Analyzing the Effect of the Collaborative Interactions on Performance of Requirements Validation

Nelly Condori-Fernández[1], Sergio España[1], Klaas Sikkel[2], Maya Daneva[2], and Arturo González[1]

[1] Universitat Politècnica de València
Camino de Vera 46022, Valencia, Spain
{nelly,sergio.espana,agdelrio}@dsic.upv.es
[2] University of Twente
Drienerlolaan 5,
7522 NB Enschede, The Netherlands
{m.daneva,k.sikkel}@utwente.nl

Abstract. [Context] Requirements validation is critical in the pursuit of quality software. It usually demands the collaboration of multiple stakeholders with different perspectives. [Question] Our community has reported scarce experimental studies on the role of collaborative interaction in requirements validation. The goal of this study is to explore the effect of collaborative interactions on the performance of requirements validation. [Principal ideas] We performed a quasi-experiment involving 118 bachelor students to act analysts, and 40 volunteering students from the Social Sciences department to act clients. The requirements were specified using UML activity diagrams. The overall performance is measured in terms of efficiency (missing requirements correctly identified in a time interval), and effectiveness (degree to which the validation yielded the correct result). Moreover, we measured also subjects' satisfaction on collaboration (questionnaire). [Contribution] We found that the teams composed exclusively of analysts showed better efficiency and effectiveness than mixed teams (client and analysts). However, for certain types of requirements, the mixed teams' efficiency was superior. Also, the degree of satisfaction was higher among the clients than among the analysts. We end up with identifying future research topics.

Keywords: activity diagrams, reviews-based validation, validation effectiveness, validation efficiency, requirements process performance, collaboration satisfaction.

1   Introduction

Requirements validation is a key activity in requirements engineering (RE); it aims to ensure that specifications accurately express the stakeholders' needs [1]. Through the requirements validation process, errors in a software requirements specification (SRS) are identified and corrected before it is used in further system development. Usually, a

requirements validation process demands the collaboration of multiple stakeholders with various needs and perspectives. In environments where stakeholders can freely discuss, share their opinions and resolve conflicts among them, it is often particularly challenging to facilitate the validation of the requirements in an efficient and effective manner.

Evaluation of collaboration interactions in software and non-software projects has been the subject of research in a number of contexts, be it virtual [8], [9], [11] or face-to-face [10], [11]. To evaluate the effectiveness of collaborative interactions, both quantitative and qualitative reasoning approaches have been proposed.

In software engineering, various publications analyze the effect of different human factors, such as personality or communication skills, in pair programming [2], [7], [10]. While these sources have been useful in planning our research, especially in developing awareness of the potential threats experienced by other authors in studies on collaboration and communication, we found few experimental studies focused on requirements validation in collaborative contexts [15, 16, 17]. Also, whatever work has been done, it is conducted from the requirements analysts' perspective. E.g. Gemino [15] carried out an experiment to compare the effect of animation and narration techniques on requirements validation. Furthermore, He et al. [17] report results of a post-task survey on the application of an inspection technique during requirements validation. Both studies were conducted only from the analyst's perspective. We note however that RE textbooks (e.g. [12]) treat requirements validation as a client-focused activity. It is therefore surprising that we as a community lack any in-depth understanding of how collaboration between clients and analysts possibly affects important outcomes of the requirements validation process. This gap of knowledge motivated us to initiate an empirical study in order to collect and analyse evidence to systematically address the gap.

The goal of this paper is to provide better understanding of how collaborative interactions affect the performance of requirements validation. We achieve this by conducting an exploratory experiment that was set up in three different scenarios: i) individual review from a client viewpoint; ii) requirements validation from an analyst viewpoint, working in pairs; and iii) requirements validation from combined client and analyst viewpoints.

The following sections provide a detailed account of our study. Section 2 provides a background on reviews-based requirements validation and positions our exploration within a requirements validation process that relies on reviews. Section 3 describes our experiment plan and Section 4 reports our results. Section 5 discusses the validity threats and Section 6 concludes the paper.

2   Reviews-Based Requirements Validation

There is a wide range of requirements validation techniques [12], [23]. This section briefly describes reviews-based validation. We opted for this class of techniques because it has been widely applied in the software industry [5], [20]. A requirements review process is usually performed by a number of people from different

backgrounds, meeting together to detect conflicts, omissions, inconsistencies and errors in specifications.

The steps of the process are the following: 1) Planning of review in terms of participants. 2) Distributing documents to the review team members. 3) Individuals read the relevant documents searching for inconsistencies, conflicts, omissions and other problems before the review meeting (pre-review). 4) Individual comments and problems are discussed; a set of actions to address the problems is agreed upon. 5) Follow-up actions to check if the agreed actions have been performed. 6) Final document is revised and the team members either accept it or plan further review iterations.

As stated in [5], a significant disadvantage of this process is its resource-intensiveness as review meetings span across several sessions and this in turn hinders the involvement of people from different departments at the same time. Hence, resource availability may easily become an issue.

In this empirical study, we, the experimenters, carried out the first two steps of the requirement review process. Steps 3 and 4 were carried out by the experimental subjects. The subjects had to check the completeness of a long SRS, represented by UML activity diagrams. In contrast to previous empirical studies ([15], [17], [21], [22]), we involve the viewpoint of the client, who had to review the SRS both by himself/herself and in collaboration with a team of analysts.

3   Experiment Planning

The goal of our experiment is to analyze the effect of the collaborative interaction on the performance of the requirements validation; in the context of bachelor students majoring in various IT sub-fields at University of Twente (UT), the Netherlands. We set out to answer three research questions (RQ):

RQ1: When teams are validating SRSs, is their efficiency affected by the type of collaborative interaction?

RQ2: Is the validation effectiveness affected by the type of collaborative interaction? We are also interested in assessing whether the type of requirements that are identified as missing in the SRS has an effect on validation effectiveness.

RQ3: Is the efficiency of teams in validating requirements affected by the degree of collaboration satisfaction?

3.1   Variables and Hypotheses

In our attempt to investigate how the collaboration among stakeholders with different background can affect in the performance of requirements validation, we considered the following independent variables:

- The *collaborative interaction type*. Given the lack of prior studies that involve the client viewpoint, we considered two types of collaborative interactions: interaction among participants with a different role (client and analyst), and interaction among participants with the same role (analyst).

- The *type of missing requirement*. Given the exploratory nature of the experiment, we limited the investigation to three types of requirements:
  - Missing *business activities* that are part of the company's work practice.
  - Missing *constraints* that apply to the business activities.
  - Missing *business forms* that are filled in/created/used by company's staff.

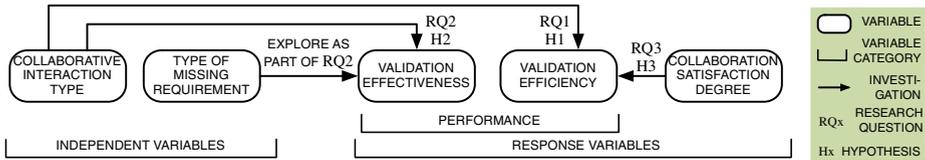Fig. 1 overviews the variables of the experiment and their hypothesised relationships.

Fig. 1. Overview of the relationships among variables

We identified the following response variables (a.k.a. dependent variables):

Validation efficiency indicates how well a team used time to correctly identify missing requirements. In this study, 12 requirements were removed from the original specification (For more detail, see the description of the experimental object in sub-section 3.2). Also, it is important to note that apart from correctly-identified missing requirements, we considered two types of errors that might be committed by the teams: 1) A functional fragmentation error, which means a functional requirement is correctly identified, yet incompletely specified and therefore appears in the form of two or more fragments (encapsulations). 2) A functional aggregation error, when two or more missing functional requirements were aggregated to a single 'higher-level functional' requirement.

Regarding the validation time, as reviews-based requirements validation usually requires several sessions to be completed, in order to increase control, we limited the time to a single 2-hour session for all teams that participated in the experiment.

Validation effectiveness is the degree to which the teams execute the validation task correctly. The test subjects were given a SRS from which the researchers had removed a set of requirements. Effectiveness is measured as the number of missing requirements correctly identified, divided by the total number of missing requirements.

Collaboration satisfaction degree. A questionnaire was designed for each of the two roles (i.e. client and analyst) that participate in the different interaction scenarios (described in Sect. 3.4). We consider two key points for the formulation of the questions [4]: (i) the members' satisfaction is the basis of the team satisfaction; (ii) the degree of satisfaction derives from the working relationship. Based on these points, we identified the different collaboration relationships in the Scenarios II and III
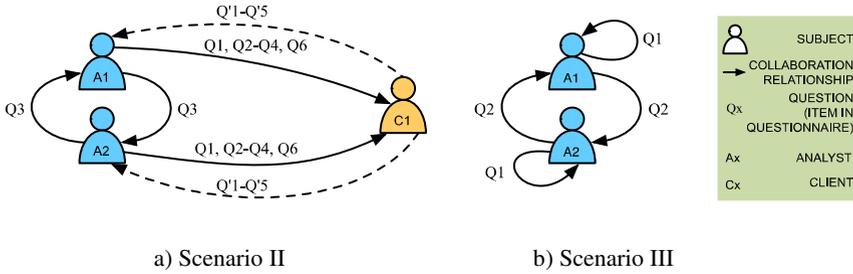
a) Scenario II   b) Scenario III

Fig. 2. Collaboration relationships identified in the scenarios

(see Fig. 2) and we formulated specific questions for each relationship. We measured the satisfaction degree by means of these 5-point Likert scale questions[1].

Thus, from our research questions the following hypotheses were derived.

- **H1$_0$:** The efficiency of teams in validating requirements is the same independently of the collaborative interaction type.
- **H2$_0$:** The requirements validation effectiveness is the same independently of the collaborative interaction type.
- **H3$_0$:** The efficiency of validating requirements is the same independently of the degree of collaboration satisfaction.

3.2   Experimental Context

Experimental Subjects. As we wanted to account for two different viewpoints in requirements validation, we included two different profiles of subjects. 1) Client role: 40 students from the UT Social Sciences Department, without any background in modeling languages, volunteered to be trained in the business domain. Participants were invited by sending them a flyer offering 50€ for participation. 2) Analyst role: 118 first-year bachelor students enrolled in the Information Systems (IS) course at UT were selected by convenience sampling and trained to play the analyst role. The students were majoring in IT sub-fields such as Computer Science, Business Information Technology, and Technical Management Science. The IS course objective is to train the students in UML-based IS requirements specification and is taught by the 3rd author of this paper.

A demographic questionnaire revealed that the group of analysts was quite homogeneous. 93% of the students had not participated in any previous course on dynamic- or static-oriented modelling techniques (also, see their perceived knowledge in Fig. 3, left). Their level of English was good; only 9% of them had obtained a grade 5 or lower out 10 points in English language (see Fig. 3, right).

---

[1] The question numbers refer to the satisfaction questionnaires available at
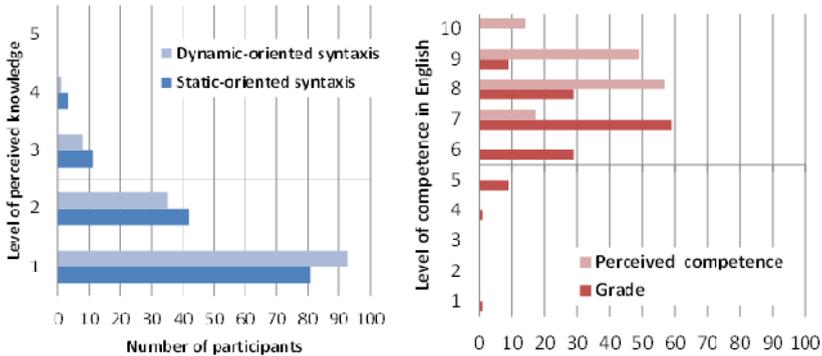http://users.dsic.upv.es/~nelly/valid.htm

Fig. 3. Demographic results of the participants playing the analyst role

Experimental Objects. The SRS describes the information system needed by a Photography Agency that manages illustrated reports provided by photographers and distributes them to publishing houses. The SRS was created by the 2nd author using UML activity diagrams. We chose this type of diagrams because it is commonly used in industry to interact with clients during review and requirements validation [13]. The other authors checked the appropriateness of the SRS for the experiment. The SRS is 49 pages long. We decided to use the whole specification and remove 12 requirements – the missing requirements that the subjects had to identify. Table 1 (in Sect. 4.1) classifies the requirements in three types: missing business activities, missing constraints, and missing business forms. The explanation of each missing requirement is omitted for the sake of brevity and can be found at `http://users.dsic.upv.es/~nelly/valid.htm`

3.3   Experimental Instruments

Demographic questionnaires. The demographic questionnaire the students acting analysts aims at assessing the subjects' English language proficiency and their background in RE modelling. The latter is operationalized by means of 7-point Likert-scale questions about their knowledge and experience with 8 IS modelling techniques that deal with static (e.g. Class Diagram) and dynamic aspects (e.g. Activity Diagram) of the system. The results are shown in Fig. 3. The questionnaire distributed to the clients also corroborated that their proficiency in English was very good.

- *Satisfaction questionnaire.* It uses 5-point Likert-scales to elicit the personal satisfaction of both the client and the analysts with their interaction during the collaborative validation; and the interaction between analysts when using a textual description.
- *Post-task survey.* It gathers information about the difficulties encountered during validation with respect to the reviewed SRS.

Moreover, a validation form was implemented to get details on the missing requirements that the subjects identified. For each missing requirement identified, the

analysts ought to offer a rationale and a textual description of the requirement identified. A link to the web version of the experimental instruments and is available at `http://users.dsic.upv.es/~nelly/valid.htm`

3.4   Experiment Design

The experiment was conducted in three different scenarios.

- *Scenario I*. (Pre-review). The subjects acting clients read the requirements specification and identified the missing functional requirements individually. The subject with the analyst role read the requirements specification in order to get familiar with the Photography Agency system.
- *Scenario II*. Groups of three subjects – two 'analysts' and one 'client' identify the missing functional requirements cooperatively.
- *Scenario III*. Pairs of subjects with the analyst role identify missing functional requirements, by using an additional textual description of Photography Agency.

Considering the two types of collaborative interaction (interaction among participants with and without different role), our experiment adopted a between-subjects design (scenarios II and III). Scenario I was also considered as part of the study, because according to the review-based requirements validation process an individual review (Pre-review) is required.

3.5   Experimental Procedure

Training Process. The subjects acting clients received 6 hours of training in the business domain (the Photography Agency). For this, we used the demonstration/practice method [3]. First, slides about the Photography Agency were used to present the Agency and its main activities. Then, in order for the subject to acquire some practice in the agency's domain, they were given three exercises to solve. After a break of 15 min, a test (a questionnaire with 12 closed questions) about the problem domain was completed by the students with the purpose of verifying their acquired knowledge. Using the grade points average ([0-1]), we found that subjects had more difficulty to correctly answer Question 7: *"Who establishes the yearly rates of the agency?"* (mean = 0.26; std dev = 0.44); this question expects the subject to identify the organizational role in charge of a given business activity. The rest of the questions had an acceptable average, which varied between 0,68 and 0,95.

The subjects acting analysts were trained on the UML activity diagrams and requirements validation as a regular part of the course. It took four two-hour sessions spread over multiple days, which consisted of lectures, exercises with multiple choice questions, and supervised exercises in the computer lab. We assessed their competence level (high:[10, 7[; medium:[7, 5[; low:[5,0])[2] in UML activity diagraming and requirements validation. We found that 70 % of the 118 subjects demonstrated a

---

[2] Reverse square brackets are used in (semi-)open intervals. For instance [10, 7[means "any real number with a value ranging from 10 to 7, including the value 10 but excluding the value 7".

medium level to assume their role as analyst. The subjects with a low competence level in modeling and validating (6.5% and 8.6%, respectively) were not considered.

Execution. Afterwards, following a guideline that specified what to do and which forms to use, the subjects, once assigned to one of the three scenarios, proceeded to identify the missing requirements applying the reviews-based requirements validation technique. The pre-review (Scenario I) was carried out individually by the clients (to identify missing requirements) and the analyst (to get familiar with the problem domain). Time expected was about 1 hour. Two days after this session, each client discussed with a team of two analysts his/her individual comments and problems that he/she had during the requirements pre-review (Scenario II). Besides, this requirements review was also carried out by 22 pairs of analysts at the same time, but in a different building of the university (Scenario III). For both scenarios (II and III) the reviews took 2 hours. Subsequently, a questionnaire was distributed with the purpose of eliciting the personal satisfaction of the collaborative validation process.

4 Analysis and Interpretation of Results

Once data collection was over, two evaluators (the first two authors) reviewed the validation forms completed by all the teams. From this review, three possible values were considered for our list of missing requirements: i) identified correctly, ii) identified with error, iii) not identified.

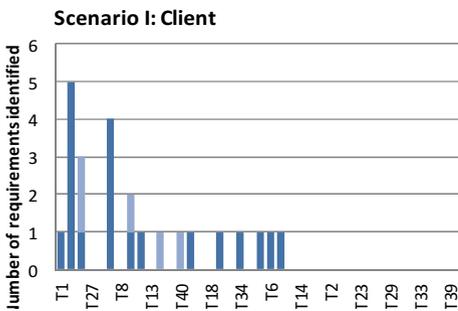### 4.1 Analyzing the Effect of Type of Interaction on Efficiency

Fig. 4. Missing requirements identified by client

To calculate the efficiency of the requirements validation, first the size of the output of the validation process is calculated, by analyzing the data collected in the evaluation forms. The requirements identified but described with any type of error were grouped and counted separately.

As Fig. 4 shows, during the pre-review, a maximum of 5 out of 12 missing requirements were identified by one of the 'clients'. Most 'clients' found at least 1 missing requirement.

26 out of 40 'clients' were not able to identify any missing requirement (the majority of them invented new requirements). However, when they interacted with the team of analysts (Scenario II), this number was reduced to 17 subjects, meaning an improvement of 22% was observed (see Fig. 5, left). We found that teams of analysts that carried out the validation task without the client support (Scenario III) showed a better efficiency than the teams of analysts supported by a client (see Fig. 5, right).

We applied the Mann-Whitney U test to verify our first hypothesis ($H1_0$), by using the data in Scenario II and III, considering only the requirements correctly identified

by the teams. We found that the two groups differed significantly from each other with U(61) = 147.5; Z=-4.536; p = .000. This suggests that the interaction strictly among analysts, using an additional textual description of Photography Agency, had a beneficial effect on efficiency of requirements validation.

Analyzing the post-task survey, we found that the pre-review of the specifications based on activity diagrams was very difficult for the clients, thus the communication does not seem to help when the clients only know their business and the analysts only know the modeling language. Communication and interaction difficulties among 'clients' and 'analysts' ' could have affected in the validation efficiency.
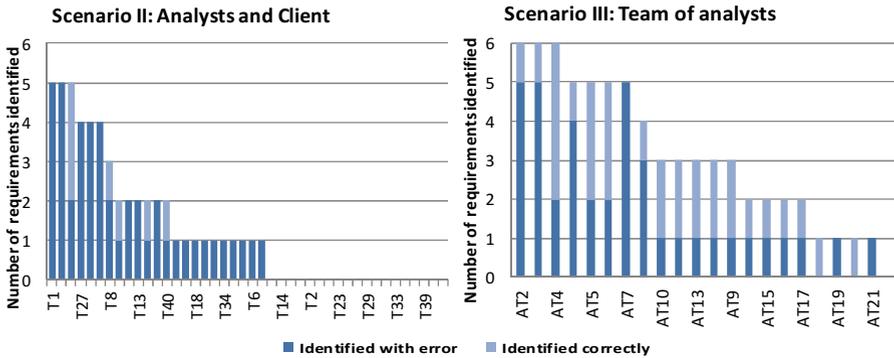

Fig. 5. Number of missing requirements identified in scenarios II and III

4.2   Analyzing the Effect of Type of Interaction on Validation Effectiveness

As validation effectiveness is an indirect measure consisting of two measures, we first discuss the results related to frequency of identification for each one of the twelve expected missing requirements to be identified (success level). Then, the completeness rate by type of missing requirements is calculated.

According to the Table 1, the requirements that were most frequently identified are R4 (36 hits) and R12 (28 hits). Both requirements are *business forms*; that is, related to documents that the company sends to external parties (e.g. a letter). Other requirements were also identified with an acceptable degree of success across all the scenarios, such as R2 and R3 (*constraints*), or R1 (*business activities*). Conversely, there were requirements that were difficult to identify, as the results show. First, for R6, R11, although some teams of analysts in Scenario III (A) were able to completely identify them, the participants of Scenario II (C+A) were completely unable to identify this type of requirements. Similar results were obtained for the second type of requirements, where R3 and R5 could not be correctly identified by any team. However, for requirements R8 and R10 (*business forms*), we observed that analysts of the third scenario score less than analyst from the second scenario (A+C). What makes them different? Both requirements involve processes that do not change any data in the

Table 1. Success level (total of hits) for each one of the twelve missing requirements (MR)

Type	MR	Scenario	Identified w. error	Correctly identified	Total
Business activities	R1	A+C	6	1	7
		A	7	1	8
	R6	A+C	2		2
		A		4	4
	R11	A+C	1		1
		A	2	4	6
Constraints	R2	A+C	6		6
		A	7	1	8
	R3	A+C	6		6
		A	7		7
	R5	A+C	1		1
		A	2		2
Business forms	R4	A+C	11	2	13
		A	5	10	15
	R7	A+C	5	1	6
		A	7	3	10
	R8	A+C	1		1
		A			0
	R9	A+C			0
		A	3	2	5
	R10	A+C	1	1	2
		A			0
	R12	A+C	8	2	10
		A	10	3	13

system. It has to be remarked, however, the scores for teams of this scenario (A+C) are low, so that doesn't say much.

For requirement R9, we observed that no analysts with the client support (A+C) could ever find it, while analysts without clients did score fairly well. R9 is a business requirement related to the acquisition process, but it involves no system interaction. So, the analysts from the second scenario (A+C), who looked primarily at the diagrams, had to rely on the clients (who all missed it) and would not spot this from the diagrams. The analysts' teams without clients did look into the requirements document, saw that R9 was defined as part of the process, and identified it as missing in the specification.

Now, by calculating the completeness rate by type of missing requirements, we found that the second group of requirement (missing constraints) were the less scored by the teams from both scenarios (Scenario II and Scenario III). Applying the Mann-Whitney U test, we corroborated that the collaborative interaction type had a clear effect on validation effectiveness ($H2_0$), but only for the missing requirements of the type *business activities* and *business forms* (see Table 2).

The interaction only among analysts with the support of a textual description had some beneficial effect on the validation effectiveness. This result indicates that requirements validation done by comparing documents is more effective than validation by means of meetings with clients.

Table 2. Mann-Whitney U statistics for completeness rate by type of missing requirements

	Business activities	Constraints	Business forms
Mann-Whitney U	306.500	400.500	180.000
Wilcoxon W	1167.500	1261.500	1041.000
Z	-2.962	-1.724	-4.110
Asymp. Sig. (2-tailed)	.003	.085	.000

A possible reason for this could be that documents comparison can be done systematically by checking if each textual statement in the description is represented in the requirements specification. On the contrary, there is no systematic way of recalling and checking each and every piece of domain knowledge in the mind of the client. Moreover, we consider that the way in which requirements were documented, by using an activity diagramming technique, played an important role in ensuring that analysts-only teams could more easily read and validate them than analyst teams with the participation of clients.

4.3   Analyzing the Effect of Collaboration Satisfaction on Efficiency

As collaboration is fundamentally a social activity relying on interaction between two or more individuals, it is inevitable that some degree of task-related effort remains at the individual level. Although we did not measure the individual performance, we evaluated the personal degree of satisfaction of the respective members in each team. To do this, we analyzed the data collected from the questionnaires that were applied in the Scenarios II and III. First we averaged the answers to the items of the questionnaire to obtain a representative value of the (client's or analyst's) satisfaction of collaborative requirements validation. Then, a 3-points ordinal scale (not at all satisfied, somewhat satisfied, and satisfied) was used in order to interpret better the results obtained (see Fig. 6). The low significance values obtained ($p < 0,05$) throughthe chi-square test suggests that the average rate of the subjects does differ in terms of satisfaction degree. In Fig. 6, (left) we observe that 80% of the clients were satisfied with the analysts' performance. Clients considered the interaction with analysts to be very helpful in identifying missing requirements. However, analyzing the effect of their personal satisfaction with the efficiency of requirements validation ($H3_0$), we observe (Fig. 6, right) that the effect is not significant. A possible explanation is that the clients were unaware of the correctness of the missing requirements identified and their satisfaction did not affect the efficiency in the requirements validation.

Regarding the analysts of Scenario II, Fig. 7 (left) shows the results to questions related to degree of satisfaction about (i) the information provided by the clients to facilitate the understanding of the problem domain (60% not at all satisfied) or the identification of missing requirements (60% not at all satisfied) and (ii) working with their partner (83% satisfied).

On the other hand, when clients were asked about the individual-proactive participation of the analysts, we found that 32% of the clients indicated that only one of the two analysts was proactive. For the purpose of verifying the consistency of the answers given by clients and analysts of scenario II, we carried out a correlation analysis.
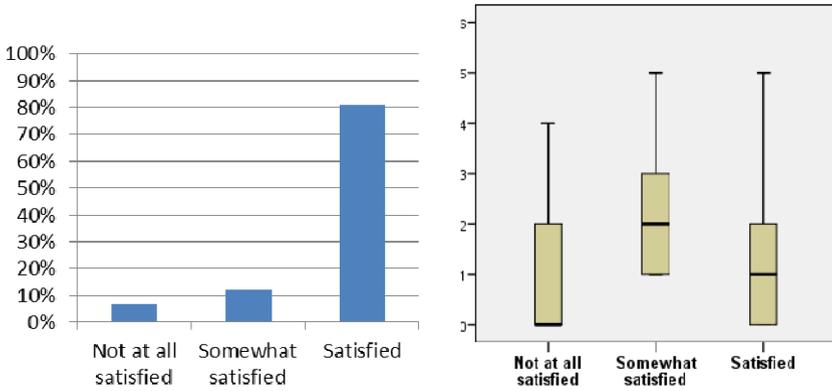
Fig. 6. Distribution of effect of the collaboration satisfaction degree (left) and its effect (right) on the efficiency of clients in Scenario II

Due to lack of space, we show only the box plot for the degree of satisfaction on the client's feedback to understand the problem domain (Fig. 7, right).
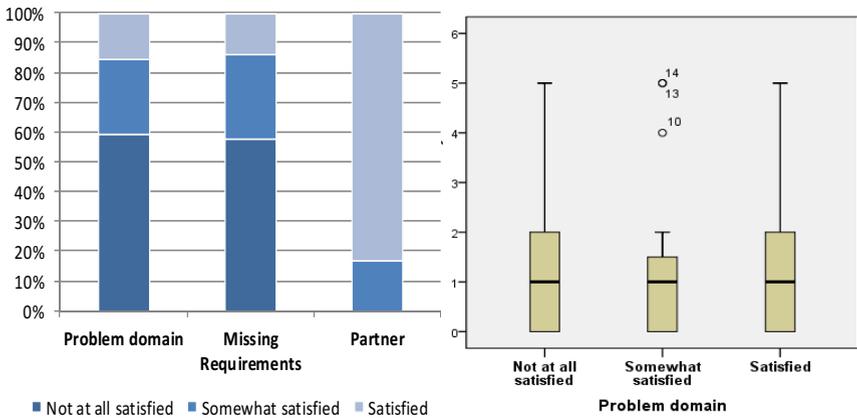

Fig. 7. Distribution of effect of the collaboration satisfaction degree (left) and its effect (right) on the performance of analysts in Scenario II

Regarding the effect of the degree of collaboration satisfaction on the analysts' efficiency in Scenario III (H3$_0$) (Fig. 8, right), we observed that the analysts that were 'somewhat satisfied' showed a greater efficiency than those who were 'satisfied'. However, only 35% of these analysts were satisfied with their partners (Fig. 8, left).
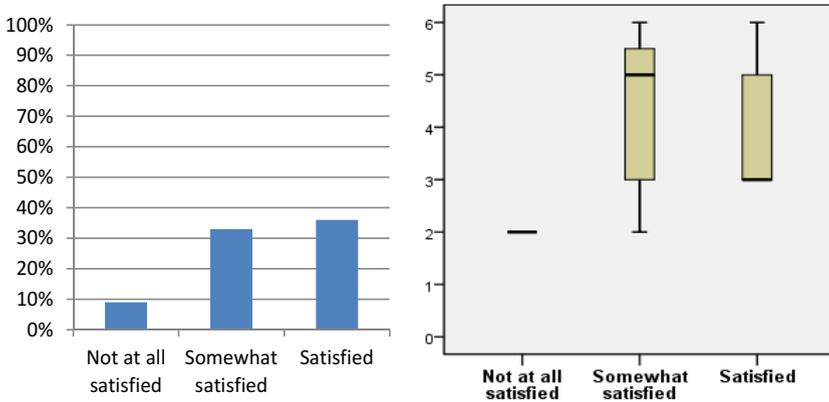
Fig. 8. Distribution of effect of the collaboration satisfaction degree (left) and its effect (right) on the efficiency of the analysts in scenario III

5 Threats and Lines for Further Empirical Research

In this experiment, we have balanced exploration – which offers the chance of gathering new knowledge, and control – which minimizes risks. Below, we discuss the threats to the validity of our results and provide rationale for some of our experiment design decisions.

Internal Validity Threats. We could not mitigate the *instrumentation threat* of "false answers" to the satisfaction questionnaire. However, we ascertained that the answers by members of the same team were correlated. Regarding the preparation of the incomplete requirement specification (by seeding faults), we tried to make a homogeneous distribution of the 12 requirements that were removed along the specification. We minimized also the threat of *subject selection* by forming the teams randomly. We had a slight *mortality*; that is, two analyst subjects in scenario II dropped out in the last minute. Thanks to a contingency plan we reassigned two subjects from scenario III to Scenario II.

Construct Validity Threats. In explorative study on collaboration in requirements validation, we chose the two basic types of stakeholders involved in the validation process (clients and analysts) and two types of interactions (client-analyst and analyst-analyst). Other roles and types of interactions exist in practice, so this construct is under-represented. However, we opted for increasing complexity gradually (see considerations below). The same applies to the types of requirements. We deliberately did not use a complex SRS, to keep the experiment under control, both in terms of time to complete the training and the experimental task. To minimize the threat of an *inadequate preoperational explication of constructs*, the collaboration satisfaction degree was operationalized according to two key issues proposed by Jun et al [4].

External Validity Threats. We think that our sample of subjects was quite representative for real clients because we involved students with no modeling competence to act as clients. In order to guarantee enough knowledge about the problem domain, a

training session was conducted with the 40 participants. As for subjects acting analysts, they received intensive training on requirements engineering and UML modeling as part of their education, prior to the experimental task. Additionally, we did tests to assess the competence of both types of subjects: i) clients: domain knowledge test; ii) analysts: UML and validation competence tests.

The use of students as surrogates for real practitioners is common practice ([14], [6]) and, given the exploratory nature of the experiment, we preferred having a large number of subjects (the chances of interesting effects to appear is increased) than using fewer practitioners. Also, the length of the Photography Agency SRS (537 IFPUG functions points) intended to be manageable for experimentation but realistic enough to include some complexity.

Considerations for Further Research. The exploratory experiment provides preliminary results, on top of which further experiments can be designed. To deepen into the mechanisms of the validation process, it may be interesting to consider additional types of stakeholders, not only based on their roles but also in their characteristics.

Similarly, other types of requirements can be included in the SRS. Our experience while trying to classify requirements and compare the way students treated the missing requirements indicates that a requirements taxonomy is needed. We found that in confronting a process-aware information system, it is not enough to distinguish between functional and non-functional requirements. As a starting point, we opted for classifying missing requirements into activities, constraints and business forms. Adopting an existing requirements taxonomy or classification (e.g. [22]) or proposing a suitable one is a recommended line for future research. We expect this to help formulate more precise hypotheses concerning the performance of subjects while validating SRSs.

An advanced experiment would require a detailed requirements classification that has proved to be valuable (even if just in research settings) and a domain case that includes both a greater variety of types of requirements and more requirements of each type. This way, the researchers can remove several requirements of each type and test whether there exist differences in their identification during validation, and whether the collaborative interaction has a significant effect. We expect such results to shed light on the mechanisms of the validation process.

Our results also show that comparing a textual description of the domain with the more technical SRS (activity diagrams) is more productive and effective than validating requirements by reviewing the with clients. However, a textual description is itself an SRS. In an industrial setting, a stakeholder needs to create this document. Also, it is subjected to the same risks as the technical SRS; namely, incompleteness and invalidity. If a textual description is to be used for this purpose, it needs to be validated first. What is the difference in terms of performance between validating a textual specification and a diagram-based one? how does the language of the specification impact the collaborative interaction? To answer these questions, further empirical research needs to be done to better understand the practical use of activity diagrams (and other notations, e.g. BPMN) in validating requirements with different stakeholders. It is also challenging to investigate how the interaction between stakeholders (clients and analysts) actually takes place during the interviews. For instance, time devoted to each task (e.g. phatic communication, answers, responses), disruptions (e.g. misunderstandings and time devoted to solve them), attitudes (e.g. proactivity).

This requires audio recording and transcribing the most relevant information. It would allow comparing the actual interaction with the perception of the subjects.

Last but not least, once more precise hypotheses have been defined for improved experiments, the use of real practitioners can be considered.

6   Conclusions and Future Work

This paper has explored the collaborative interaction effect during requirements validation and has revealed certain relationships that deserve future investigation. We found that clients validating a requirements specification on their own are limited by their knowledge of the technical languages; their performance increases when they work collaboratively with a team of analysts. However, as per our results, the most successful scenario has been a team of analysts checking the specification against a textual description of the domain. Since this scenario entails certain difficulties and risks when applied in industrial settings (e.g. the textual description is a specification itself and may be incomplete or unavailable as one monolithic document but exist in the form of scattered interview proceedings) it is necessary to investigate deeper the collaborative interaction of clients working hand by hand with analysts.

We also found that clients were more satisfied with the collaboration during requirements validation, than analysts, which can be due to the fact that analysts feel more responsible towards to outcome of the interview since they often lead it. Other interesting outcomes have appeared, but they need to be contrasted with an observation of the behavior of subjects during the interview; this is planned as future work.

We are aware of some risks due to removing certain requirements from the full SRS, e.g. too much emphasis on one type of requirements can cause problems of allocating equal time/schedule and resources to other requirements types. We, therefore, plan to propose and evaluate a requirements taxonomy that applies to business information systems in general and aids in requirements validation. On the other hand, as the selected validation technique could have had also an effect on response variables (e.g. satisfaction), we seek to replicate the experiment by including others interesting review-based validation techniques, such as the checklist-based reading technique.

References

1. Cheng, B., Atlee, J.M.: Research Directions in Requirements Engineering. In: Future of Software Engineering (FOSE), ICSE 2007, pp. 285–303. IEEE CS Press (2007)
2. Panagiotis, S., Ioannis, S., Lefteris, A., Ignatios, D.: An experimental investigation of personality types impact on pair effectiveness in pair programming. Empirical Softw. Eng. 14(2), 187–226 (2009)

3. DOE Handbook, Alternative Systematic Approaches to Training, 1074-95 (January 1995)
4. Jun, L., Ya-Feng, L.: A Method of Evaluating Collaboration Satisfaction Degree of NPD Team. In: Proc. CSCWD 2010, pp. 156–160 (2010)
5. Saqib, B., Sheraz, A.: Requirements Validation Techniques practiced in industry: Studies of six companies, Blekinge Institute of Technology, Sweden. Master thesis, Software Engineering (Octotber 2008)
6. Condori-Fernandez, N., Daneva, N., Sikkel, K., Herrmann, A.: Practical relevance of experiments in comprehensibility of requirements specifications. In: EMPIRE 2011 Collocated at the RE Conference, Trento-Italy, Italy, pp. 21–28 (August 2011)
7. Walle, T., Hannay, J.E.: Personality and the Nature of Collaboration in Pair Programming. In: Proc. ESEM 2009, pp. 203–213 (2009)
8. Dwyer, P.: An Approach to Quantitatively Measuring Collaborative Performance in Online Conversations. Computers in Human Behavior 27, 1021–1032 (2011)
9. Lin, C.-P., Wang, Y.-J., Tsai, Y.-H., Hsu, Y.-F.: Perceived Job Effectiveness in Cooperation: A Survey of Virtual Teams within Business Organization. Computers in Human Behavior 26 (2010)
10. Choi, K.S., Deek, F., Im, I.: Pair Dynamics in Tem Collaboration. Computers in Human Behavior 25, 833–852 (2009)
11. Patel, H., Pettit, M., Wilson, J.: Factors of Collaborative Working: a Framework for a Collaboration Model. Applied Ergonomics 43, 1–26 (2012)
12. Lauesen, S.: Software Requirements: Styles and Techniques. Addison-Wesley (2002)
13. Dobing, B., Parsons, J.: How UML is Used. Commun. ACM 49(5), 109–113 (2006)
14. Runeson, P.: Using Students as Experiment Subjects - an Analysis on Graduate and Freshmen Student Data. In: 7th Int. Conf on EASE, Staffordshire, UK, pp. 95–102 (2002)
15. Gemino, A.: Empirical comparisons of animation and narration in requirements validation. Requir. Eng. 9(3), 153–168 (2004)
16. Condori-Fernandez, N., Daneva, M., Sikkel, K., Wieringa, R.J., Dieste, O., Pastor, O.: A Systematic Mapping Study on Empirical Evaluation of Software Requirements Specifications Techniques. In: ESEM 2009, pp. 503–505. CS Press (2009)
17. He, L., Carver, J.C., Rayford, B.: Using Inspections to Teach Requirements Validation. CrossTalk: The Journal of Defense Software Engineering 21(1) (2008)
18. Basili, V.R.: The Empirical Investigation of Perspective-Based Reading. J. of Empirical Softw. Eng. 1(2), 133–164 (1996)
19. Leite, J.C.S.P., Freeman, P.A.: Requirements Validation through Viewpoint Resolution. IEEE Transactions on Software Engineering 17, 1253–1269 (1991)
20. Raja, U.A.: Empirical Studies of Requirements Validation Techniques. In: 2nd International Conference on Computer, Control and Communication, vol. 1(9), pp. 17–18 (2009)
21. Albayrak, O.: An Experiment to Observe the Impact of UML Diagrams on the Effectiveness of Software Requirements Inspections. In: ESEM 2009, pp. 506–510 (2009)
22. Walia, G., Carver, J.: Using Error Abstraction and Classification to Improve Requirements Quality: Conclusions from a Family of Four Empirical Studies. J. of Empirical Softw. Eng. 18(4), 625–658 (2013)
23. Aurum, A., Petersson, H., Wohlin, C.: State-of-the-art: Software Inspections after 25 Years. Journal of Software Testing Verification and Reliability 12(3), 133–154 (2002)