

User Preference and Performance with UML Interaction Diagrams

Jennifer Swan
University of
Hertfordshire

Maria Kutar
University of
Salford

Trevor Barker
University of
Hertfordshire

Carol Britton
University of
Hertfordshire

j.swan@herts.ac.uk m.kutar@salford.ac.uk t.l.barker@herts.ac.uk c.britton@herts.ac.uk

Abstract

During the requirements process it is of key importance that all representations used are clearly understood by those who must use them. Therefore it is essential to ensure that those representations are presented as effectively as possible. User preference is one area that may influence the effectiveness of the representations presented. This paper describes a study that was carried out to explore the relationship between user preference for UML sequence and collaboration diagrams. Results show that participants who preferred sequence diagrams showed improved performance when using sequence diagrams. This was true for preferences expressed both before and after the task. However, participants who did not prefer sequence diagrams showed an overall improved performance for both types of diagram over the group that preferred sequence diagrams.

1. Introduction

The selection of appropriate tools for use in systems development can influence the success of the development process. Diagrams are widely used as a tool to aid software development. The choice of diagram for particular projects often reflects the experience or preferences of the development team more than objective consideration of possible alternatives [1]. Representations are used throughout the requirements elicitation and validation process and can prompt users to contribute information about the problem and the intended system as well as check that a developer has understood the specified requirements. It is essential that all those involved, including users who may be untrained in the use of languages used in software development, have access to a representation they can readily understand.

The Unified Modelling Language (UML) [2] provides two diagrams for representation of scenarios known collectively as interaction diagrams. Interaction diagrams, which are isomorphic, consist of sequence and collaboration diagrams. Sequence diagrams predominantly display time and ordered interactions whereas collaboration diagrams predominantly display activities and links between objects.

One factor that may have a strong influence on the success of the use of diagrams during the requirements process is the users' preference for a particular type of technique. Intuitively, it would be expected that any tools and techniques users prefer will help them to perform their job better. As Petre [3] points out "The importance of sheer likeability should not be underestimated; it can be a compelling motivator." This intuition is, generally, supported by research, which has shown that overall, if users prefer one way of solving a problem to another, they will perform better with the technique that they prefer [4].

The purpose of this paper is to report on the relationship between user preference and performance with diagrams used in software development. The main research question underlying this area of the study is to try to discover whether a relationship exists between a participant's preference for a diagram type and their performance whilst using that diagram. The study described examines whether participants' showed a difference in response times when understanding information contained in sequence diagrams and collaboration diagrams. This information is then compared to the participant's responses regarding their preference. It was expected that subjective user preference after using the diagrams would be reflected in objectively measured performance. This hypothesis derives from the author's previous study of this relationship with these diagrams [5], and earlier studies of the relationship between preference and performance.

The rest of this paper is structured as follows

- In section 2 we discuss research relating to user preference and performance in this area
- In section 3 we describe the design of the study
- In section 4 we present our analysis and results
- In section 5 we discuss the findings and implications
- In section 6 we give our conclusions and identify directions for future work

2. Preference and Performance

There is a body of research that has examined the link between subjective preference and objective performance with respect to different aspects and types of software systems. For example, websites [6], [7] moving map systems [8] and graph layout algorithms [9] have been studied in this context. The relationship between subjective preference and objective performance measures remains unclear after these studies, as the results have not been consistent. Given this mixed approach to evaluation of software systems, which may focus on objective performance measures, subjective ratings, or a combination of the two, a clearer understanding of this relationship between objective preference and subjective performance is desirable. The suggestion of Frokjaer [10] that effectiveness, efficiency and satisfaction be considered as *independent* aspects of usability is sensible. However, it is not an approach that is widely taken. Gaining a better understanding of the relationship between various performance measures may enable us to draw conclusions from a more focused approach.

A number of factors have been suggested as possible explanations for the mixed results of earlier studies. It has been noted that familiarity of users with the object under evaluation may influence subjective measures [11]. This must be considered in the context of the findings of a meta-study [5] which found that user preferences correlate to improved objective performance once users had some experience with the artefact. User preferences prior to gaining this experience did not show such a correlation. Other research has indicated that (in the context of web pages) preference is more strongly affected by interest in content than successful achievement of user goals [12].

Previous studies by the authors [5], [13] into the

relationship between user preference for UML sequence and collaboration diagrams and objective performance with the diagrams mirrored the findings of Nielsen and Levy: Results showed that user preference for one of the two diagram types before carrying out the task was not reflected in improved performance with that type of diagram compared with the other. However, after carrying out the task, user statements about which type of diagram they preferred working with were matched by improved performance with that type of diagram. Nevertheless, the objective performance measure used in these studies was that of accuracy alone. The study described below provides a measure of performance, which covers both speed and accuracy, with differing results to both the Nielsen and Levy study and earlier studies carried out by the authors of this paper.

3. Design of the Study

This study was carried out using bespoke software developed specifically for this investigation. The aim was to time participant's responses when answering questions on the information contained within various interaction diagrams. The software was designed to gather data on both the time taken to respond and the accuracy of responses, with participants asked to answer questions relating to information contained in sequence and collaboration diagrams. This software was subjected to pilot testing using heuristic evaluation with five experts answering questions relating to each screen and the overall study.

3.1 Experimental Setting

The study took place in the same room for all participants although only one person participated in the study at any one time. This was to ensure that all conditions for the study were the same for everyone. All participants were asked to adjust their environment to make themselves comfortable and then read the same sheet of information. The 40 participants were a mixture of final year Computer Science undergraduates, Ph.D. students from the department of Computer Science and members of staff from either the Departments of Computer Science or Psychology at the University of Hertfordshire. All participants had some previous experience with UML diagrams during their studies or work. Each participant was introduced to the task and given an explanation of what was expected. Personal data was gathered and participants

were given the option of not disclosing their contact details.

3.2 Pre-test preferences

An example of a sequence diagram and corresponding scenario was then displayed and participants were asked to read and try to understand. The same scenario shown as a collaboration diagram followed this, again the participant was asked to read and try to understand. For the purpose of this study the diagrams were called Type A and Type B respectively fig. 1. (a) and fig. 2. (b). Each of the participants was asked to state their experience and then select their preference for a diagram. There were three possible options: sequence, collaboration or no preference. Table 1 shows the number of responses to each option.

Table 1:
No. of Preferences for Diagram Type Pre-test/ Post-test

	No. of Participants Pre-test	No. of Participants Post-test
Sequence Diagram	26	27
Collaboration Diagram	7	4
No Preference	7	9

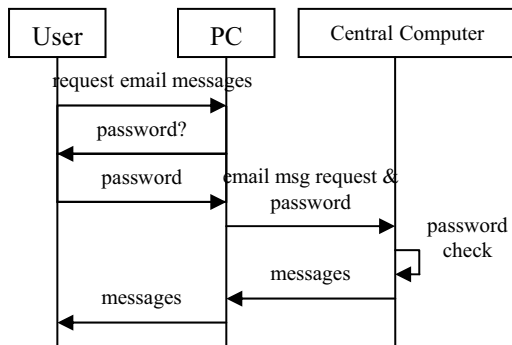


Fig. 1. (a) An example of diagram Type A

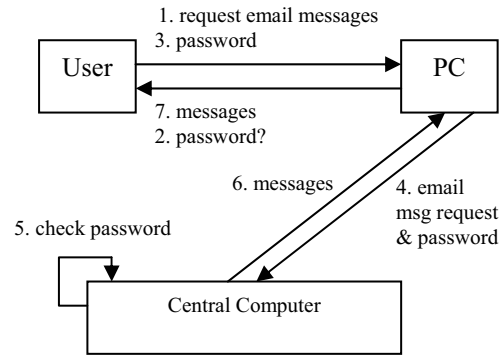


Fig. 1. (b) An example of diagram Type B

It can be seen that there was an overwhelming preference for sequence diagrams. There was space provided on the screen for the participants to leave comments about their preference, although this was optional. Comments from participants who preferred sequence diagrams include “it was easier to follow”, “it is easier to understand because it was in order”, “it does not seem as confusing as the other one”. One participant who preferred the collaboration diagram thought “it was easier to follow as it was numbered”. Due to the small sample size of participants who preferred collaboration diagrams it was necessary to reclassify the results as (i) ‘participants who prefer sequence diagrams’ and (ii) ‘participants who do not prefer sequence diagrams’. ‘Participants who do not prefer sequence diagrams’ is an amalgamation of those who prefer collaboration diagrams and those who had no preference. When participants were asked their preference the two thumbnail diagrams fig. 2. (a) and fig. 2. (b) which outline the structure of a sequence and a collaboration diagram were always visible.

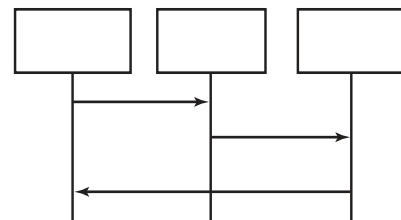


Fig. 2. (a) Thumbnail of a Sequence Diagram

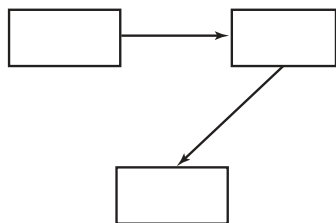


Fig. 2. (b) Thumbnail of a Collaboration Diagram

3.3 Experimental task

A series of four diagrams were displayed to each participant, two of which were sequence diagrams (fig. 3.) and two of which were collaboration diagrams (fig. 4.). The order in which the diagrams were displayed was randomised to ensure there was no learning effect. The diagrams were comprised of approximately 30 interactions each and were of similar complexity. Two different scenarios were modelled in the diagrams – an ATM scenario and a lift scenario.

Each diagram had six questions associated with it relating to the information contained in it and the diagram was visible throughout the time the participants were answering questions. The questions asked related to either ordering information or activity information. To ensure that the information in the diagrams was read carefully, the questions asked about information that was specific to the particular scenario represented in the diagram, rather than the general case of using a lift or ATM machine. The answers were usually a numeric value as opposed to a simple yes, no or don't know. Participants could only continue once they had input a correct answer to a question, an additional measure to try to ensure the information was read carefully. Below is an example of one of the sets of questions used in the study.

1. How many times did the user call the lift?
2. Which floor did the lift start at?
3. How many times was the number 15 displayed on the internal lift panel?
4. Which floor did the lift stop at before it went to floor 2?
5. Which statement is true?

The internal lift panel displayed floor 2
 twice during the series of events
 before floor 12
 after floor 23
 directly after audio alert

before floor 23

6. How many times is an audio alert sounded?

The participants were not told that their answers were timed as it was felt that this may change any strategy they adopt to answer the questions. All participants were asked to work quickly and accurately.

3.4 Post-test preferences

Once all the questions had been answered correctly, the participants were again asked to select their preference. The same thumbnails were used as before, and the participant could again choose from sequence, collaboration or no preference (Table 1 shows the number of post-test responses to each). The participants were thanked and informed that this was the end of the study.

4. Results

Table 2 shows the mean scores for task completion times obtained by participants for those who expressed a pre-test and post-test preference for sequence diagrams and those who did not prefer sequence diagrams (i.e. either preferred collaboration diagrams or had no preference).

Table 2: Mean task completion time for sequence and collaboration diagrams and pre-test and post-test preference (N=40)

Condition	Task completion time: Sequence (seconds)	Task completion time: Collaboration (seconds)
Preferred sequence diagram (pre-test)	200.29	225.33
Did not prefer sequence diagram (pre-test)	177.39	180.54
Preferred sequence diagram (post-test)	201.19	232.48
Did not prefer sequence diagram (post-test)	179.11	184.15

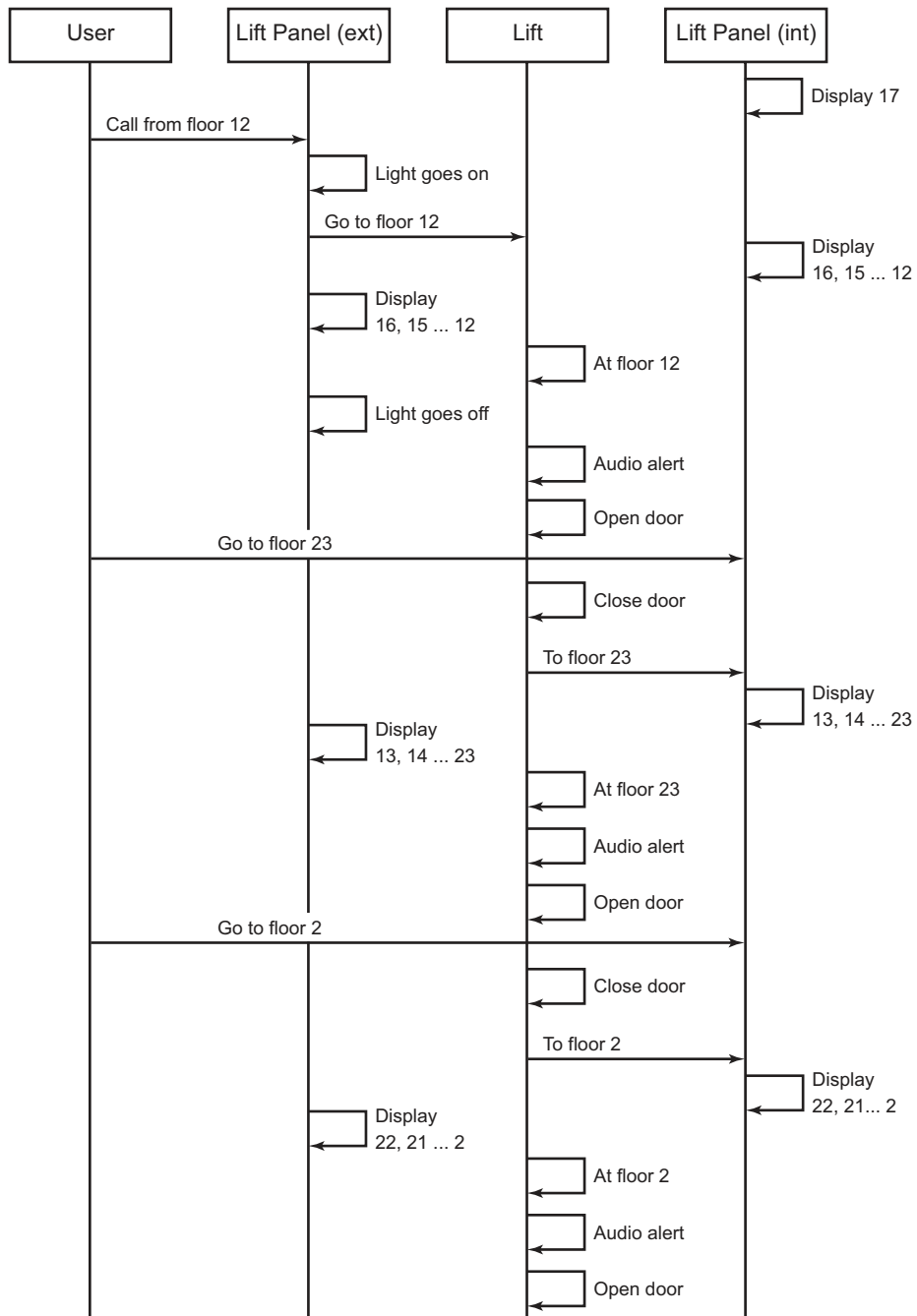


Fig. 3. An example of one of the sequence diagrams used in the study

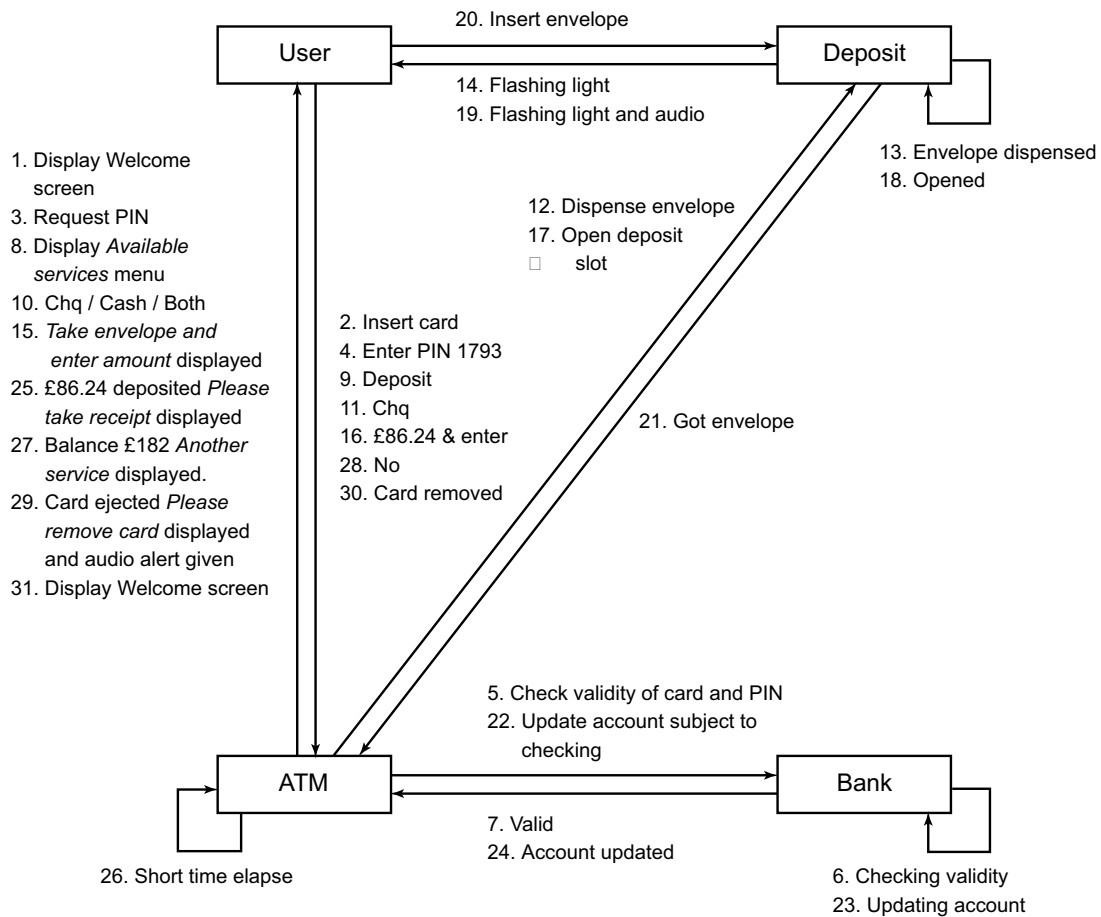


Fig. 4. An example of one of the collaboration diagrams used in the study

In order to test the differences in the mean task completion times, an analysis of variance (ANOVA) was performed, using the SPSS software package. This test was performed in order to test the significance of the differences shown in table 2 above. Table 3 shows the results of this ANOVA and presents the significance of the differences in the mean task completion times for both diagram types (sequence and collaboration combined) obtained for those expressing a pre-test and post test preference for sequence diagrams.

This level of significance obtained in the ANOVA ($p < 0.05$) supported the view that the difference in the mean values shown in table 2 were due to the effect of the independent variable, preference, and not due merely to chance. This result suggests that those expressing a pre-test preference for sequence diagrams perform the tasks significantly slower than those who express no preference or prefer

collaboration diagrams.

Table 3: ANOVA performed on data in table 2 showing the significance in mean task completion times for those expressing a pre-test and post-test preference for sequence diagrams

Condition	N	F	p
Pre-Test preference	40	4.55	0.040
Post-Test preference	40	3.187	0.082

There is a similar effect for post-test preference, although the difference is only approaching significance ($p = 0.082$).

An interesting comparison from the data presented in table 2 relates to the difference between task completion times for the two types of diagram, for those expressing a preference or no preference for sequence and collaboration diagrams. Table 4 below shows the results of a repeated measures ANOVAs performed on the data summarised in table 2, to show the significance of any differences in task completion times on sequence and collaboration diagrams, for those expressing a preference against those who expressed no preference, or preferred collaboration diagrams.

Table 4: Repeated measures ANOVA showing the significance in task completion times for sequence and collaboration diagrams and preference for diagram type

Condition	N	F	p
Pre-test Preferred diagrams	sequence 26	4.52	0.044
Pre-test Did not prefer sequence diagrams	14	0.120	0.734
Post-test Preferred diagrams	sequence 27	4.05	0.050
Post-test Did not prefer sequence diagrams	13	0.53	0.48

The results show that for those expressing a preference for sequence diagrams their task completion times were significantly faster for that type in both pre-test and post-test preference ($p \leq 0.05$). Those that had no preference or expressed a preference for collaboration diagrams performed equally well on both types of diagram ($p > 0.05$).

This analysis is interpreted as follows:

- Those participants that had a pre-test preference for sequence diagrams performed better with them
- Those participants that had a post-test

preference for sequence diagrams performed better with them

- Those participants that did not have a pre-test preference for sequence diagrams did not perform significantly better with either diagram type
- Those participants that did not have a post-test preference for sequence diagrams did not perform significantly better with either diagram type
- Pre-test analysis shows those participants that did not have a preference for sequence diagrams performed significantly better overall than those that did
- Post-test analysis shows those participants that did not have a preference for sequence diagrams performed significantly better overall than those that did

There was no significant difference observed due to the effect of previous experience with these diagrams.

5. Discussion

The findings from this study differ from the previous study carried out by the authors [5], and also with the meta-study reported by Nielsen and Levy [4] – it shows a relationship with pre-test preference which the previous studies did not find. It may be that when shown an example of the different diagram types each participant was able to judge which diagram they would perform better with, and therefore chose this as their preference. The results of the post-test analysis suggest the participants were aware of which diagram they performed better with.

There is an interesting finding that participants who did not prefer sequence diagrams performed better overall than the participants who did prefer sequence diagrams. It may be that people who did not prefer sequence diagrams were more ‘technically able’, as the collaboration diagram seems the more complex of the two. This might suggest that if they found collaboration diagrams present few problems with regard to comprehension, then they would find sequence diagrams very easy to comprehend. This is merely conjecture at this stage but is a potential area for future research.

One implication of these findings for the selection of techniques used in the requirements process is that where users prefer sequence diagrams we are able to state that they are more readily understood and would

therefore make a more suitable choice. However, further studies would need to be carried out to identify an effective approach where users prefer collaboration diagrams.

It should be noted that the results are in line with the hypothesis that all participants can understand that the diagrams given enough time, but some will take longer to comprehend them than others.

6. Conclusions

This study has produced unexpected findings: a relationship has been found between pre-test and post-test preference for sequence diagrams and performance when using interaction diagrams.

The study design was intended to overcome perceived weaknesses of earlier studies of the use of these diagrams, such as participants guessing, or providing answers that relate to the general case instead of the scenario represented. In addition, the measure of performance is richer, with the inclusion of timing information. As a result we have stronger confidence in these findings.

A significant volume of data has been gathered in the study, enabling analysis of further areas, such as whether certain types of information can be more readily extracted from one diagram type over the other. Future work will examine both this wider range of factors, and also whether the relationship found between pre-test and post-test preference and performance can be replicated with different representations.

10. References

- [1] McCluskey, T., Porteous, J., Naik, Y., Taylor, C. & Jones, S. (1995). A requirements capture method and its use in an air traffic control application. *Software Practice and Experience*, 25 (1).
- [2] Booch, G., Jacobson, I. and Rumbaugh, J. (1999) *The unified modelling language user guide*, Addison-Wesley.
- [3] Petre, M. (1995) *Why looking isn't always seeing*, Communications of the ACM, Vol 38, No. 6.
- [4] Nielsen, J. and Levy, J. (1994) Measuring usability, preference vs. performance, Communications of the ACM, Vol. 37, No. 4, April 1994.
- [5] Britton, C., Kutar, M., Anthony, S., Barker, T., Beecham, S. and Wilkinson, V. An Empirical Study of User Preference and Performance with UML Diagrams In, Proceedings of IEEE 2002 Symposia on Human Centric Computing Languages and Environments (HCC02), Arlington, Virginia, September 02, IEEE Press.
- [6] Benbunan-Fich, R. (2001) Using protocol analysis to evaluate the usability of a commercial web site. Information and Management, Vol. 39, No. 2.
- [7] van Schaik, P. and Ling, J. (2001) The effects of frame layout and differential background contrast on visual search performance in web pages. Interacting with Computers, Vol. 13, NO. 5.
- [8] Ruffner, J.W., Lohrenz, M.C. and Trenchard, M.E. (2000) Human factors issues in advanced moving-map systems. Journal of Navigation, Vol. 53, No.1.
- [9] Purchase, H.C., Alder, J-A and Carrington, D. (2000) *User preference of Graph Layout Aesthetics: a UML stud*, Proceedings of the Graph Drawing Symposium, J. Marks (Ed.), Lecture Notes in Computer Science 1984, pp 5 – 18, Springer Verlag, 2000
- [10] Frokjaer, E., Hertzum, M. and Hornbaek, K. (2000) Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated? In Proceedings of the SIGCHI conference on Human factors in computing systems, The Hague, The Netherlands.
- [11] Kissel, G. V. (1995) *The effect of computer experience on subjective and objective software usability measures*. In Conference companion on Human factors in computing systems, Proceedings of ACM SIGCHI Conference on Human Factors in Computing Systems, Denver, Colorado, United States
- [12] Spool, J. et al. (1999) *Web Site Usability. A Designer's guide*. Morgan Kaufman Publishers Inc, San Francisco, California.
- [13] Kutar, M., Britton, C., Barker, T. and Swan, J. (2003) Investigating the Influence of Structure on User Performance with UML Interaction Diagrams. Proceedings of EASE and PPIG 03, Keele University April 03.