An online corpus of UML Design Models : construction and empirical studies

Karasneh, B.H.A.

Chapter 8

Using Examples for Teaching Software Design

In this chapter, we present a research that is positioned in the field of software design method and teaching thereof. The aim of this research is to study the effects of using a collection of examples for creating a software design. We ran a controlled experiment for evaluating the use of a broad collection of examples for creating software designs by software engineering students. In this study, we focused on software designs as represented through UML class diagrams. The treatment is the use of the collection of examples. These examples are offered via a searchable repository. The outcome variable we study is the quality of the design (as assessed by a group of experts). After this, all students were offered the opportunity to improve their design using the collection of examples. We ran a post-assignment questionnaire to collect qualitative data about the experience of the participants. Considering six quality attributes measured by experts, our results show that: 1) the models of the students who used examples are 18% better than those of who did not use examples. 2) the models of the students who did not use examples for constructing became 19% better after updating their models using examples. We complement our statistical analysis with insights from the post assignment questionnaire. Also, we observed that students are more confident about their design when they use examples. Students deliver better software designs when they use a collection of example software designs.

This chapter is based on the following publications:

- Bilal Karasneh, Rodi Jolak and Michel R. V. Chaudron. **Using Examples for Teaching Software Design.** *In Proceedings of the 22st Asia-Pacific Software Engineering Conference (APSEC2015). New Delhi, India. 2015*
- Bilal Karasneh, Michel R. V. Chaudron. **Online Img2UML Repository: An Online Repository for UML Models.** *In Proceedings of the 3rd International Workshop on Experiences and Empirical Studies in Software Modeling (EESS-MOD@MODELS 2013), pages 61-66, Miami, USA. 2013.*

Examples could guide students who are novices in UML modeling to create better designs. In fact, modeling is much more difficult for students compared to programming tasks. This difficulty occurs because for most programming languages, students can get feedback immediately from the harsh compiler on the code they produce, e.g. through compile and run-time errors. However, such feedback is poorly available in modeling, and current CASE tools do not give hints for users related to models quality. Therefore, students need to have some feedback from their teacher, or at least from each other to trust their designs. As a consequence, students cannot get the kind of self-learning facility when they study modeling as they can get for learning programming.

We see that a variety of examples could improve students' performance for creating designs. From another side, a variety of examples could make students confused and make the modeling task more difficult to perform. We conducted a controlled experiment and asked students to give their feedbacks about using examples for creating software designs.

We offered students our UML Repository [111][112], to find examples of class diagrams. We asked students to compare between using the model repository, and searching for useful examples on the internet. We also asked the students how they preferred to search for example class diagrams.

The results show that using examples helps most of the students for creating and improving their designs, and it makes them more confident. A few students felt confused when they used examples because they could not distinguish between good and bad examples. The results also show that most of the students prefer using the UML Repository more than searching the internet to find models. This is because the repository produces more specific results because it supports searching based on details of the contents of UML models, e.g. names of classes and operations. It is difficult to search for these features using generic internet search engines. As a result, this specialized search saves students' time and effort.

8.1 Related Work

Van Gog et al. [113] showed that using examples decreases the mental effort required to understand problems. On the other hand, Goldstone and Son [114] showed that examples capture intuition. Seater [115] stated that although examples may include incorrect solutions beside the correct ones, understanding them gives additional insight and helps to reason why and when the model is correct. Many studies come up with the fact that solving problems is more effective when having multiple examples [114][115][116].

Bkak et al. [117] presented the Example-Driven Modeling (EDM) approach, which uses explicit examples for creating and validating business knowledge. They present many questions related to EDM, like: How useful are the examples for building models? What is the impact of examples on the comprehension of models? What kind of tool support is needed to work with examples?

Many empirical studies were conducted to evaluate comprehension techniques in UML. Zayan et al. [118] constructed a controlled experiment for empirical evaluation of EMD. They represented the abstraction part as UML class diagram and the examples as object diagrams. The results showed that EDM is better than having model abstraction only. Many studies take the impacts of the layout on model comprehension into account, and most of them use controlled experiments with different methods, like eye tracking, questionnaires and surveys [119]. Störrle [120][121] studied impacts of models layout from different perspectives. Nugroho [122] illustrated the effectiveness of the levels of detail on the UML model comprehension.

Basing on models comprehension, and by this experiment, we want to see whether multiple examples having different layouts and different levels of detail can help students to create better models and use elements of UML in the right way. In addition, we want to know if they could learn from examples how many details to include in their designs, make them more confident and optimize their model.

Our UML Repository [111][112] contains more than 810 UML class diagrams collected from the internet, open source projects, collaborations and via scientific literature. There are no other published UML models yet. Models are available as images and XMI. The available class diagrams vary in size and complexity. This repository is built using our Img2UML tool [52][53] that converts UML diagrams stored in image formats into XMI.

In our experiment, we use the UML Repository where models are searchable. Users can search for models using class-, attribute-, and operation names. In addition, users can search for models based on design metrics, for example: the number of classes.

8.2 Research Questions and Hypotheses

In this section, we explain our research questions and hypotheses that we are going to test. We address the following research questions:

- RQ1: Does the aid of using examples improve the quality of software design models created by novices?
- RQ2: Does using a variety of examples affect the quality of models created by novices?
- RQ3: What is the best way to offer examples? Can a model-repository be considered a good way of offering model examples?

Accordingly, we formulate the following *null*/**alternative** hypotheses:

- $H1_0$/**H1**: Constructing models with the aid of examples [*does not improve*]$_0$ /[**improves**] the quality of models created by novices.
- $H2_0$/**H2**: Improving models with the aid of examples [*does not improve*]$_0$ /[**improves**] the quality of models created by novices.
- $H3_0$/**H3**: Using a variety of examples [*does not improve*]$_0$ /[**improves**] the quality of models created by novices.
- $H4_0$/**H4**: Using a variety of examples [*does not improve*]$_0$ /[**improves**] the quality of models improved by novices.
- $H5_0$/**H5**: Constructing models with the aid of examples [*does not improve*]$_0$ /[**improves**] the students perceived confidence in their created models.
- $H6_0$/**H6**: Using model repositories can be considered as [*normal*]$_0$ /[**better**] way of offering model examples.

In the hypotheses, we differentiate between constructing/creating and improving models with aid of examples. Improved models are models updated by students later. We consider that the students who do not improve their models are confident about these models, which they created during the experiment. Quantitative and qualitative analyzes were performed to accept or reject the hypotheses.

8.3 Experiment Design

The aim of the study is to investigate the effect of using model examples in creating designs. We followed the guidelines for software engineering experimentation [123]. Figure 8.1 shows the experimental approach.

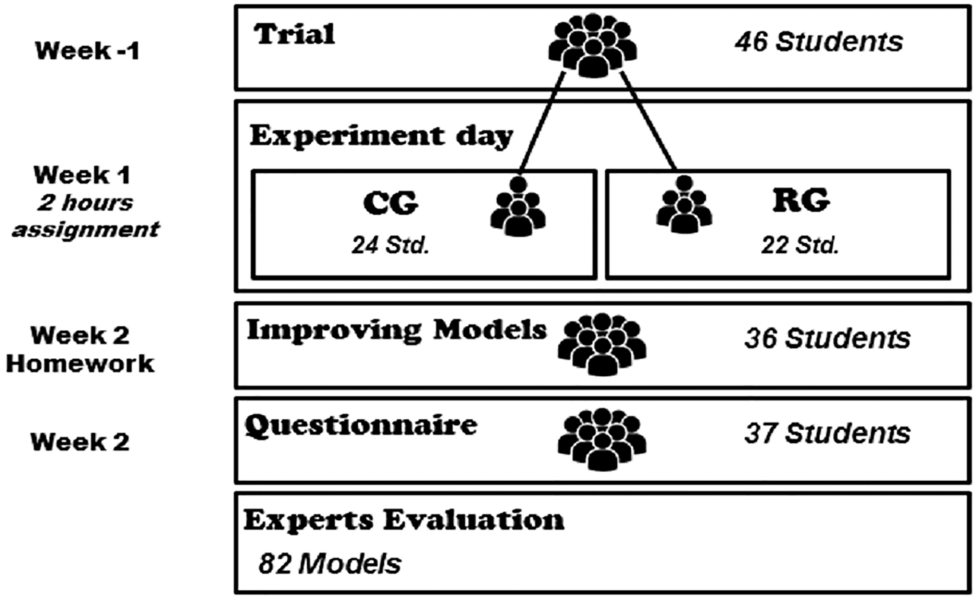

Figure 8.1: *Experimental Approach*

8.3.1 Method

The experiment was conducted at Leiden University; both experts and students participated. First, students were given a modeling task, and they used the StarUML CASE tool [58], to create their models. The students were separated into two groups: one group was allowed to search for models available in the UML Repository, and we call this group the Repository Group (RG). The other group was not allowed to use any examples, and we call it the Control Group (CG). Before starting with the experiment, the students did not know in which group they were. Students also had a two weeks trial before the experiment with another modeling task. The trail was important to get them familiar with using the repository. Students were given two hours to create their models. After that, we gave all students from both groups two days to improve their models using the repository. For CG, they had the opportunity to use the repository to improve their models. For RG, this simply meant that they had more time to use the repository to improve their models. After that, we asked the students to answer an evaluation questionnaire.

8.3.2 Operation

The entire experiment including the trail, constructing models, improvements, the questionnaire and experts' evaluation covered a period from the middle of November

2014 until the middle of February 2015. The experiment had been scheduled at the end of a software engineering course.

8.3.3 Evaluation

Five experts from Leiden and Chalmers Universities joined this experiment; they are working in academia and have some experience in the industrial domain. Two of them are also expert in teaching courses like software engineering and software modeling. We asked them to evaluate the students' models basing on six quality attributes (scale from 1 8): understandability, layout, extensibility, modifiability, completeness and correctness. Experts evaluated all of the 82 models individually. Experts did not know whether a certain model belongs to the RG or CG group. Moreover, they did not know which models were created or improved during which phase of the experiment.

8.3.4 Participant

Forty-six students from the Master program "ICT in Business" from Leiden University participated in the experiment. All the students, at that time, had less than one year of experience in UML modeling. The students were randomly divided into two groups, 22 students were assigned to the RG, and 24 students to the CG.

8.3.5 Data Collection

Students models, questionnaire answers, and models evaluation were collected. All of these data were available at the repository. First, the students uploaded their models to the repository, and they answered two questions: 1) about their background (academia, industry or both), 2) about their experience in UML modeling whether it is less than one year, <1-2>, <2-5> or more than 5 years. Second, students could update their models and upload them again. Third, the students answered a post-experiment questionnaire. The questionnaire contained 12 questions: seven questions with a scale (-4,4) or (1,8), four open questions, and one multiple choice. Finally, the experts used an online evaluation form. The evaluation form shows a class diagram created by a student and six quality attributes, each one of them has a scale from 1 8:

- For Understandability, Extensibility and Modifiability: (1) is difficult, (8) is easy.
- For Layout: (1) is complex, (8) is simple.
- For Completeness: (1) is not complete, (8) is complete.
- For Correctness: (1) is not correct, (8) is correct.
- Uploading, evaluation, and questionnaire forms are available in the supplemental materials of the experiment [108].

8.4 Results

In this section, we describe the results of a quantitative analysis based on the experts' evaluation. We also describe the outcomes of qualitative analysis based on the post-experiment questionnaire.

To check the normality of both populations (RG and CG), we took the average of experts' evaluation per each quality attribute, and then we used the Shapiro-Walk test [124]. When both populations are verified to be normally distributed, the independent samples students t-test [123] was used to check whether there is a significant difference between the mean of both populations. However, when the populations are not normally distributed; the non-parametric Mann-Whitney test [125] was used for the same comparison of the mean of both populations.

We used the statistical package R [126] to perform all tests. We chose a significance level of 0.05, which corresponds to a 95% confidence interval.

8.4.1 Experts Evaluation

In this section, we show the results of the experts' evaluation for both RG and CG, and the experts evaluation for CG before and after they use the repository.

8.4.1.1 Comparison between the RG and CG

Normality of all evaluations was checked, and the assumption was met for all cases. RG models have in average an 18% better evaluation of all quality attributes than CG models. Table 8.1 shows the one-tail t-test results for all quality attributes evaluations. All p-values in Table 8.1 show that there is a significant difference between RG and CG, and the RG has better results. Figure 8.1 shows the distribution of scores of RG and CG as boxplots.

8.4.1.2 Analyzing the improvement by the control group

Only 20 students from 24 improved their models in the CG, so we made the comparison based on the experts' evaluation for the models created by those 20 students. The normality of all evaluations was checked, and the assumption was met for understandability, modifiability, and correctness. After updating their models with the aid of examples, the CG received 19% better evaluation of all quality attributes compared to what they got for their old models created before using examples. Table 8.2 shows the results of the one tail test for all quality attributes. From Table 8.2, we observe that all p-values, except for the correctness case, indicate a significant difference for CG before and after using the repository, and there are better results after using the repository. For correctness, there are better results after they used the repository in

Table 8.1: *Results of Students t-test one tail*

Quality attribute	Groups	Mean	Median	p-value
Understandability	RG	5.56	5.58	0.001
	CG	4.61	4.83	
Layout	RG	6.12	6.25	0.003
	CG	5.09	5.42	
Extensibility	RG	5.64	5.67	0.009
	CG	4.66	4.83	
Modifiability	RG	5.53	5.58	0.00
	CG	4.52	4.67	
Completeness	RG	5.77	6.13	0.006
	CG	4.76	5	
Correctness	RG	5.25	5.42	0.024
	CG	4.46	4.67	

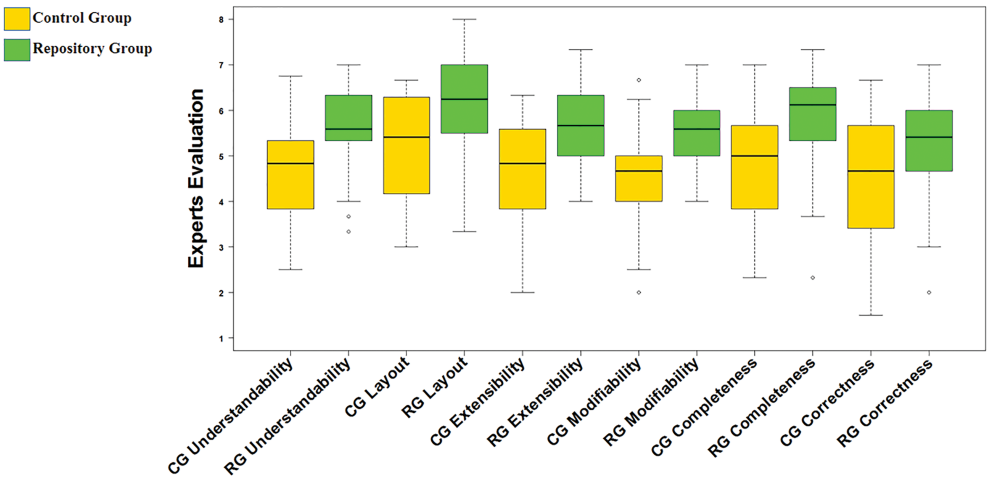

Figure 8.2: *Evaluation of RG models and CG models*

terms of the median. Figure 8.3 shows the evaluation of CG before and after they used the repository.

8.4.1.3 Comparing RG and CG after the CG Improves Their Models

We could not find any significant difference between evaluations of RG models created during the experiment, and improved models of the CG. It makes us conclude that

Table 8.2: *Results of students t-test and Mann Whitney test (One tail) of the CG*

Quality attribute	Groups	Mean	Median	p-value
Understandability	After	6.03	6.33	0.000
	Before	4.71	4.83	
Layout	After	6.05	6	0.024
	Before	5.25	5.58	
Extensibility	After	5.89	6	0.000
	Before	4.69	5.13	
Modifiability	After	5.78	5.83	0.000
	Before	4.62	4.67	
Completeness	After	5.94	6.17	0.000
	Before	4.83	5	
Correctness	After	5.18	5.33	0.068
	Before	4.63	4.67	

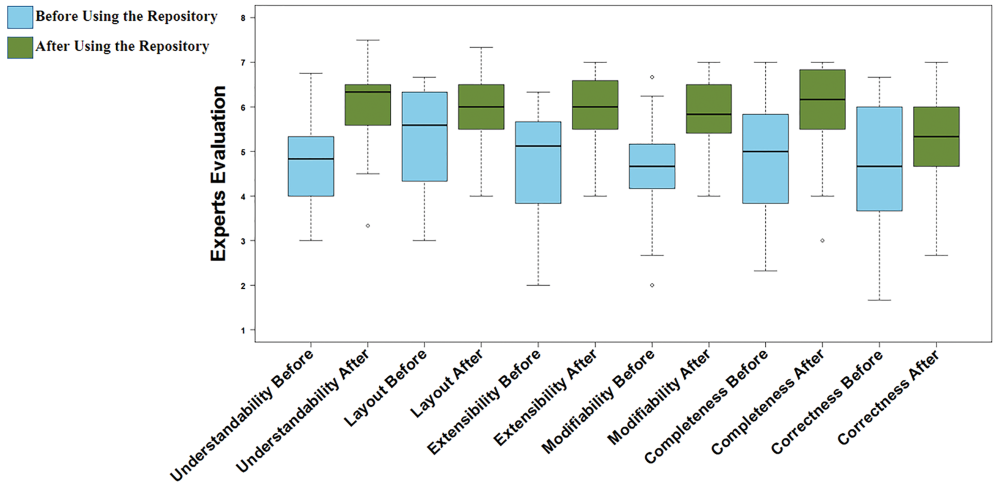

Figure 8.3: *Evaluation of CG models created during the experiment and after their improvement using the repository*

there is no difference of using examples during the construction of models or during the improvement.

8.4.2 Models Improvements

We instructed the students not to improve/update their models when they already think that they have good design. We assume that the students who did not improve their models were confident about their models. From the 46 students that participated in the experiment, only 36 students improved their models: 16 from RG and 20 from CG. Quantitatively: 27% of the students from RG were confident about their models, and 16% of the students from CG were confident about their models.

8.4.3 Post-assignment Questionnaire

In this section, we discuss the highlights of the responses of the students to the post-assignment questionnaire. 37 students responded to the questionnaire, 17 are from RG, and 20 are from CG. Next, we discuss the questions and responses of the post-assignment questionnaire:

1. *Do you think that using examples of class diagrams helps you to make a better design?*
   Type of the Answer: A Likert scale from -4 = not help to 4 = helps a lot without a (0) option.
   Analysis: Figure 8.4 shows the histogram of students' answer to this question. 86% of the students stated that using examples helps them in creating their design.

2. *Do you find it useful that the repository contains multiple examples for the same application/domain?*
   Type of the Answer: A Likert scale from 1 = not useful to 8 = very useful.
   Analysis: Figure 8.5 shows a graph of the student responses: 89% of the students stated that having multiple examples is very useful.

3. *How do you rate the relevance of class diagrams you found in the repository to your own design assignment?*
   Type of the Answer: A Likert scale from -4 = not relevant to 4 = very relevant without a (0) option.
   Analysis: Figure 8.6 shows that 94% of the students found relevant class diagrams in the repository related to their assignment.

4. *How do you rate the quality of class diagrams you found in the repository?*
   Type of the Answer: A Likert scale from -4 = poor quality to 4 = good quality without a (0) option.
   Analysis: Figure 8.7 shows that 86% of the students said that models in the repository have good quality. From their comments for this question, some students measured the quality based on the information within the assignments. Some students found some mistakes in some diagrams, but most of the models that they had explored have good quality.

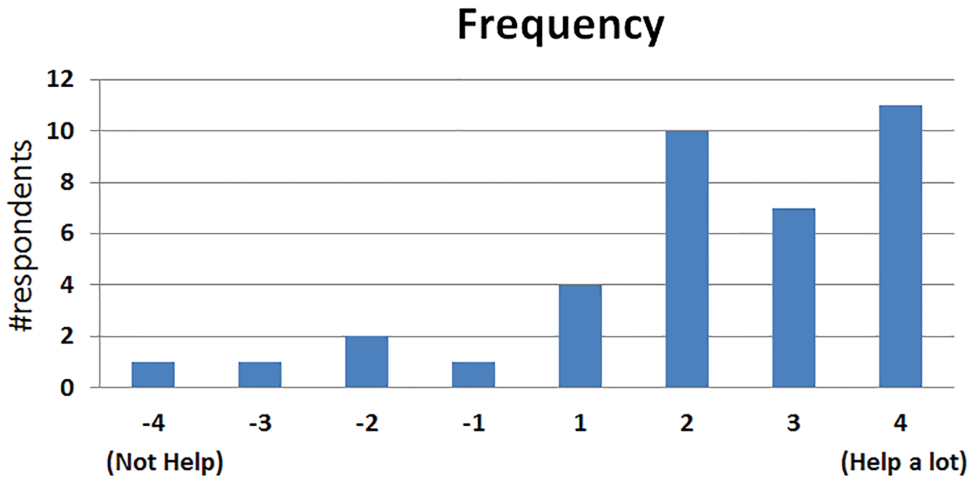

Figure 8.4: *Using examples of class diagrams helps to create better design*

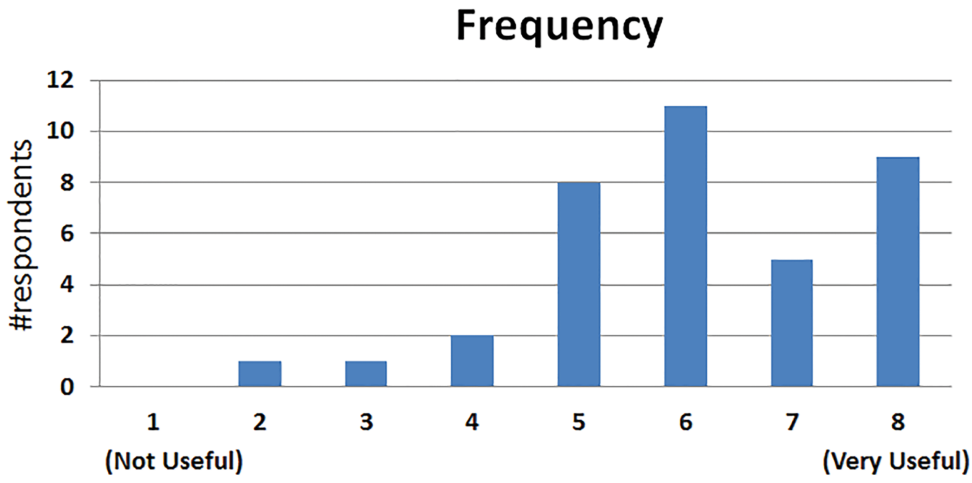

Figure 8.5: *Usefulness of having multiple examples for the same application domain in the repository*

5. *Do you think using the UML repository is more helpful to you than searching for examples of UML diagrams on the internet?*
   Type of the Answer: A Likert scale from -4 = not help to 4 = help a lot without a 0 option.
   Analysis: Figure 8.8 shows that 92% of the students said that using the model repository is more helpful than searching for models on the internet.

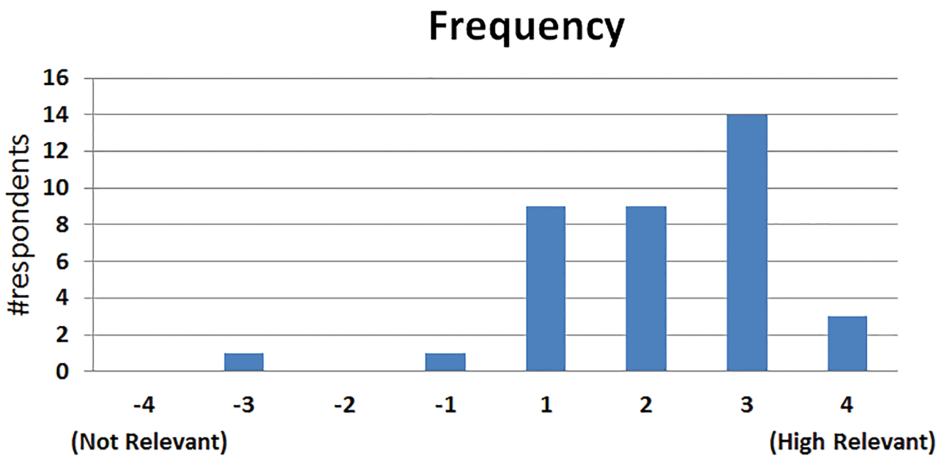

Figure 8.6: *Rate of the relevant class diagrams found in the repository to the design assignment*

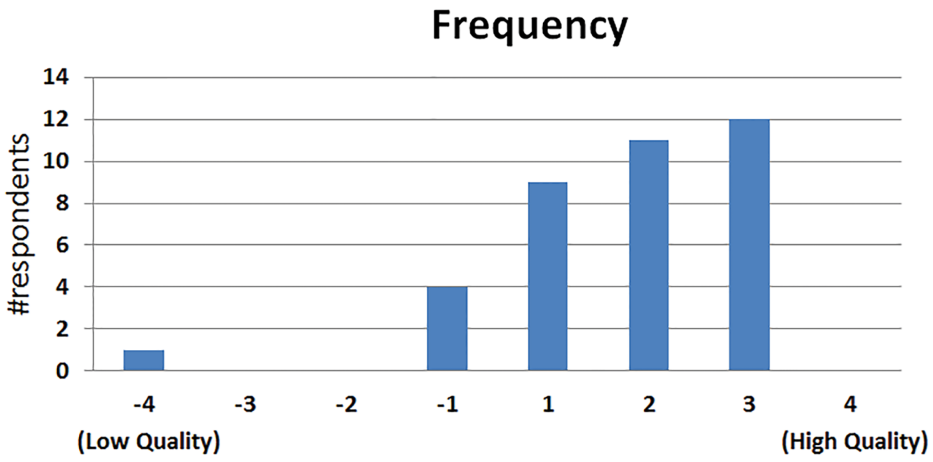

Figure 8.7: *Rate of the quality of class diagrams found in the repository*

6. *How do you rate the usefulness of searching based on class-, attribute- and operation-names to find relevant class diagrams?*
   Type of the Answer: A Likert scale from 1 = not useful to 8 = very useful.
   Analysis: Figure 8.9 shows that 95% of the students said that searching based on class-, attribute- and operation names is very useful.

7. *Did you find the available search techniques efficient to help you to find class diagrams?*
   Type of the Answer: A Likert scale from -4 = not efficient to 4 = very efficient without a (0) option.

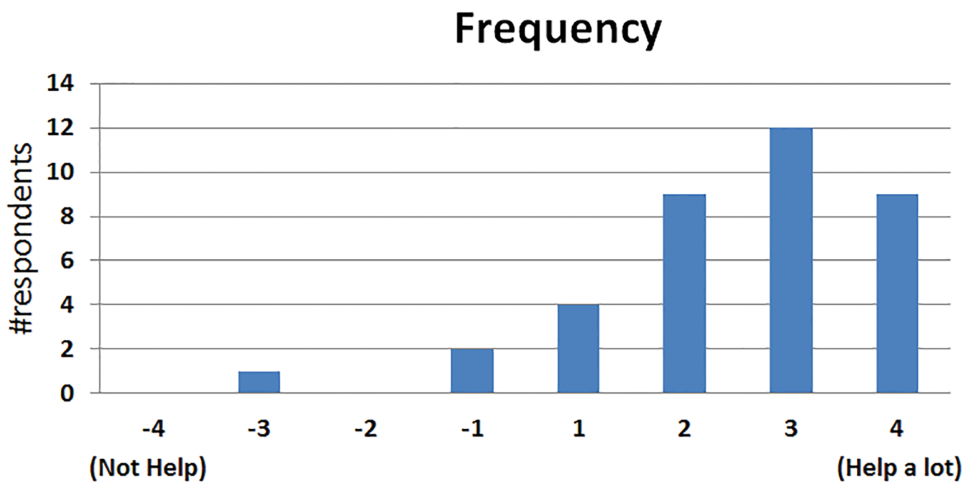

Figure 8.8: *Using UML Repository is more helpful than searching for examples on the internet*

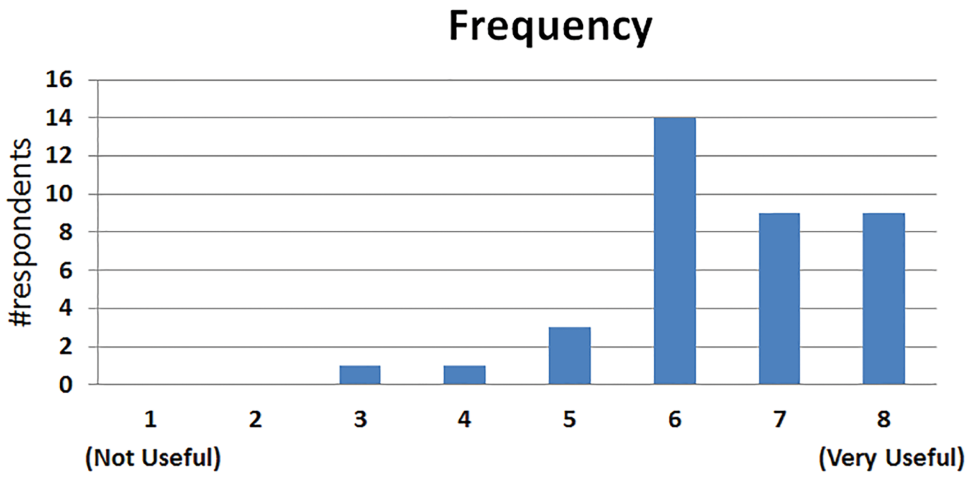

Figure 8.9: *Usefulness of searching based on class-, attribute- and operation names for finding relevant models*

Analysis: 89% of the students said that the available searching techniques in the repository are efficient.

8. *Are there another types of searching for class diagrams that you would like to see?*
   Type of the Answer: Open question.
   Analysis: 62% of the students stated that there is no need for other type of search-

ing mechanism. 22% of the students prefer to add a category-based searching mechanism. 14% prefer to add a search mechanism based on the description of class diagrams. 2% of the students asked to add a search mechanism based on relationship labels.

9. *Which aspects of the repository did you find easy to use in creating your own class diagram?*
   Type of the Answer: Open question.
   Analysis: 76% of the students stated that the search mechanism, in general, is the easiest aspect to use in the repository. 15% stated that multiple examples are the aspect that is the easiest to use. 9% stated that the search mechanism using class names is the easiest aspect to use.

10. *Which aspects of the repository did you find useful in creating your own class diagram?*
    Type of the Answer: Open question.
    Analysis: 49% of the students stated that searching is the most useful aspect in the repository. 44% of them considered the examples in the repository are the most useful aspect. 7% of students said that the easiness in using the repository is the most useful aspect.

11. *Which information from the class diagrams that you found in the repository did you use for making your own design?*
    Type of the Answer: Open question.
    Analysis: 64% of the students stated that relationships are the information they found and used for the design. 43% stated naming as the most useful information they found and used. 40% of the students stated that the syntax is the most useful information they found and used. Finally, 21% of the students said that roles and the structures of class diagrams are the most useful information they found and used.

12. *During the experiments, how much time did you spend using the repository?*
    Type of the Answer: Multiple choice, (<10 Min), (<10,20>), (<20,30>) and (>40).
    Analysis: The answer of this question is regarded to models improvement/updating, where both RG and CG use the repository. Figure 8.10 shows that 24% of the students said that they used the repository less than 10 minutes. 32% of the students said that they used the repository from 10 to 20 minutes. 30% of the students said that they used the repository from 20 to 30 minutes. 14% used the repository for more than 40 minutes.
    Figure 8.11 shows that 76% of students of RG spent less than 40 minutes using the repository. On the other hand, Figure 8.12 shows that 95% of students of CG spent less than 40, and 30% of the students used the repository for less than 10 minutes. Therefore, we conclude that students spent less time using examples when they use it only for updating their models in comparison with using it for creating and updating.

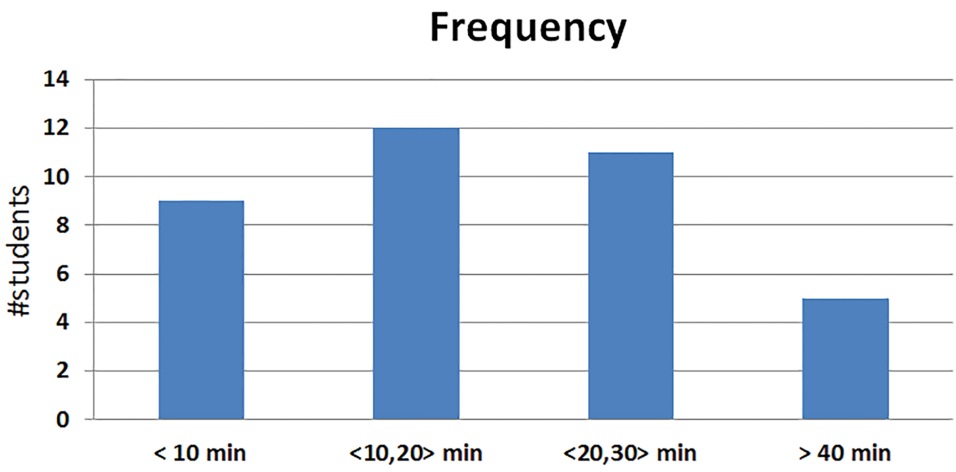

Figure 8.10: *Time spent by students using the repository*

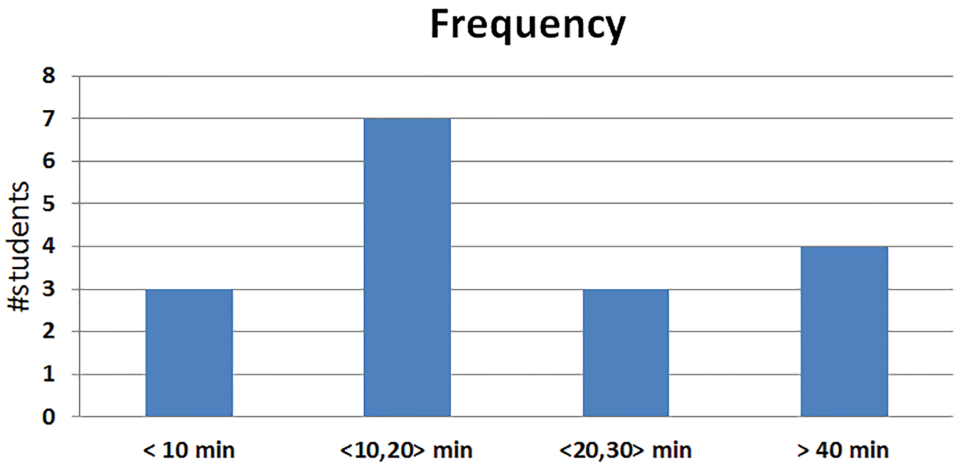

Figure 8.11: *Time spent by students using the repository*

8.5 Discussion

In this section, we discuss the results based on the aforementioned research questions, and we explain the results with more details considering the answers of the students to the questionnaire.

Table 8.1 shows that students from RG achieved higher evaluations than CG. This guides us to accept that constructing models with the aid of examples improve the

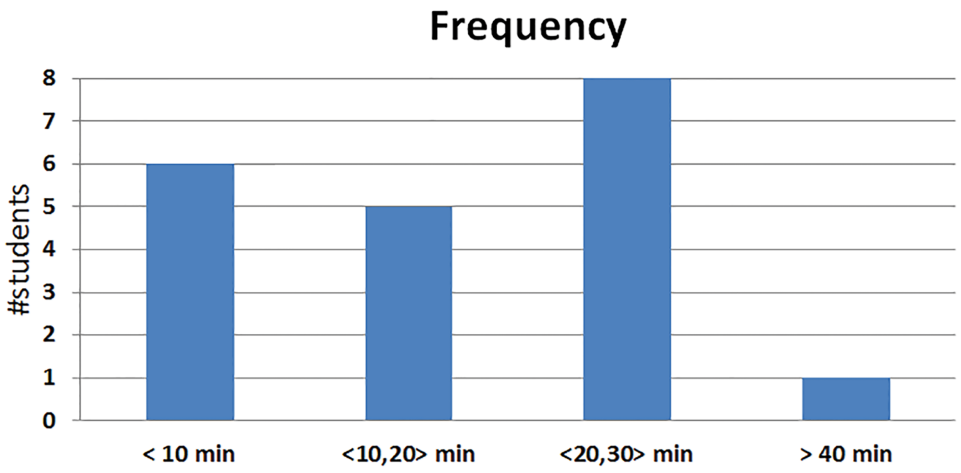

Figure 8.12: *Time spent by students using the repository*

quality of models created by novices compared with others who do not use examples. From Table 8.2, students of the CG achieved higher evaluation for their improved models using the repository compared with their old models. This also shows that using examples helps novices to improve qualities of their models.

Regarding the repository questionnaire, most of the students stated that using examples is really helpful. From their comments, we summarized that the repository of examples helps in the following manners:

- It gives the students a direction or starting point for creating their models.
  "Yes, even with some errors in the models (of course I didn't expect we could find the answer in the repository) it gave me a direction"

- It gives students many ideas from a different perspective as there is no optimal solution.
  "Yes, by seeing other examples and others models. And maybe see it from a new or different perspective which improves your diagram"

- It helps students to correct mistakes in their design.
  Especially when at a beginner level, using the repository helps with the design, ideas that I may have missed like level of abstraction, or solving a certain problem. Also it helps to spot mistakes and minor flaws"

- It helps them to complete their diagrams and reminds them regarding some details.
  It helps me to make a more complete diagram. It reminds me of some details"

- It helps in creating models more easily and more quickly.
  You can make your model more easily and quicker with the help of the repository"

A few students disagreed that using examples is helpful. We summarize their comments into two main categories:

- Constructing models with the aid of examples in general makes them lazy, but they agree that using examples improves their models.
  I think that using examples does not really help. It makes you lazy and does not force you to really learn about UML and Software Engineering"

One possible lesson could be to choose the assignment such that there are no models in the repository that are close to the same application domain nor close to the solution. If the repository grows, it may be more practical to restrict the diagrams that students get access to; hence to exclude some diagrams from being found.

- For some students, using examples confuses them. They prefer working without examples because they cannot distinguish whether the examples are good for their task or if they are of good quality or not.
  "It helped me a bit. I don't have a lot of experience with UML and this helped me to see and use it as a reference. It made me also doubt myself, because they were all so different"

Most of the students prefer to see multiple examples, because:

- They can pick out elements from different models, and make a new one.
  "Yes I did, in that way you have multiple perspectives. You can pick out the best elements from each model and make a personal one with this information and my basic knowledge (there is not one solution)"

- Compare models and align:
  "You can see what is usual for this application/domain to align your own model with a kind of standard"

- Multiple examples give multiple depictions of a certain subject, and show important functionalities which are not taken into account.
  "A variety of examples regarding a specific topic is always helpful, as the user has access to multiple depictions and implementations of a certain subject and he may trace functionalities that he might not have foreseen as necessary"

A few students do not like to see multiple examples, because it makes them more confused:
"It is useful to see multiple examples for the same domain but it also creates some doubt because you don't know if your modeling way is the right way"

For rating the relevance of class diagrams in the repository to the design assignment, students stated that they can find relevant examples to their assignments. Students

stated that it is enough to give a direction and to create their model.
"Helpful enough to give directions to use for our design"

One student commented that he dislikes using examples because it makes students lazy as they make use of examples to create their models.
"Everyone designs software according to his own view. Therefor I think you should try to model as much as possible by yourself. I also consider taking someone else's UML design as lazy"

For rating the quality of class diagrams in the repository, students stated that models that they found in the repository are helpful to give them directions for their design. Some examples are not correct and have some mistakes. Nevertheless, students learned how to skip or correct these mistakes. A few students stated that it is not easy to find good examples. We infer this from the keywords they used for searching, some keywords help to retrieve good relevant models.
"Most of the diagrams have a clear and understandable structure, which makes it easy for the user to extract any information that he wants"
"Some were of really high quality but there were also some models that contained some errors"
"A lot of diagrams made no sense, were incomplete or were not usable"

In response to the searching in the UML Repository is more helpful that searching on the internet, students stated that they prefer using the UML repository. Their preference is motivated as follows: Multiple examples are easily accessed, and the possibility to direct the searching mechanism to target class-, attribute- and operation-names gives advantages to the repository.
"Having the ability to search by keywords in the UML repository is as effective as using a search engine, plus it has the advantage that it has a variety of examples implemented in different ways and styles"

For the usefulness of searching based class-, attribute- and operation names, students stated that the search based on class, attribute and operation names is very useful. From their comments, they said it was easy to find relevant examples and specific information that they need. They stated that this kind of search for models could answer many questions in their mind, for examples: what should this class look like? How many attributes and operations should be present in the class? How could these relationships be used? Etc.
"It is essential. A very useful feature. I think it is easy to find relevant diagrams when searching for class names"

For easy and useful aspects that student find in the repository or they prefer to be available, most of the students stated that searching is the easiest and most useful aspect of the repository, and they do not prefer other kinds of searching mechanisms at the moment. Other students suggested searching by category or title of class diagrams. Also, some students suggested searching on the description of models. One student suggested searching by relationship name to find how people use it.
"Searching for relevant diagrams based on the classes in my diagram"
"I'd like to be able to search on the description of a model" "Of course we can come up with many different criteria to be added to the repository, but the combination of the class specifications with

the design metrics properties is more than enough in order to find a desirable class diagram"

For which information from class diagrams they found in the repository they use to make their own design, students used the repository to find the following:

- Syntax: *"what a diagram should graphically look like, and how to use class diagram notation"*
- Naming: *"Naming, specific names of classes, operations"*
- Relationships: *"Relations, I got some inspiration for making correction to my class diagram relation after I saw the examples in the repository"*
- Roles and structures: *"Relations, roles and structure"*, *"operations and how to connect entities"*

For the time students spent using the repository, 86% of the students use the repository for 20 30 minutes maximum, which could be considered as a short period compared to two hours experiment time. 76% of students of RG spent less than 40 minutes using the repository. On the other hand, 95% of students of CG spent less than 40 minutes. In the CG group, 30% of the students used the repository for less than 10 minutes. So we conclude that students spent less time using examples when they use it only for updating their models in comparison with using it for creating and updating.

8.6 Threats to Validity

This section discusses the threats to validity.

8.6.1 Internal Validity

We ensured that students are familiar with UML class diagrams, UML assessment, and the UML repository before the experiment. They had the experiment at the end of the software engineering course, and they had a trial two weeks before the experiment.

To avoid participants' expectations from biasing the results, we did not inform them about our experimental hypotheses nor what results we are aiming to know. Before conducting the experiment, the students did not know whether they belonged to RG or CG. We assigned students randomly to RG and CG. We gave all students the same amount of time to finish the task of creating and improving their models.

8.6.2 External Validity

There are two main concern: participants and materials used. Regarding participants, the goal of our research is to study how to learn creating software designs. For this students are an appropriate population. We cannot know if the same finding holds for professionals.

Regarding materials used, the task we gave participants is quite typical in size and complexity compare to the assignments found in UML textbooks. This was confirmed by several lecturers in this field. Also, the modeling task is in a domain that is familiar to students. We cannot know if our findings apply to larger systems.

Models in the repository are related to a wide range of different application domains and vary in size and complexity.

8.7 Conclusion and Future Work

We presented a controlled experiment that aimed at evaluating the effects of using examples for teaching students software design. The main result of our experiment is that model examples aid students in constructing and improving their models. Students appreciated this approach, and they also like the idea of using the repository for finding relevant examples which assist them during the construction of their design. Experts ranked the models produced by the students in the repository group higher than those of the control group for all quality attributes. Students, who were given access to the repository to improve their solution, increased the quality of their solution as assessed by experts for all quality attributes. We observe that using examples makes students more confident in their models: 27% of the repository group see no need to improve their models compared to 16% of the control group. Most of the students stated that using examples is helpful for them in order to create and improve their designs. Students used examples to understand the relationships between classes, naming of class diagram elements, and roles and structure of the class diagram. The model repository is a suitable environment for offering examples. Students preferred the repository over using internet search engines. The main reason of that preference is that it allows them to search for models based on model-contents e.g. class names, which is not possible via internet search.

The time spent for searching examples in the UML Repository is less when students use it only for updating and improving their models. Students suggested some features that could be added to the repository, for example having a model/system name and the possibility to search for models by category and their descriptions.

For future work, we want to study the changes made by the students to improve their models, and then study what they learned from examples. We are going to enrich the repository with some of the features suggested by students. In addition, we think about other useful features e.g. models comparison and plagiarism detection.