visuell representasjon. Men jeg følte at når du har en tabell som er under der og er såpass ryddig, så følte jeg ikke at jeg trengte statistikken. Men i utgangspunktet er jeg mer glad i det visuelle. Jeg liker grafer bedre. Det gjør jeg.

I: Så bra. Var det noen andre ting du syns var vanskelig?

P: Nei, ellers syns jeg det var bra. Jeg likte også spørsmålstegnene, som forklarte fra effort betydde og sånt. Så får du en forklaring på det. Det er hjelpsomt.

I: Ja sant. Nei men så bra. Følte du at all denne statistikken hjalp deg å forstå hvordan studentene gjorde det?

P: Jeg vil si ja, men det er et par ting som jeg ikke så veldig mye på. Blant annet den "Score per attempt" husker jeg ikke så mye av, og ikke "Time spent per question". Mer da det stolpediagrammene som viste tid mer spesifikt. Ellers så var det det at jeg følte ikke det store søylediagrammet ga meg så mye, men det var mest fordi tabellen under var ryddig.

I: Ja, interessant. Nei men bra, noe mer du vil tilføye?

P: Mmmm... Jeg synes det var interessant på hovedsiden at du hadde både notifications og si noe om systemet osv. Og så var det veldig fin sånn boks "dette er quizene", og så er det tydelig du kan få flere quizer på siden, som går til andre fag. Så det ville vært én ting kanskje, er dashbordet for et fag, eller er det for alle fag. Kan du ha flere quizer på samme fag, eller hvordan fungerer det?

I: Nei det kan du ikke. Systemet lager quizer i faget. Så det tar av de 119 spørsmålene og lager quizer. Så da ville du hatt andre fag eventuelt.

P: Ja skjønner, det er jo veldig kult. Nei ellers synes jeg det er veldig bra

I: Så bra. Supert, takk for hjelpa!

P: Bare hyggelig!



## **Appendix B**

# **Design, Implementation & Product**

## **B.1 User tests**

The following section presents the user test guide used for this study, and the observational results from the user tests. The observations were written in Norwegian, as it is the native language of the researcher, and the notes had to be written rapidly as the user test was conducted.

### B.1.1 User test guide

## User test

1. You are a teacher in the Web development course, and you are now in the middle of the semester. You are using SmartU to monitor the adaptive tests you have created for your students. With SmartU, you can manage the quiz content and monitor the usage of questions with graphs and other visualizations, and you can find problematic questions and improve your quizzes. You log in to the system to assess your content. First, you are curious as to whether there are any announcements for you.
2. You want to find out how many students are taking quizzes, how many they have taken, and how long time they spend.
3. You want to see how the quizzes in the course TDT4110 ITGK are doing.
  - What do you see? Please explain the features.
4. You would like to see how many questions there are in the course, and how they are distributed based on difficulty.
5. You are unsure as to what Effort really means. How would you proceed to gain more information about this?
6. What can you tell about the students' performance in the course?
7. You notice the suggestions. The first suggestion you decide to not do anything with, and just let the question be as it is.
8. The second suggestion you decide to check out.
  - What can you read from the page?
9. Which alternative do you believe is set to be the correct one? What gave you this indication?
10. You notice that the answer is wrong, it should be 16. You decide you want to correct this error. You realize that the statistics of this question may be faulty due to the error in the question, you want to reset the statistics about the question. How would you do that?
11. You feel finished with fixing the question and decide to go back.
12. You would like to get more information about all the questions in the course.
  - What do you see?
13. You would like to see some other metrics displayed on the bar chart.
14. You notice there is a typing error at Question 2, and would like to look further into it.
15. You would like to correct the error. It does not seem to have affected the statistics, so you do not want to reset them for this question.
16. You feel finished assessing the questions in the course, and would like to go back to the homescreen

B.1.2 Observations

Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7	Subject 8	Subject 9	Subject 10	Subject 11	Subject 12	Subject 13	Subject 14	Subject 15	Subject 16	Subject 17	Subject 18	Subject 19	Subject 20	
1	Uklart hvor mange som får quizzen. Generelt uklart. Finner det til slutt. Finner fort frem. Misforstår vanskelighetsgrad/ mastery level. Effort uklart. Vanskelig å si om prosent eller antall.	Finner ikke helt informasjonen. Bruker tid, men finner det til slutt. Finner med en gang Ser på ? intuitiv. Drøsker hover effect	Finner det. Ville hatt bjelle opppe til høyre. Liker å se dom i dom	Usikker på hvem avsender er	Bruker litt tid. Får ikke så mye ut av overview. Burde stå avg							Its forvint over announcements/notifications	Enkelt, ser også opppe til høyre.	Enkelt, ser etter mer announcements inne på emnet.		Enkelt, tenker bjelle opppe til høyre. Skjønner ikke at overview er det han trenger. Gir sagt ut av at han ikke får noen total		Ser med en gang		
2								Bruker litt tid, men skjønner fort		ser ikke ?, som forklarer	Enkelt, burde forklart y aksen. Aksene smelter litt sammen.	Er inne på ITGK. Finner det til slutt.	Ser det.	Syns ikke det er så intuitivt. Mer dårlig formulert spørsmål	Ser det fort, blir forvint av X-aksen til ratoen.				Ser med en gang. Syns det er tydelig	Knoter litt. Blander aksar
3					Finner forskanngene. Misforstår farger.					Skjønner det mest	Starter på summary. Uklart om times used. Uklart effort. Fargekoden.		Får god oversikt, skjønner hva ting representere.	Får god oversikt Misforstår High/hard. Liker suggestions	Får god oversikt. Forvint over times used.			Stusser litt på avgTTA		
4																				
5														Ser for. Usikker på definisjonen.	Ser for. Skjønner ikke helt effort					
6														Ser for. Samvrer å se om normalforvint.						
7	Dårlig formulert task. Finner fort frem																			
8	Skjønner fortsatt ikke effort.	Usikker på hva times used er. Får god oversikt, skjønner det.	Misforstår times used	Usikker på times used				Misforstår times used.												
9	Ser det, basert på performance, og så bold	Ser på bold. Prosent burde ikke vært bold.	Ser på bold	Ser på bold	Ser bold, men ser det sent	Ser ikke.	Usikker. Resonnerer ut fra performance.	Ser uthvett med en gang. Og koder performance	Ser Bold med en gang.	Usikker på spm. Ser på perf.	Ser på performance.	Performance	Ser på Bold	Ser på performance	Ser på performance. Ser bold til slutt, vil ha farge	Ser bold, såvidt	Ser på bold, burde vært mer tydelig. Forstår ikon	Ser på bold, chance of success	Ser på bold, såvidt, og performance etter hvert	Ser ikke
10	Finner fort, fikser feilen											Enkelt. Stusser over at ting blir satt til 50%		Syns reset statistics ble litt skjøvt						Ser på bold, burde vært mer tydelig
11	Går helt hjem	Går helt hjem	Går helt hjem	Enkelt	Går helt hjem	Går helt hjem	Går helt hjem	Går helt hjem	Går helt hjem.	Går tilbake	Går helt hjem	Går tilbake	Går helt hjem	Går tilbake	Går helt hjem	Går tilbake	Går helt hjem	Går helt hjem	Går helt hjem	Går helt hjem
12																				
13																				
14																				
15																				
16										Ser ikke at loggen er hjem knapp										
Other comments	Skjønner ikke effort.	Skjønner ikke times used og farger	Skjønner ikke effort og times used	Usikker på times used og farger	Misforstår farger			Misforstår times used og farger	Misforstår times used	Liker den store oversikten.	Forvint av fargebruk, times used og effort			Liker oversikten i laget. Får med header i scroll. Syns det er mye data. Usikker på effort	skjønner ikke times used og effort, misforstår farger		Savner navn på aksar. Skjønner ikke times used	Skjønner ikke at man er delt inn i grupper		

## B.2 Final User Interface

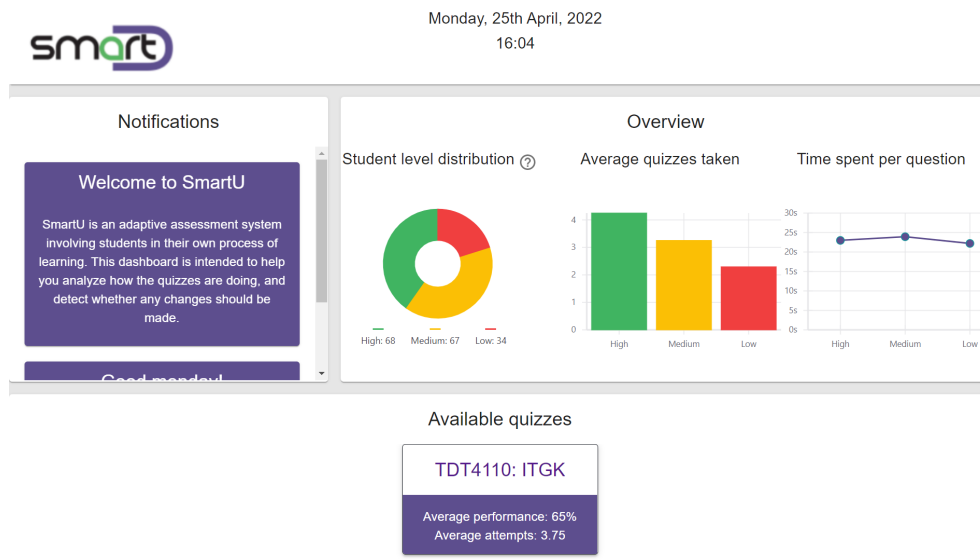


Figure B.1: Home screen

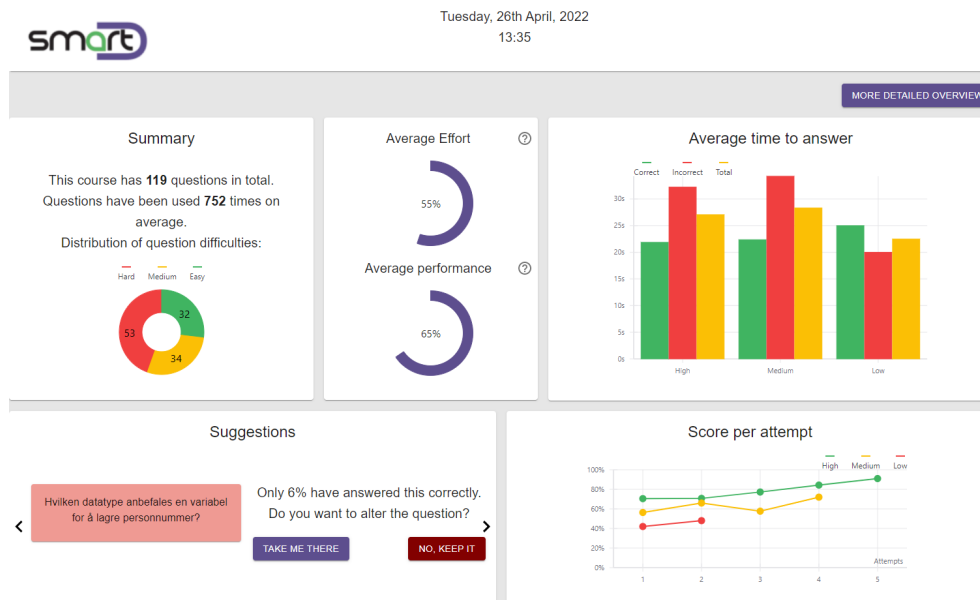


Figure B.2: Course screen

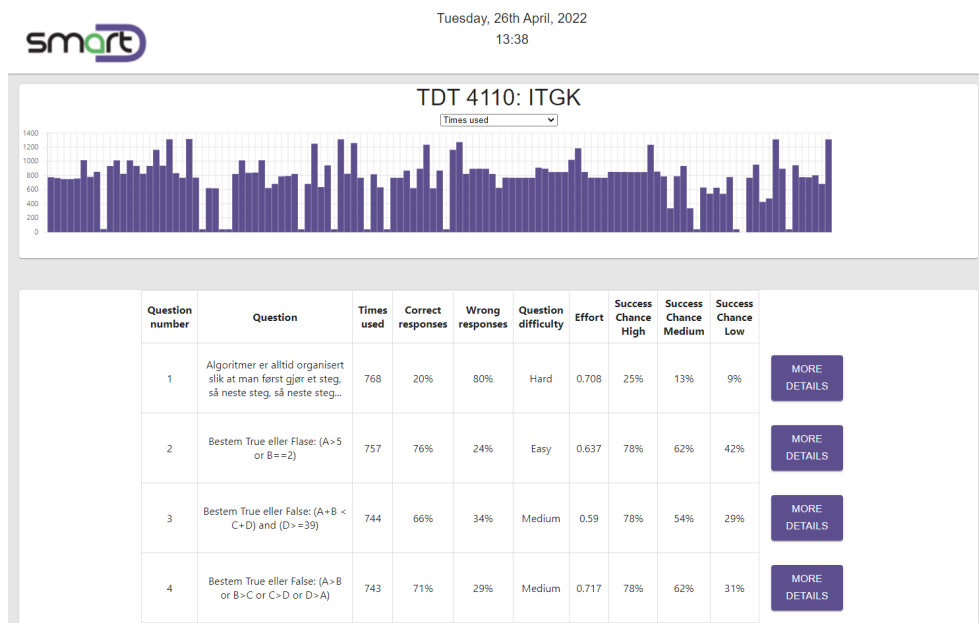


Figure B.3: Detailed overview screen



Figure B.4: Question screen

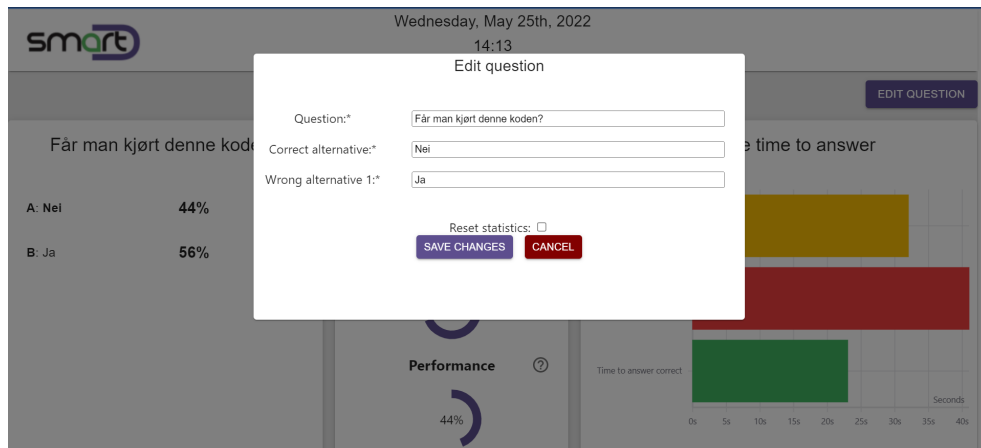


Figure B.5: Edit question screen





## Appendix C

# Results

### C.1 Questionnaire & Descriptives

**Table C.1:** Overall evaluation of the usability of the Dashboard (OEUD)

*Mean is based on Likert Scale from (Strongly Disagree) 1 to 5 (Strongly Agree)*

<b>Questionnaire question</b>	<b>N</b>	<b>Mean</b>	<b>Std Dev</b>
I think that the navigation in the dashboard was nearly effortless.	20	4.20	.696
I think that when I needed help to learn how to use the dashboard, the system provided me with sufficient information.	20	3.85	.933
I think that a user who has never seen the dashboard before can learn how to accomplish basic tasks fast.	20	4.65	.489
I think that experienced users should be able to complete tasks using the dashboard quickly.	20	4.85	.366
I think that it is not frequent that users make errors while using the dashboard.	20	3.90	.553
I think that the interaction with the dashboard is clear and understandable.	20	4.45	.510
Using the dashbaord makes me happy to accomplish my assessment tasks.	20	3.95	.686
Using the dashbaord gives me enjoyment for my continuous teaching.	20	4.00	.794
Using the dashbaord leads me to exploration in my continuous assessment.	20	4.40	.681

**Table C.3:** Usefulness of the Dashboard (UOD)

*Mean is based on Likert Scale from (Strongly Disagree) 1 to 5 (Strongly Agree)*

<b>Questionnaire question</b>	<b>N</b>	<b>Mean</b>	<b>Std Dev</b>
It was easy to understand what the graphs reflected.	20	4.05	.605
I was able to make sense of the graphs and extract information.	20	4.32	.749
I think the dashboard helped my understanding of how the students were doing.	20	4.60	.503
I believe that a graphical representation is more fitting than a textual representation of the data displayed in the dashboard.	20	4.55	.686
I found it useful to view stats of the students' performance over time	20	4.50	.513
I understood that the mastery-level reflected level of knowledge as a result of previous performances	20	3.95	.911
I found it easy to understand that the color green was used to represent a positive association.	20	4.74	.733
I found it easy to understand that the color yellow was used to represent a mediocre/neutral association.	20	4.58	.838
I found it easy to understand that the color red was used to represent a negative association.	20	4.84	.688
I think that the graphical representations in the dashboard provided all the information I needed.	20	4.25	.550

## C.2 SUS-Score

**Table C.5:** Attitude towards the Dashboard (ATD)

*Mean is based on Likert Scale from (Strongly Disagree) 1 to 5 (Strongly Agree)*

<b>Questionnaire question</b>	<b>N</b>	<b>Mean</b>	<b>Std Dev</b>
I believe that the provided visualizations in the dashboard made information easy to extract.	20	4.05	.686
I was motivated to further use the dashboard to see progression unfold in the visualizations.	20	4.20	.696
I felt more informed and confident conducting an assessment when utilizing the dashboard.	20	4.30	.571
The design and colors in the dashboard made me feel confident and comfortable while using it.	20	4.15	.813
The design and colors in the dashboard made me feel happy.	20	4.00	.882
I would like to continue to use this dashboard to follow up on students' progress.	20	4.45	.510
All the available statistics and suggestions made me more motivated to monitor the course questions.	20	4.55	.510

Table C.7: System Usability Score (SUS)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Total Score
P1	4	3	4	2	4	3	5	2	3	2	70
P2	5	2	5	1	5	1	4	1	4	2	90
P3	5	2	4	2	3	2	2	1	4	2	72.5
P4	5	2	5	2	4	1	5	1	2	3	80
P5	4	3	4	3	4	1	4	1	4	2	75
P6	3	1	5	1	4	2	5	2	5	1	87.5
P7	4	1	4	1	4	2	5	1	5	1	90
P8	5	2	5	2	4	1	4	1	5	2	87.5
P9	4	2	4	1	4	1	5	2	4	1	85
P10	5	1	5	2	5	1	5	1	4	2	92.5
P11	4	3	4	2	4	2	4	1	5	1	80
P12	4	2	5	2	4	3	5	3	2	3	67.5
P13	4	2	4	1	5	1	4	1	3	1	85
P14	4	1	5	1	4	1	5	1	4	2	90
P15	4	2	5	1	5	1	5	2	4	2	87.5
P16	4	2	4	1	4	1	5	1	4	2	85
P17	4	2	4	1	4	1	5	2	3	2	80
P18	4	3	4	1	4	2	4	1	4	1	80
P19	4	1	5	1	5	2	5	1	4	1	92.5
P20	5	2	4	2	5	2	5	2	4	3	80
Average:											82.875
Positive result			Neutral result			Negative result					

