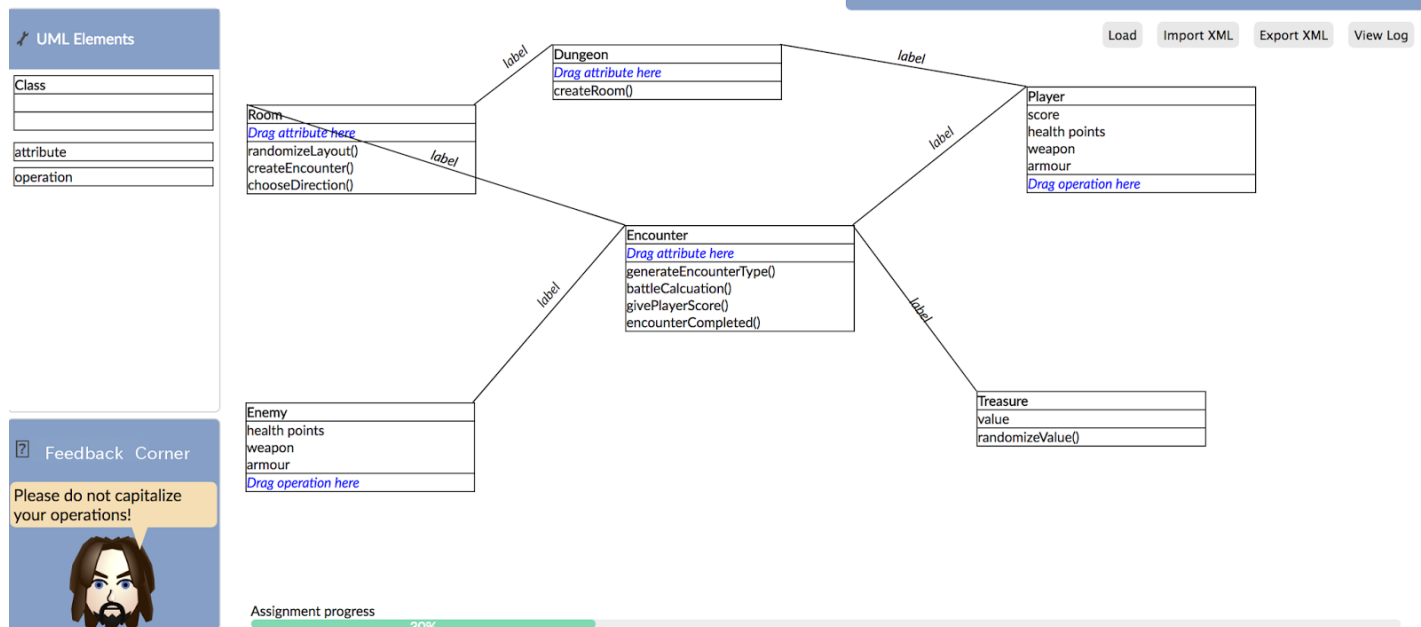




WebUML - an Online UML Editor



Improving the Learning Process of Students with the Use of Automated Feedback for Software Design in E-learning Environments

Bachelor of Science Thesis in Software Engineering and Management

AMAN GHEZAI
ROB BRUINSMA



CHALMERS
UNIVERSITY OF TECHNOLOGY



UNIVERSITY OF GOTHENBURG

The Author grants to University of Gothenburg and Chalmers University of Technology the non-exclusive right to publish the Work electronically and in a non-commercial purpose make it accessible on the Internet.

The Author warrants that he/she is the author to the Work, and warrants that the Work does not contain text, pictures or other material that violates copyright law.

The Author shall, when transferring the rights of the Work to a third party (for example a publisher or a company), acknowledge the third party about this agreement. If the Author has signed a copyright agreement with a third party regarding the Work, the Author warrants hereby that he/she has obtained any necessary permission from this third party to let University of Gothenburg and Chalmers University of Technology store the Work electronically and make it accessible on the Internet.

Improving the Learning Process of Students with the Use of Automated Feedback for Software Design in E-learning Environments

Aman Ghezai
Rob Bruinsma

© Aman Ghezai, June 2017.

© Rob Bruinsma, June 2017.

Supervisor: Dave Stikkolorum

Examiner: Abdullah Mamun

University of Gothenburg
Chalmers University of Technology
Department of Computer Science and Engineering
SE-412 96 Göteborg
Sweden
Telephone + 46 (0)31-772 1000

Improving the Learning Process of Students with the Use of Automated Feedback for Software Design in E-learning Environments

Aman Ghezai

Department of Computer Science and Engineering
University of Gothenburg
Gothenburg, Sweden
gusghezam@student.gu.se

Rob Bruinsma

Department of Computer Science and Engineering
University of Gothenburg
Gothenburg, Sweden
gusbruiro@student.gu.se

Abstract—In education feedback is generally regarded as crucial for improving knowledge and is a significant factor in motivating learning, but the process of providing timely and relevant feedback in software design studies can be challenging. In this paper, we aimed at implementing an automated feedback agent in a web-based UML class diagram editor for novice designers. In order to collect requirements and to provide a relevant feedback agent, students were interviewed. To gain further insight in the automated feedback agent we conducted an experiment and compared students that used the feedback agent with those that did not. Based on statistical analysis and a questionnaire we learned that i) students had experienced improvement in their learning process. ii) automated feedback has no statistically significant effect on students' learning outcome. iii) automated feedback has a statistically significant effect on students task performance. The statistical result indicates that having a feedback agent does not necessarily improve students grades. On the other hand, our results from the interviews and questionnaire show that a feedback agent can play a significant role in improving the learning process and performance of students in software design. We suggest follow-up studies to investigate the results in larger educational contexts.

Keywords-class diagram; UML; WebUML2; Feedback agent; E-learning;

I. INTRODUCTION

Today's e-learning models of higher education are based on conventional distance education, initially intended for individuals to gain access to higher education [1]. In such systems, feedback is often given after handing in an assignment, hence the process of learning is based on delayed feedback. In an educational context, feedback is generally regarded as crucial to improving knowledge and a significant factor in motivating learning [2]. Instant feedback therefore becomes an important aspect in the learning of software modelling and design. The study of software design and modelling with the aid of e-learning platforms has been evolving with the advancement of technology. These developments have improved the accessibility and reach of software design studies while at the same time creating challenges in providing a suitable learning platform with an instant feedback mechanism to support learning.

The Unified Modeling Language (UML) is the standard modelling language in the field of software engineering, and is intended to provide a standard way to visualize the design of a system [3]. UML has been used as a modelling language to teach students software design with the aid of e-learning systems. As the result of advancement in technology there has been a variety of UML modelling tools and systems that have been developed to enable the learning of software design. While these tools and systems provide a vast variety of support, the learning curve of designing advanced UML diagrams such as class diagrams and sequence diagrams can be a challenging process for students [4]. There have been efforts in assisting students in learning UML design through a variety of computer supported collaborative online learning systems. MOOC (massive open online courses) [5] is such an example, where the most widely used approach in providing feedback is to use peer assessment. However, because the assessment is unmoderated, it lacks credibility [6].

Furthermore, most of these systems are not dedicated to providing automated feedback, as they rather focus on collaborative work. Hence, in this research we aim to integrate an automated feedback mechanism (hereafter referred to as "feedback agent") into the existing UML editing tool WebUML2 [7]. This thesis project is part of a current research by Dave R. Stikkolorum and his colleagues in the area of software design learning processes. Stikkolorum's previous works in identifying students' common difficulties and strategies during the design of class diagrams [7], as well as his work in revealing students' UML class diagram modelling strategies with WebUML2 and LogViz [8], inspire us to perform this study.

A. Purpose of the Study

In this research we will be integrating a feedback agent into the existing WebUML2 tool in order to provide formative feedback on software design to novice designers and students. Our research strategy will be of a design science methodology [9], and includes a qualitative and quantitative data collection.

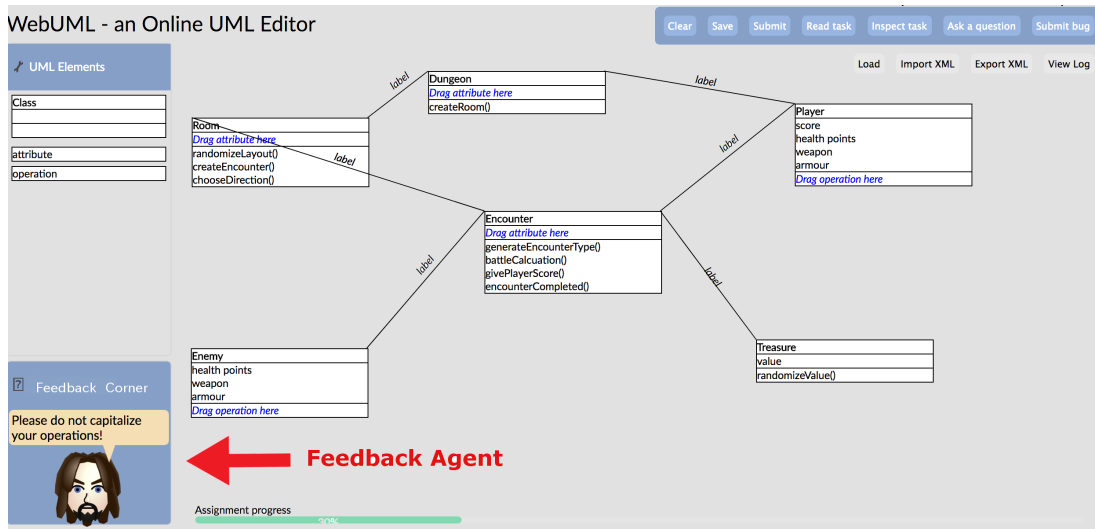


Figure 1: WebUML2 user interface with feedback agent

B. Research Questions

In this section we present our research questions. RQ1 is our main research question, followed by our sub-questions RQ2 and RQ3.

RQ1. Does the use of automated feedback mechanisms in e-learning systems improve the learning process of students of UML class diagram design?

RQ2. Does such a mechanism improve the learning outcome of the students? (quality of class diagrams created.)

RQ3. Does the students' performance (measured in time) in solving the task improve by the automated feedback?

Past research has shown that timely feedback is crucial in improving knowledge and in motivating students to learn [10], [11], [12]. However, it is less clear how automated feedback in software modelling and design affect students' learning process of class diagram design and their performance in solving design tasks.

In this paper, we want to determine whether automated feedback can improve students' learning process, learning outcome and performance. This research proposes that having a feedback agent can improve students' learning process and performance. This research is therefore intended to make contributions to the literature on automated feedback and the learning process of class diagram design.

In section II, we explore related work and in section III we explain the research method used. The results are presented in section IV and discussed in section V. In section VI the validity threats are discussed. Finally we conclude and identify future work in section VII.

II. BACKGROUND

A. WebUML2

WebUML2 is a web based UML editor and research tool created by Dave R. Stikkolorum and his associates. WebUML2 is used to design UML diagrams, specifically class diagrams. The tool has been used in researches related to the study of software design. Currently the tool allows students to design class diagrams and it is still evolving for further studies. In order to perform our study we have integrated a feedback agent into WebUML2.

B. Our Version of WebUML2

The current version of WebUML2 is mainly an editor and hence does not provide feedback to students. Students can create and edit class diagrams which they can export for later use. In order to implement the feedback agent that would assist students in their class diagram design, we constructed a list of requirements which were discussed with an expert. These requirements were later iterated on through two rounds of interviews. The following requirements were developed to support students in the learning process of the basic conventions of UML class diagram design and by considering the students requirements.

- 1) Feedback on class diagram naming conventions.
 - All classes created should have a name.
 - Class names should start with a capital letter.
- 2) Feedback on missing attributes in class diagrams.
 - All the attributes created should have a name.
 - Attribute names should start with a lower case.
- 3) Feedback on missing operations.
 - All the operations created should have a name.
 - Operation names should start with a lower case.
- 4) Feedback regarding different type of associations.
 - All associations should have a meaningful label.

The above requirements were chosen because they address the basic and general aspects of class diagram and UML. Hence allowing students at all levels to design and implement the basic concepts of class diagram and UML. Figure 1 shows the new WebUML2 editor with the feedback agent integrated.

C. Feedback

Both observational and experimental research has shown that feedback is one of the most powerful factors that influence learning, for instance Hattie & Gan [11] and Hattie & Timperley [12] have shown that appropriate feedback given in a timely fashion can help improve the learning outcomes of students. Based on a study of over 500 meta-analyses representing over 20 million students [13] they conclude that feedback is one of the most powerful influences on learning. The general concept of how feedback can help a learning process can be related to our study, but the practical implementation of a feedback agent in a learning platform and how that affects a learning process is not covered in this study. Therefore, creating an opportunity to further study the topic of feedback in e-learning platforms.

In educational and instructional contexts formative feedback can be provided through different external agents to the learner (i.e., teachers, parents, peers or computer-based systems). While task processing to enhance understanding are identified as internal sources of information, the other sources of information that can be comprehensible by the learner are external.

Formative feedback is an important factor in improving the learning process and has received much attention in instructional research. Feedback can be provided in several ways and the complexity of the information provided can vary from simple evaluative instructions to more complex elaborated instructions. According to Narciss [14] and Shute [15], any sort of information that is provided to students about the state of their performance or their state of learning in order to guide the learners' thinking in the direction of the learning standards is considered feedback.

Even though there is large amount of research in formative feedback and its implications, research on interactive tutoring systems shows that designing and investigating formative feedback strategies for digital learning environments is challenging (i.e., Arroyo et al., [16]; Goldin, Koedinger, & Alevan, [17]; Mitrovic, Ohlsson & Barrow, [18]). Another implication that is important to consider and understand is that the effect of formative feedback may differ depending on two major factors, namely contextual factors like task complexity and individual factors like motivation. Through integrating both the individual and contextual factors, Narciss (2006, 2008, 2012a, 2012b) [19] has developed the Interactive Tutoring Feedback Model (ITF-Model) to provide theoretical and empirical framework for designing and evaluating feedback strategies.

D. Related Work

Krause et al. studied the influence of feedback and cooperative learning on example-based e-learning. They concluded

that feedback clearly supports learning, by helping students reflect on both their own knowledge and the presented material [20]. This study deals with the topics of e-learning and effects of feedback on the learning process. The concepts of how a feedback intervention in e-learning supports learning are explored, making the study relevant to our study. The difference between our study and [20] is that this study particularly focuses on the influences of feedback on cooperative learning. Another difference is the field of study; while our study is focused in the field of software engineering the study by Krause, Stark and Mandl is focused in the field of statistics.

The role of formative feedback to support students' learning in an interactive learning environment is broadly addressed by Goldin et al. [21]. This study focuses on the role of formative feedback in learning in general, as well as conditions and effects of different kinds of formative feedback strategies. According to this study, a feedback agent needs to be designed by considering the learners characteristics and the level of task complexity. Hence giving insight on how formative feedback is used in different technological platforms and how a formative feedback agent can be implemented effectively. This will help students better understand their learning process, what the goals are and how to reach the goals [21].

Narciss suggests, using the ITF-Model [19], that there are three groups of factors that influence the advantages and limitations of formative feedback in instructional frameworks. The first group of factors accounts for the requirements of learning tasks and the knowledge or skills needed to meet the tasks requirements. This means that the formative feedback provided to the learner needs to be suitable for the level of the learning task requirement, while catering for the level of knowledge or skill of the learner. Hence, feedback intervention should be tailored to the requirements of the tasks and the instructional context, in order to support the learning process more effectively. To do so, an analysis of the competencies needed to meet the task requirements is required in order to deliver the desired standards of formative feedback [22]. In relation to our feedback agent, it is designed only to address the basic concepts of class diagram design. The task we used in our experiment is also designed by considering the level of knowledge or skill of the learner.

In the second group of factors the ITF-Model addresses the individual learner, as well as how the different types of learner strategies and learners' motivation promote or constrain the extent of the students improvement in attaining the desired standards with formative feedback. For instance, in a situation where the learner is not actively attending the provided feedback, even the most suitable and well designed feedback intervention will not help the learner. Hence, identifying learners motivation and the factors that influence the processing of feedback is a critical aspect. Most feedback models encourage the investigation of the learner characteristics at least on motivational, cognitive and meta-cognitive levels [11]. As this study suggests, besides providing the feedback, in our study we have also considered the aspects of learners motivation and the factors that influence the processing of feedback.

Consequently, we have interviewed students regarding what about the feedback motivates them and involved them in the design process of the feedback agent.

The third group of factors concerning formative feedback in the ITF-Model relate to the characteristics of the feedback’s strategy and its message in terms of its communicational and informational value [21]. For feedback to be useful it needs to be a reliable and correct assessment of the learners’ current state of task completion as well as based on the representation of task requirements. The feedback message generated should provide clear information that would help the student to close the gaps between their current and intended state of learning.

Dolonen, Chen and Mørch [23] present a similar approach to our study through integrating a software agent with FLE3 (Future Learning Environment) - a distributed computer supported collaborative learning environment. The software agent system presents feedback to both to students and instructors. The feedback agent to students is generated based on principles of collaboration and knowledge building [23], hence putting an emphasis on collaboration. Furthermore, the software agent computes statistics used to detect possible problems and to present advice to the instructor [23]. Nonetheless the main difference with our study is that Dolonen, Chen and Mørch focused on encouraging collaborative learning and providing feedback regarding to students collaborative work while we focus on giving task related feedback in software modelling and design.

Anckars work on “Providing automated feedback on software design for novice designers” [24] has a similar approach to how we performed the study but the difference is that the feedback agent “was built around a specific UML modelling task” [24] and uses an example solution. Hence, the difference with how our feedback agent works. Our feedback agent is not dependent in a specific task and example solution rather the feedback agent depends on UML class diagram design notations. Thereby the feedback mechanism approach and the study method used closely relate to our problem domain but the solution approach differs.

Vasilyeva, Pechenizkiy, and De Bra discuss the adaptation and personalisation of feedback in e-learning systems [16]. For example, scaling the feedback frequency to the amount of mistakes the student makes. Adaptation and personalisation of feedback in order to provide effective feedback can be identified with the first group of the ITF-Model discussed before, where Narciss [19] advises that feedback needs to be suitable for the level of the learning task requirement.

However, none of the methods and models address the relation of the feedback agent and its implication in the learning process of software design. Nonetheless, in our design of the feedback agent we have considered the ITF-Model by Narciss [19] and the suggestions of Vasilyeva, Pechenizkiy, and De Bra. The students will have an option to generate feedback by pressing a button at any time while designing or at the end of the task. Thus giving students the option to get the current state of their task completion by actively requesting feedback or to get general feedback after performing the task.

III. RESEARCH METHODOLOGY

In this section we introduce the method we used. First we present the overall research strategy and framework. Then we present the design and development of the feedback agent, followed by the data collection.

A. Research Strategy

A design science methodology will be used in the development of the feedback agent for the WebUML2 tool. This methodology is chosen because the study involves the implementation and evaluation of a designed software artefact to address a specific problem domain [17]. The study was conducted iteratively and incrementally through involving students in the design and development stages of our prototype. Moreover, an experiment was conducted with students using our feedback agent in order to collect quantitative data. This was done through comparing students’ grade and performance, where students used WebUML2 both with and without feedback agent.

The grades were given by experts in the field of software design and were given based on a five point scale rubric [25]. The grades reflect the students’ overall understanding of the assignment. The performance is hence measured in terms of how well the students solved the tasks in a given time with the help of the feedback agent.

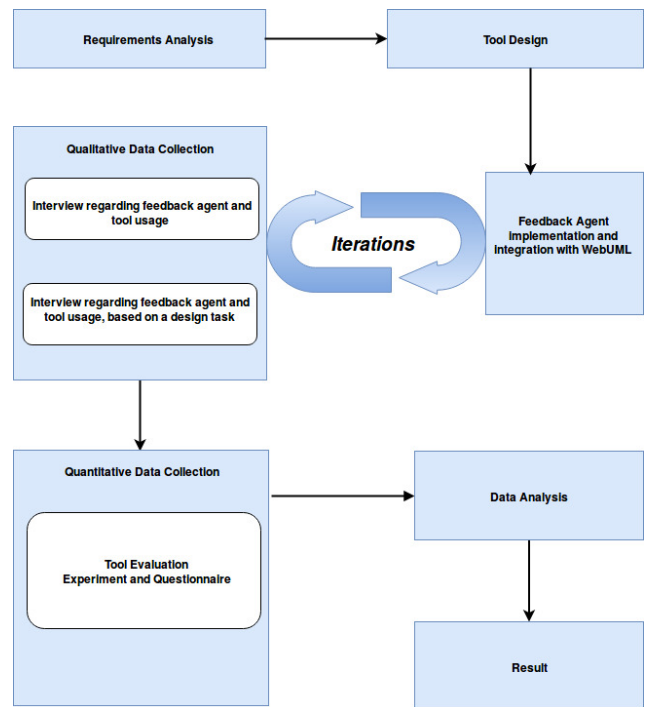


Figure 2: Research Framework

The qualitative data collection was done through interviewing students in regard to the feedback agent and the tool in general. The data collected was then used to analyze and improve the artefact. In order to address the research questions, the research is divided into two stages where the feedback agent is continuously developed and evaluated. The

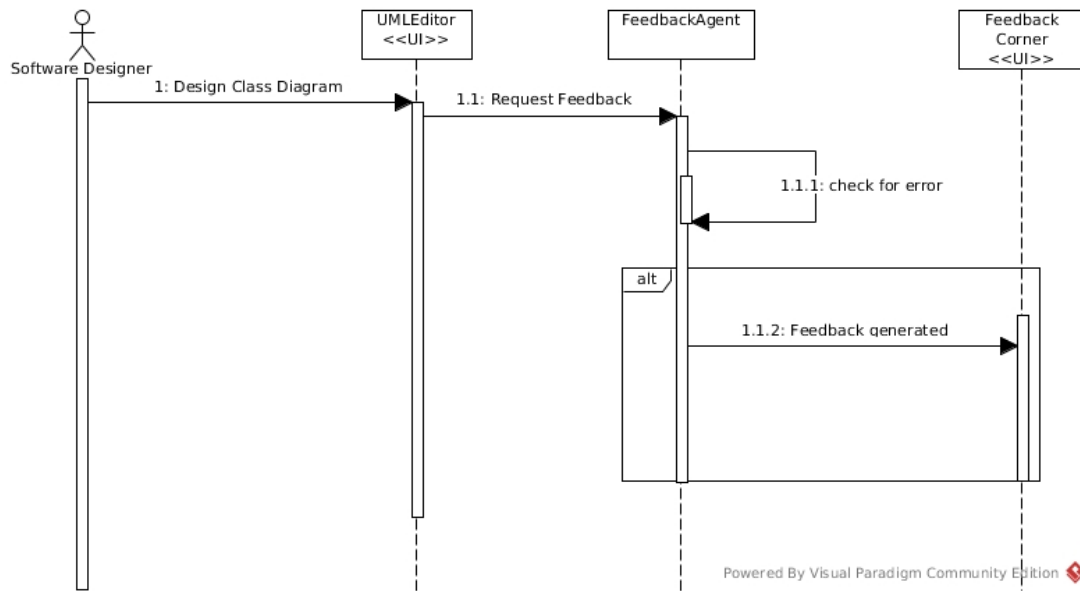


Figure 3: Feedback Mechanism Sequence Diagram

artefact is developed using iterative and incremental design, where the testing and evaluation is formative, part of the development process [18]. The quantitative data that was collected is evaluated to determine how students perform in the tasks through grading their solutions. The results of the experiment are assessed through statistical analysis. Figure 2 visualizes the different stages of our research process.

B. Design and Development

The feedback mechanism developed will be in the form of a feedback agent which provides instruction and feedback to students during a software design session. Figure 3 shows the sequence diagram of the feedback agent. It analyses the students' input and checks the different software design components like the existence, placement and naming conventions of classes, attributes, operations and associations. If the solutions provided are not adhering to the standard UML notations, the student gets automated feedback. For instance, in the case of class names, the first letter in the name always needs to be a capital letter (upper case) while for attributes and operations the first letter in the name always needs to be in small letters (lower case). The feedback agent evaluates the naming conventions of classes, attributes and operations based on the standard UML class diagram notations and provides appropriate feedback.

C. Interviews

During the first stage, two interviews were conducted, focused on getting the participants' opinions on various aspects of the feedback agent. We chose a semi-structured interview format. Semi-structured interviews use open-ended questions with which we elicit relevant information from the participants, while allowing them the freedom to elaborate and discuss their ideas [26]. The results of the qualitative data

collection are used to answer RQ1, with further support of the results from the quantitative data.

In the first round 4 third year BSc. students from the Software Engineering and Management (SEM) program at the University of Gothenburg were interviewed. They were given a short verbal introduction on using the tool before being asked to create a UML class diagram of their choice and to request feedback from the tool while working on the design. They were also asked to experiment by making various errors to further explore the feedback agent. Afterwards the following questions were asked:

- What did you like about the feedback agent?
- What didn't you like about the feedback agent?
- Is there any feedback missing that you would have found helpful?
- What did you think about the location of the feedback?

In the second round, two experienced developers with theoretical knowledge from past studies and three first year BSc. students from the SEM program were interviewed. They were asked to create a UML class diagram based on a task (Appendix 1) and to request feedback from the tool while working on the design. Afterwards the following questions were asked.

- In what way was the feedback you received relevant to what you were doing?
- How did the feedback influence you in finding a correct solution?
- What do you think about the manner in which the feedback was presented? What did / didn't you like?
- Was your motivation affected by the feedback and in what way? What about your performance?
- Is there any feedback you didn't get that might have been helpful?

The analysis of both rounds of interviews was used to itera-

tively improve the feedback agent and to validate its design and implementation.

D. Experiment

In this subsection we describe the steps for our quantitative data collection process, which will be in the form of an experiment, followed by a questionnaire.

1) *Experiment Preparation*: An experiment is conducted with students and after the experiment a questionnaire is used in order to answer our research questions. The experiment is performed with 20 BSc. students from the SEM program at the University of Gothenburg. The students were asked to design a class diagram based on a task (Appendix 1). We evaluate the class diagrams in terms of quality and task completion. This is done through the help of expert grading.

2) *Subjects*: The participants of the experiment are SEM Bachelors students at the University of Gothenburg in Sweden. The students were randomly selected from first year to third year students, in order to broaden the data and reduce bias. The subjects have experience in software design and as part of their bachelor degree the students have obtained practical training in software design principles and UML class diagram design.

The scope of this experiment is specified in terms of the goal definition template presented in [27]. The aim of the goal definition template is to make sure that all the important aspects of our experiment are identified before the planning and execution of the experiment:

*Analyze the use of automated feedback mechanism in e-learning systems
for the purpose of evaluating the effect of automated feedback in software design learning
with respect to the quality of class diagram designed and performance of students
from the point of view of novice software designers/students
in the context of class diagram design in software development.*

3) *Experiment Planning and Operation*: In this experiment three types of variables are defined: independent, controlled and dependent variables [27].

a) *Independent variables*: The independent variables are the variables that we can control and change in the experiment [27], hence the independent variable in this experiment is “feedback agent availability”, measured on a nominal scale with ranges: available and unavailable. The variable has two levels: experimental (available) and control (unavailable).

b) *Controlled variables*: The control variable in this experiment is the information provided in the task, which is reviewed by an expert.

c) *Dependent variables*: The dependent variables measure different aspects of the class diagrams created, such as (i) task grades (the quality of class diagrams created) and (ii) task performance (time taken to perform the task). The task grades follow a range of 1-5 based on expert grading, while the task performance is measured in minutes and seconds. We expect to use (i) and (ii) to answer RQ2 and RQ3.

4) *Task*: The students were asked to create a UML class diagram based on a task they were given. The task was a short description of 173 words written in English (Appendix 1). Afterwards, they were given a short tutorial on the tool. The students were then asked to perform the task. Half of them were asked to use the feedback agent throughout their design process, while the other half did not have the feedback agent available. Their solutions to the task were recorded by taking a screenshot of the final state of their design. Afterwards, the students were asked to fill out a questionnaire (Appendix 2).

5) *Instruments*: The instruments that we used in this experiment consist of the following:

- The WebUML2 tool.
- Introduction to WebUML2 and the feedback agent.
- A questionnaire to be filled after the experiment is done (Appendix 2).

6) *Hypotheses*: The main hypothesis of the experiment is that the grades scored by students with help of the feedback agent are better than the grades scored by students without the feedback agent. Hence the task grades (the quality of class diagrams created) with a feedback agent (FBon) and task grades without the feedback agent (FBoff) are explained. The hypothesis regarding task performance (time taken to perform the task) with and without a feedback agent are also explained below:

a) *Hypothesis for task grades (TG)*: The null hypothesis states that the task grade with a feedback agent is equal to the task grade without the feedback agent.

$$H_0TG : Grade(FBon) = Grade(FBoff)$$

The alternative hypothesis states that the task grade with a feedback agent is greater than the task grade without the feedback agent.

$$H_1TG : Grade(FBon) > Grade(FBoff)$$

b) *Hypothesis for the task performance (TP)*: The null hypothesis states that the task performance (time to perform the task) with a feedback agent is equal to the task performance grade without the feedback agent.

$$H_0TP : Time(FBon) = Time(FBoff)$$

The alternative hypothesis states that the task performance with a feedback agent is smaller than the task performance without the feedback agent.

$$H_1TP : Time(FBon) < Time(FBoff)$$

TABLE I
INTERVIEW ANSWERS

Tool Functionality
Extending the functionality is definitely needed though to make a class diagram. More functionality could definitely make this tool to be useful.
I want to be able to delete classes.
Help about the type of associations. Didn't remember what's composition and what's aggregation.
The Feedback Agent
It told me I shouldn't capitalize the methods, that was a good feedback. Feedback was good as far as I could test it.
If it is a standard to keep everything lowercase for example. Like it complained about the attributes. Then it's nice to make the person abide by the conventions. So that's nice.
It's fine, and works well.
It's a nice touch to have tests for capital letters and such, especially for learning purposes.
The feedback helped to adhere to consistent naming.
The feedback helped not to miss any associations or leave labels and attributes empty.
The feedback helped in avoiding mistakes later on. It taught me the style guide of writing UML diagrams.
Implemented Suggestions
I dislike the positioning of the button though. I have to go back and forth with my eyesight.
Class can be called Class. Should be reserved.
I don't know if <i>attribute</i> should be reserved. I can call an attribute <i>attribute</i> if I want to.
There are no warnings or mentions about duplicates.

7) *Design*: The subjects of the experiment were randomly divided into two groups, one group solving the task with a feedback agent and the other group without. Both the groups were given the same instructions on how to use WebUML2 and the same task to solve. After the task both groups were given a questionnaire in order to enable qualitative analysis of the subjects' experience.

8) *Execution*: The experiment was run over three days during the second week of May 2017. The participants were given a brief instruction on how to use WebUML2 and the feedback agent and then were randomly assigned to one of the two groups, to solve the task with support of the feedback agent or without. Moreover, the students were introduced to the task that they would have to solve.

In order to separate the time needed to understand the task from the time needed to solve the actual task, the students were instructed to wait with designing until they understood the task description and were ready to start. Then time was taken from the time the students started to design until they finished their design. After the experiment was done, the collected data was graded, analyzed and then used to measure the dependent variables.

IV. RESULTS

In this section we present our findings from the interviews, the experiment and the questionnaire we have conducted. We start off showing the results of the interviews. We then present the results of the experiment and the statistical analysis of the results. Finally the findings of the questionnaire are presented.

A. Interviews

During the analysis of the interview data, three distinct subjects were touched upon. While our questions didn't focus on the functionality of the WebUML tool itself, but rather on the feedback agent, it did get brought up. We also received opinions and thoughts about the feedback agent and feature suggestions as a result of our interview questions. Therefore, we sorted the results of the interviews into the categories Tool Functionality, Feedback Agent and Implemented Suggestions (table 1).

TABLE II
GRADES

With Feedback		Without Feedback	
Grade	Time (M:S)	Grade	Time (M:S)
3	17:44	2	17:28
4	13:06	2	14:35
3	11:26	1	10:25
3	19:31	4	25:17
3	17:43	2	18:05
3	19:22	2	17:10
4	15:45	3	25:30
3	13:36	2	23:45
2	10:44	4	19:34
2	9:28	3	17:05

B. Experiment

To collect our quantitative data we asked 20 students to design a class diagram based on the task found in appendix 1. This experiment yielded 20 UML class diagram solutions

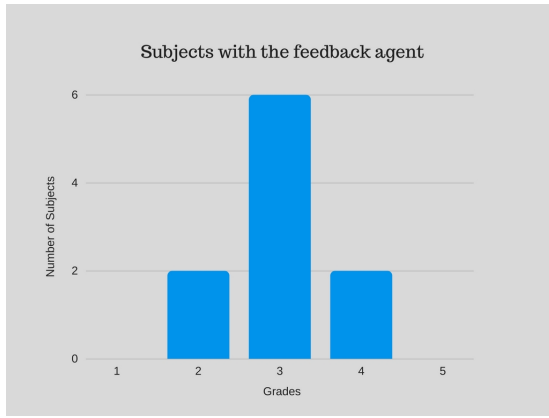


Figure 4: Grades - Sample with feedback

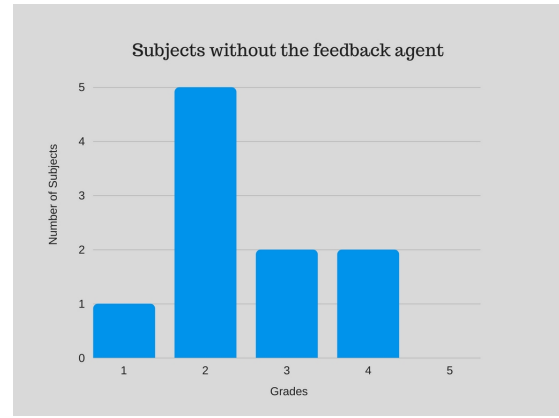


Figure 5: Grades - Sample without feedback

which were graded according to a five point scale rubric [25] by two experts who came to a consensus by discussion. The grades thus range from 1 (lowest) to 5 (highest). Table 2 lists the grades scored and the time taken to design the class diagram for both groups of students. To give a better overview of the grades earned by both groups, Figure 4 visualizes the grades for students with the feedback agent and figure 5 shows the grades for those without.

1) *Testing the null hypothesis for task grades:* We ran the Shapiro-Wilk test [28] to test our samples for normality using the following null hypothesis:

$$H_0 : \text{data is sampled from a normal distribution.}$$

The test result is significant with a p-value of 0.025, which is less than our alpha of 0.05, causing us to reject the null hypothesis.

Now we know the data is not normally distributed, the sample size is relatively small and because we want to compare data from two independent samples, we used the Mann Whitney U test to test our hypothesis [29].

$$H_0TG : \text{Grade}(FBon) = \text{Grade}(FBoff)$$

The test resulted in a U-value of 33. The critical value of U at $p < 0.05$ is 27. So the result is not significant. The p-value of 0.10 is not less than our alpha of 0.05 either, confirming the result is not significant. Thus we were not able to reject the null hypothesis.

2) *Testing the null hypothesis for task performance:* Again we ran the Shapiro-Wilk Test to test our samples for normality using the null hypothesis that the samples are normally distributed.

$$H_0 : \text{data is sampled from a normal distribution.}$$

The test resulted in a p-value of 0.42 and 0.46 for the two respective samples. Both are larger than our chosen alpha of 0.05, which means we cannot reject the null hypothesis and will assume the data is normally distributed. Based on the normal distribution and our data being 2 independent samples,

we chose an independent measures t-test [30] to test our null hypothesis:

$$H_0TP : \text{Time}(FBon) = \text{Time}(FBoff)$$

The resulting t-value is 2.11. The resulting p-value is 0.02. The result is significant with a p-value of 0.02 which is less than our alpha of 0.05, meaning we reject our null hypothesis.

C. Questionnaires

1) *General questions:* After the students designed the UML class diagram we asked them to fill out a questionnaire. The first two questions asked all 20 students about their experience in both UML modelling (see figure 6) and designing Class Diagrams (see figure 7). All the students had some familiarity with both topics, with the majority feeling somewhat to moderately familiar with them.

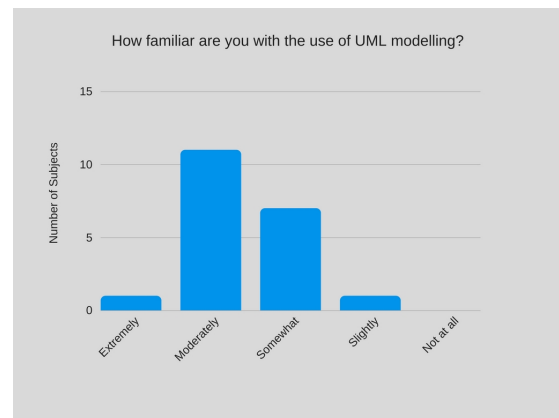


Figure 6: Familiarity with UML modelling

2) *Feedback agent questions:* The remainder of the questions were related to feedback and were only asked to the 10 students who had the feedback agent enabled during their task.

When asked about their learning process, opinions were divided (figure 8). While some students agreed that the feedback increased their learning process, a few disagreed and a few were undecided.

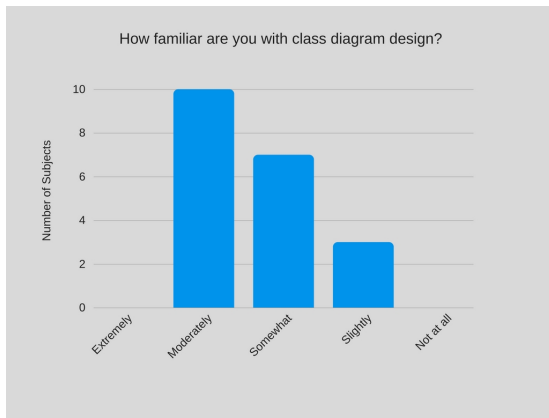


Figure 7: Familiarity with class diagram design

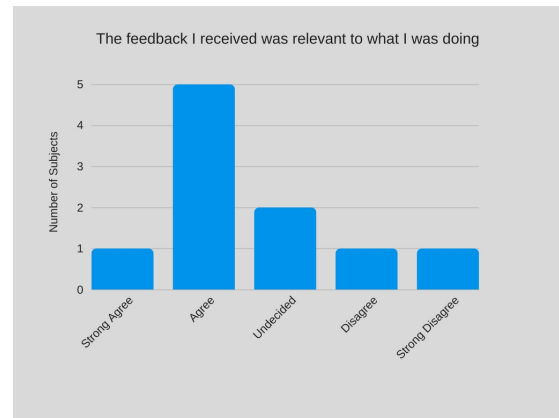


Figure 9: Relevance of feedback

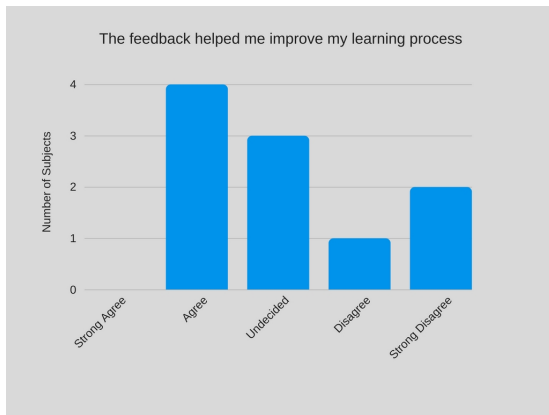


Figure 8: Learning process

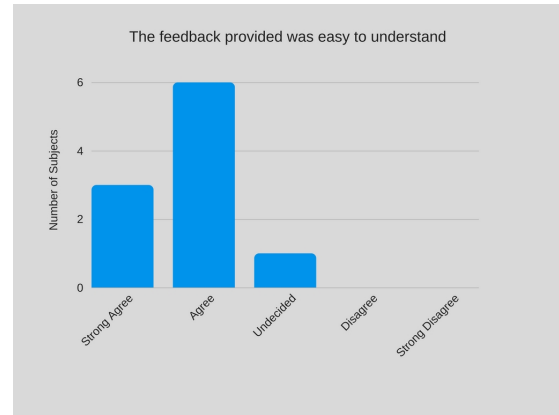


Figure 10: Understandability of feedback

Another question asked them if the feedback they received during the task was relevant (figure 9). The majority of students agreed.

When asked if the feedback was easy to understand (figure 10), everyone agreed, apart from one student who was undecided.

V. DISCUSSION

In this section we discuss the research questions posed in section III based on the results presented in section IV. Our findings suggest that i) Automated feedback has no statistically significant effect on students' learning outcome. ii) Automated feedback has a statistically significant effect on students' task performance. iii) Furthermore, our qualitative data suggests that students had experienced improvement in their learning process, however the findings vary between students.

A. RQ1. Does the use of automated feedback mechanisms in e-learning systems improve the learning process of students of UML class diagram design?

In order to answer RQ1 we have analyzed both our qualitative data collected through the interviews and the questionnaire that was conducted after the experiment. The qualitative data

suggests students had experienced improvement in their learning process of UML class diagram design. In the interviews we conducted most of the students give a positive response regarding their learning process with the feedback agent. "Like it complained about the attributes. Then it's nice to make the person abide by the conventions. So that's nice." said one student, expressing the support the feedback agent gave in terms of UML naming conventions. Another one addresses how the feedback agent improves their learning process by helping them avoid unnecessary mistakes, "The feedback helped in avoiding mistakes later on. It taught me the style guide of writing UML diagrams.". Moreover in the questionnaire when students were asked if the feedback helped them improve their learning process, most of them responded positively as shown in figure 8. Hence, we can say that the feedback agent had a positive effect in the learning process of UML class diagram.

However, there were also findings that indicate otherwise, mostly regarding the tools' functionality and behaviour. For instance, one student said, "Extending the functionality is definitely needed though to make a class diagram. More functionality could definitely make this tool to be useful." indicated the need to extend the functionality of the tool in order to be able to create more complex class diagrams thereby demanding more complexity from the feedback agent. "I want

to be able to delete classes.”, others focus more specifically on WebUML2’s behaviour and not the feedback agent. Since the editing tool is in its development stages there were bugs and limitations that limited what students could do, therefore affecting the learning process.

Another interesting finding was that providing feedback on aspects like inheritance realization, composition, aggregation and generalization cannot easily be addressed in general terms. To give specific and relevant feedback on these aspects, a build-in example solution designed around a single scenario, would likely be needed. However, that was not in line with the goals for our research, because this approach limits the usability of the feedback agent as the solution needs to be hard-coded.

B. RQ2. Does such a mechanism improve the learning outcome of the students? (quality of class diagrams created)

According to our statistical analysis we found out that the feedback agent that we integrated into WebUML2 did not have a significant effect on the students learning outcome. These findings are based on the expert gradings that we compared statistically. The grades reflect how well the students did in the modelling task; the group with the feedback agent had an average grade of 3.0 points while the group without the feedback agent had 2.5 points. This shows that the students with the feedback had slightly better results. As Hattie and Timperley suggested, feedback influences the learning process [13], the grades scored by students with a feedback agent indicates that as shown in Figure 4. However, from the statistical analysis we cannot conclude that automated feedback improves the learning outcome of the student, thereby rejecting the null hypotheses and answering RQ2. Even though there is no statistically significant difference in the result, we are not convinced that automated feedback did not have an effect on the outcome of the student’s grade (quality of class diagrams created). A more thorough investigation needs to be performed with a larger group of subjects and varied tasks in order to determine the effects of automated feedback both in the learning outcome and learning process of students. Hence a similar study like Hattie and Timperley did with a large group is advised.

C. RQ3. Does the students’ performance (measured in time) in solving the task improve by the automated feedback?

Running a statistical analysis on the recorded time it took students to complete the task showed that students with the feedback agent took shorter time(s) to complete the task, while students without took longer time. Comparing the mean values of the group with feedback (890.5s) versus those without (1133.4s), we observe that the difference is in favor of the students that had the feedback agent available to them. Moreover, we could reject the null hypothesis with a 95% confidence level and say student’s performance in solving the task is improved by the automated feedback, thereby answering RQ3. This could be a result of students having to worry less about making mistakes while working on the

task, therefore spending less time to complete it. In contrast to students without the feedback where they have reconsider all their steps. The feedback provided was simple to understand by the students, as shown in table 1, and hence students could process them fast. When asked about the relevance and the easiness of the feedback, the majority of students agreed that the feedback provided was easy to understand and relevant. This was reflected in the questionnaire as shown in figure 9 and 10.

Another factor could be that the task being simplistic and straightforward in nature in line with Narciss [19] suggestion with the ITF-model. Where the learners characteristics and the level of task complexity is considered in designing the questions. Having a simple task supported with feedback can result in students performing better than they would have in a normal case, hence we advice for a future study where bigger groups of students are exposed to different levels of task and feedback complexity. Moreover, the students level of skill can also be tested before the test and after the test with the feedback agent, in order to see their learning yield.

VI. THREATS TO VALIDITY

In this study, the threats to internal and external validity are most critical, but the threats to construct and conclusion validity are also addressed. Concerning the internal validity, even though the subjects of our experiment were from the software engineering and management program at the University of Gothenburg and had attended a course in software design and UML design as part of their study, their confidence and motivation in solving the task varied. In order to address this student were informed that the task was designed for beginners and that they don’t need to design advanced class diagrams. Furthermore to mitigate the threat of having students with very distinct knowledge in class diagram and background in UML design, all the participants were informed about class diagrams, WebUML2 and the task. In addition students were rewarded with homemade banana bread for completing the experiment in order to motivate them.

The use of a limited number of subjects for our experiment raises the issue of external validity, since subjects with different backgrounds and different levels of experience would contribute to generalization of our findings. While the subjects are chosen from each year of the bachelor program in Software Engineering and Management, with different levels of experience, they have the same background. However, we have included two experienced developers with real-world experience in software design and development in our interview stages. Nonetheless, we are fully aware of that we cannot generalize the conclusion based on our subjects.

Concerning construct validity, in our experiment we used a single task, in order to test how the automated feedback affects students. Because the test was very simplistic and the fact that the feedback agent being tested in one task, makes it difficult to measure the actual effect of the feedback agent. The reason for this risk is the time constraint of the experiment and development phase. In order to fully address this risk

several experiments with different tasks need to be performed. Moreover, a more complex feedback agent that can deal with different task requirements needs to be developed.

Furthermore, some students were afraid of being evaluated, hence introducing evaluation apprehension. This risk was mitigated through ensuring students anonymity in the experiment. Finally the threats to conclusion validity are considered to be under control. The normality in the sample is checked in order to consider the correct type of test, between parametric and non-parametric test.

VII. CONCLUSION AND FUTURE WORK

In this paper we aimed to reveal if a feedback agent can improve a students' learning process, learning outcome and performance. We integrated a feedback agent into the web-based UML editing tool WebUML2. The feedback agent was iteratively developed through interviewing students regarding to the design and content of the feedback agent. The result from our interview with the students and the questionnaire after the experiment indicate that a feedback agent can, to some extent, play a significant role in improving learning process.

The WebUML2 with the feedback agent was used to run an experiment, where two groups of student designed a class diagram for a simple task. One group had access to the feedback agent while the other group did not. For each group we recorded the solution and the time it took to complete the task. This allowed us to statistically analyze the effects of the feedback on learning outcome and performance. From a statistical point we could conclude that the automated feedback has no significant effect on the learning outcome, but that it does have a significant effect on students' performance measured in time. Meaning the students' grade in terms of class diagram quality did not have a significant difference between the students with the feedback agent and those without it, while in terms of performance students with the feedback agent performed significantly better in terms of time than the students without feedback agent.

Our study gives insight into the benefits of automated feedback in software design e-learning environments, while realizing the challenges related to giving automated feedback for complex class diagrams designs. On the contrary, our research showed that a feedback agent helped to improve performance and helped improving the learning process.

While this thesis has demonstrated the potential of a feedback agent for software design in an e-learning environment, we also see that there is a need for further research on the following aspects:

i) How to design a feedback agent which can address more complex concepts like inheritance realization, composition, aggregation and generalization? Having such a feedback agent will allow to perform a better experiment with more complex tasks. Thereby giving a greater perspective and understanding on students' learning process.

ii) Performing the study with a larger subjects and varied tasks in order to determine the effects of the feedback agent

both in the learning outcome and students performance.

ACKNOWLEDGEMENT

We would like to thank all interviewees and participants in the experiment, who made this research possible. We would also like to thank Francisco Gomes for his constructive feedback, and our supervisor, Dave Stikkolorum, for his support and guidance.

REFERENCES

- [1] N. Wagner, K. Hassanein, and M. Head, "Who Is Responsible for E-Learning Success in Higher Education? A Stakeholders' Analysis." *Journal of Educational Technology & Society*, vol. 11, no. 3, pp. 26–36, 2008. [Online]. Available: <http://www.jstor.org/stable/jeductechsoci.11.3.26>
- [2] R. Moreno, "Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia." *Instructional Science*, vol. 32, pp. 99–113, 2004. [Online]. Available: <http://dx.doi.org/10.1023/B:TRUC.0000021811.66966.1d>
- [3] J. Rumbaugh, I. Jacobson, and G. Booch, *The Unified Modeling Language Reference Manual*. Addison-Wesley, 2010, vol. 2, pp. 496–497.
- [4] W. Chen, R. H. Pedersen, and O. Pettersen, "CoLeMo: A collaborative learning environment for UML modeling." *Interactive Learning Environments*, vol. 14, no. 3, pp. 233–249, 2006. [Online]. Available: <http://dx.doi.org/10.1080/10494820600909165>
- [5] H. K. Suen, "Peer Assessment For Massive Open Online Courses (Moocs)." *The International Review of Research in Open and Distributed Learning*, vol. 15, no. 3, 2014. [Online]. Available: <http://www.irrodl.org/index.php/irrodl/article/view/1680>
- [6] U. M. Krause, R. Stark, and H. Mandl, "The effects of cooperative learning and feedback on e-learning in statistics." pp. 158–170, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0959475208000376>
- [7] D. R. Stikkolorum, T. Ho-Quang, B. Karasneh, and M. R. V. Chaudron, "Uncovering students' common difficulties and strategies during a class diagram design process: an online experiment." in *MoDELS*, 2015.
- [8] D. R. Stikkolorum, T. Ho-Quang, and M. R. V. Chaudron, "Revealing students' uml class diagram modelling strategies with webuml and logviz." in *2015 41st Euromicro Conference on Software Engineering and Advanced Applications*, Aug 2015, pp. 275–279.
- [9] I. Koskinen, J. Zimmerman, T. Binder, J. Redstrom, and S. Wensveen, "Design research through practice: From the lab, field, and showroom." *IEEE Transactions on Professional Communication*, vol. 56, no. 3, pp. 262–263, Sept 2013.
- [10] D. J. Nicol and D. Macfarlane-Dick, "Formative assessment and self-regulated learning: a model and seven principles of good feedback practice," *Studies in Higher Education*, vol. 31, no. 2, pp. 199–218, 2006. [Online]. Available: <http://dx.doi.org/10.1080/03075070600572090>
- [11] J. Hattie and M. Gan, "Instruction based on feedback." in *Handbook of Research on Learning and Instruction*, R. Mayer and P. Alexander, Eds. Routledge, 2011, pp. 249–271.
- [12] J. Hattie and H. Timperley, "The power of feedback." *Review of Educational Research*, vol. 77, no. 1, 2007. [Online]. Available: <http://journals.sagepub.com/doi/abs/10.3102/003465430298487>
- [13] J. Hattie, "Influences on student learning." p. 81–112, 1999. [Online]. Available: <http://teacherstoolbox.co.uk/downloads/managers/Influencesonstudent.pdf>
- [14] S. Narciss, "Feedback strategies." in *Encyclopedia of the Sciences of Learning*, N. Seel, Ed. Springer, 2012, vol. 6, pp. 1289–1293.
- [15] V. J. Shute, "Focus on formative feedback." *Review of Educational Research*, vol. 78, no. 1, pp. 153–189, 2008. [Online]. Available: <http://dx.doi.org/10.3102/0034654307313795>
- [16] I. Arroyo, B. P. Woolf, D. G. Cooper, W. Burleson, and K. Muldner, "The impact of animated pedagogical agents on girls' and boys' emotions, attitudes, behaviors and learning," in *2011 IEEE 11th International Conference on Advanced Learning Technologies*, July 2011, pp. 506–510.
- [17] I. M. Goldin, K. R. Koedinger, and V. Aleven, "Learner differences in hint processing," in *Proceedings of the 5th International Conference on Educational Data Mining*, K. Yacef, O. Zaïane, H. Hershkovitz, M. Yudelson, and J. Stamper, Eds., 2012, pp. 73–80.

- [18] A. Mitrovic, S. Ohlsson, and D. K. Barrow, "The effect of positive feedback in a constraint based intelligent tutoring system," *Computers & Education*, vol. 60, no. 1, pp. 264–272, 2013. [Online]. Available: <https://doi.org/10.1016/j.compedu.2012.07.002>
- [19] S. Narciss, "Conditions and effects of feedback viewed through the lens of the interactive tutoring feedback model." in *Scaling up Assessment for Learning in Higher Education.*, D. Carless, S. M. Bridges, C. K. Y. Chan, and R. Glofcheski, Eds. Springer Singapore, 2017, pp. 173–189. [Online]. Available: http://dx.doi.org/10.1007/978-981-10-3045-1_12
- [20] U. M. Krause, R. Stark, and H. Mandl, "The effects of cooperative learning and feedback on e-learning in statistics." *Learning and Instruction*, vol. 19, no. 2, pp. 158–170, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0959475208000376>
- [21] I. Goldin, S. Narciss, P. Foltz, and M. Bauer, "New Directions In Formative Feedback In Interactive Learning Environments," *International Journal of Artificial Intelligence in Education*, pp. 1–8, 2017.
- [22] R. Clark, D. Feldon, J. Van Merriënboer, K. Yates, and S. Early, "Cognitive task analysis," in *Handbook of research on educational communications and technology*, J. M. Spector, M. D. Merrill, J. J. G. van Merriënboer, and M. P. Driscoll, Eds. Mahaw: Lawrence Erlbaum Associates., 2008, p. 577–593.
- [23] J. Dolonen, W. Chen, and A. Mørch, "Integrating software agents with fle3," in *Designing for Change in Networked Learning Environments: Proceedings of the International Conference on Computer Support for Collaborative Learning 2003*. Kluwer Academic Publishers, 2003, pp. 157–161. [Online]. Available: http://dx.doi.org/10.1007/978-94-017-0195-2_21
- [24] H. Ankar, "Providing Automated Feedback on Software Design for Novice Designers." Gothenburg University, 2015. [Online]. Available: https://gupea.ub.gu.se/bitstream/2077/39973/1/gupea_2077_39973_1.pdf
- [25] D. Stikkolorum, "Online Experiment with WebUML : recording difficulties and strategies of students." 2015. [Online]. Available: <http://liacs.leidenuniv.nl/~stikkolorumdr/LIACS/%20Technical/%20Report/%20Stikkolorum/%202015-07.pdf>
- [26] C. B. Seaman, "Qualitative Methods in Empirical Studies of Software Engineering," *IEEE Transactions on Software Engineering*, vol. 25, no. 4, pp. 557–572, 1999. [Online]. Available: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=799955>
- [27] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, *Experimentation in Software Engineering: An Introduction*. Kluwer Academic Publishers, 2000.
- [28] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality," *Biometrika*, vol. 53, no. 3-4, pp. 591–611, 1965. [Online]. Available: <http://www.jstor.org/stable/2333709>
- [29] H. Mann and D. Whitney, "On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other," *The Annals of Mathematical Statistics*, vol. 18, no. 1, pp. 50–60, 1947. [Online]. Available: <http://www.jstor.org/stable/2236101>
- [30] W. Mendenhall and T. Sincich, *Statistics for Engineering and the Sciences, Sixth Edition*. Taylor & Francis, 2015.