

# An Empirical Study of Factors that Affect User Performance when Using UML Interaction Diagrams

Jennifer Swan University of Hertfordshire j.swan@herts.ac.uk	Trevor Barker University of Hertfordshire t.1.barker@herts.ac.uk	Carol Britton University of Hertfordshire c.britton@herts.ac.uk	Maria Kutar University of Salford m.kutar@salford.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. The research reported in this paper relates to an empirical study carried out to investigate factors which might affect user performance when using UML interaction diagrams. Several variables were investigated in the study; these were identified from the related literature and earlier research by the authors as being important in understanding interaction diagrams. The independent variables investigated in the study were diagram type, user pre-test and post-test preference, individual's cognitive style, text direction, scenario type and question type. Time taken to formulate the correct answer was the dependent variable used as the measure of performance. Statistical analysis of data showed significant differences for several variables, including diagram type, preference, and scenario type ( $p < 0.05$ ).*

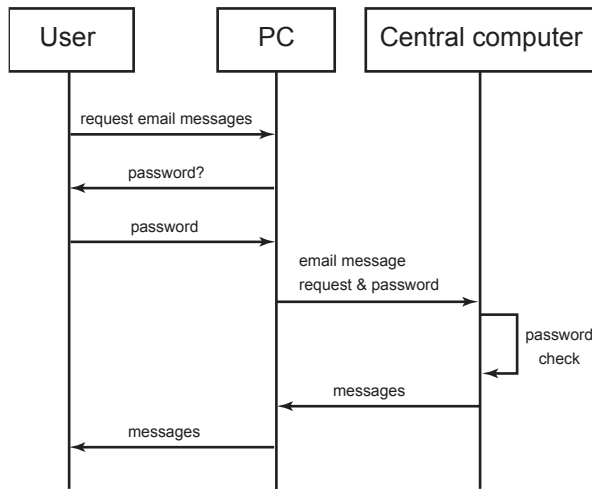
## 1. Introduction

'The UML (Booch, 1999) is a general-purpose visual modeling language that is used to specify, visualize, construct, and document the artefacts of a software system.' [1]. The UML is independent of any software development process as long as the process is object-oriented. In recent years it has emerged as 'the de facto standard for the representation of software engineering diagrams' [2].

Diagrams are often used in the specification of systems during the development process, often by stakeholders with little or no experience of their use.

By forming part of the specification they can be viewed and used by a variety of people, all of whom may have differing roles in the development process. Choice of diagram for particular projects often reflects the experience or preferences of the development team more than objective consideration of possible alternatives [3]. It is imperative that the diagrams used can be clearly understood by all who must use them. The research reported here relates to understanding the underlying factors that influence how we interpret diagrams. We hope that this understanding will have practical and theoretical applications in the use of diagrams in the software development process.

The UML provides a number of diagram types to support the specification, design and implementation of software systems. Two of the most important models contained within the UML are the *use case* model which enables description of the tasks that must be performed by, or with the assistance of the system, and the *class* diagram which describes the classes that will be used to achieve this and the relationships between them [4]. Interaction diagrams provide a bridge between the use case model and the class diagram; they are used to represent scenarios and provide a detailed record of the way in which objects interact to perform a task. The interaction diagram shows the sequencing of messages between collaborating objects [5]. There are two types of interaction diagram, sequence (figure 1.) and collaboration (figure 2.), which can be used interchangeably. These diagrams represent the same information but use differing structures and syntax. The difference is in emphasis. Sequence models emphasise time whilst collaboration models emphasise object relationships [1]. This means the diagrams make an excellent vehicle for investigation into the effects different factors may have on user performance when trying to understand these diagrams.



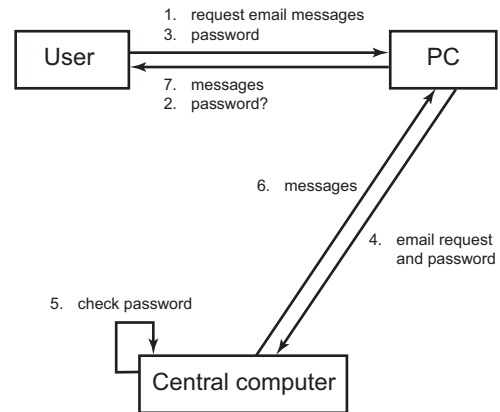
**Figure 1. Sequence Diagram**

The Diagrams Research Group (DRG) is based in the UK at the University of Hertfordshire and Salford University. The DRG's main focus is to conduct research into software engineering diagrams. Interaction diagrams were used in a series of studies carried out by the DRG. These studies investigated different areas that may affect the ease with which a user could understand each of the interaction diagrams. The areas investigated were diagram type, preference and performance, cognitive style, scenario, text direction and question type.

## 2. Background

The study makes comparisons between the two types of interaction diagram to try to improve the ease with which people can understand them. As a first stage to this work an empirical study was developed which investigated factors identified as being important in the related literature, previous research and feedback. These factors are diagram type, preference and performance, cognitive style, text direction, scenario and question type. Our findings on the effect of cognitive style on interpreting diagrams are beyond the scope of this paper. The work is currently in press and will be presented elsewhere [6]. The remainder of this section will introduce the independent and dependent variables used in the study in detail.

If a representation is clearly structured, it will involve less effort on the part of readers to find, decompose and abstract information, and thus be easier to understand [7], [8], [9], [10]. As can be seen in figures 1 and 2, sequence and collaboration diagrams contain the same information structured in different



**Figure 2. Collaboration Diagram**

ways. It was an aim of this research to investigate the importance of structure in interaction diagrams and whether it had an effect on user performance. Our hypothesis was that both diagram types can be readily understood, but that sequence diagrams would be more easily intelligible to the reader and this would be reflected with improved user performance times. Previous studies carried out by the DRG did not support this hypothesis, with no differences found in performance for sequence and collaboration diagrams. The results showed that structure did not seem to be an important factor in these studies [11], although this was tested using accuracy alone. In order to investigate the effect of structure more precisely, we analysed the time taken to correctly answer questions relating to information contained in each of the diagram types.

Another 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 [12] 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 [13]. It was our intention to discover whether or not a relationship exists between a participant's preference for a diagram type and their performance whilst using that diagram. Our hypothesis was that a pre-test preference for one type of diagram over the other would be reflected in improved performance when answering questions relating to information contained

within that type of diagram. It was also thought the participant would perform better with the diagram he or she chose as his or her post-test preference. Previous studies by the DRG and colleagues [14], [11] mirrored the findings of Nielsen and Levy with no correlation for pre-test preference and performance, although users' post-test preferences were matched with improved performance for that diagram type. However, the performance measure used in these studies was accuracy alone. In order to investigate this further, we gathered data on pre-test and post-test preferences. We then compared this to participants' response times when understanding information contained in sequence diagrams and collaboration diagrams.

Purchase et al. [15], found horizontal text to be an important factor when looking at preferences for text direction on class diagrams, although they did not find any significant effect of text direction for collaboration diagrams. Text labels should be horizontal rather than a mixture of horizontal and vertical according to Petre [12]. With this in mind, we decided to investigate whether or not text direction had any effect significant on performance when using interaction diagrams. Our hypothesis was that participants would show improved performance when using diagrams containing horizontal text than diagrams containing a mixture of horizontal and vertical text. In order to investigate this, participants answered questions relating to information on diagrams displaying horizontal text and diagrams displaying a mixture of horizontal and vertical text. The performance measure used was speed and accuracy.

Scenarios are used in many different ways to aid the development of software systems such as specification generation [16] and object oriented design [17], [18], [19] and Requirements Engineering [20], [21]. A scenario may be defined as representing "one instance of a use case, describing a particular sequence of events that may occur in trying to reach the use case goal" [22]. Scenarios are recognised as playing an important role throughout the development process, and "can be used as a tool in requirements gathering, interface design and evaluation" [23]. When carrying out empirical studies it is of key importance to know which, if any, of the independent variables might be inadvertently influencing the outcome of the results. Therefore it is essential that measures are taken to ensure these variables are controlled as carefully as possible. Our hypothesis was that scenario would have no significant effect on participant's performance providing the scenarios were of similar complexity and size. Previous studies carried out by the authors was able to show that a 'scenario effect' might have influenced the results due to differing levels of familiarity participants had with scenarios used. In

order to try to combat the 'scenario effect' the two scenarios used were chosen carefully and the research team deemed them to be of similar size, complexity and familiarity.

In response to suggestions from other researchers in this area, question type was included as an independent variable in this study. As interaction diagrams emphasise both 'ordering' and 'activity' type information, the reviewer felt it was important to look at the questions in the study carefully, to ensure an even balance of questions asking for 'ordering' information and 'activity' information. 'Ordering' information was gathered by asking questions referring to the order or sequence in which events took place, whilst 'activity' information was gathered by asking questions such as how many times an event took place. We want to investigate the importance of question types when using interaction diagrams and whether it had an effect on user performance. Our hypothesis was that performance would be improved for both 'ordering' questions used in sequence diagrams and 'activity' questions used in collaboration diagrams. In order to achieve this, each diagram asked for an even balance of ordering information and activity information.

The dependent variable chosen was 'time to interpret diagrams correctly', rather than 'the number of correct first responses', used in previous studies. It was hoped that this better reflected the way diagrams were likely to be used in the real world in realistic contexts. In such complex real world scenarios, incorrect assumptions are usually corrected. It is the time taken to formulate the correct interpretation that is important in such cases, not simply the first guess at a solution. It was also suggested that the new approach would be likely to reduce the amount of variance due to systematic error caused by guessing.

### 3. Design of the Study

A software application was developed for use in this study using Microsoft Visual Basic. This application was able to present diagrams, scenarios and questions and also to time responses accurately and record the number of attempts. Eight different diagrams were used in the experiment, although each participant would only see a randomised sub-set of four of these. This was to enable randomisation of the independent variables to ensure, as far as possible, that bias due to learning effect was eliminated. Each of these randomisations ensured the participant saw four diagrams which (i) comprised of two each of the sequence and collaboration diagrams, (ii) used both the lift and ATM scenarios twice and (iii) contained

horizontal text for two of the diagrams and mixed text for the other two diagrams. Table one, in section 3.3 shows the independent variables used in each of the eight diagrams.

### 3.1 Pilot Study

Prior to the first full run of the experiment, the software prototype was subjected to pilot testing using the method of heuristic evaluation to ensure that the user interface was error free. The evaluation was based on Nielsen's method [24] and related to the overall usability of the application interface, as well as issues related to the presentation of the diagrams. Five usability experts from the University of Hertfordshire used the application in the same environment as the participants. Each of the experts were given a set of instructions and evaluation guidelines prior to the use of the application. Experts reported on the usability of each screen using Nielsen's heuristics. In addition to the interface, the experts were asked to report on the clarity of instructions to be given to participants.

All five experts reported that they found the study to be aesthetically pleasing and the fonts used were large enough and easy to read. All five experts commented on the difficulty of reading a mixture of horizontal and vertical text. Feedback regarding the instructions given was that they were clear, concise and easy to understand. None of the experts ever felt 'lost' in the application and thought when information was needed it was always visible. Despite one or two minor changes that were recommended and made to the application, the experts reported that they found the application usable, presenting no barriers to the study with regards to the instructions given within the application and to participants as an introduction to the study.

### 3.2 Experimental Setting

Volunteers were a mixture of final year Computer Science students, postgraduate students from the Department of Computer Science and members of staff from either the Department of Computer Science or the Department of Psychology. 40 people participated in the study and all had some previous experience with UML diagrams during their studies or work.

A standard presentation procedure was employed for each participant in the study to remove bias as far as possible. The study took place in a controlled laboratory environment, with each participant completing the study individually, with only the experimenter in the room. This was to ensure as far as possible that the conditions for the study were the same

for everyone. All participants were given the same introduction, setting out the experiment and presenting an overview of what they might expect in the presentation.

### 3.3 The Experiment

After adjusting the environment to suit themselves, each participant was shown a series of introductory screens providing examples of both diagram types. Once the participant had studied both types of diagram he or she was asked to select their preference for one of the diagram types. During this process a small image of each of the diagram types was visible to assist the choice. Participants could also report that they had no particular preference. In the study, a series of four diagrams were displayed to each participant in a random order, two of which were sequence diagrams (figure 3) and two of which were collaboration diagrams (figure 4). 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 associated questions relating to the information presented in the diagram. The diagram was visible throughout period when the participants were answering questions. The questions asked related to either 'ordering' information or 'activity' information. To be as sure as possible that the diagrams were 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. Presented 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 informed that their answers were timed as it this might potentially influence any strategy adopted in answering questions. All participants were asked to work quickly and

accurately. The duration of the study was approximately 30 minutes on average.

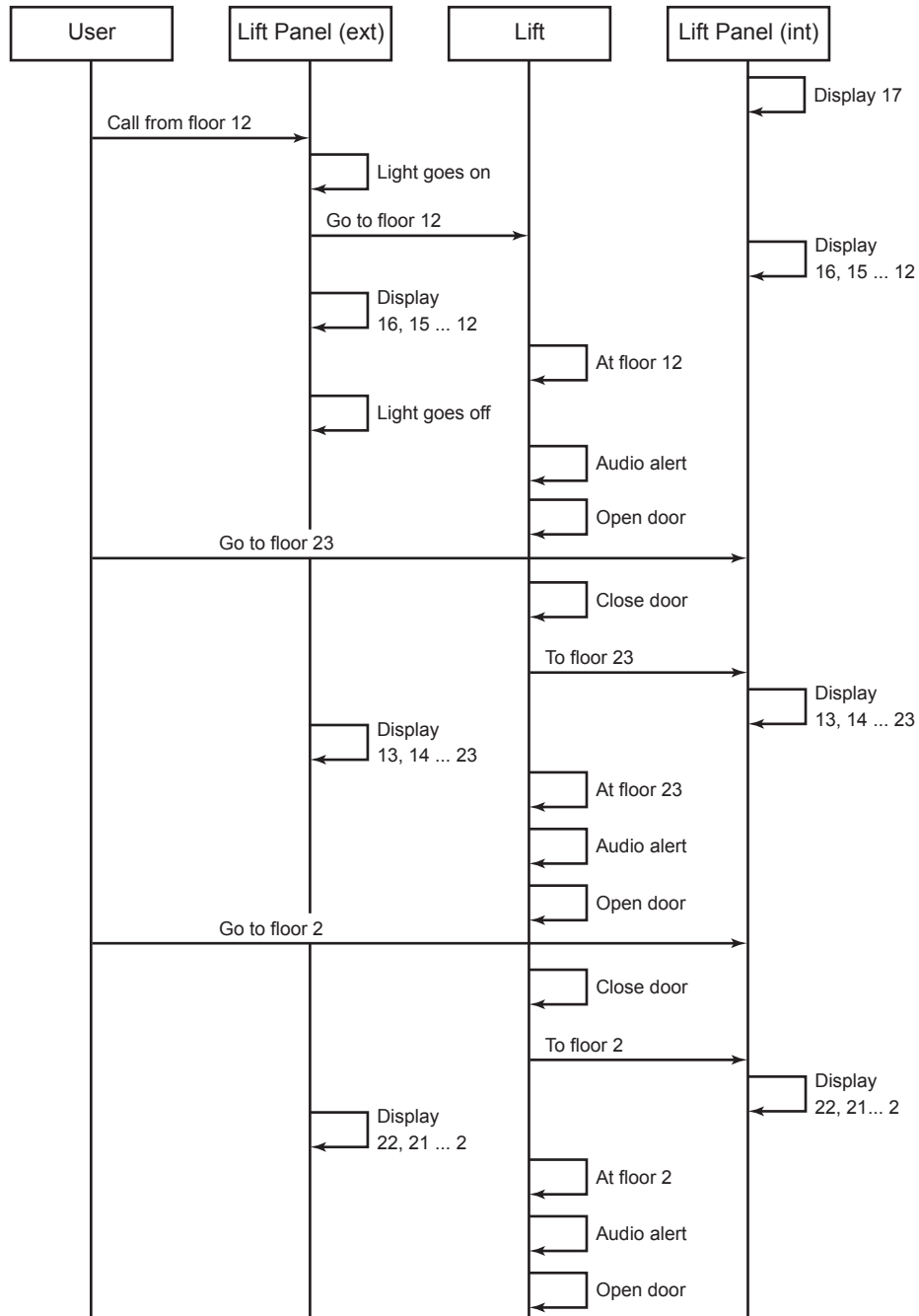
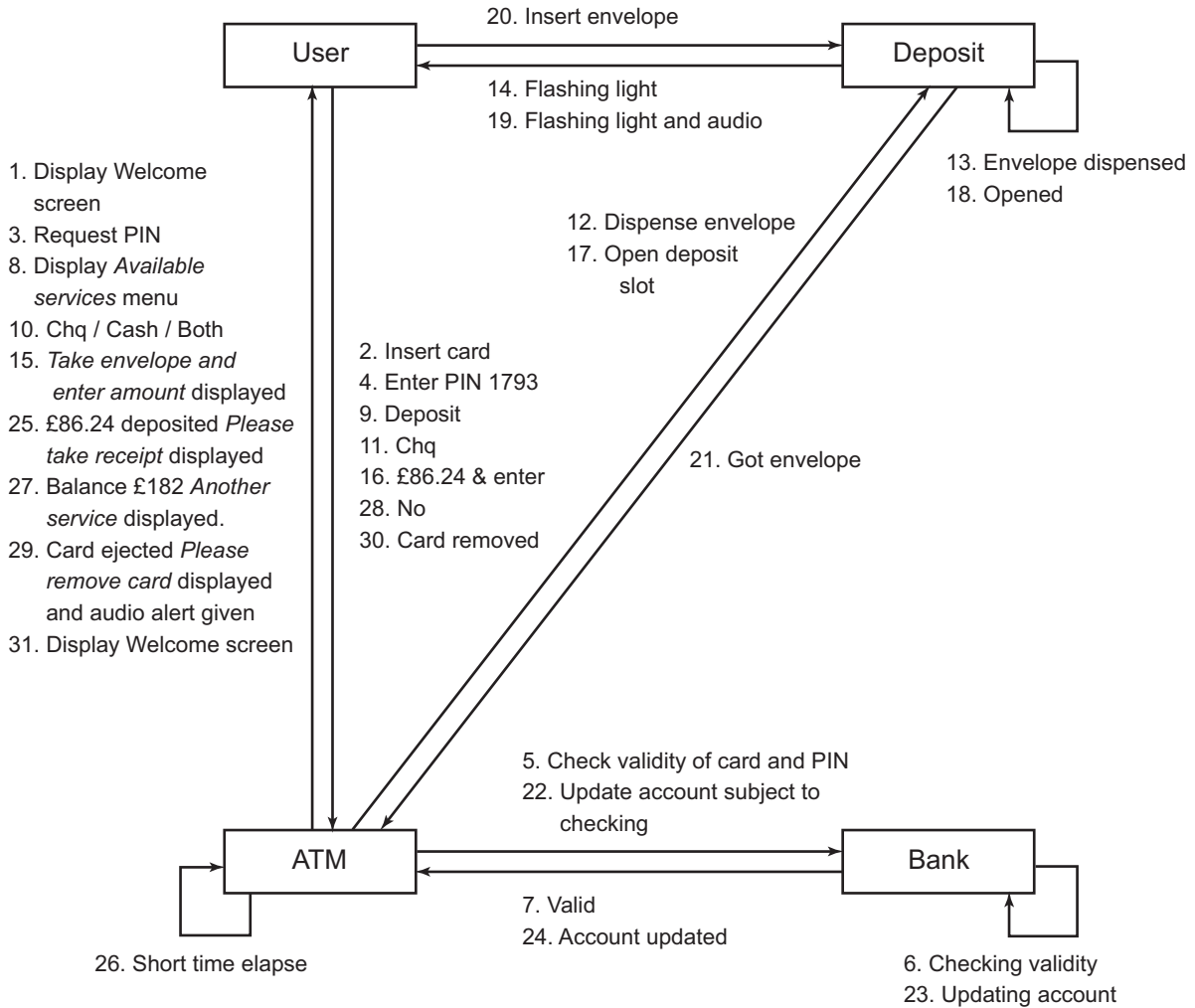


Figure 3. An example of a sequence diagrams used in the study



**Figure 4. An example of a collaboration diagram used in the study**

Table 1 below shows the eight diagrams used in the study and the independent variables used in each of the diagrams.

Diagram number	Diagram Type	Scenario	Text Direction
1	collaboration	lift	mixed
2	sequence	lift	horizontal
3	collaboration	ATM	horizontal
4	sequence	ATM	mixed
5	sequence	lift	mixed
6	collaboration	lift	horizontal
7	collaboration	ATM	horizontal
8	sequence	ATM	mixed

## 4. Results

In the following section the data obtained in the experiment is presented and analysed. Statistical analysis was performed using the SPSS software application version 11.5. All timings are measured in seconds.

### 4.1 Diagram Type

In table two below we present the mean times taken for participants to provide correct answers to questions for different diagram types.

Table 2: below shows the mean time obtained in the experiment for the correct interpretation of sequence and collaboration diagrams

**Table 2: Time taken to complete tasks in the experiment and diagram type**

Diagram Type	Mean Time	SD
Sequence Diagram	194.01	65.42
Collaboration Diagram	216.77	72.52

In order to test the significance of the means obtained in the experiment presented in table 2 above, a repeated measures Analysis of Variance (ANOVA) was performed on the data summarised there.

Table 3 below shows the results of an ANOVA performed on data in table 2 showing the significance in mean task completion times for sequence and collaboration diagrams.

**Table 3 Repeated measures ANOVA performed on the data summarized in Table 2 (N=40)**

Variable	F	P (one tailed)
Diagram Type	4.418	0.021

These results were taken to indicate that the significant difference in performance was due to the effect of diagram type ( $p < 0.05$ ). The analysis of data related to diagram type supported the hypothesis that both diagrams could be understood adequately, but the collaboration diagram took longer, questions relating to the sequence diagrams being answered significantly faster. All participants were able to answer questions correctly for both types, as no questions were left unanswered.

## 4.2 Preference

Table 4 below shows task completion time for sequence and collaboration diagrams and pre-test and post-test preference 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).

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 4 below. Table 5 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 4 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 4: The effect of diagram type on the time taken to complete tasks (N=40)**

Condition	Task completion time:	Task completion time
	Sequence	Collaboration
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

Table 5 below shows the results of an 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.

**Table 5: Time taken to complete tasks in the experiment and pre-test and post-test preference**

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 4 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 6 below shows the results of a repeated measures ANOVAs performed on the data summarised in table 4, 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 6 below shows the results of a repeated measures ANOVA showing the significance of any effect of preference on the differences in mean task completion times for sequence and collaboration diagrams

**Table 6: Repeated measures ANOVA performed for pre-test and post-test preference for diagram type.**

Condition	N	F	P
Preferred sequence diagrams (pre-test)	26	4.52	0.04
Did not prefer sequence diagrams (pre-test)	14	0.120	0.73
Preferred sequence diagrams (post-test)	27	4.05	0.05
Did not prefer sequence diagrams (post-test)	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.

### 4.3 Text Direction

Table 7 below presents the mean time taken for sequence and collaboration diagrams when displaying different text directions.

**Table 7: The effect of text direction on the time taken to complete tasks for both diagram types (sequence and collaboration together).**

Text Direction	Mean Time	SD
Horizontal Text	210.28	77.52
Mixed Text	200.50	70.50

In order to test the significance of the means obtained in the experiment presented in table 7 above, a repeated measures Analysis of Variance (ANOVA) was performed on the data summarised there. Table 8 below shows the results of a repeated measures ANOVA showing the significance in task completion times for sequence and collaboration diagrams and text direction

**Table 8: Results of repeated measures ANOVA performed on the data summarized in table 7 (N=40)**

Condition	F	P
Text Direction	0.506	0.24

No difference could be attributed to the effect of text direction

### 4.4 Scenario

In table 9 presented below, the mean time obtained in the experiment for sequence and collaboration diagrams and scenario.

**Table 9: Time taken to complete tasks in the experiment and scenario**

Scenario	Mean Time	SD
ATM scenario	229.79	79.59
Lift scenario	181.00	75.60

Table 10 below presents the results of a repeated measures ANOVA on the significance of differences in task completion times for sequence and collaboration diagrams and scenario.

Past studies carried out by the DRG have shown a significant difference in mean scores due to the effect of the scenario on performance. The technical report detailing these findings is the work of Kutar and colleagues [11].



**Table 10: Results of repeated measures ANOVA performed to test the significance of any differences in mean times for different scenarios.**

Condition	F	P
Scenario	15.05	>0.001

#### 4.5 Question Type

Table 11 below shows the effect of question type on the mean time recorded for sequence and collaboration diagrams.

**Table 11: The effect of question type for sequence and collaboration diagrams.**

Question Type	Mean Time	SD
Ordering Information	194.01	83.61
Activity Information	216.77	104.47

An ANOVA was performed on the data summarised in table 11. No significant difference was found ( $p>0.13$ ).

### 5. Discussion

In this research, we investigated several factors related to the interpretation of sequence and collaboration diagrams, using time to produce a correct response as the dependent variable. The independent variables investigated were diagram type, pre-test and post-test preference for a diagram type, nature of scenario, text direction and question type.

It was interesting that our findings relating to performance and diagram type were different from those of previous studies carried out by the authors [11], [14]. The results suggest a significant effect of diagram type on performance, which was not found in our earlier studies. This finding is well supported by in the literature [7], [8], [9], [10] which presents the view that more clearly structured representations will be easier to understand. Our results also support our hypothesis that time to understand information is a better measure than simply recording correct responses. All participants were able to solve the problems, but sequence diagrams took less time to comprehend than collaboration diagrams. It is possible that this is because sequence diagrams appear less complex to users than collaboration diagrams. We suggest that the reason we were unable to see this effect in the past was the use of correct response as the dependent variable introduced greater variance due to guessing. The use of ‘time to obtain a correct response’ reduced systematic error and led to improved statistics.

Our findings on the effect of pre-test and post-test preference support our hypothesis. This is also in

accordance with the literature [12], [13]. Our findings suggest that if users prefer one type of diagram for solving a problem to another, they will perform better with the technique that they prefer. There is an interesting finding that participants who did not prefer sequence diagrams performed better overall than the participants who did prefer sequence diagrams. We tentatively suggest that participants that preferred collaboration diagrams were more ‘technically able’, as collaboration diagrams are generally accepted to be the more complex of the two, all other things being equal. We intend to investigate this more fully in the future.

We were unable to find any significant effect of text direction and therefore we must reject our experimental hypothesis that participants would perform better using the diagrams with horizontal text. No significant effect on user performance could be attributed to text direction for either sequence or collaboration diagrams ( $p>0.05$ ). A review of the relevant literature in this area presents a somewhat mixed picture [12], [15]. It seemed likely to us that text direction would be a significant factor in understanding diagrams, and that failure to find the effect was due to a feature of our experimental settings. We therefore intend to undertake further research of this area in order to understand more fully the effect of text direction on the interpretation of sequence and collaboration diagrams.

An important finding in this study was the significant effect of scenario on user performance. This result was contrary to our hypothesis that scenario would have no significant effect on participant’s performance, providing the scenarios were of similar complexity and size. The result is however in line with findings from previous studies carried out by the authors. Careful measures were taken to control many aspects of scenario type, such as familiarity, size and complexity, and although unlikely, for these reasons, it is possible the precise choice of scenarios in this study was responsible for this significant effect. Scenario type was controlled for within the experimental design, so that our other findings were not related to differences in scenario. It is probably true to say that scenario familiarity and complexity influence how well we understand diagrams and we are not surprised by this. It is important then, in experimental design that this factor is adequately controlled for. By using a design that presented both scenarios with both diagram types, we were able to control for the scenario factor. The importance of scenario familiarity in interpretation of diagrams in the real world is an interesting potential for the future. It is likely that different scenarios might map better to different diagram types, as they might be expected to map the structure of the real word situations differently. This will be investigated in the future. Question type is another potential source of

systematic bias. For this reason, questions used in the experiment were as far as possible designed to have an equal balance of ‘ordering’ information and ‘activity’ information.

## 6. References

- [1] Rumbaugh, J., Jacobson, I., and Booch, G. (1999) *The Unified Modelling Language Reference Manual*, Addison-Wesley.
- [2] Purchase, H.C., Colpoys, L., McGill, M., Carrington, D. and Britton, C. (2001) UML class variations: an empirical study of comprehension. Presented at the Australian Visualisation conference, Sydney, 3 – 4 December 2001.
- [3] 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).
- [4] Pooley, R. & Stevens, P. (1999), *Using UML*, Addison Wesley.
- [5] Maciaszek, L. (2001), Requirements Analysis and System Design. Developing Information Systems with UML, Addison-Wesley.
- [6] Swan, J., Barker, T., Britton, C. and Kutar, M. The Interpretation of Complex Diagrams: The Importance of the Wholist-Analytic Cognitive Style Dimension. Proceedings of the European Learning Styles Information Network Conference, University of Surrey, 2005.
- [7] Green, T. (1983) Learning Big and Little Programming Languages, in A. Wilkinson (Ed.), Classroom Computers and Cognitive Science, Academic Press, New York.
- [8] Sengler, H. (1983) A model of program understanding, in T. Green, S. Payne. and G. van der Veer, (Eds.), The Psychology of Computer Use, Academic Press.
- [9] Eysenck, M. & Keane, M. (1990). Cognitive psychology: A student’s handbook.. Lawrence Erlbaum Associates.
- [10] Winn, W. (1993) An account of how readers search for information in diagrams, Contemporary Educational Psychology, 18.
- [11] 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.
- [12] Petre, M. (1995) *Why looking isn’t always seeing*, Communications of the ACM, Vol 38, No. 6.
- [13] Nielsen, J. and Levy, J. (1994) Measuring usability, preference vs. performance, Communications of the ACM, Vol. 37, No. 4, April 1994.
- [14] 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.
- [15] Purchase, H.C., Carrington, D.A. and Allder J-A. (2002) Empirical evaluation of aesthetics-based graph layout, *Empirical Software Engineering*, 7(3), pp233-255, 2002.
- [16] Anderson, J. S., and Durney, B. “Using Scenarios in Deficiency driven Requirements Engineering”, *Requirements Engineering’93*, IEEE Computer Society Press, 1993, pp. 134-141.
- [17] Booch, G. *Object Oriented Analysis and Design with Applications*, Benjamin/Cummings Publishing Company Inc., Redwood City, CA, 1994.
- [18] Jacobson, I., Christerson, M., Jonson, P. and Overgaard, G. *Object-Oriented Software Engineering: A Use Case Driven Approach*, Addison-Wesley, 1992.
- [19] Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F. and Lorensen, W. *Object-oriented Modeling and Design*, Prentice-Hall, Inc., 1991.
- [20] Hsia, P., Samuel, J., Gao, J., Kung, D., Toyoshima, Y. and Chen, C. “Formal approach to scenario analysis”, *IEEE Software*, (11)2, March 1994, pp. 33-41.
- [21] Potts, C., Takahashi K. and Anton, A. *Inquiry-Based Scenario Analysis of System Requirements*, Technical Report GIT-CC-94/14, Georgia Institute of Technology, 1994.
- [22] Britton, C. & Doake, J. (2000), Object-Oriented System Development: A Gentle Introduction, McGraw-Hill.
- [23] Bennett, S., McRobb, S. & Farmer, R. (1999), *Object-Oriented Systems Analysis and Design Using UML*, McGraw-Hill.
- [24] Nielsen, J. (1994). Heuristic evaluation. In Nielsen, J., and Mack, R.L. (Eds.), [\*Usability Inspection Methods\*](#). John Wiley & Sons, New York, NY.