Experimental Comparison of Sequence and Collaboration Diagrams in Different Application Domains

Chanan Glezer, Mark Last, Efrat Nahmani , Peretz Shoval*

Department of Information Systems Engineering,
Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel
*Corresponding author: shoval@bgu.ac.il; Fax. +972-8-6477527

Abstract. This article reports the findings from a controlled experiment where both the comprehensibility and the quality of UML interaction diagrams were investigated in two application domains: management information system (MIS) and real-time (RT) system. The results indicate that collaboration diagrams are easier to comprehend than sequence diagrams in RT systems, while there is no difference in their comprehension in MIS. With respect to quality of diagrams constructed by analysts, in MIS collaboration diagrams are of better quality than sequence diagrams, while in RT there is no significant difference in their quality.

1 Introduction

UML defines twelve types of artifacts in the form of diagrams which are divided into the following three categories: class, object, component, and deployment diagrams represent the application's *static* structure; use case, interaction (sequence and collaboration), activity and state charts represent the application's *dynamic* behavior; and packages, subsystems, and models represent the application modules' *organization* [7].

The focus of this study is on UML's interaction diagrams, which depict a pattern of interaction among objects. Interaction diagrams come in two forms emphasizing different aspects of an interaction: *sequence diagrams* and *collaboration diagrams*. Our goal is to evaluate and compare these two types of diagrams. *Sequence Diagram* depicts an explicit sequence of stimuli messages exchanged among object instances participating in the interaction. Sequence diagrams include lifelines for the instances used to portray the temporal dimension of the modeled pattern. *Collaboration Diagram* is a directed graph where nodes represent communicating entities and edges represent communications. The edges are numbered to represent the order of communications.[1]

Sequence diagrams and collaboration diagrams express similar information but depict it in a different way. Both diagrams are considered symmetric and it is therefore

[1] Due to space limitation we do not show examples of sequence and collaboration diagrams.

possible to convert a sequence diagram to a collaboration diagram and vice versa [1]. The main difference between the diagrams is that a sequence diagrams emphasizes the temporal dimension, exploiting the lifelines artifact, and is therefore assumed to be better in depicting an order of events or pause between events [1], [7], [10] and [11]. On the other hand, a sequence diagram does not portray the interaction among objects exploited by a system. If a system is complex, it might therefore be difficult to infer the mutual relationship and messages relayed between the objects using a sequence diagram.

A collaboration diagram addresses the loopholes of a sequence diagram by depicting relationships among involved objects. A collaboration diagram is therefore recommended for supplementing both class and use-case (static-view) diagrams because it depicts the interactions among objects (dynamic-view) [3]. Moreover, a collaboration diagram is claimed to enable better modeling of complex branching and concurrent activation of multiple processes [4], or control of multiple threads [10]. A collaboration diagram, however, does not capture the temporal dimension, and the relative order of messages exchanged between objects needs to be enumerated explicitly.

2 Related Work

Most of the extant work on UML interaction diagrams focused on conceptual analysis and comparison of the features of collaboration and sequence diagrams. The only empirical research that we know of comparing interaction diagrams was performed by Otero and Dolado [8], [9] who performed a set of experiments in an attempt to investigate the comprehensibility of interaction diagrams in UML.

In their first study [8], eighteen students of Informatics analyzed three types of diagrams: sequence, collaboration and state diagrams, within three different applications and application domains. Their main conclusion was that the comprehension of the dynamic models in OO designs depends on the diagram-type and on the complexity of the application. In a subsequent study, Otero and Dolado [9] performed an experiment comprising of two parts. The first part was a repetition of their earlier study. (The repetition study may be considered more powerful because of a better experimental design which eliminated the effect of learning caused by practice or sequence.) The second part of the 2004 experiment examined which combinations of dynamic diagrams (sequence-collaboration, collaboration state, or sequence-state) improve the understanding of a system. The main conclusion of the second part of study was that regardless of the application domain, a higher semantic comprehension of the application is achieved whenever the dynamic behavior is modeled by using the pair sequence–state diagrams.

The experiments described above suffer from several limitations. First, they compare non-equivalent types of diagrams because state diagrams are analyzed as equivalent to interaction diagrams (sequence and collaboration). Second, the 2004 study evaluated pairs of models without providing sufficient rational and proving that the pairs are in fact interchangeable. Third, they did not address the issue of building diagrams and their quality with regards to different application domains.

The purpose of our study is to fill the research gaps on UML interaction diagrams. As we have seen, prior research has mainly focused on comprehension of diagrams, but the issue of the quality of the diagram types has not been addressed yet. By investigating performance in terms of comprehension, quality, time and user/analyst preference in different application domains, the findings of this study are expected to provide a wide-angle view on the UML interaction modeling aspect, and hopefully contribute to the productivity of analysts, designers and programmers in complex information technology environments. Such environments are more likely to comprise of both heterogeneous types of applications (i.e., real time reactive/non reactive, managerial) and tight interaction between system designers and analysts.

3 The Experiment

3.1 Experimental Design and Variables

The goal of this study is to evaluate and compare sequence diagrams and collaboration diagrams from the two main perspectives (analysts and users) by conducting two controlled experiments. From the users' perspective, we are interested in the *comprehensibility*, i.e. understandability of the diagram, while from analysts' perspective, we are interested in the *quality*, i.e. correctness and completeness of the diagrams. The two types of diagrams are evaluated in two types of applications: a management information system (MIS) and a real-time reactive system (RT). Figure 1 depicts the experimental design of the study (it actually describes both experiments that were carried out as a part of a student test.)

In the "comprehensibility" experiment we compare users' comprehensibility of diagrams, the time it takes them to comprehend the diagrams and the perceived (subjective) ease of comprehension. We measured the following three dependent variables:

C – total score for questions on diagrams comprehension

T_c – total time spent to answer questions on diagrams comprehension

P_c –ranking of subject's perceived comprehensibility of a certain diagram

In the "quality" experiment we compare the quality of diagrams as created by analysts, the time it takes to create the diagrams and the perceived ease of constructing them. We measured the following three dependent variables:

Q – total score for quality of an interaction diagram constructed

T_q – total time spent on diagram construction

P_q – ranking of subject's perceived ease of construction of a certain diagram

In both experiments the independent variables are the two types of diagrams: sequence (Seq) and collaboration (Col), and two types of systems: a security system, representing a real-time reactive system involving time dimension and concurrency (RT); and a library system, representing a management information system (MIS). These type of systems represent a substantial portion of prominent industry applications and were also used extensively in previous research [8], [9], thus enhancing the practical implications and validity of our findings.

The controlled variables are the tasks and the subjects: in the "comprehension" experiment the task was to answer a questionnaire which measures comprehension of a

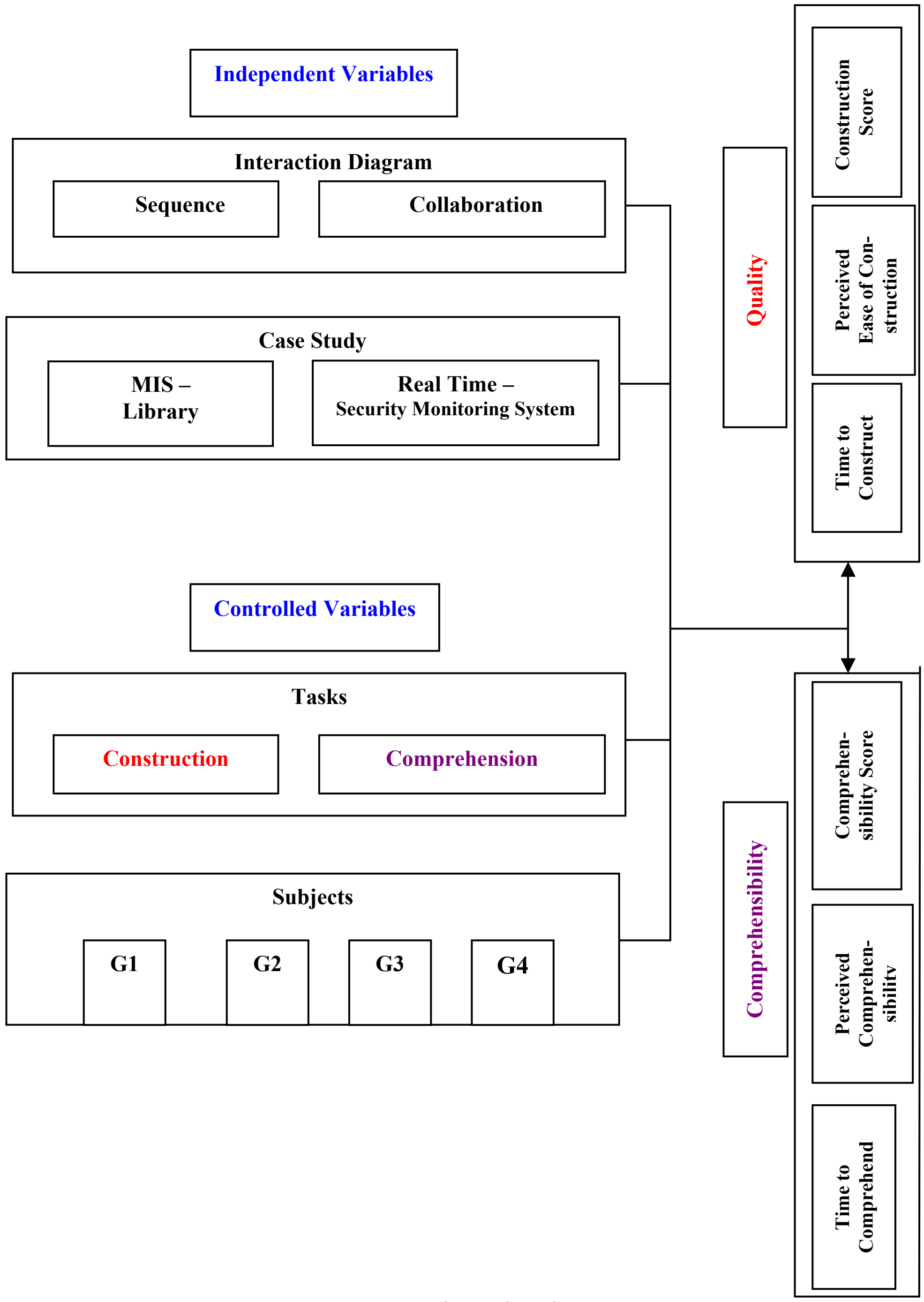

Figure 1. Experimental Design

given diagram; in the "quality" experiment the task was to construct a diagram. The subjects in the two tasks were randomly divided into four groups, as explained below.

The experiment was divided into two sessions. In Session 1 we conducted the "comprehension" experiment: subjects were asked to express comprehensibility of a given interaction diagram by answering five multiple-choice questions. In Session 2 we conducted the "quality" experiment: subjects were asked to construct an interaction diagram based on a brief narrative specification of the system, a class diagram and a use case diagram. Quality was measured by the correctness of the created diagrams with respect to the correct solutions. In addition to performing the tasks, the subjects were required to record the overall time spent to complete each session. Finally, they were also asked to express their subjective opinions on diagrams comprehension and construction.

3.2 Hypotheses

Our objective is to address the following questions:
(1) is there a difference between a sequence and a collaboration diagram (in terms of quality and/or comprehensibility) for dynamic modeling of a RT system?
(2) is there a difference between a sequence and a collaboration diagram (again, in terms of quality and/or comprehensibility) for dynamic modeling of a MIS?

As we have seen, previous research indicated that sequence diagram has an advantage over collaboration diagram in the representation of the temporal order, which is particularly important in real-time systems where the processes take pre-defined time slots and may be processed concurrently [7], [8], [9]. On the other hand, collaboration diagram better depicts static relationships between objects. These relationships are important in management information systems, since they are used to exchange messages between objects.

Since we are comparing between the two types of interaction diagrams using six dependent variables (as listed above) and the comparison is performed separately for two types of applications (RT vs. MIS), we have a total of 6x2 = 12 statistical tests. For each test, the null and the alternative hypotheses are defined in Table 1 below.

3.3 Subjects

The experiment was carried out as a midterm test in a course on OO Analysis and Design. Each subject was randomly assigned to one of four groups, differing by diagram-type (collaboration vs. sequence) and application-type (MIS vs. RT system). The theoretical material and the practical examples have been taken from the course textbook [2]. Before participating in the experiment, the students have studied two types of interaction diagrams, as well as other UML diagrams The case studies presented in the class covered a variety of applications including management information and real-time systems. Building interaction diagrams from use cases and class diagrams was part of the course homework assignments.

Seventy-six subjects participated in the experiment. This sample size enabled us to reach conclusions at the significance level of 0.05 and higher. In the first session of the experiment, which dealt with comprehension of diagrams, each subject has received a class diagram, a use case diagram and an interaction diagram (collaboration or sequence) of one of the systems. In the second session of the experiment, which dealt with quality of constructed diagrams, each subject has received the system requirements in the form of a brief narrative specification of the system, a class diagram and a use case diagram of one of the systems. This session assignment was to construct a sequence diagram or a collaboration diagram for the corresponding system.

Table 1: Statistical Tests Performed in the Experiment

Dependent Variable	MIS	RT system
C: score for diagram comprehension	Ha0: $\mu_{c, MIS, col}=\mu_{c, MIS, seq}$ Ha1: $\mu_{c, MIS, col}\neq\mu_{c, MIS, seq}$	Hb0: $\mu_{c, RT, col}=\mu_{c, RT, seq}$ Hb1: $\mu_{c, RT, col}\neq\mu_{c, RT, seq}$
Q: score for quality of diagram constructed	Hc0: $\mu_{q, MIS, col}=\mu_{q, MIS, seq}$ Hc1: $\mu_{q, MIS, col}\neq\mu_{q, MIS, seq}$	Hd0: $\mu_{q, RT, col}=\mu_{q, RT, seq}$ Hd1: $\mu_{q, RT, seq}\neq\mu_{q, RT, col}$
T_c: time (min.) spent on diagram comprehension	He0: $\mu_{tc, MIS, col}=\mu_{tc, MIS, seq}$ He1: $\mu_{tc, MIS, col}\neq\mu_{tc, MIS, seq}$	Hf0: $\mu_{tc, RT, col}=\mu_{tc, RT, seq}$ Hf1: $\mu_{tc, RT, seq}\neq\mu_{tc, RT, col}$
T_q: time (min.)spent on diagram construction	Hg0: $\mu_{tq, MIS, col}=\mu_{tq, MIS, seq}$ Hg1: $\mu_{tq, MIS, col}\neq\mu_{tq, MIS, seq}$	Hh0: $\mu_{tq, RT, col}=\mu_{tq, RT, seq}$ Hh1: $\mu_{tq, RT, seq}\neq\mu_{tq, RT, col}$
P_c: subjective comprehensibility of each diagram type	Hi0: $\mu_{pc, MIS, col}=\mu_{pc, MIS, seq}$ Hi1: $\mu_{pc, MIS, col}\neq\mu_{pc, MIS, seq}$	Hj0: $\mu_{pc, RT, col}=\mu_{pc, RT, seq}$ Hj1: $\mu_{pc, RT, seq}\neq\mu_{pc, RT, col}$
$P_{q:}$ subjective ease of construction of each diagram type	Hk0: $\mu_{pq, MIS, col}=\mu_{pq, MIS, seq}$ Hk1: $\mu_{pq, MIS, col}\neq\mu_{pq, MIS, seq}$	Hl0: $\mu_{pq, RT, col}=\mu_{pq, RT, seq}$ Hl1: $\mu_{pq, RT, seq}\neq\mu_{pq, RT, col}$

3.4 Assignment of Subjects to Treatments

In this experiment we have manipulated two factors (diagram and system), each having two possible levels (types). Consequently, we have used a “2x2” factorial design. Each subject was assigned randomly to one of the resulting four groups (“cells”), denoted as G1, G2, G3, and G4 respectively, and then measured two times, first in the comprehension session and then in the construction session. Both the type of diagram and the system were different from one session to the next. In other words, each subject performed a comprehension task using a certain type of diagram and a certain system; and later on he/she performed a task of constructing another type of diagram for another system. The treatment conditions of each group are shown in Table 2.

Table 2: Factorial Design: Assignment of Groups to Treatments

Group	Size	Session 1 (Comprehension)	Session 2 (Quality of Construction)
G1	19	System = RT Diagram = Sequence	System = MIS Diagram = Collaboration
G2	19	System = RT Diagram = Collaboration	System = MIS Diagram = Sequence
G3	20	System = MIS Diagram = Sequence	System = RT Diagram = Collaboration
G4	18	System = MIS Diagram = Collaboration	System = RT Diagram = Sequence

The total score for each session was calculated as follows:

- Each correct answer to a multiple-choice question has received one point. No points were deducted for a wrong answer. Hence, the maximum total score for the first session was five points.
- The maximum total score for the second session was 60 points and it was calculated over all diagram components. The number of points deducted for each mistake in a diagram component was based on the severity of mistake, depending on the different components of the diagram. We identified the following components of interaction diagrams: objects, messages, message sequences, links, and time line. The number of points deducted for each mistake was 2, 4, 6, 8, or 10, depending on the severity of a particular mistake. Totally omitting an object was considered the most severe mistake resulting in a deduction of 10 points. Some mistakes were relevant only to one type of diagram. For example, omitting message number was relevant only to collaboration diagrams, while not showing activity duration was relevant only to sequence diagrams.

4 Results

4.1 Analysis Strategy

In order to evaluate the hypotheses stated in Table 1 with respect to the six dependent variables, we used two-sided *t-test*, which is based on the following assumptions regarding the two samples:

- *Independence*. Based on our experimental settings, there is no reason to believe that the results of the subjects in the first group (e.g., those doing sequence diagrams) are in any way related to the results of the subjects in the other group (e.g., those doing collaboration diagrams).
- *Normal distribution*. The scores of every group in each session were tested for normality using the chi-square test. The results varied between $p = 0.054$ (Group 4, Session 1) and $p = 0.976$ (Group 1, Session 2). This means that no statistically significant departure from the normality assumption was found.
- *Equal variance*. We have compared the variances of every group in each session using F-test (see Table 3). Most groups were not found significantly different in terms of their variance. Based on the results of the F-test, we have applied equal variance (homoscedastic) or unequal variance (heteroscedastic) *t-test* as appropriate. In general, violation of the equal variance assumption is not problematic unless the two samples are quite different and one of the samples is small. In our study, we use nearly equal-size and relatively large groups (see Table 2 above).

The results of two-sided *t-tests* performed in the experiments are summarized in Tables 4-15 below. The two *t-tests* corresponding to each dependent variable were analyzed independently of each other (and not as a series of *t-tests*), since we are not interested in the overall differences between diagrams across the diverse application types. The minimum significance level for rejecting a null hypothesis was $\alpha = 0.05$. Detailed explanation and interpretation are provided in the sub-sections below.

Table 3: F-Test for Difference of Variances

Groups Compared	Comprehen-sion	Quality	Comprehen-sion Time	Construc-tion Time	Perceived Compre-hen-sibility	Perceived Ease of Construc-tion
1-2	0.686	0.137	0.343	0.278	0.722	0.863
1-3	0.276	0.750	0.348	0.205	0.556	0.637
1-4	0.009	0.791	0.514	0.291	0.517	0.927
2-3	0.135	0.071	0.060	0.833	0.808	0.758
2-4	0.003	0.227	0.776	0.961	0.746	0.799
3-4	0.102	0.561	0.114	0.879	0.925	0.591

4.2 Comprehension of Diagrams

The results of two-sided t-tests evaluating the difference in diagrams comprehension for the MIS and the RT systems are shown in Tables 4 and 5, respectively. According to Table 4, the difference in comprehension of the two diagram types is not statistically significant in the case of the MIS. On the other hand, Table 5 shows that in the case of the RT system, the comprehension score of the collaboration diagram is significantly higher than the score of the sequence diagram. Cell means for diagram comprehension in each system are shown graphically in Figure 2.

Table 4: T-Test of Diagrams Comprehension for MIS

Diagram Type	Mean	Standard Deviation	t	P value	Power	Effect Size
Sequence	4.250	0.966	0.103	0.919	0.156	0.143
Collaboration	4.222	0.646				

Table 5: T-Test of Diagrams Comprehension for Real-Time System

Diagram Type	Mean	Standard Deviation	t	P value	Power	Effect Size
Sequence	2.315	1.249	2.345	**0.025**	**1.000**	**3.227**
Collaboration	3.315	1.376				

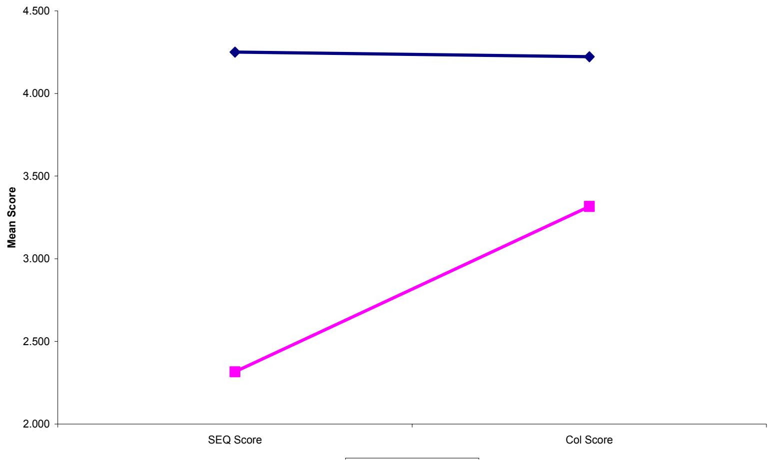

Figure 2. Cell Means for Diagram Comprehension

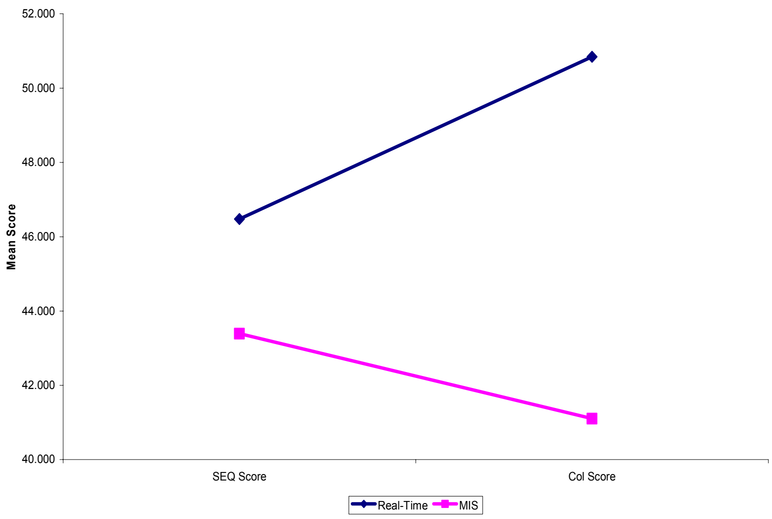

Figure 3. Cell Means for Diagram Quality

4.3 Quality of Diagram Construction

The results of two-sided t-tests evaluating the difference in diagram quality for the MIS and RT systems are shown in Tables 6 and 7, respectively. Table 6 depicts that in the case of the MIS the scores of the collaboration diagram are significantly higher than the scores of the sequence diagram. On the other hand, Table 7 indicates that in a RT system, the difference in the quality of the two diagram types is not statistically significant. Cell means for diagram quality in each system are shown graphically in Figure 3.

Table 6: T-Test of Diagram Quality for MIS

Diagram Type	Mean	Standard Deviation	t	P value	Power	Effect Size
Sequence	46.473	5.337	2.043	**0.048**	**1.000**	**2.813**
Collaboration	50.842	7.639				

Table 7: T-Test of Diagram Quality for Real-Time System

Diagram Type	Mean	Standard Deviation	t	P value	Power	Effect Size
Sequence	43.388	7.154	0.909	0.369	0.985	1.258
Collaboration	41.100	8.245				

4.4 Time Spent on Diagram Comprehension

According to the results of two-sided t-tests evaluating the difference in the time (number of minutes) spent on diagrams comprehension for MIS and RT systems (shown in Tables 8 and 9, respectively) there is no statistically significant difference between the average times spent on each diagram type in either application.

Table 8: T-Test of Comprehension Time for MIS

Diagram Type	Mean (minutes)	Standard Deviation	t	P value	Power	Effect Size
Sequence	28.450	6.328	0.962	0.342	1.000	1.313
Collaboration	26.000	9.235				

Table 9: T-Test of Comprehension Time for Real-Time System

Diagram Type	Mean (minutes)	Standard Deviation	t	P value	Power	Effect Size
Sequence	22.473	7.890	0.199	0.843	0.447	0.274
Collaboration	23.052	9.907				

4.5 Time Spent on Diagram Construction

According to the results of two-sided t-tests evaluating the difference in the time (number of minutes) spent on diagrams construction for MIS and RT systems (shown in Tables 10 and 11, respectively); there is no statistically significant difference between the average times spent on each diagram type in either application.

Table 10: T-Test of Construction Time for MIS

Diagram Type	Mean (minutes)	Standard Deviation	t	P value	Power	Effect Size
Sequence	56.812	19.613	0.101	0.299	1.000	1.584
Collaboration	65.128	26.157				

Table 11: T-Test of Construction Time for Real-Time System

Diagram Type	Mean (minutes)	Standard Deviation	t	P value	Power	Effect Size
Sequence	42.389	18.728	0.964	0.362	1.000	1.488
Collaboration	49.136	18.249				

4.6 Perceived Comprehensibility

In the post-test questionnaire, each subject has expressed the perceived comprehensibility of the diagram he/she analyzed. Comprehensibility was expressed using a 1-5 Likert ordinal scale, where the score of 1 indicated that the diagram was very comprehensible, while the score of 5 indicated that the diagram was absolutely incomprehensible. The questionnaires were summarized with respect to the test versions done by the subjects, i.e. by diagram types and application types. According to the results of two-sided t-tests evaluating the difference in comprehensibility scores for MIS and RT systems (shown in Tables 12 and 13, respectively), there is no statistically significant difference between the scores of each diagram type in either application.

Table 12: T-Test of Perceived Comprehensibility for MIS

Diagram Type	Mean	Standard Deviation	t	P value	Power	Effect Size
Sequence	2.889	1.131	0.398	0.693	0.743	0.591
Collaboration	2.733	1.099				

Table 13: T-Test of Perceived Comprehensibility for Real-Time System

Diagram Type	Mean	Standard Deviation	t	P value	Power	Effect Size
Sequence	2.778	1.308	0.538	0.594	0.655	0.750
Collaboration	3.000	1.201				

The correlation coefficient between the comprehension score and the perceived comprehensibility score of *the same* subject is 0.207. This correlation coefficient is not significantly different from zero (p-value = 0.086). It means that no relationship was found between subjects' perception and their actual performance in the test.

4.7 Perceived Ease of Construction

In the above post-test questionnaire, each subject has also estimated the perceived ease of constructing the diagram he/she has built in the test. The easiness score was given on the 1-5 scale, where the score of 1 indicated that the diagram was very easy to build, while the score of 5 meant that it was extremely difficult to build the diagram. The questionnaires were summarized with respect to the test versions done by the subjects, i.e. by diagram types and application types. According to the results of two-sided t-tests evaluating the difference in easiness scores for MIS and RT systems (shown in Tables 14 and 15, respectively), there is no statistically significant difference between the scores of each diagram type in either application.

Table 14: T-Test for Perceived Ease of Construction in MIS

Diagram Type	Mean	Standard Deviation	t	P value	Power	Effect Size
Sequence	3.368	0.895	0.699	0.489	0.899	0.977
Collaboration	3.166	0.857				

Table 15: T-Test for Perceived of Ease of Construction in Real-Time System

Diagram Type	Mean	Standard Deviation	t	P value	Power	Effect Size
Sequence	2.533	0.833	1.121	0.271	1.000	1.674
Collaboration	2.888	0.963				

The correlation coefficient between the construction score and the perceived ease of construction of *the same* subject is 0.054. This correlation coefficient is not significantly different from zero (p-value = 0.657). It means that, like in the case of comprehension, there is no relationship between subjects' perception and their actual performance in the test.

5 Discussion

With respect to comprehension of diagrams, the results of this study indicate that collaboration diagrams are easier to comprehend in the case of RT applications, but there is no difference in comprehension of the two diagram types in the case of MIS. Our findings contradict the findings of Otero and Dolado [9] who found that sequence diagrams are easier to comprehend in modeling synchronous real time systems. The fact that comprehension scores obtained in our study for the MIS were higher than those obtained for the RT system can be explained by the fact that our participants lacked adequate skills and training for comprehending real-time applications since students majoring in Information Systems Engineering are trained more in business-oriented applications than in real-time and embedded systems. Since the MIS students subjectively perceived the real-time system to be more complex, and using a collaboration diagrams the real time system appeared more comprehendible than using sequence diagrams, we may conclude the following: while it makes no difference which diagram type to use in the case of a simple system, there is an advantage in using collaboration diagram in the case of a more complex system for the population of Information Systems students and, probably, graduates as well.

With respect to quality of diagrams constructed by analysts, we found that in MIS, collaboration diagrams are significantly better than sequence diagrams, but there is no significant difference in the quality of diagrams produced for RT systems.

Our findings were not able to confirm a significant difference between the diagrams with regards to the time needed to construct them. But as we have seen, subjects spent more time on constructing interaction diagrams of an MIS than of a RT system. This result is consistent with the equivalent result on time spent on comprehending diagrams. As explained for the equivalent results, a possible reason for this may be that participants prefer spending time on a type of system which they are more familiar with (because they see a better chance to obtain a higher test grade on quality, as the amount of time spent on the task did not affect their test grade).

With respect to the perceived comprehensibility, the results we obtained are congruent with the expectations based on related work [7], [10], [11]. For the RT system, sequence diagrams yielded a better mean score than collaboration diagrams, and for an MIS, collaboration diagrams yielded a better mean score than sequence diagrams, though these differences were not statistically significant. Surprisingly, the results of perceived ease of construction were not congruent with our finding on comprehensibility; that is, collaboration diagrams yielded a better mean score for a RT application and sequence diagrams yielded a better mean score for MIS. Again, the differences were not statistically significant. Interestingly, the perceived ease of construction for real-time applications in general was better than for MIS applications. In practice, however, scores obtained by participants in constructing the MIS were better than for the RT system. This contradiction could be explained by the fact that participants were not very familiar with the RT application domain, thus underestimating the difficulty of modeling interaction in RT applications.

6 Conclusions

The implications of our study are in further investigating the contingency of UML interaction diagrams in terms of quality and comprehension for various application settings. The results from our controlled experiments suggests a rule of thumb for employing collaboration and sequence diagrams in modeling RT and MIS applications for both analysis and design stages. In some cases, however, our results contradict earlier studies [8], [9].

Our study suffers, however, from several limitations. First, the applications tested in the study (a library system and a security monitoring system) cannot be considered full-scale commercial applications. Second, the fact that only two applications were evaluated limits the external validity of this study. Third, the problems had a very slight difference in complexity and the results are assumed to depend, to a certain extent, on the problem description which might have caused some bias in the results; however for a given problem (MIS, RT) this should not have an effect on the findings with regards to the recommended interaction diagram. A common limitation of experimental research on model/method evaluation refers to the students participating, who in our case played both the role of users (in the "comprehension of diagrams" part of the experiment) and analysts (in the "quality of diagram construction" part).

Future research on this matter should first attempt to address some of the limitations mentioned above. Thus, a repeat experiment should use more experienced participants, possibly with professionals from the IT industry working on RT and MIS projects. The bias in favor of the MIS applications could be remedied by using a heterogeneous population of participants, including students majoring in Software

Engineering, who are more acquainted to RT systems than Information Systems Engineering students. It is also recommended to diversify and control the complexity of the applications in the experiment, in a similar method used by Otero and Dolado [9]. Regarding the diagram construction session, the quality of the diagrams produced by the participants could be further validated in practice by developing two full-scale versions of the same system: one based on a collaboration diagram and the other based on a sequence diagram. Another important issue for future research is to verify the fit (contingency) of sequence and collaboration diagrams for modeling various types of applications which incorporate a substantial amount of interaction with their clients, i.e., computer games, multimedia information kiosks, and customer relationship management (CRM) applications.

References

1. Booch, G., Rumbaugh, J. and Jacobson, I: *The Unified Modeling Language User Guide.* Reading, Mass: Addison-Wesley (1999).
2. Maciaszek, L.: *Requirements Analysis and System Design: Developing Information Systems with UML*. Harrow, Essex, UK: Pearson Education (2001).
3. Martin, R.C.: *UML Tutorial: Collaboration Diagrams*. Engineering Notebook Column, Nov./Dec. (1997).
4. McNeish, K.: *UML Collaboration Diagrams*. CoDe May/June (2002), 18-22.
5. Miller, J. and Mukerji, J.: *MDA Guide Version 1.0*. Available from http://www.omg.org/mda/mda_files/MDA_Guide_Version1-0.pdf (2003). (date accessed: 3/9/2004.)
6. Minium, E.W., Clarke, R.C. and Coladarci, T.: *Elements of Statistical Reasoning*. New York: John Wiley & Sons, Inc. 2nd Ed. (1999).
7. OMG *UML 1.5: The Current Official Version*. Available from http://www.uml.org/#UML1.5 (2003). (date accessed: 3/9/2004.)
8. Otero, M.C. and Dolado, J.J.: An initial experimental assessment of the dynamic modeling in UML. *Empirical Software Engineering* 7 (1) (2002), 27–47.
9. Otero, M.C. and Dolado, J.J.: Evaluation of the comprehension of dynamic modeling in UML. *Information & Software Technology* 46 (2004), 35-53.
10. Øystein H.: From MSC-2000 to UML 2.0 – The Future of Sequence Diagrams. SDL 2001, LNCS 2078 (2001), 38-51.
11. Stevens, P. and Pooley, R.: Using UML Software Engineering with Objects and Components. Addison-Wesley: NY. (1999).