An Experimental Comparison of Use Case Models Understanding by Novice and High Knowledge Users

Balsam A. MUSTAFA*

School of Engineering, Computing, and Science
Swinburne University of Technology, Sarawak, Malaysia,

Abstract. Use cases were introduced into the Unified Modeling Language to capture the functional requirements in object oriented systems development. This work reports the results of comparing two controlled experiments conducted on use case models with different subjects, in which the effect of use case format on users understanding of systems requirements is assessed. Replication with subjects of different knowledge in use case technique allowed us to investigate whether subjects experience play any role in the comprehension of use case models. The results of the controlled experiments showed that for the comprehension tasks, which required only surface understanding of the Use Case model, the provision of diagrams along with the textual use case descriptions significantly improved comprehension performance of both novice and high knowledge users. However, diagrams had no effect on users performance in the deep understanding tasks. Moreover, there was no evidence that prior experience with use cases has influenced subject’s performance in surface and deep understanding tasks in familiar and unfamiliar application domains.

Keywords. Use Case model, UML, Prior knowledge, Controlled experiment, User comprehension

Introduction

Use cases are a key technique in the Unified Modeling Language (UML), the standard modeling language of the object oriented approach. Use cases are used during requirements gathering to model the functional requirements of the software system from the perspective of its users. Use cases are introduced by Jacobson [20] to describe the interactions between the system and its users, and to aid in communication between system analysts and users. Larman [26] argued that use cases are functional requirements that indicate what the system will do rather than the internal working of the system, its components, or design. He also indicated that the wide adoption of use cases may be due to its simplicity and that they are easy to understand intuitively. The choice of representational format of a Use Case model varies in research [9, 40, 8]. Although use cases are text based, some researchers use the UML diagram with structured or unstructured scenarios. The diagram is introduced in UML to visualize the use cases, the relationships between them and with actors who interact with the system. An example use case diagram is shown in Figure 1 which is one of the models used in this study. A simple use case diagram shows only the use cases that represent the main

*Corresponding author. Tel +6 082 416353 (7837); fax +6 082 260813
Email address: bmustafa@swinburne.edu.my
functionality of the system, users who request these functions and initiate use cases. While the diagram could be more complicated when it shows advanced relationships (<<include>>, <<extend>>, <<generalization>>) between use cases and actors. However, there is a lack of research on the role of UML diagrams in supporting user understanding when accompanying the text in a use case model.

Figure 1  HSS case detailed diagram
The motivation behind this work is that researchers in the field of information systems have realized the need to pay attention to the human dimensions in systems development. In particular, when identifying information systems requirements, systems analysts and users must communicate effectively and share a common understanding of the work problems and the required solutions [49, 18, 48]. Bostrom and Thomas [3] argue that “the use of effective communication patterns will enhance the ability of users to specify requirements and developers to elicit and evaluate requirements”. Researchers contend that models used in communication should be understood by people involved in the development because it is necessary that they reach a shared understanding of the system requirements. In this social process, human factors have significant implications because analysts, users, customers, each bring their own cognitive abilities, personality characteristics, and prior knowledge about the problem domain [43]. The effect of prior knowledge for example has been of interest in a number of related contexts such as program comprehension [4], systems analysis and design methodologies [1] and conceptual schema development [42]. Fewer studies examined the effect of prior knowledge in the modeling technique directly. Khatri et al. [23] used entity-relationship (ER) and extended entity-relationship (EER) schemas in examining the effects of prior knowledge in both modeling method and the application domain. Ricca et al. [38] investigated the role of experience with the use of UML stereotypes on comprehension level.

In previous research, we investigated the effect of use case format on novice users [33]. An experiment is conducted and the results suggested that, for novice users, supporting the text description of a use case model with diagrams seems to facilitate better comprehension of the model semantics, and that the provision of simple diagrams along with textual use case descriptions might be sufficient to support the negotiation and communication of requirements between novice end-users and system analysts. To gain a comprehensive view, the experiment was replicated [32] in which high knowledge users was the target population. Both experiments were done in university environment, students were used as end users and professors as high knowledge users but not real parishioners have been involved. This study presents the first attempt to compare the performance of novice and high knowledge users of use case model and try to find out if there is any difference in understanding use case models that can be attributed to the effect of user’s experience in both the modeling method (use cases) and the application domain under study.

This paper is structured as follows. In Section 2 we review the empirical works related to this study. Sections 3, and 4 present the cognitive theories that suggest why different formats of use case model might affect user understanding. Two experiments to test the suggested hypotheses are described in Sections 5 and 6. The effect of experience is discussed in Section 7, and Section 8 discusses the threats to the validity of the study. Finally, conclusions and future research are presented in Section 9.

2. Related Work in UML

There are many kinds of research in the context of UML that have investigated the comprehension of UML diagrams. Genero et al. [16] conducted a family of experiments to evaluate the relation between a set of measures for the structural properties of UML class diagrams and their understandability. They also provided prediction models for such quality characteristics. Their results indicted that measures related to associations and generalization affect the understanding of UML class diagram. Xie et al. [52] investigated the effect of their proposed synchronization
adorned UML sequence diagram notation on students understanding of concurrent execution and concurrency. The results showed statistically significant advantage of the notation. Ricca et al. [38] presented a family of three experiments to analyze and assess the use of UML stereotypes to evaluate their usefulness in web application comprehension for different categories of user. They found that stereotypes significantly improved the performance of low experience subjects. Manso et al. [27] carried out a family of five experiments to validate the effect of the size and structure complexity of UML class diagram on the cognitive complexity and consequently on the comprehensibility of the diagrams. The results revealed strong relationship between the structure complexity and cognitive complexity which affects the comprehensibility of the diagrams. The experiments performed by Tilley and Huang [47], highlighted the limitations of UML to support program understanding. They found that UML does not provide sufficient support to represent domain knowledge but stereotypes can improve.

Cruz-Lemus et al. [11] study of state chart diagrams showed that composite states improve the understandability of state charts provided that subjects had prior experience in using them. They also investigated in a family of three experiments [12] the impact of structural complexity on the understandability of UML state chart diagrams. They found that the size and control flow complexity of the state chart in terms of features such as the number of states, events, guards and state transitions and the sequence of actions within a state affect the understandability of the diagrams.

3. Learning with Text and Illustration

Many studies in educational psychology [7, 35, 41, 17, 13] focused on the advantage of using pictures to supplement the text in teaching complex materials and how different presentations interact. Studies [51] on the benefits of diagrams in learning contend that graphs can support learning and enable the learner to acquire meaningful information and serve as an important instructional function. The Conjoint Retention hypothesis [25] explains that when students view material that contains both diagram/map in conjunction with the related text, the student recall the text information better than when they are provided with text only. Robinson and his colleges [39] indicated that when students are provided with adjunct displays that appear outside the text such as pictures, geographic maps, concept maps, graphs, diagrams, and so forth, they perform better on tests to measure knowledge of text information referenced in the displays. Mayer’s research [28] in multimedia learning showed that learners lacking domain relevant background knowledge benefit more from illustrations in teaching materials. The use case model under consideration comprises a diagram and text. The diagram itself consists of both graphic symbols and text. Thus, the model can be considered as a multimedia presentation [29].

Mayer [28, 30] pointed out that multimedia learning occurs when students use information presented in two or more formats, such as visually presented pictures and verbally presented narration to construct knowledge. He explains that the verbal and pictorial information are processed in different channels in working memory. Construction of different mental models in different channels may reduce the cognitive load on working memory, freeing up more resources to process new information so active learning could take place. Active learning results from selecting, organizing relevant information from different channels into coherent representations, and integrating them with existing knowledge in long term memory.

The above theories suggest that different forms of visualizations will lead to different understanding of the learners and also to different patterns of performance, when
individuals have to solve tasks after learning on the basis of their previously gained knowledge. We proposed the following research question:

**Research Question 1:** Does the format of use case model influence the understanding and the patterns of performance, when individuals have to solve tasks on the basis of their previously acquired knowledge? and which use case format, text only or text accompanied with diagram better supports user understanding of the domain requirements?

On the basis of theories discussed above we hypothesize that using a diagram to accompany the text descriptions in a use case model may improve viewers understanding of what the suggested software would do, specifically as it focuses their attention on the functions that the system would provide and make the relationships among model elements more obvious. Two representations of the domain may help the viewer to use one to aid his understanding of the other and integrate both sources to better understand the domain. We formulated the following hypothesis:

**Hypothesis 1:** Novices and high knowledge individuals who receive a use case model consisting of both diagrams and text may develop a higher level of understanding of the system requirements faster than individuals who receive a use case model consisting only of text.

### 4. Problem Solving and Cognitive Load Theory (CLT)

Cognitive Load Theory initially developed by John Sweller in the early 1980s to provide instructional designs that facilitate understanding, learning and problem solving [45, 46, 36, 50]. The theory is based on Miller’s assumption [31] that human working memory has a limited storage and processing of novel information. In contrast an unlimited long term memory stores huge amount of information in the form of schemas. A schema is a cognitive construct that can be used to overcome the limitations of working memory. Many elements that are organized in schema are treated as one element which requires less processing in working memory. When schemas used often it then automatically processed rather than consciously reducing the load on working memory.

The theory assumes two sources of cognitive load imposed on working memory, “intrinsic cognitive load” which can not be altered by instructional manipulation because it depends on the extent of interactivity between the elements of the material being learned and the expertise of the learner. The “extraneous cognitive load” is generated by the manner in which information is presented to learners and is under the control of instructional designers. The theory focuses on how instruction should be structured to reduce unnecessary extraneous working memory load. If several items in a material can be learned individually without reference to each other then the material is low in element interactivity [37], and it imposes low cognitive load when processing it in working memory. In contrast, high-element interactivity material imposes high cognitive load on working memory. With expertise, many interacting elements can be incorporated in a schema that acts as a single element and imposes a minimal working memory load.

Conceptual models are very complex material which consists of many interacting elements that must be processed simultaneously in working memory in order to be understood [50]. It contains words and pictures with different kinds of relationships to represent the domain of interest. It is necessary for the viewer to be able to communicate with the model developer to process all elements at the same time which
may become difficult for users with little experience in these models. We proposed the following research question:

**Research Question 2:** How does the degree of detail in a UML diagram that accompanies text in a use case model affect end user comprehension of the domain requirements?

A detailed diagram of a use case model contains more elements (use cases, actors, relationships) than a simple diagram. When exposing novices with little experience in modeling to such material at once, this may imposes a high, intrinsic cognitive load because many elements must be processed in working memory simultaneously. To help them understand, they must be able to develop cognitive schemata that incorporate the interacting elements. Because novices lack these schemata, this would hinder their understanding. We formulated the following hypotheses:

- **Hypothesis 2:** Novice users who receive a text with simple diagram use case model may develop a higher level of understanding of the system requirements faster than novice users who receive a text with detailed diagram use case model (Table 1).

- **Hypothesis 3:** High knowledge users who receive a text with detailed diagram use case model may develop a higher level of understanding of the system requirements faster than high knowledge users who receive a text with simple diagram use case model (Table 1).

**Table 1** Sub hypotheses (Predictions)

<table>
<thead>
<tr>
<th>Hypothesis 1</th>
<th>Novice users</th>
<th>High Knowledge users</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1A</td>
<td>Participants using text with diagram model will perform the comprehension task more accurately than those using text with detailed diagram.</td>
<td></td>
</tr>
<tr>
<td>H1B</td>
<td>Participants using text with diagram model will perform the comprehension task faster than those using text with detailed diagram.</td>
<td></td>
</tr>
<tr>
<td>H1C</td>
<td>Participants using text with diagram model will perform the verification task more accurately than those using text with detailed diagram.</td>
<td></td>
</tr>
<tr>
<td>H1D</td>
<td>Participants using text with diagram model will perform the verification task faster than those using text with detailed diagram.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hypothesis 2</th>
<th>Novice users</th>
<th>High Knowledge users</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2A</td>
<td>Participants using text with simple diagram model will perform the comprehension task more accurately than those using text with detailed diagram.</td>
<td></td>
</tr>
<tr>
<td>H2B</td>
<td>Participants using text with simple diagram model will perform the comprehension task faster than those using text with detailed diagram.</td>
<td></td>
</tr>
<tr>
<td>H2C</td>
<td>Participants using text with simple diagram model will perform the verification task more accurately than those using text with detailed diagram.</td>
<td></td>
</tr>
<tr>
<td>H2D</td>
<td>Participants using text with simple diagram model will perform the verification task faster than those using text with detailed diagram.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hypothesis 3</th>
<th>Novice users</th>
<th>High Knowledge users</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3A</td>
<td>Participants using text with detailed diagram model will perform the comprehension task more accurately than those using text with simple diagram.</td>
<td></td>
</tr>
<tr>
<td>H3B</td>
<td>Participants using text with detailed diagram model will perform the comprehension task faster than those using text with simple diagram.</td>
<td></td>
</tr>
<tr>
<td>H3C</td>
<td>Participants using text with detailed diagram model will perform the verification task more accurately than those using text with simple diagram.</td>
<td></td>
</tr>
<tr>
<td>H3D</td>
<td>Participants using text with detailed diagram model will perform the verification task faster than those using text with simple diagram.</td>
<td></td>
</tr>
</tbody>
</table>
5. Experiment 1 (Novices)

5.1. Design
The experiment had one independent variable with three levels which are the three presentations of use case model (text only, text with simple diagram, text with detailed diagram) to explore whether it could improve model comprehension. All models were made equivalent where the information contained in one model was also inferable from the other. The evaluation was based on comparing the three different formats of use case model. Two cases were used in this experiment from two separate sources to increase the generalization of the results. The dependent variable was the level of user understanding of the use case models. The experiment aimed to measure two levels of understanding, surface and deep. The use case model is understandable if it allows users to recognize problem domain information and extend their understanding in problem solving. The rationale of this investigation is that, users and analysts communicate to improve their understanding of the problem domain. Deep understanding of the domain is crucial to ensure the correct requirements of the proposed solution are clearly identified. Thus, we used two tests of understanding, a comprehension test consists of 12 multiple choice questions (example in Appendix A), and a verification test to find out 8 inconsistencies in the functional requirements of the system depicted in the model. The dependent variables were the (accuracy), the correct answers to the questions of the comprehension and verification tests and Time taken to answer the questions. The accuracy scores were used to determine the measure of effectiveness, while the time is used to determine the measure of efficiency or the degree of effort exerted to understand the model. Two moderator effects have been considered, the effect of prior knowledge in use case modeling and the prior knowledge in the two case domains used in this study. This was to ensure no significant differences in those factors between subjects that might affect the results of the experiment. Subjects were randomly assigned to three groups of 28 persons each corresponding to three experimental conditions (text only, text with simple diagram, text with detailed diagram).

5.2. Participants
The participants were 84 undergraduate students (49 female and 35 male), from the Faculty of Computer Science and Information Systems at the University of Technology in Malaysia, who had completed a course of software engineering of which use cases were taught. They lacked substantial knowledge in use case modeling as assessed by a pretest questionnaire.

5.3. Material and Procedure
The paper-and-pencil materials consisted of a pre-test questionnaire to rate the participants’ experience, familiarity, confidence and competence with use case models (text/diagram). Other questions were to rate individual’s experience with the two domain cases used in the experiment, Simulation of an Automated Teller Machine (ATM) and Home Security System (HSS). Data collected from these questions were used to create the variable scales (Prior knowledge with modeling method (KMETHOD), and prior knowledge with problem domain (KDOM) which has been used as covariates in the statistical analysis method (MANCOVA). The computerized materials (Appendix B) were two tests; a comprehension test consists of 12 questions for each case (ATM&HSS) to measure the viewer surface understanding of the model. To answer these questions, subjects received use case models that depict the correct
functionality of the systems. After completing the comprehension test, the correct models have been taken away and subjects given another use case models of the same cases that contain different faults. The verification test was to measure deeper understanding of the models. This test asked the subjects to identify any inconsistencies between a given model (that was seeded with such faults) and their knowledge of the system requirements gained from the comprehension task, and to explain why they think it is incorrect. Time to complete the tests was collected by the same computerized package. After completing the pretest questionnaire, each subject completed two cases in computerized tests. For each case the subjects completed two tasks in the following order: comprehension task, verification task. The order in which subjects started the two case domains was counter-balanced to control for any learning between the two domains. The sequence of tasks was applied in a particular order to ensure internal validity of the experiment. Finally each subject completed a post-test after the verification task of the second case to measure the perceived ease of interpretation associated with the method.

6. Experiment 2 (High Knowledge Users)

This experiment is a replication of experiment1. In experiment 2 we consider high knowledge users as our target population to gain a comprehensive view of how stakeholders of different degrees of knowledge in the modeling technique may understand the use case models for the purpose of identifying system requirements.

By comparing the performance of both kinds of users on the same tasks in the two experiments we may understand the effect of experience in dealing with use cases on the understanding of system requirements depicted in the models.

6.1. Method

The experimental design, material, and procedure were similar to experiment 1.

6.2. Participants

The participants were 30 members of the Faculty of Computer Science at the University of Technology in Malaysia (17 females and 13 males). Their experience with Use case modeling is obtained from teaching the subject for Software Engineering students in the faculty, workshops, and training courses, and it ranges between 2-10 years. Subjects were randomly assigned to three groups of 10 persons each corresponding to the three experimental conditions (text only, text with simple diagram, text with detailed diagram).

6.3. Analyses and Results of Two Experiments

In this study statistical significant of differences has been tested using multivariate analysis of covariance (MANCOVA). The three treatments (text, text with simple diagram, text with detailed diagram) represented three levels of the independent variable. The dependent variables were the comprehension, verification test’s accuracy scores, and time taken to complete each test. Data were screened for extreme or
missing values, and statistical assumptions for MANCOVA including the normality distribution of the dependent variables were considered. Alpha was set at .05 when evaluating tests for statistical significance. The MANCOVA analysis between the treatment groups (Table 2) is made after considering the effect of the intervening variables: (1) Knowledge of use case model (KMETHOD) and (2) domain knowledge (KDOM) for the two cases ATM and HSS, respectively. The measure of scale variables (KMETHOD & KDOM) was not to determine the level of subject’s prior knowledge, rather it was to ensure no significant differences between the three groups in those variables which may affect the results. However, those variables were excluded from the final MANCOVA analysis as it is found no significant effect of them on the dependent variables. Table 2 shows significant multivariate effect of the treatments used as independent variables on the dependent variables. It shows high values of power and Eta square ($\eta^2 > .171$ in experiment 1) and ($\eta^2 > .34$ in experiment 2) which indicate large treatment effect sizes according to Cohen’s definition of effect size [9].

In analyzing the data of both experiments, the MANCOVA was followed by pairwise comparisons (Table 3) to test for sub hypotheses (Table 1) and to see which dependent variable affects the MANCOVA results.

Hypothesis (1): The comprehension task scores (accuracy) and time taken to complete the task confirmed sub hypotheses H1A and H1B (Table 1). The groups of novices and high knowledge subjects who received a text with diagrams (simple or detailed) use case models scored significantly higher correct answers and completed the test in less time than did the text only group, although for the high knowledge subjects (experiment 2) the difference in the comprehension test scores of (HSS) case was in slightly lesser statistical degree (Sig. = 0.055).

For the verification task which requires deep understanding, the results in Table 3 indicated no significant differences between the text with diagram (simple or detailed) and text only groups. Thus, there was insufficient support for sub hypotheses H1C and H1D. The same results obtained in the two experiments might indicate higher complexity of the verification task, which demands finding any inconsistencies in the use case models based on the participant’s understanding of the original models received with the comprehension test that show the correct functionality of the system. A possible reason for the failure to find significant differences between the three groups of people is that Complex tasks usually require more background knowledge in both the domain under study and the method used to model the domain, to be accomplished. A second reason is that the experiments might have suffered from a “floor effect”, which may have biased the results. This means that the task of finding inconsistencies was just too complex for the participants. In experiment 2 (high knowledge subjects), the groups that received text with diagrams use case models completed the verification task in significantly more time than that taken by the text only group. A plausible reason is that the inconsistencies in the models with diagrams were distributed both in the text and the diagram; therefore, finding them may take longer time than when they are in the text only. This may also indicate that taking more time in completing a task does not mean doing it with higher accuracy.
Hypothesis (2): According to the cognitive load theory, we assumed that the schemata that the experts already have obtained would help them make use of the detailed diagram and incorporate the more interactive elements in the detailed diagram. That is, they would be able to use their knowledge to overcome ambiguities and inaccuracies in the diagrams, be more accurate and take less time in answering comprehension questions and finding inconsistencies. This is based on the assumption that, as for each segment of time they may exert ‘less’ mental effort to process the information because the information that they have is clearer and more complete. The results from Table 3 did not support hypothesis 2 in experiment 1 (novices), may be because the manipulation of the difference between the two types of diagrams in the models was not strong enough.

Hypothesis (3) in experiment 2 was not supported. An analysis of the effect size (d) between the two treatments (simple vs. detailed diagram) indicated a large effect size for all dependent variables. Thus, the failure to find significant results may be due to lack of power, which in turn is caused by group sizes being too small. A post-hoc power analysis to investigate whether the non-significant results likely reflect true non-significance rather than Type 2 error indeed showed little power. Future research may reveal the effect of the level of detail in the diagram on high knowledge users understanding by using larger sample sizes.
7. The Effect of Experience

Current research in cognitive psychology has shown that human perception and learning of new concepts depends on a person’s knowledge about the world and past experiences. In the study of Sutherland et al. [44] on the impact of prior event information on memory, they claim that what we know and believe influence how we remember information. They indicated that individuals with high knowledge in a particular domain perform significantly better on memory tasks related to that domain than those with less knowledge. Hinds and Patterson [19] explain that “experts organize and access knowledge differently than those with less expertise, expertise comes with experience and with the acquisition of tacit knowledge in the domain”. Kalyuga et al. [21] who studied the effects of level of experience on learning from text and diagram indicated that learning from different kinds of presentation depends on the level of expertise of the learner and that differences in expertise provide the largest and most reliable differences in performance among individuals. In studies that compared experts and novices performance [34], experts were considered to be those with both knowledge and experience in applying this knowledge to a variety of problems within the domain. In contrast, beginners are those with some exposure to, but little experience and less performance in tasks within the domain.

Our final procedure was to find out the differences in performance between novices and more knowledgeable users of use case model as presented in the two experiments. Therefore, an ANOVA analysis was conducted followed by pair wise comparisons to test if there were considerable difference in performance across the same groups in the two experiments, which may be attributed to the effect of prior knowledge in using the models. The analysis was done for each of the two cases (ATM & HSS) separately.
Each pairwise comparison was tested at the .05 significance level. The ANOVA results indicated no significant differences between the means of dependent variables in both experiments which led us to conclude that there is no evidence for an experience effect and future research may reveal more information in this issue. Moreover, when testing for the interaction effect between treatment and experience, the ANOVA analysis indicated an instance of non significant “expertise reversal effect” particularly across the groups which received a text with detailed diagram model in both experiments as shown in Figure 2. These results may suggest that novices and high knowledge users react in different ways to detailed diagrams. Expertise reversal effect is explained in educational psychology as that the relative advantage of an instructional procedure designed for novices can turn into a disadvantage to learners at higher level of expertise [50, 22]. The lack of statistical significance in performance between novices and knowledgeable users may be caused by several factors. There may be an insufficient difference in expertise on the use case modeling between novices and experts, or the amount of learning during the learning period may have been minimal - if learning is minimal then effects also will be minimal.

8. Validity Threats

The two experiments have been able to differentiate between the presentations with text only and that with text and diagrams in two cases at least in the comprehension task. Differences were consistently observed in the two cases used in the experiments which maximize the internal validity of the results.

Construct validity primarily addresses the instrument used in the experiment (comprehension and verification) whilst the other threats address concerns with the experiment.

Construct Validity. This threat considers whether the design of the instrument is flawed or that the outcomes were biased by the design of the instrument. In this study we concentrate on measuring the “understanding” of individuals viewing the models. Therefore, our concern was: do the operationalized measurements of “comprehension” and “verification” used in the study accurately reflect the underlying construct of “understanding”. An important threat to the construct validity is the lack of underlying theoretical basis for the empirical research [14]. Our strong emphasis on existing theories that support our hypotheses counters this threat partially. However, since “understanding” is a cognitive process, it is difficult to directly observe it, and tests to measure participants’ performance were conducted to assess the level of understanding cognitively developed by each participant. The insignificant finding of the effect of experience explained in Section 7 might be a consequence of the design of the tests because there may have been sufficient learning from the models but the tests may not have picked up what was learned to a sufficient degree. Furthermore, our moderator variables of prior knowledge of domain and the modeling method used in the MANCOVA analysis are quite subjective and have not been formally validated in information systems research as “valid” measures of experience. Thus, the difference in prior knowledge between the participants in the three groups in each experiment may have not been measured precisely.
Statistical Validity. The most common threats to statistical conclusion validity include violations in the assumptions underlying statistical procedures, low statistical power, and low effect size. The randomization of participants in addition to the design of the experiments which was balanced, independent with equal group sizes, contributed in reducing the impact of violations in the assumptions. MANOVA assumptions were considered before the analysis started and the normal distribution of each dependent variable in each group has been verified using graphical and non-graphical tests. The tests conducted are, homogenous covariance matrices (Box’s M), (Levene’s test) of homogeneity of Covariance and (Bartlett-Box) test of homogeneity of variance associated with the MANOVA method which are sensitive to the normal distribution assumption.

External Validity. The restriction imposed on the study was to limit the complexity of the cases. The cases were designed to be small enough to be understood in a time reasonable for the study. The results, therefore, may not be extended to real world problems that are extremely complex. The fact that the cases are not as complex

Figure 2 The Interaction Effect of Treatment and Experience on Dependent Variables
as those in the real world, does not discount the differences observed, but might limit the extension of the results to more complex situations. Although this is a limitation of the study, it was necessary to conduct the experiments. Since the subjects in experiment 2 are teaching staff of faculty the results are potentially not generalizable to practitioners. Students as subjects could be a threat for the external validity of the experiment. The subjects in the first experiment were all undergraduate students with little experience in use case modeling. According to Kitchenham et al. [24], students can be used as subjects.

9. Conclusion and Future Research

This paper presented findings of an experimental study to measure the effectiveness of accompanying the textual descriptions of use case models with diagrams when communicating functional requirements with users of different levels of knowledge with this modeling method. Since the level of understanding developed from viewing use case models is the focus of this study, we draw on the cognitive sciences to provide theoretical grounding for this research. We explain the results in the context of Multimedia theory and Cognitive load theory to fully understand the implications of this research. The focus of theory may improve our ability to design and refine effective modeling techniques. The results of both experiments (novices and high knowledge users) confirm the hypotheses generated based on the cognitive theory of multimedia learning and the conjoint effect [25].

This research has both implication for research and practice. For practice, our results of the comprehension test confirm those from earlier studies in information systems area [5, 6] and suggest that it would be useful for modelers to ensure that they use multimedia principles when preparing use cases in the early stages of requirements elicitation, by accompanying a text with a diagram of the domain. Second, our results suggest that modelers should not rely too much on users’ feedback about how easy it was for them to understand the domain represented in a use case model. Other studies by Burton-Jones and Meso [5, 6], found that the correlation between users’ perceptions of how easy it was to understand the domain and their actual understanding of the domain was low. We obtained similar results in this study. For example, although the analysis of perceived ease-of-use showed no significant difference between the three treatment groups in both experiments (novices and high knowledge users), the text with diagrams groups took longer time than the text only group in the deep understanding task, which may indicate larger mental effort. This highlights the difficulty in measuring “actual mental effort” that determines the level of expertise of participants. Also, lack of the effect of experience in the final analysis indicates the need to assess participants’ expertise before conducting the empirical tests. Van Merrienboer and Sweller [50] argued that an appropriate assessment of expertise should at least include measures of mental effort and performance because that may reflect the quality of cognitive schemata, and most performance measures are not particularly designed with this end in mind. Future research is needed to develop valid and reliable measures of effort exerted in understanding these models and to inform practitioners about the most effective way to develop both accurate and easy to understand models. This comparative information would aid practitioners in choosing which technique to use in requirements elicitation based on their performance.

The results with regards to the effect of having simple versus detailed diagrams did not support our hypotheses derived from cognitive load theory in both experiments. However, in experiment 2, the large effect sizes calculated for all dependent variables associated with no significance may be due to the tests having little power, which in turn was caused by sample sizes being too small. A post-hoc power analysis indeed showed little power. Thus, future research should verify these results by recruiting
larger sample sizes and increase the difference in degrees of detail between the diagrams.

(Appendix A) Example of the multiple choice questions of the ATM case
(Comprehension test)

Which of the following describes what takes place when the customer wants to withdraw cash and there is not enough money in the customer’s account
(1) The transaction fails and the ATM asks the customer to choose another account with sufficient funds
(2) The transaction fails and the ATM returns card to customer
(3) The ATM informs customer of insufficient fund and aborts transaction
(4) The ATM informs customer and asks him/her to enter a different amount of money

When the customer finishes a withdraw transaction from savings account and select not to perform another transaction
(1) The system dispenses cash and prints a receipt
(2) The system dispenses cash, prints a receipt, and prompts the user to press ‘cancel’ to return his/her card
(3) The system ejects the card, dispenses cash, and prints receipt
(4) Not clear in the model

Example of the multiple choice questions of the HSS case (Comprehension test)

Which of the following sensors will be triggered by an intruder when the security system is armed in away mode?
(1) Access point monitor, smoke detector, and motion detector
(2) Access point monitor and motion detector
(3) Motion detector and smoke detector
(4) Access point and smoke detector

The home security system is armed at either owner at home or away mode. Home owner will be notified of
(1) Intruder detection via local /mobile intruder alarm
(2) Intruder and smoke detection via local, mobile and remote intruder/smoke alarm
(3) Smoke and intruder detection via mobile intruder and smoke alarm
(4) Smoke and intruder detection via local /mobile smoke or intruder alarm
(Appendix B) Computerized test interface

References


