In Search of Requirements Analyst Characteristics that Influence Requirements Elicitation Effectiveness: a Quasi-experiment

Alejandrina Aranda
DLSIIS, Facultad de Informática
Universidad Politécnica de Madrid
Boadilla del Monte, Spain
am.aranda@alumnos.upm.es

Oscar Dieste
DLSIIS, Facultad de Informática
Universidad Politécnica de Madrid
Boadilla del Monte, Spain
odieste@fi.upm.es

Natalia Juristo
DLSIIS, Facultad de Informática
Universidad Politécnica de Madrid
Boadilla del Monte, Spain
natalia@fi.upm.e

***Abstract—Context*: Elicitation effectiveness depends on non-well-understood analyst's skills and abilities. Identifying which analysts' characteristics have stronger influence on elicitation may help to improve requirements quality. *Objective*: Identify the analysts' characteristics that influence on the elicitation effectiveness. *Method:* We analyzed the impact of: the analyst's experience in interviews, elicitation and requirements; their academic qualifications, the familiarity with problem domain and the time spent during the elicitation session in the effectiveness of the elicitation and subsequent consolidation of requirements, using a quasi-experiment. *Results:* The knowledge of the problem domain, the analysts' academic qualifications and the elicitation time do not appear to influence the effectiveness of the elicitation sessions. The analyst's experience exerts a slight negative influence on the effectiveness of the elicitation session. The analyst's experience and familiarity with problem domain adversely affect the consolidation process. Finally, the analyst's academic qualifications have a strong positive impact (statistically significant) on the effectiveness of the consolidation process. *Conclusions:* Although the evidence is still scarce, it seems the analyst's confidence on his own experience may be harmful in some cases. Specific training in software requirements may yield much higher gains than non-specific analyst experience.**

Keywords-Quasi-experiment; Requirements elicitation; Effectiveness; Experience

I. INTRODUCTION

NOWADAYS, there are a wide range of elicitation techniques for gathering requirements, including brainstorming [1] protocol analysis [2], JAD [3], etc. In practice, interviews are the most commonly used method [4] and just about the de rigueur software requirements elicitation technique [5]. There are several types of interview: structured and unstructured [6], cognitive [7], semantic [8], etc. However, there can be no doubt that the unstructured (also frequently referred as to open) interview is the most often used [9]. The primary characteristic of an unstructured interview is that it has no previously designed schedule (set of questions, list of items to be explored) [4]. As the analyst decides on the course of the questioning, the effectiveness of unstructured elicitation is strongly dependent on the person doing the interviewing.

In the field of software engineering, relatively little is known about how analysts work with requirements [10]. Some papers do, however, explore analysts' influence on interview effectiveness. Experience is the mostly commonly researched concept [8], [11]. The empirical evidence gathered is far from conclusive, but it has been consistently observed that experience affects requirements elicitation, albeit contrary to the dictates of commonsense: inexperienced analysts are equally or more effective than experienced analysts.

Apart from experience, other aspects, such as introversion/extroversion or independent/dependent [12] have also been studied, though to a lesser extent. In the latter two cases, effects are not as clear and, contrary to the case of experience, are not enough to venture any conclusion. Finally, other characteristics (like academic qualifications, familiarity with the problem) that might have a bearing and have not yet been explored spring to mind.

Software system quality largely depends on requirements quality. Therefore, all aspects capable of improving interview effectiveness are of utmost importance. The aim of this article is to empirically check which requirements analyst characteristics are likely to influence requirements elicitation. To do this, we have conducted a quasi-experiment with Universidad Politécnica de Madrid master's students. These students used the unstructured interview technique to complete

an elicitation session on a problem with which they were unfamiliar. At the end of the quasi-experiment, we administered a demographic questionnaire which we used to gather data related to student characteristics, such as interview, elicitation and requirements experience, academic qualifications, analyst familiarity with the problem domain and time taken to complete the elicitation session. We measured subject elicitation effectiveness in two different ways: by the amount of information mentioned during the elicitation session and by the amount of information reported in writing in the later interview consolidation.

We applied statistical correlation to identify the relations between analyst characteristics and elicitation effectiveness. The results indicate that analysts' academic qualifications influence (are highly correlated and statistically significant) requirements consolidation process effectiveness; that is, analysts trained in systems engineering or computing science retain and consolidate more information than analysts from other backgrounds. On the other hand, we have found negative trends, that is, characteristics that common sense dictates should have a positive influence actually have a negative bearing. For instance, we find that less experienced subjects are slightly more likely to mention more information in the elicitation session than subjects more experienced in requirements activities, whereas experience and familiarity with the domain are liable to have a negative bearing on the consolidation process and information retention capability, that is, inexperienced subjects that are unfamiliar with the problem consolidate and retain more information than more experienced analysts that are familiar with the problem. Finally, aspects that should ostensibly have a bearing, such as subject familiarity with the problem domain and academic qualifications, have no influence whatsoever on elicitation session effectiveness.

The article is structured as follows. Section II discusses work related to the research reported here. Section III describes the quasi-experiment run and the analysis procedure. Section IV describes the results. Section V lists the experiment validity threats. Finally, Section VI outlines the conclusions.

II. RELATED WORK

In view of its importance and complexity, many papers have addressed the requirements elicitation activity. However, recent systematic reviews [13], [14] have shown that there are few empirical studies aimed at studying the effectiveness and efficiency of interview techniques. [15], [16], [17], [8], [11] report empirical studies focused on verifying whether analysts' characteristics have a bearing on the elicitation and consolidation of customer needs.

Pitts and Browne [11] designed an experiment in the information systems field to examine the cognitive stopping rules used by analysts to determine when the elicited requirements are sufficient for system development to go ahead. As far as we are concerned, they also studied the influence of analyst experience on requirements elicitation. Fifty-four professional analysts with at least two years' experience in systems development participated in the experiment. Subjects had to be as experienced as possible to make the effects more perceptible. Experimenters analysed the influence of experience in terms of the number, breadth and depth of elicited requirements. As a result, they reported that analyst experience does not influence the determination of the requirements, that is, the number, breadth and depth of requirements do not depend on how many years' experience analysts have. This quite surprising result is not, however, an exceptional case, as we shall see later.

Studies by Burton, Shadbolt, Hedgecock and Ruggn. [18] and Corbridge, Rugg, Major, Shadbolt and Burton [12] empirically compared the effectiveness of four elicitation techniques (including the open interview technique) in terms of number of clauses gathered and time taken to complete, transcribe and codify the elicitation session. They also analysed the effect of the expert's personality and cognitive style on the result of the elicitation session. Thirty-two senior geology and medical students, respectively, participated in each experiment. Their experience was confined to their academic training in the identification of igneous rocks and the diagnosis of abdominal conditions. The EPI (Eysenck Personality Inventory) and EFT (Embedded Figures Test) psychometric tests, respectively, were used to evaluate the individual's personality characteristics (introverted /extroverted subject) and cognitive style (dependent or independent). Among other results, researchers reported in the first experiment that introverted people elicited more information than extroverted people in less time [18]. This again is quite remarkable, as common sense again dictates just the opposite, that is, that extroverted people should be more effective. This effect is not observed in the second experiment, however [18]. Additionally, subject cognitive style has no significant effects in either study, that is, interview effectiveness does not rely on whether the subject is dependent or independent.

Marakas and Elam [8] designed and executed an experiment in the information systems field with the aim of evaluating the effectiveness of the semantic interview technique (a type of semi-structured interview) and the unstructured interview. Twenty inexperienced and experienced subjects participated in the experiments. Experience was measured by how many years the subject had been working in systems analysis and software development. Inexperienced subjects were final-year master in software engineering students, whereas experienced subjects were professionals: systems analysts and software developers. Experimenters managed to identify differences of effectiveness between the two interview types (specifically, subjects were more effective using the semantic interview). However, experienced subjects were only marginally better (about 3%) than novice subjects, irrespective of the interview type used, and, in any case, the differences were not statistically significant.

Finally, Agarwal and Tanniru [15] experiment in the expert systems field with the aim of comparing the effectiveness of structured and unstructured interviews in terms of number of elicited rules, among other measures. Thirty subjects with different levels of experience participated in the experiment: a group of novice knowledge engineers trained in the

unstructured interview, a group of expert knowledge engineers that applied the unstructured interview. Novice subjects were students, whereas experts were professionals with knowledge engineering experience and experience working on at least one expert system or analysts with at least three years' systems analysis experience. For reasons of experimental design, subjects only used unstructured interviews. Researchers reported that experienced subjects (who used unstructured interviews) were slightly better (17%) than inexperienced subjects (using unstructured interviews), but the statistical differences were not significant. In sum, the studies conducted revealed that experience did not have any positive bearing on the effectiveness of the interviews held by the analysts. This means that either there is no such effect or it is very small (which would explain why no significant differences have been obtained in any case). As regards other factors, such as personality, what little evidence is there is contradictory.

III. Experimental Design

A. Aim, Research Question, Variables and Metrics

The aim of this research is to further analyse how some analyst characteristics influence requirements elicitation effectiveness. We study the following characteristics: analyst familiarity with the problem domain, academic qualifications and experience. We also examine whether the time taken by analysts to complete the elicitation session somehow influences the amount of gathered and consolidated information. We have not studied the influence of personality factors, like introversion and extroversion, in this quasi-experiment, because they are intrusive.

In view of the above points, **the research question** that we will try to answer in this paper is:

- RQ: Which requirements analyst characteristics influence elicitation effectiveness?

To be able to answer this research question we have conducted an empirical study in the shape of a **quasi-experiment.** Quasi-experiments are conducted when subjects cannot be randomly assigned to an experimental condition or, alternatively, a treatment cannot be assigned to a group. This applies here, because the variables are inherent characteristics of the experimental subjects and cannot be randomized or blocked.

Each experimental subject in the quasi-experiment has elicited the requirements on a specific problem. Requirements elicitation was divided into two major stages: the elicitation session and the consolidation process. **Elicitation session** means the process during which the requirements analyst interacts and confers with customers to gather information about their needs and reach agreement on what the software system to be developed should do. The **consolidation process** is the process by which requirements analysts take in and report the information gathered in the elicitation session, that is, specify customer needs in formal terms.

So, elicitation effectiveness will be determined in two different ways. We will take into account, first, the effectiveness of subjects during elicitation sessions and, second, the effectiveness of subjects during the process of consolidating the elicited information. Additionally, although the requirements elicitation and consolidation processes are consecutive, they do not have to be enacted in quick succession. The consolidation process can be postponed for hours or even days. Even though this is contrary to generally accepted recommendations about how to perform interviews because there is a risk of analysts forgetting important information, there can be no denying that such delays do occur. On this ground, despite the fact that it is not strictly speaking a measure of effectiveness, our study includes retention capability. Retention capability refers to how much valid information requirements analysts are able to consolidate from the information mentioned during the elicitation session.

Thus, we were able to further refine the research question, which we divided into several sub-questions:

- $RQ_{1.1}$ to $RQ_{1.4}$: Does the analyst's [Familiarity with the Problem | Academic Qualifications | Experience | Session Time] influence the effectiveness of the elicitation session?
- $RQ_{2.1}$ to $RQ_{2.4}$: Does the analyst's [Familiarity with the Problem | Academic Qualifications | Experience | Session Time] influence the effectiveness of the information consolidation process?
- $RQ_{3.1}$ to $RQ_{3.4}$: Does the analyst's [Familiarity with the Problem | Academic Qualifications | Experience | Session Time] influence information retention?

In our study we have evaluated three dependent and four independent variables. The selected **dependent variables** or response variables should answer the research question, as shown in Table I. These variables refer to effectiveness, **measured** according to the amount of information mentioned, retained and identified (or, at least, specified in the reports resulting from the consolidation process) by subjects during the elicitation process. These variables are obtained by analysing the elicitation sessions that have been recorded during the quasi-experiment, as well as the reports submitted by the subjects as a result of the requirements consolidation process that they have enacted.

TABLE I. Dependent Variables

Dependent Variable	Metric	Evaluated Item
Eliciation session effectiveness (SessionEffec)	Number of domain elements mentioned during the elicitation session (over total number of domain elements)	Interviews (recorded elicitation sessions)
Information consolidation process effectiveness (ConsolidationEffec)	Number of domain elements described by subjects during their eliciation session consolidation (over total number of domain eleements).	Reports (list of requirements and consolidated information)
Retention Capability (RetentionCapa)	Number of consolidated elements over number of elements mentioned in the elicitation session	Elicitation sessions and reports

The **independent variables** or factors, listed in Table II, refer to the above-mentioned analyst characteristics. Except for time, which was determined from the elicitation session recordings, the values of these variables have been taken from a demographic questionnaire administered to each subject at the end of the quasi-experiment.

TABLE II. INDEPENDENT VARIABLES

Independent Variables	Description
Familiarity with the problem	Subjects' familiarity with the problem domain.
Academic qualifications	Subjects' theoretical knowledge of the software development field
Interview eliciation and requirements experience	Subjects' professional experience measured in years
Elicitation time	Elicitation session duration

B. Subjects

The quasi-experiment participants were 12 Master in Software Engineering students from the Universidad Politécnica de Madrid's School of Computing (Spain) enrolled in the requirements engineering course unit. They are all professionals working in computing and related areas from several Latin American countries (Argentina, Brazil, Colombia, Ecuador, and Peru) and Spain. The subjects were experienced in systems development and also in requirements elicitation and analysis tasks. They stood to benefit by executing the experimental task properly, as the quasi-experiment overlapped with the practical assignments for course unit assessment. Through this practical assignment, students had to demonstrate that they were capable of eliciting information about a problem using the open interview technique. They also had to demonstrate that they were capable of analysing and specifying the information gathered from the elicitation session as functional and non-functional requirements for the system to be developed.

C. Preparation of the Experiment

The subjects were set the problem of eliciting requirements for a battery recycling plant control system. We intended to explore the effect of familiarity with the problem domain, thus the selected problem had to be quite out of the ordinary to assure that a fair number of subjects had little or no experience in the domain. The problem description is available in the web appendix at http://www.grise.upm.es/sites/extras/8/. The full description of this problem covers four types of elements: business goals, system requirements, concepts and processes. The elements into which the problem domain was divided have been used as a baseline checklist against which researchers measured the effectiveness of the elicitation (either during the elicitation session or later consolidation).

All the elicitation sessions were carried out according to the open interview technique within a set time limit. All sessions lasted at most 25 minutes, where the first five minutes of the session were used to explain the goals of the session and the problem to be dealt with to the subjects and the other 20 minutes was the interview proper. In the elicitation sessions, subjects played the role of the requirements analyst and the experimenter the role of the customer.

The elicitation sessions were scheduled over a two-day period so as not to overtax experimenters. The experimental subjects could not be randomly assigned to sessions because of their work commitments.

D. Execution

The experiment was executed according to the planned schedule. A total of 12 interviews were conducted over the two-day period. Some sessions has to be rescheduled because some subjects failed to respect the schedule.

All the interviews were conducted individually, recorded and delivered to each experimental subject. At the end of the quasi-experiment, the subjects completed a demographic questionnaire (available in the web appendix at http://www.grise.upm.es/sites/extras/8/) from which we were able to gather the values of the independent variables. Finally, the subjects submitted a consolidation report on the elicited information a week after the interviews were held. Students were not asked to specify the time between elicitation and consolidation, as there was a high risk of them supplying biased information.

Note that two subjects failed to complete the questionnaire designed to gather demographic data and three subjects did not to submit the consolidation report. This meant that the analysed experimental population was reduced to 10 subjects for the elicitation session and seven subjects for the consolidation process.

E. Analysis Procedure

We have calculated descriptive statistics and built illustrative charts: box and scatter plots based on the experimental data (from the recordings of the elicitation sessions, the consolidated reports and the demographic questionnaires). Although the sample size is very small, we have also calculated correlations and applied the Mann-Whitney test with a significance level set at $\alpha = 0.05$ to evaluate the dependency and significance of the results. The statistical analysis was conducted using the SPSS package.

F. Data Preparation

To analyse **elicitation session effectiveness**, we first transcribed all 12 interviews to facilitate the later analysis of the information mentioned during the elicitation session.

At the end of the process, we identified the problem domain elements in the text of the transcriptions (for more details on the domain elements, see the web appendix). The annotations made in the transcriptions of the interviews were condensed into a summary table.

As for the procedure enacted in the elicitation process, we have gathered the data from the reports submitted by the subjects to evaluate the effectiveness of the **consolidation process**, following almost the same reference criteria as above. To evaluate subjects' **retention capability**, we have calculated

the difference between the results obtained in the elicitation and consolidation sessions.

We have dichotomized the independent variables and associated a dummy variable (0 and 1) with each group to facilitate the analysis and comprehensibility of the results. The dichotomizations are useful for comparing groups and building box plots (as we will see in Section IV). These are generally more easily interpretable methods than scatter plots and correlations, and the analysis of the results is no less rigorous [19]. The dichotomization procedure was as follows:

- With respect to their familiarity with the problem, subjects had the options of saying that they were unfamiliar or familiar with or knowledgeable about the problem. Most answered that they were unfamiliar or familiar with the problem, although one answered that he or she was knowledgeable. For the purposes of statistical analysis, we grouped the subjects that were familiar with and knowledgeable about the problem domain in one group, reducing problem familiarity to two levels: familiar (1), unfamiliar (0).
- Regarding academic qualifications, subjects with degrees in systems engineering and computing science were considered as computing subjects (1), whereas the holders of all other degrees (civil engineering, industrial engineering, mathematics, economics and others) were classed as non-computing subjects (0).
- As regards subjects' interview, elicitation and requirements experience, we have grouped the experimental population into: experienced subjects (1) and inexperienced subjects (0). Experienced subjects are subjects that have been working on such activities for more than two years, whereas inexperienced subjects have worked in interviewing, elicitation and requirements-related activities for less than one year.

The data gathered are available in Appendix A. The same data after the dichotomization process is available in the web appendix.

IV. RESULTS

In this section we will answer each of the research questions stated in section III. For each of the elicitation aspects (session effectiveness, consolidation effectiveness, retention capability), we have analysed the influence of the subjects' characteristics (familiarity with the problem, academic qualifications, professional experience in requirements activities), as well as time taken to complete elicitation. The analysis was primarily conducted informally, using box plot diagrams built by dichotomizing the empirical data. We have not placed too much emphasis on the statistical analysis as the sample is very small (12 students) and is therefore susceptible to small sample effects that can cause false-positives and false-negatives. However, we have calculated the correlation coefficients and Mann-Whitney test. They are available in appendix B. When appropriate, we will reference the corresponding p-values in the text.

1) Elicitation Session

$RQ_{1.1}$ *Does an analyst's familiarity with the problem influence elicitation session effectiveness?*

Subjects familiar with the problem are likely to mention more information in elicitation sessions than subjects that are unfamiliar with the problem, as shown in Figure 1. Subjects those are familiar with the problem mention on average 51% of domain elements. Subjects unfamiliar with the problem mention on average 48%.

As Appendix B shows, the results were not statistically significant. The significance value (p-value = 0.46 > 0.05) is greater than the specified level. Therefore, the between-group difference in effectiveness is not significant. We conclude that **familiarity with the problem does not influence elicitation session effectiveness**.

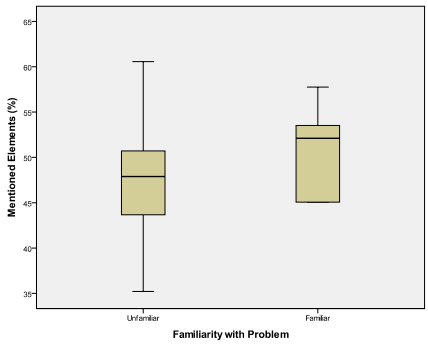

Figure 1. Boxplot: effect of familiarity with problem on number of mentioned elements.

$RQ_{1.2}$ *Do an analyst's academic qualifications influence elicitation session effectiveness?*

As Figure 2 shows, computing subjects and non-computing subjects both gather on average a similar amount of information, that is, the same amount of information (49%) is mentioned on average in elicitation sessions led by computing subjects and non-computing subjects. The interquartile range is almost the same and the fact that the maximums and minimums are greater for computing subjects may be simply due to small sample effects (for example, the median for non-computing subjects is, contrary to what one might think, greater than for computing subjects).

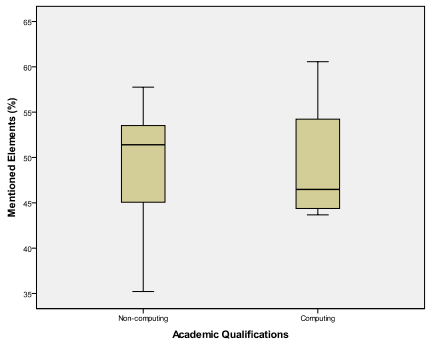

Figure 2. Boxplot: effect of academic qualifications on the number of mentioned elements.

The statistical tests (p-value = 0.75 > 0.05) corroborate that there are no significant differences between both groups. It can be said therefore that the **subjects' qualifications have no effect on the number of domain elements mentioned during the elicitation sessions**, that is, an elicitation session by a computing subjects is not necessarily more effective than a session by a non-computing subjects.

$RQ_{1.3}$ *Does an analyst's professional experience influence elicitation session effectiveness?*

Inexperienced subjects are likely to mention more information during the elicitation session than subjects experienced in requirements activities, as shown in Figures 3 and 4.

Figure 3 shows that the value of the median, the interquartile range and the maximums and minimums are greater in the group of inexperienced subjects. These results clearly indicate that more information is likely to be mentioned in elicitation sessions by subjects without interview experience than in sessions by inexperienced subjects. The associated statistical tests suggest that there really is such a tendency (p-value= 0.09), although it is not significant.

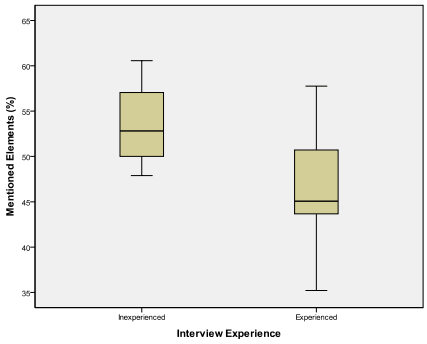

Figure 3. Boxplot: effect of interview experience on the number of mentioned elements

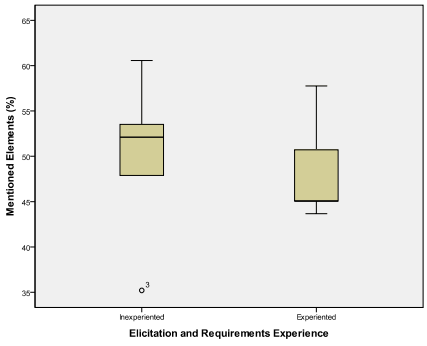

Figure 4. Boxplot: effect of eliciation and requirements experience on the number of mentioned elements

Figure 4 illustrates that the performance of subjects with elicitation and requirements experience was similar, that is, there is a slight tendency for more elements to be mentioned in elicitation sessions by subjects without elicitation experience than in sessions where subjects had elicitation and requirements experience. As above, these results are not statistically significant (p-value = 0.46 > 0.05).

We can conclude that **subjects' professional experience has a slightly negative influence on elicitation session effectiveness.**

$RQ_{1.4}$ *Does elicitation time influence elicitation session effectiveness?*

Although elicitation time is not a personal characteristic (like familiarity with the problem, professional experience and academic qualifications), it has been considered as another independent variable because, although a time limit was set for the elicitation session, only two subjects used up the allotted time (20 minutes), as shown in Appendix A. The implication is that the experimental subjects decided of their own accord how much time they required for elicitation (from 8 to 24 minutes). Consequently, this is a completely legitimate independent variable. Time to some extent denotes the elicitor's impatience or self-assuredness, an issue whose effect on the number of elements mentioned during the elicitation session and later in the consolidated information we believe to be worthwhile studying.

The results show, that there is no significant direct relation between the number of elements mentioned during and the time taken to complete the elicitation session. Specifically, the correlation coefficient is very small and far from being significant (r=0.21; p-value=0.57 > 0.05). In other words, the number of mentioned domain elements does not appear to increase with elicitation session duration. Therefore, the **interview duration does not determine elicitation effectiveness.**

2) Consolidation Process

$RQ_{2.1}$ *Does an analyst's familiarity with the problem influence information consolidation process effectiveness?*

As Figure 5 shows, subjects unfamiliar with the problem are more likely to be able to consolidate more domain elements than subjects familiar with the problem. If we compare this tendency with the result for the elicitation session, the effect observed during the consolidation process is contrary: **subjects unfamiliar with the problem domain consolidate more information than subjects that are familiar with the domain**. On average, subjects unfamiliar with the domain consolidate 39% of the domain elements mentioned in their elicitation sessions, whereas subjects familiar with the problem consolidate 24% of elements. These results are not significant, as shown by the statistical tests presented in Appendix B.

$RQ_{2.2}$ *Do analyst's academic qualifications influence information consolidation process effectiveness?*

We have found that there is a marked dependency (statistically significant) between the subjects' academic qualifications and the number of consolidated elements. Despite the small sample size, the effect of qualifications is evident: computing subjects gather on average more problem domain elements (43%) than non-computing subjects (19%).

Note that not only is this effect visually noticeable, but it is also verified experimentally.

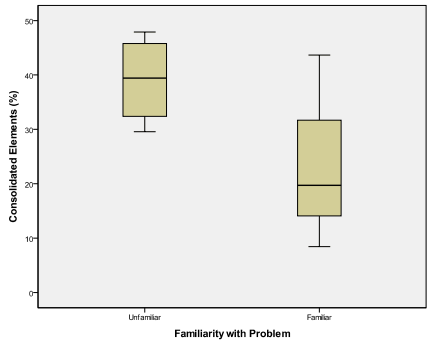

Figure 5. Boxplot: effect of subject familiarity with the problem on number of consolidated elements

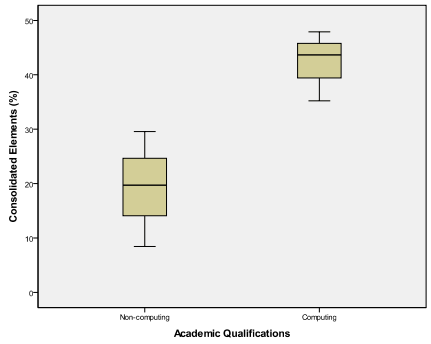

Figure 6. Boxplot: effect of academic qualifications on number of consolidated elements

Specifically, the significance value (p-value = 0.03 < 0.05) is smaller than the established significance level; this result indicates that, statistically speaking, the differences between means is significant. Therefore, we can conclude that **there is a relation between academic qualifications and the percentage of information gathered during the consolidation process**.

$RQ_{2.3}$ *Does the analyst's professional experience influence information consolidation process effectiveness?*

Figure 7 shows the number of domain elements consolidated depending on subject experience in requirements activities: interviews, elicitation and requirements.

Figure 7 shows that while most subjects have professional experience in requirements activities, their experience is highly variable and widely dispersed; however, the maximum and minimum values and medians obtained are greater for inexperienced subjects. Inexperienced subjects consolidate on average 46% of the elements mentioned in their elicitation sessions. Experienced subjects consolidate 27% of the elements. This is indicative of a clearly negative tendency where **inexperienced subjects consolidate more information than experienced subjects**. As the significance value (p-value = 0.08 > 0.05) is greater than the established value, the difference between the means is not statistically significant. Note that more experienced subjects would be expected to consolidate more information. However, the results suggest just the opposite.

Figure 7. Boxplot: effect of professional interview, elicitation and requirements experience on the number of consolidated elements

$RQ_{2.4}$ *Does elicitation time influence information consolidation process effectiveness?*

Note that we cannot compare the time that subjects spent on the consolidation of the domain elements, as we have not gathered that information during the experiment. However, we have analysed the influence of elicitation time on the number of consolidated elements, that is, we have checked whether subjects that took longer to complete their elicitation session consolidated more information.

The results show, that there is no direct relation between the duration of the interviews and the number of consolidated elements. In other words, the correlation coefficient indicates that there **is a slight tendency for subjects to consolidate more information in less time**. However, this result is not significant as the significance level (p-value = 0.39 > 0.05) is greater than the established level.

3) Retention capability

The empirical study was designed to widen the distance between elicitation and consolidation. Specifically, subjects were given a week to submit the interview consolidation. As most subjects were working at the time of the study, this surely encouraged subjects to postpone consolidation. Note that subjects had access to a recording of their interviews for the purposes of consolidation. This did not, however, prevent them from forgetting information, as we will see now.

We have used the retention capability metric, that is, how many problem domain elements mentioned in the interview was the subject capable of consolidating, to quantify subject effectiveness.

$RQ_{3.1}$ *Does an analyst's familiarity with the problem influence information retention?*

Analysing the values in Figure 8, we find that the group that is unfamiliar with the problem recalls on average more information than the group that is familiar with the problem domain. And, although the maximums are higher for the group that is familiar with the problem, the median of the group that

is unfamiliar with the problem is, contrary to what one might think, higher than for the other group. This behaviour may be simply due to small sample effects (7 subjects). Note that not all subjects (12 students) can be analysed as they did not all submit the final report with the consolidated information. Additionally, some subjects did not complete the demographic questionnaire.

Although the results are not statistically significant (p-value = 0.48 > 0.05), we conclude that there is tendency for **subjects that are unfamiliar with the problem domain to consolidate more information than subjects familiar with the domain.**

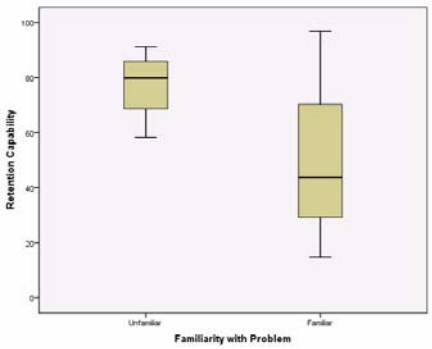

Figure 8. Box-plot: effect of familiarity with the problem on retention capability

$RQ_{3.2}$ *Do an analyst's academic qualifications influence information retention?*

Despite the small sample size, qualifications have a patent effect as they did on the information consolidation process: computing subjects are capable of retaining on average 87% of the domain elements mentioned in their interviews, whereas non-computing subjects retain 39% of the information, as shown in Figure 9.

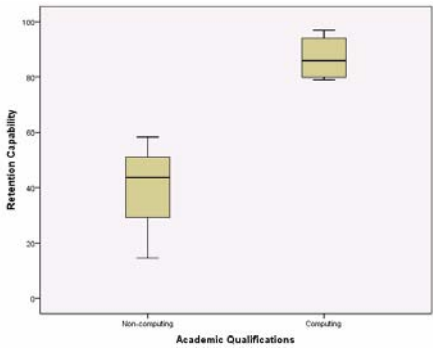

Figure 9. Boxplot: effect of academic qualifications on information retention capability

This result is statistically significant, that is, the difference between the means is significant because the significance value (p-value = 0.03 < 0.05) is smaller than the established significance level. Therefore, we can say that **there is a relation between academic qualifications and subjects' retention capability.**

$RQ_{3.3}$ *Does an analyst's professional experience influence information retention?*

As for elicitation session and information consolidation effectiveness, we found, contrary to expectations, that subjects inexperienced in requirements activities consolidate more of the information mentioned in the elicitation session.

Figure 10 shows that inexperienced subjects are likely to retain more information than experienced subjects. Although the experienced group is more disperse and scores the maximum value; a visual inspection shows that inexperienced subjects recall on average more information, and both maximum and minimum values are high. However, this result is debatable because of the small sample size. As the significance value (p-value = 0.43 > 0.05) is greater than the established value, the difference between means is not statistically significant. However, we conclude **analyst experience is likely to have a negative influence on information retention.**

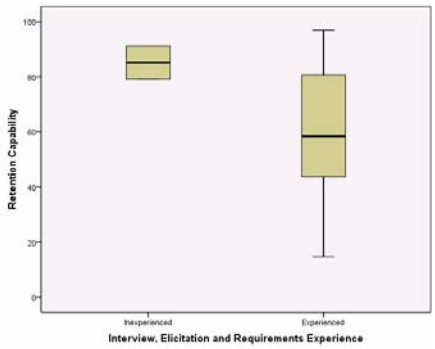

Figure 10. Boxplot: effect of interview, elicitation and requirements experience on subject retention

$RQ_{3.4}$ *Does elicitation time influence information retention?*

There is no relation between retention capability and elicitation time. This means that subjects' effectiveness at recalling elements during information consolidation does not depend on interview duration. This result is not statistically significant, as the significance value (p-value = 0.35 > 0.05) is greater than the established significance level.

V. VALIDITY THREATS

There are three main threats to the validity of the results:

1. The first and most evident is the type of design. Subjects are not randomly assigned to treatments in a quasi-experiment. This increases the risk of extraneous variables being confounded with independent variables, meaning that the observed effects may be wrongly attributed to the independent variables.
2. The small sample size only adds to this problem by exaggerating any influence of confounding variables over and above what it would probably be in a larger population.

3. The subjects are taken from a convenient rather than a random sample (that is, subjects are students enrolled for a particular course unit rather than having been recruited from a larger population). Generalization of our results to analysts as a whole can be done only with due caution.
4. The experimental setting was quite different to what professional analysts may be used to: simulated customer, limited time and no previous knowledge about the problem to solve (that is, the battery recycling plant control system).
5. The consolidation of the information gathered in the elicitation session was not under experimental control. It was possible that students prepared their consolidation reports at different times (from 1 hour to 7 days) or even cheat.

There first three threats cannot be addressed in any other way than changing the study design and the population where the subjects were drawn. Threat #3 is mitigated by the fact that the master course where the quasi-experiment was performed had a professional orientation. The experienced students enrolled in the master were practitioners (developers, analysts), and therefore their results can be considered representative of their kind.

We believe that threat #4 has a marginal effect at most. Students thought they were carrying out a practical assignment (debriefing happened after the consolidation reports were submitted). They were highly motivated and performed professionally. Time was not an obstacle to elicit information (most of the students finished the elicitation session before running out of time).

Threat #5 is delicate. Retention capability cannot be studied unless a relatively long period between elicitation and consolidation takes place, but this period may cause confusion of variables. For instance, maybe the subjects with less experience didn't have a demanding job, so they could meet to compare their reports. This may be a good explanation of the apparent poor performance of experienced subjects. We need to repeat the quasi-experiment with different design (e.g.: submission of the consolidation report within 2-hours after elicitation) to find out.

VI. Conclusions

This paper reports a quasi-experimental empirical study intending to check whether an analyst's elicitation effectiveness is influenced by particular requirements analyst characteristics. Most of the results turned out not to be statistically significant, but, even so, we have attempted to identify trends.

It is surprising that contextual variables that should apparently have a positive influence on the results went against common sense. We have found that there is a tendency for subject experience to have a slight and negative influence on the amount of information gathered in the elicitation session, whereas subjects' familiarity with the problem domain and academic qualifications have no influence whatsoever. Subjects' familiarity with the problem and professional experience are likely to have a negative influence on consolidation processes and information retention, that is, subjects that are unfamiliar with the problem domain and inexperienced in requirements activities consolidate and retain more information than subjects that are familiar with the problem and have experience. Finally, an analyst's academic qualifications do have a significant positive influence on the interpretation and consolidation of the information mentioned in the elicitation session. We have found that elicitation session duration has no bearing on any of the three measures of elicitation process effectiveness.

Note that because the number of experimental subjects is small, the results may simply be due to small sample effects. More experiments are necessary to check the validity of our findings.

Likewise, we need to keep in mind that the current study has been designed with the intention of minimising the potential experience's positive effects. In particular, the problem to solve (the battery recycling plant control system) was atypical. We found that experience does not constitute an advantage in such situation (actually, it has pernicious effects). However, it is possible that experience has strong positive effects in more familiar problems. Again, we need to carry out more experiments to find out.

Transferring these results to industry, requirements should preferably be elicited by analysts or engineers trained in computing science, as analysts with these characteristics can be expected to be better able to understand customer needs, plan more effective elicitation sessions and, consequently, consolidate quality software requirements. On the other hand, experienced analysts should preferably not rely on their own personal skills experience when eliciting or consolidating information, but should make the deliberate effort to understand the problem, even if it is similar to other cases that they have come across before.

In the future we aim to conduct further empirical studies and experiments, again comparing elicitation effectiveness against these and other characteristics, such as subjects' and experts' command of expression and quickness of understanding, fatigue, applied elicitation technique, consolidation time, etc., in order to verify and check the validity of the results.

Acknowledgment

This work has been funded in part by the Itaipu Binacional Postgraduate Grant / Fundación Parque Tecnológico Itaipu and the Spanish TIN2011-23216 project.

References

[1] A. F. Osborn, *Applied Imagination: Principles and Procedures of Creative Problem-Solving.* New York, NY: Charles Scribner's Sons, 1963.

[2] K. A. S. H. A. Ericsson, *Protocol Analysis: Verbal Reports as Data.* Cambridge, MA, USA: MIT Press, 1996.

[3] J. Wood and D. Silver, *Joint Application Design: How to Design Quality Systems in 40% Less Time.* New York: Wiley, 1989.

[4] S. E. Hove and B. Anda, "Experiences from conducting semi-structured interviews in empirical software engineering research," in *Software Metrics, 2005. 11th IEEE International Symposium,* 2005, pp. 10 pp.-23.

[5] N. J. Cooke, "Knowledge elicitation," in *Knowledge Elicitation*, F. Thomas Durso and R. S. Nickerson, Eds. Chichester, England: John Wiley & Sons Ltd, 1999, pp. 479-510.

[6] G. Kotonya and I. Sommerville, *Requirements Engineering.* New York , NY , USA: Wiley, 1998.

[7] J. W. Moody, J. E. Blanton and R. P. Will, "Capturing expertise from experts: The need to match knowledge elicitation techniques with expert system types," *Journal of Computer Information Systems,* pp. 89-95, 1999.

[8] G. M. Marakas and J. J. Elam, "Semantic structuring in analyst and representation of facts in requirements analysis," *Information Systems Research,* vol. 9, pp. 37-63, 1998.

[9] W. P. Wagner, J. Otto and Q. B. Chung, "Knowledge acquisition for expert systems in accounting and financial problem domains," *Knowledge-Based Syst.,* vol. 15, pp. 439-447, 11/1, 2002.

[10] N. Maiden, "Exactly How Are Requirements Written?" *Software, IEEE,* vol. 29, pp. 26-27, 2012.

[11] M. G. Pitts and G. J. Browne, "Stopping Behavior of Systems Analysts During Information Requirements Elicitation," *Journal of Management Information Systems,* vol. 21, pp. 203-226, 2004.

[12] B. Corbridge, G. Rugg, N. P. Major, N. R. Shadbolt and A. M. Burton, "Laddering - technique and tool use in knowledge acquisition," *Knowledge Acquisition,* vol. 6, pp. 315-341, 1994.

[13] A. Davis, O. Dieste, A. Hickey, N. Juristo and A. M. Moreno, "Effectiveness of requirements elicitation techniques: Empirical results derived from a systematic review," in 2006.

[14] O. Dieste and N. Juristo, "Systematic Review and Aggregation of Empirical Studies on Elicitation Techniques," *IEEE Transactions on Software Engineering,* pp. 304, 2011.

[15] R. Agarwal and M. R. Tanniru, "Knowledge Acquisition Using Structured Interviewing: An Empirical Investigation," *Journal of Management Information Systems,* vol. 7, pp. 123-141, 1990.

[16] G. J. Browne and M. B. Rogich, "An Empirical Investigation of User Requirements Elicitation: Comparing the Effectiveness of Prompting Techniques," *Journal of Management Information Systems,* vol. 17, pp. 223-249, 2001.

[17] D. Carrizo, O. Dieste and M. López, "Diseño de una nueva replicación de experimentos sobre entrevistas en elicitación de requisitos utilizando datos de las amenazas a la validez," in A Coruña, España, 2011 .

[18] A. M. Burton, N. R. Shadbolt, A. P. Hedgecock and G. A. Rugg, "A formal evaluation of knowledge elicitation techniques for expert systems: Domain 1," in 1987 .

[19] R. MacCallum, S. Zhang, K. Preacher and D. Rucker, "On the practice of dichotomization of quantitative variables," in 2002, pp. 10-40.

APPENDIXES

A. Demographic Data of Experimental Subjects

Interview/Subject	*Academic Qualifications*	Experience (years)			Subjects' familiarity with the problem (*)	Total no. of mentioned elements	SessionEffec	Total no. of consolidated elements	ConsolidationEffec	Subject Effectiveness – Retention Capability	Time (min:sec)	Time (sec)
		Interview	*Elicitation*	*Requirements*								
E01	MS in Systems Eng.	3	5	6	0	31	44%	25	35%	81%	17:15	1035
E02	MS in Electrical Eng.	0	0	0	1	38	54%	-	-	-	19:59	1199
E03	MS in Mathematics	3	1	1	0	25	35%	-	-	-	16:01	961
E04	MS in Systems Eng.	0	0	0	0	43	61%	34	48%	79%	14:39	879
E05	MS Civil/Industrial Eng.	3	2	2	1	41	58%	6	8%	15%	17:44	1064
E06	-	-	-	-	-	31	44%	33	46%	106%	23:04	1384
E07	BS in Microprocessors	0	0	0	1	37	52%	-	-	-	24:09	1449
E08	MS in Economics	10	2	2	2	32	45%	14	20%	44%	12:09	729
E09	BS in Data Processing	5	2	2	0	36	51%	21	30%	58%	10:40	640
E10	-	-	-	-	-	39	55%	18	25%	46%	18:34	1114
E11	MS in Systems and Computing	1	1	1	0	34	48%	31	44%	91%	8:29	509
E12	MS in Systems Engineering	3	3	3	1	32	45%	31	44%	97%	12:40	760

* Familiarity with problem: 0- Unfamiliar, 1-Familiar, 2- Knowledgeable

B. Statistical Significance

			Elicitation Session Effectiveness			Consolidation Effectiveness			Retention Capability		
			Number of mentioned elements			*Number of consolidated elements*			*Number of elicited elements*		
			r	*Sig.*	*U Test*	*r*	*Sig.*	*U Test*	*r*	*Sig.*	*U Test*
Subject Characteristics	Familiarity with problem		0.22	0.54	0.463	-0.56	0.19	0.212	-0.47	0.29	0.480
	Academic qualifications		0.02	0.96	0.748	0.87	0.01	0.032	0.88	0.01	0.034
	Experience	*Interview*	-0.51	0.13	0.087	-0.63	0.13	0.079	-0.44	0.33	0.439
		Elicitation	-0.10	0.78	0.463	-0.63	0.13	0.079	-0.44	0.33	0.439
		Requirements	-0.10	0.78	0.463	-0.63	0.13	0.079	-0.44	0.33	0.439
	Elicitation Time		0.21	0.57	-	-0.39	0.39	-	-0.41	0.35	-