University of Twente

Information System Group

Technical Report

Preliminary Survey on Empirical Research Practices in Requirements Engineering

Nelly Condori-Fernandez
Maya Daneva
Roel Wieringa

Technical Report nr: TR-CTIT-12-10
Date: 24-04-2012

University of Twente (The Netherlands)

Preliminary Survey on Empirical Research Practices in Requirements Engineering

Nelly Condori-Fernandez, Maya Daneva, Roel Wieringa
University of Twente,
Enschede, The Netherlands.
{n.condorifernandez, m.daneva, r.j.wieringa}@utwente.nl

Abstract. [Context and Motivation]. Based on published output in the premium RE conferences and journals, we observe a growing body of research using both quantitative and qualitative research methods to help understand which RE technique, process or tool work better in which context. Also, more and more empirical studies in RE aim at comparing and evaluating alternative techniques that are solutions to common problems. However, until now there have been few meta studies of the current state of knowledge about common practices carried out by researchers and practitioners in empirical RE. Also, surprisingly little has been published on how RE researchers perceive the usefulness of these best practices. **[Objective]** The goal of our study is to improve our understanding of what empirical practices are performed by researchers and practitioners in RE, for the purpose of understanding the extent to which the research methods of empirical software engineering are adopted in the RE community. **[Method]** We surveyed the practices that participants of the REFSQ conference have been using in their empirical research projects. The survey was part of the REFSQ 2012 Empirical Track. **[Conclusions]** We found that there are 15 commonly used practices out of a set of 27. The study has two implications: first it presents a list of practices that are commonly used in the RE community, and a list of practices that still remain to be practiced. Researchers may now make an informed decision on how to extend the practices they use in producing and executing their research designs, so that their designs get better. Second, we found that senior researchers and PhD students do not always converge in their perceptions about the usefulness of research practices. Whether this is all right and whether something needs to be done in the face of this finding remains an open question.

Keywords: empirical research, survey, requirements engineering

1 Introduction

In recent years, there has been an increased interest in empirical research in Requirements Engineering (RE). This increase is not only reflected in the number of published empirical studies but also in the growth of methodological advise on empirical software engineering (SE). For example, we observe an increasing diversity of proposed checklists concerning the planning, execution and reporting on empirical SE studies [4][5][6][7].

Although the majority of these checklists have not yet been sufficiently evaluated (in terms of their usability and usefulness), they list several interesting recommended practices for guiding empirical research in SE. However, what do we know about the usage of these recommended practices in RE? With the purpose of improving our understanding of what empirical practices are commonly performed by RE researchers and practitioners, we conducted a survey with the participants of the REFSQ 2012 conference. This survey was designed, considering the most recommended practices listed in the Unified Checklist, which was recently proposed by Wieringa [1]. This unified checklist is based on a logical analysis of the empirical research cycle [2] and a comparison of existing checklists inside and outside SE.

The paper is structured as follows. Section 2 describes the research method. Section 3 presents the survey results obtained and provides the discussion over these results. Finally, in section 4 we provide our final conclusions.

2 Method

This section describes how our study was conducted. We begin by listing the research questions. Next, we present the survey instrument we used. Then, we characterize the respondents for our survey. Finally, we explain our data collection and analysis techniques applied in this study.

2.1 Research questions

With the purpose of obtaining a better understanding of empirical research practices applied currently in Requirements Engineering (RE) community, we aim to address the following research questions:

RQ1: What are common practices in designing and reporting empirical research carried out by researchers and practitioners?

RQ2: What recommended practices reported in the literature do researchers and practitioners consider useful for designing and reporting empirical research?

2.2 Questionnaire Design

By following the guidelines provided by Kitchenham and Pfleeger in [3], we created a web-based survey consisting of 50 questions (summarized in Table 1). 30 out of 50 questions were formulated to discover which of the recommended practices in the literature are performed by the respondents. Each of these questions was rated on a 3-point nominal scale [‘yes’, ‘no’, ‘unsure I understand what you ask’].

The remaining questions were formulated in order to understand the usefulness perceived of the most recommended practices for empirical research (case studies and

experiments). A 5-point Likert scale was used for this set of questions, where 1 = not useful and. 5 = very useful.

The questions focus on different recommended practices to be considered through six phases of the empirical cycle [2]: research problem investigation, research design, research design validation, execution and results evaluation. The questions were adapted from the unified checklist proposed by Wieringa [1], as guidelines for the empirical research design and report. We tested the questionnaire with 1 PhD student and 1 Post doc researcher, who have experience in designing experiments. The questionnaire testing discovered the unclear questions, and it helped us to remove some ambiguities.
Moreover, in order to gather information about the respondents, five closed-ended questions were asked at the beginning of the survey. The information included the sector of their current job (e.g. academia); their role in the organization, experience years in requirements engineering, experience level in designing experiments or case studies. The survey was implemented using the Surveygizmo tool[1], and was configured to be used in laptops (computers), tablets and mobile platforms.

2.3 Data collection

The survey was electronically distributed by the REFSQ 2012-participants mailing list, which was established to facilitate communication among the organizers of the conference, researchers and practitioners participating in the 18th International working conference on Requirements Engineering: Foundation for Software Quality[2]. From 110 participants that were registered at the REFSQ conference, 36 of them completed our survey, 6 participants answered partially and 7 participants abandoned the survey after reading the instructions. We collected survey data during two weeks, from 19 to 30 March 2012. Actually, the data collection was originally planned to be carried out only during the conference week, but with the purpose of increasing our response rate this was extended to one week more. Two reminder emails were sent to encourage participants who had not yet responded the survey to reply.

[1] http://www.surveygizmo.com/
[2] http://www.refsq.org/2012/

Table 1. Summary of survey questions

	ID	Question	Scale
Research context	Q1	Is your empirical research usually motivated by the goal to improve some artefact ?	Nominal
	Q2	Do you usually define a top-level knowledge goal for your empirical research?	
	Q3	Do you usually review the current state of knowledge related to your empirical research?	
	Q4	Do you think that the following practices would be useful to have a better contextualization of your research?	Likert
	Q4.1	Definition of improvement goal	
	Q4.2	Definition of knowledge goal	
	Q4.3	Review of the current state of knowledge	
Research problem	Q5	Do you usually define a conceptual framework for the phenomena to be investigated in your research?	Nominal
	Q6	Do you usually operationalize the concepts of this framework?	
	Q7	Do you validate these operationalizations?	
	Q8	Do you usually formulate the research questions in your empirical research?	
	Q9	Do you usually describe the population in your empirical research?	
	Q10	Do you think that the following practices would be useful to improve the understanding of your research problem?	Likert
	Q10.1	Definition of relevant concepts of the phenomena to be investigated	
	Q10.2	Operationalization of the concepts defined	
	Q10.3	Validation of the operationalization of concepts	
	Q10.4	Formulation of research questions	
	Q10.5	Description of population	
Research design and justification	Q11	Do you usually justify the acquisition process of the object of study for your empirical research?	Nominal
	Q12	Do you consider any ethical issue in your research involving human subjects?	
	Q13	Do you usually justify the representativeness of the object of study for the population in your empirical research?	
	Q14	Do you usually consider all the assumptions of inference techniques to be used in your empirical research?	
	Q15	Do you usually plan the procedures to be followed in the experimental treatment?	
	Q16	Do you usually specify any instruments needed to apply the treatments of your experimental research?	
	Q17	Do you usually specify any instruments needed for measurement?	
	Q18	Do you usually specify procedures to be followed when performing measurements?	
	Q19	Could you indicate whether you usually consider the validity of the following issues?	
	Q19.1	Measures	
	Q19.2	Measurement procedure	
	Q19.3	Measurement instrument	
	Q19.4	Treatment	
	Q19.5	Treatment procedure	
	Q19.6	Treatment instrument	
	Q20	Do you think that the following practices would be useful to improve your research design?	Likert
	Q20.1	Justification of the acquisition process of the objects of study	
	Q20.2	Ethical issues	
	Q20.3	Representativeness of the objects of study selected	
	Q20.4	Consideration of all assumptions of the inference technique to be used	
	Q20.5	Specification of the treatments planning	
	Q20.6	Design of the instruments and procedures to apply the treatments	
	Q20.7	Design of the measurement instruments and procedures	
Execution	Q21	Do you think that is necessary to report what actually happened during the execution of an empirical research about the following issues?	Nominal
	Q21.1	Deviations from the acquisition plan of objects of study	
	Q21.2	Deviations from the treatment plan	
	Q21.3	Deviations from the measurement plan	
Results	Q22	Do you usually explain your observations in terms of underlying mechanisms or available theories?	Nominal
	Q23	Do you usually assess the plausibility of your explanations?	
	Q24	Do you usually answer the research questions explicitly?	
	Q25	Do you usually verify that the contributions to improvement goal are described in your report?	
	Q26	Do you usually verify that the contributions to knowledge goal are described in your report?	
	Q27	Do you think that the following practices would be useful to improve the report of your empirical results?	5-item Likert
	Q27.1	The use of mechanisms or available theories to explain your observations	
	Q27.2	Plausibility assessment of your explanation	
	Q27.3	Plausibility assessment of tested hypotheses	
	Q27.4	Contributions to improvement goal	
	Q27.5	Contributions to knowledge goal	

2.4 Respondents' characteristics

As is shown in Figure 1, the survey response captured a diverse of range of roles, since Master students from academia to Senior consultants from industry. 17 out of 42 respondents were PhD candidate (40,5%), only one of them worked also in the industry sector. The other almost half of respondents were senior researchers (42,9%), where 15of them come from academia, 2 from industry and 1 from both sectors.

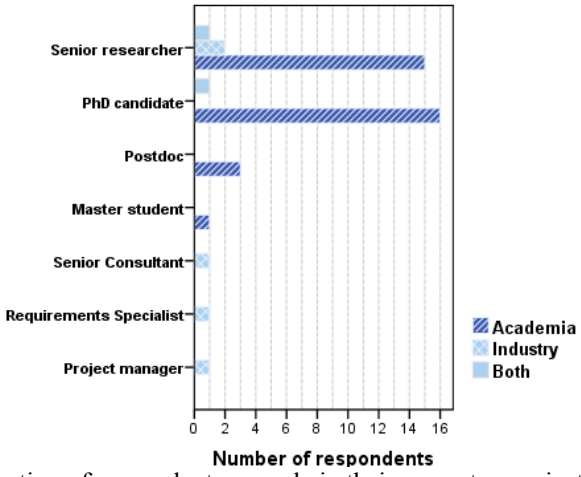

Figure 1. Distribution of respondents per role in their current organization

The survey participants also reflect a diverse range of experience with requirements engineering (See Figure 2) and empirical research (See Table 2).

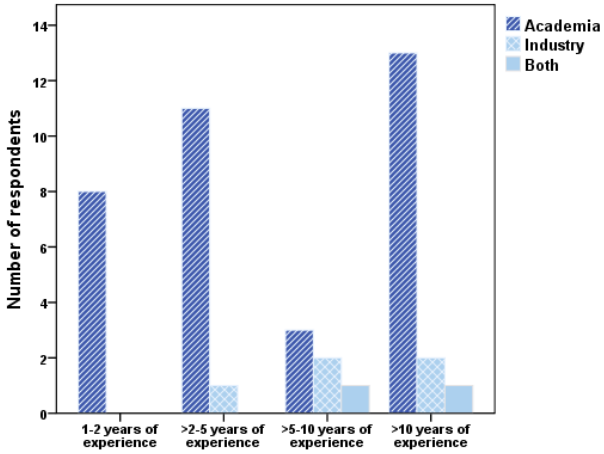

Figure 2. Distribution of respondents' experience with Requirements Engineering

Table 2. Experience in designing experiments or cases studies.

Number of times		Sector			Total
		Academia	Industry	Both	
designing experiments	>30	0	0	0	0
	>20-30	0	1	0	1
	>10-20	4	0	0	4
	>5-10	6	1	1	8
	1-5.	21	2	1	24
	0	4	1	0	5
Total		35	5	2	42

Number of times		Sector			Total
		Academia	Industry	Both	
designing case studies	>30	1	1	0	2
	>20-30	2	1	0	3
	>10-20	4	1	1	6
	>5-10	3	1	0	4
	1-5	21	0	1	22
	0	4	1	0	5
Total		35	5	2	42

3 Survey results

This section provides the most significant observations found, which are organized in five sections that corresponding to the phases of the empirical cycle, such as was mentioned in Section 2.2.

An analysis of frequencies per research question was carried out, as well as a chi-square test was applied with the purpose of knowing whether for the two groups with major percentage of participants (PhD Students, and senior researchers) there are significant differences in their respective opinions about their common practices (See Appendix, Table 7).

3.1 Research context. Common practices on the contextualization of the problem to be empirically investigated were collected from the first three questions (Q1, Q2, and Q3) of the questionnaire. This set of questions corresponds to the recommended practices listed in the unified checklist proposed by Wieringa [1].

As is shown in Figure 3, 35 out of 39 respondents (89%) acknowledge that they usually review the current knowledge related to their empirical research (Q3). 32 of them (82%) stated that they usually define a knowledge goal when investigating an engineering problem (Q2). It is important to remark that 6 respondents did not get to understand this question. 3 out of these 6 respondents were post-Docs, 2 PhD students, and 1 a senior researcher. However, 34% out of 39 responses stated that they omit the definition of improvement goals in their empirical research (Q1) as part of their practice. Only 1 respondent reported the question as not understandable. This respondent was a senior researcher with a medium level of empirical experience.

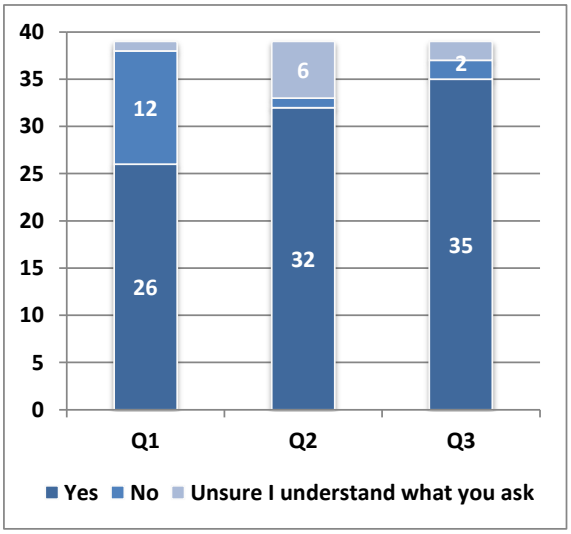

Figure 3. Distribution of practices on contextualization of empirical research problems

Applying the Chi-square test of goodness of fit, we found that the definition of improvement goal (Q1) can be considered as a common practice but only among senior researchers (p=0,004). However, for the definition of knowledge goal (Q2) and review of the current state of empirical knowledge (Q3), we corroborated enough evidence to consider them to be common practices among PhD students and senior researchers (p=0,001).

Table 3 shows that the percentage of neutral responses was higher for the first recommended practice "definition of improvement goal" than for the other two practices. This means that 23% of respondents preferred to choose a neutral position. In general terms, respondents tend to perceive the last two practices as very useful (above 50%).

Table 3. Perceived usefulness of the recommended practices for contextualizing

	Perceived Usefulness				
Question	**1 (not useful)**	**2**	**3 (neutral)**	**4**	**5 (very useful)**
Q4.1	2.9%	0.0%	22.9%	34.3%	40.0%
Q4.2	0.0%	2.9%	14.3%	25.7%	57.1%
Q4.3	0.0%	0.0%	11.1%	27.8%	61.1%

3.2 Research problem. Figure 4 shows our observations collected from the next five questions(Q5-Q9); where we can note that the practice with highest percentage of respondents (97%) is the "formulation of research questions" (Q8), followed surprisingly by the "description of the population to be investigated" practice (Q9) with a 89% of respondents. We also noted that only 57% of respondents recognized to the "definition of relevant concepts of the phenomena to be investigated" (Q5) as part of their common practices. The other half of respondents stated that they did not consider this practice in their empirical studies (22%) or simply were not able to understand the question (18%). Figure 4 also illustrates that the total of affirmative responses for question Q6 and Q7 decrease drastically. This is because the Q6 and Q7 were enabled only if respondents answered the respective previous question (Q5 and Q6) affirmatively. Thus, only 23% indicated that the validation of the most relevant concepts previously operationalized is considered in their empirical research.

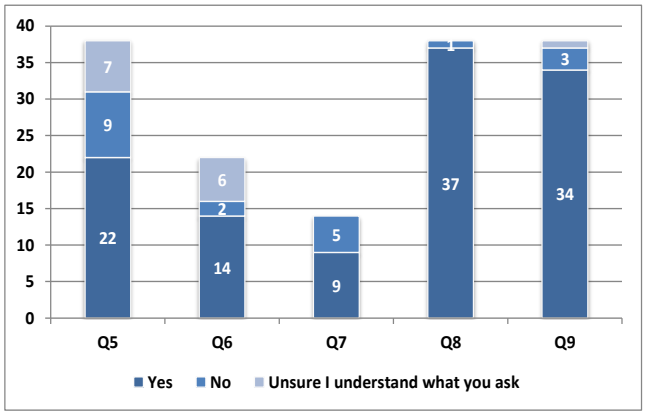

Figure 4. Practices applied to enable a better understanding of a research problem

Applying the Chi-square test for this set of questions, we found enough evidence only for the last two practices recommended for understanding better the problem to be investigated empirically: formulation of research questions and description of population. ($p<0,05$).

Analyzing the distribution of frequencies for usefulness perceived (

Table **4.** Perceived usefulness of the practices recommended for understanding the research problem

	Perceived Usefulness				
Question	**1 (not useful)**	**2**	**3 (neutral)**	**4**	**5 (very useful)**
Q10.1	0.0%	2.9%	11.4%	25.7%	60.0%
Q10.2	0.0%	6.5%	25.8%	29.0%	38.7%
Q10.3	0.0%	3.0%	27.3%	21.2%	48.5%

Q10.4	0.0%	0.0%	2.7%	18.9%	78.4%
Q10.5	0.0%	2.7%	13.5%	32.4%	51.4%

3.3 Research design and justification. In this section, we report our results collected from the questions (Q11-Q19.6) formulated in order to know which of the practices are most applied by the respondents for getting better research designs and justifications. Figure 5 shows that the practice of "justifying the acquisition process of the object of study" is the one that is least applied by the respondents (48%); followed by the practice of "considering all assumptions of inference techniques" (17 out of 37). In both cases, a considerable number of respondents found difficulties to understand these questions (Q11 and Q14). This can be due to the fact that the questions were rather ambiguous, or that respondents are not familiarized with the terminology, precisely because these recommended practices are not applied by them.

We also noted that 35% of respondents did not consider any ethical issue in their empirical research (Q12). This observation can be due to the fact that respondents are partially aware of the meaning of ethics (e.g. they can believe that ethical issues should only be considered where experiments could induce life threatening conditions in humans).), we can see that only 38.7% of respondents perceived the practice "operationalization of the relevant concepts" as very useful, while 26% chose a neutral response.

We also noted that although the "description of population" practice was considered as a common practice by the senior researchers and PhD students, only 51% of respondents perceived this practice as very useful and 32% as useful. A possible explanation could be that majority of our respondents were more familiarized with case studies, where concepts on population and operationalization are not sufficiently addressed by respondents.

Table 4. Perceived usefulness of the practices recommended for understanding the research problem

	Perceived Usefulness				
Question	**1 (not useful)**	**2**	**3 (neutral)**	**4**	**5 (very useful)**
Q10.1	0.0%	2.9%	11.4%	25.7%	60.0%
Q10.2	0.0%	6.5%	25.8%	29.0%	38.7%
Q10.3	0.0%	3.0%	27.3%	21.2%	48.5%
Q10.4	0.0%	0.0%	2.7%	18.9%	78.4%
Q10.5	0.0%	2.7%	13.5%	32.4%	51.4%

3.3 Research design and justification. In this section, we report our results collected from the questions (Q11-Q19.6) formulated in order to know which of the practices are most applied by the respondents for getting better research designs and justifications. Figure 5 shows that the practice of "justifying the acquisition process of the object of study" is the one that is least applied by the respondents (48%); followed by the practice of "considering all assumptions of inference techniques" (17 out of 37). In both cases, a considerable number of respondents found difficulties to understand these questions (Q11 and Q14). This can be due to the fact that the questions were rather ambiguous, or that respondents are not familiarized with the terminology, precisely because these recommended practices are not applied by them.

We also noted that 35% of respondents did not consider any ethical issue in their empirical research (Q12). This observation can be due to the fact that respondents are partially aware of the meaning of ethics (e.g. they can believe that ethical issues should only be considered where experiments could induce life threatening conditions in humans).

On the other hand, considering that questions Q15 and Q16 showed only whether the respondents had experience in designing experiments, we noted that 3 out of 4 respondents, who did not understand the question Q16, were senior researchers with a high level of empirical experience. However, 10 of 28 respondents who stated that they consider this practice ("specification of any instrument to apply the treatments"), were also researchers with a high level of empirical experience.

Applying the chi-square test, we found that although 28 respondents answered affirmatively to the question Q16; there is only a significant difference in the opinions given by PhD students (p=0,001) but not by senior researchers (p=0,02). For questions Q13 (justification of the representativeness of the object of study for the population), Q17 (specification of any instrument for measurement), and Q18 (specification of procedures to be followed when performing measurements), we found enough evidence to affirm that these three practices are those most applied by our respondents.

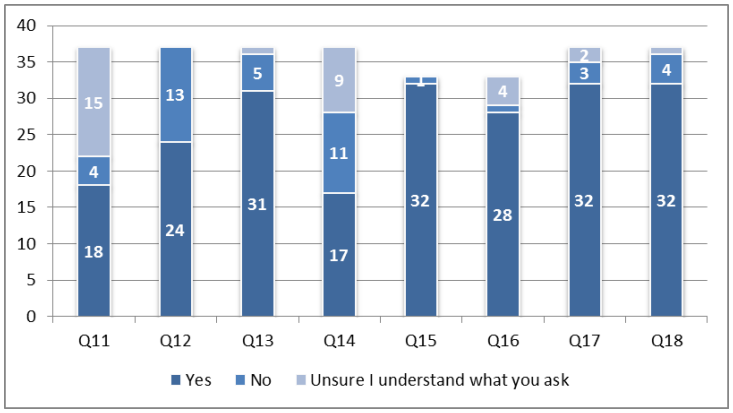

Figure 5. Practices applied to get a better research design and justification (part I)

Figure 6 shows results about the practices recommended regarding the validity of measures (Q19.1), measurement procedures (Q19.2), measurement instruments (Q19.3), treatments (Q19.4), treatment procedures (Q19.5) and treatment instruments (Q19.6). More than 70% of respondents stated that they apply the first four practices in their research. However, we corroborated that the last two practices recommended are only applied by senior researchers.

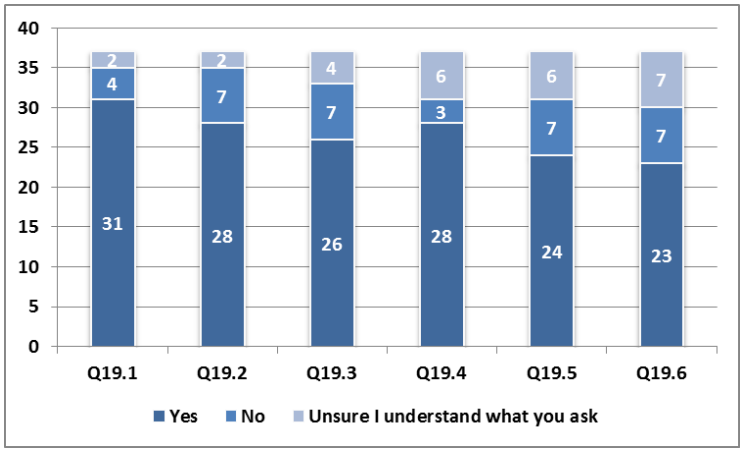

Figure 6. Practices applied to enable better research design and justification (part II)

Analyzing the distribution of frequencies for usefulness perceived (

Table **4.** Perceived usefulness of the practices recommended for understanding the research problem

	Perceived Usefulness				
Question	**1 (not useful)**	**2**	**3 (neutral)**	**4**	**5 (very useful)**
Q10.1	0.0%	2.9%	11.4%	25.7%	60.0%
Q10.2	0.0%	6.5%	25.8%	29.0%	38.7%
Q10.3	0.0%	3.0%	27.3%	21.2%	48.5%
Q10.4	0.0%	0.0%	2.7%	18.9%	78.4%
Q10.5	0.0%	2.7%	13.5%	32.4%	51.4%

3.3 Research design and justification. In this section, we report our results collected from the questions (Q11-Q19.6) formulated in order to know which of the practices are most applied by the respondents for getting better research designs and justifications. Figure 5 shows that the practice of "justifying the acquisition process of the object of study" is the one that is least applied by the respondents (48%); followed by the practice of "considering all assumptions of inference techniques" (17 out of 37). In both cases, a considerable number of respondents found difficulties to understand these questions (Q11 and Q14). This can be due to the fact that the questions were rather ambiguous, or that respondents are not familiarized with the terminology, precisely because these recommended practices are not applied by them.

We also noted that 35% of respondents did not consider any ethical issue in their empirical research (Q12). This observation can be due to the fact that respondents are partially aware of the meaning of ethics (e.g. they can believe that ethical issues should only be considered where experiments could induce life threatening conditions in humans).), we can see that only 16.7% of respondents perceived the practice "justification of the acquisition of the object study" as very useful, while 30% chose a neutral response.
We also noted that although the "specification of measurement instruments and procedures" practices were considered as a common practice by the senior researchers and PhD students, only 47% of them perceived both practices as very useful and 23.5% preferred to choose a neutral response. Once, this could be due to that majority of our respondents were more familiarized with case studies, where measurement concepts are less used than by researchers familiarized with experiments.

Table 5. Perceived usefulness of the practices recommended for research design and justification

	Perceived Usefulness				
Question	**1 (not useful)**	**2**	**3 (neutral)**	**4**	**5 (very useful)**
Q20.1	6.7%	13.3%	30.0%	33.3%	16.7%
Q20.2	8.8%	29.4%	20.6%	11.8%	29.4%
Q20.3	0.0%	3.0%	18.2%	30.3%	48.5%
Q20.4	0.0%	12.5%	18.8%	28.1%	40.6%
Q20.5	0.0%	9.1%	24.2%	30.3%	36.4%
Q20.6	6.3%	9.4%	25.0%	18.8%	40.6%
Q20.7	5.9%	2.9%	23.5%	20.6%	47.1%

3.4 Research execution

Concerning the questions on research execution, the respondents mostly declared that they understand the questions. However, it is noteworthy that in Q21.1, about the report of deviations from acquisition plan of objects study, there were a higher number of subjects who were unsure about the meaning of this practice in comparison to other questions in this section (see Figure 7).

Overall, these answers suggest that nearly 90% of the participants do consider it necessary to report what actually happened during the execution of empirical research, in terms of deviations from either the acquisition plan of objects of study (Q21.1), or the treatment plan (Q21.2), or the measurement plan (Q21.3).
Applying the chi-square test, we found that although 26 respondents answered affirmatively to the question Q21.1; there is only enough evidence to confirm that "the report of deviations from the acquisition plan of objects of study" is a common practice among PhD students (p=0,002) but not by senior researchers (p=0,041). However, reporting the deviations from the treatment and measurement plans are considered valuable information to be reported (by senior researchers and PhD students).

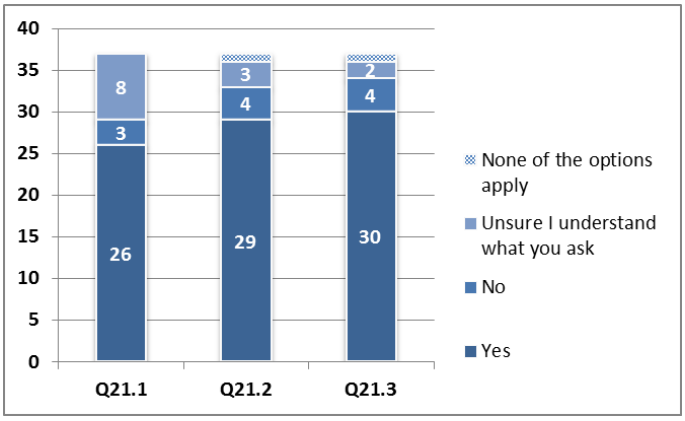

Figure 7. Research execution practices

3.5 Results analysis.

Questions Q22 through Q26 concern what the participants say that they do when analyzing their results (see Figure 8).
Regarding the terminology used, everyone understood the question Q24, but a few respondents answered that they were unsure about the meaning of "explain observations in terms of underlying mechanisms or available theories" (Q22), or "assess the plausibility of explanations" (Q23), or "verify that contributions to the improvement/knowledge goal are described" (Q25 and Q26).
According to what people usually do in their analyses, we can say that nearly 90% of the participants try to answer the research questions explicitly. In contrast, about 22% of the participants (majority of them PhD Students) affirmed that they do not usually explain their observations in terms of available theories (Q22), which suggests that these researchers follow a more descriptive analysis, simply reporting their observations without making the effort to link it with underlying mechanisms.
Applying the chi-square test, we corroborated that the first two practices (Q22 and Q23) are usually applied by senior researchers (p=0,004) but not by PhD students (p=0,04).

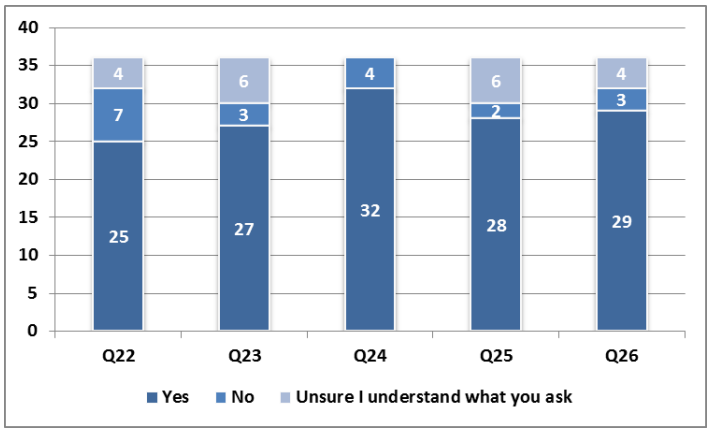

Figure 8. Result analysis practices

Prior questions dealt with what researchers do commonly when they analyze their results. However, it is also interesting to know more about the perceived usefulness on the recommended practices included in this section. Table 6 shows the results on a 5-point Likert scale of the perceived usefulness for the practices Q27.1-Q27.5. The results show that the participants mostly consider useful or very useful all the practices recommended in order to improve result analysis.

Table 6. Perceived usefulness of practices recommended for obtaining better empirical reports

	Perceived Usefulness				
Question	**1 (not useful)**	**2**	**3 (neutral)**	**4**	**5 (very useful)**
Q27.1	2.8%	0.0%	5.6%	30.6%	61.1%
Q27.2	2.9%	0.0%	17.6%	23.5%	55.9%
Q27.3	0.0%	3.0%	9.1%	21.2%	66.7%
Q27.4	6.1%	0.0%	9.1%	27.3%	57.6%
Q27.5	0.0%	0.0%	6.1%	36.4%	57.6%

4 Summary and Conclusions

This paper describes a study of the empirical research practices in the requirements engineering community. Although our survey was distributed to all attendees (researchers and practitioners) of one of the premium conferences in RE, our conclusions are drawn only from experiences of PhD Students and Senior Researchers. This is because we got much more responses from academia than industry. Next, we list our conclusions, followed by brief explanations.

- *Practices on contextualization of empirical research*: We observe that the definition of improvement goals appears to be a common practice but only among senior researchers. This can be explained with the abilities of senior researchers to put the research at hand in a perspective and connect it to a 'bigger picture', a broader scope to which the research relates. Moreover, PhD students usually start their research with exploration in mind and if their purpose is merely exploratory , the definition of improvement goals may even not be required. However, we found enough evidence that lets us consider the definition of knowledge goals and review of the current state of empirical knowledge as common practices among both RE senior researchers and PhD students. Our respondents, regardless of their professional experience levels, perceived these three recommended practices as useful.
- *Practices applied to enable a better understanding of a research problem*: Our study found that only two out of the five practices that were recommended for understanding the problem, were actually used. These are the formulation of research questions and the description of the population to be investigated. This lets us conclude that the problems investigated by researchers could not be being fully understood due to a lack of definition of relevant concepts of the phenomena, as well as their respective operationalization and validation. This lack of conceptualization could be due to a lack of many theories in our field. Also, it could be possibly explained by the relatively limited use of existing theories from other disciplines in the area of RE. We also found that about 26% of our respondents did not perceive either the operationalization of concepts or their validation as useful. They preferred to choose a neutral position for both practices. A possible explanation would be that as these two practices are not currently required for publishing case studies, the level of awareness is relatively low.
- *Practices applied to enable a better research design and justification*. We found that the justification of the representativeness of the object of study, specification and validation of measurement instrument, and measurement procedures were identified as the practices most applied by our respondents. However, we make the note that in some cases using these practices alone may not be enough for getting good enough research designs and justification. This is because a good research design usually includes thorough consideration of ethical issues, justification of study object selection, and assumptions of inferences techniques. We believe that these practices are candidates worthwhile including in research designs by RE researchers.
- *Practices applied to enable better reports on research execution*. From our observations, we concluded that deviations from the original plans of the treatment plan or the measurement plan are considered valuable information to be reported. However, we also found evidence that only for PhD students seem to be necessary to report what actually happened during the execution in terms of deviations from the acquisition plan of objects of study to the end point of their research process. Some of the senior researchers responded to this question by saying that they did not understand what was meant in the question. We think that the answer of the senior researchers can be considered a reflection of the need of senior researchers for more precision and more elaboration of the meanings embedded in the question.

- *Practices applied to enable better reports on results analysis*. The most surprising finding is that the majority of PhD Students does not usually explain their observations in terms of underlying mechanisms or available theories, which suggests that they follow a more descriptive analysis. We could find only evidence of good practices for better report on results analysis among senior researchers.
- Finally, we also observed that the respondents tend to give greater importance to the practices recommended in order to improve result analysis than to practices recommended in order to get better research design and justifications.

Acknowledgments

This work was in part funded by the Intra European Marie Curie Fellowship Grant 50911302 PIEF-2010. The authors would like also thank all the participants of this survey.

References

1. Wieringa, R.J. (2012) A Unified Checklist for Observational and Experimental Research in Software Engineering (Version 1). TR-CTIT-12-07, CTIT, UT, Enschede. ISSN 1381-3625
2. Wieringa R. J. Design science as nested problem solving. ACM 4 the DESRIST, 2009 , pp. 1–12.
3. Kitchenham, B.A., Pfleeger, S.L.: Principles of Survey Research - Part 3: Constructing a Survey Instrument. SIGSOFT Software Engineering Notes 27, 20–24 (March 2002).
4. Pfleeger S., "Experimental design and analysis in software engineering," Annals of Software Engineering, vol. 1, no. 1, pp. 219–253, 1995.
5. Kitchenham B. A., Pfleeger Sh. L., Pickard L. M., Jones P. W., Hoaglin D. C., El Emam K., and Rosenberg J. Preliminary guidelines for empirical research in software engineering. IEEE Trans. Softw. Eng. 28(8), 2002, pp. 721-734.
6. Jedlitschka A., Pfahl D., "Reporting guidelines for controlled experiments in software engineering," IEEE ISESE 2005, pp.94-104.
7. Runeson P. and Höst M. 2009. Guidelines for conducting and reporting case study research in software engineering. Empirical Softw. Engineering. 14(2), 2009, pp. 131-164.

Appendix

Table 7. Chi-square statistics for questions of the survey-common practices

Job role		**Q1**	**Q2**	**Q3**	**Q5**	**Q6**
PhD candidate	Chi-square	,529	14,58	18,47	3,87	,66
	df	1	2	2	2	2
	Asymp. Sig.	,467	,001	,000	,144	,717
Senior researcher	Chi-square	11,20	11,26	11,26	5,200	5,44

	df	2	1	1	2	1
	Asymp. Sig.	,004	,001	,001	,074	,020
Job role		**Q7**	**Q8**	**Q9**	**Q11**	**Q12**
PhD candidate	Chi-square	,333	12,25	9,000	4,625	2,250
	df	1	1	1	2	1
	Asymp. Sig.	,564	,000	,003	,099	,134
Senior researcher	Chi-square	,500		19,200	3,600	1,667
	df	1		2	2	1
	Asymp. Sig.	,480		,000	,165	,197
Job role		**Q13**	**Q14**	**Q15**	**Q16**	**Q17**
PhD candidate	Chi-square	9,000	,125	8,333	13,50	12,875
	df	1	2	1	2	2
	Asymp. Sig.	,003	,939	,004	,001	,002
Senior researcher	Chi-square	19,20	6,40		5,40	11,26
	df	2	2		1	1
	Asymp. Sig.	,000	,041		,020	,001
Job role		**Q18**	**Q19.1**	**Q19.2**	**Q19.3**	**Q19.4**
PhD candidate	Chi-square	12,87	9,875	9,875	9,875	12,85
	df	2	2	2	2	2
	Asymp. Sig.	,002	,007	,007	,007	,002
Senior researcher	Chi-square	11,26		5,400	10,80	14,80
	df	1		1	2	2
	Asymp. Sig.	,001		,020	,005	,001
Job role		**Q19.5**	**Q19.6**	**Q21.1**	**Q21.2**	**Q21.3**
PhD candidate	Chi-square	3,87	2,000	12,500	16,625	16,625
	df	2	2	2	2	2
	Asymp. Sig.	,144	,368	,002	,000	,000
Senior researcher	Chi-square	14,80	14,80	6,400	11,26	19,20
	df	2	2	2	1	2
	Asymp. Sig.	,001	,001	,041	,001	,000
Job role		**Q22**	**Q23**	**Q24**	**Q25**	**Q26**
PhD candidate	Chi-square	6,50	6,500	6,250	12,50	12,50
	df	2	2	1	2	2
	Asymp. Sig.	,039	,039	,012	,002	,002
Senior researcher	Chi-square	11,20	19,20		8,06	19,20
	df	2	2		1	2
	Asymp. Sig.	,004	,000		,005	,000