A qualitative survey of regression testing practices

Engström, Emelie; Runeson, Per

Published in:
Product-Focused Software Process Improvement/Lecture Notes in Computer Science

DOI:
10.1007/978-3-642-13792-1_3

2010

A Qualitative Survey of Regression Testing Practices

Emelie Engström and Per Runeson

Department of Computer Science, Lund University, SE-221 00 LUND, Sweden
{Emelie.Engstrom,Per.Runeson}@cs.lth.se

Abstract. *Aim*: Regression testing practices in industry have to be better understood, both for the industry itself and for the research community. *Method*: We conducted a qualitative industry survey by i) running a focus group meeting with 15 industry participants and ii) validating the outcome in an on line questionnaire with 32 respondents. *Results*: Regression testing needs and practices vary greatly between and within organizations and at different stages of a project. The importance and challenges of automation is clear from the survey. *Conclusions*: Most of the findings are general testing issues and are not specific to regression testing. Challenges and good practices relate to test automation and testability issues.

Keywords: Regression testing, Survey, Industry practice.

1 Introduction

Regression testing is retesting of previously working software after a change to ensure that unchanged software is still functioning as before the change. According to IEEE, regression testing is *Selective retesting of a system or component to verify that modifications have not caused unintended effects and that the system or components still complies with its specified requirements* [1]. The need for effective strategies for regression testing increases with the increasing use of iterative development strategies and systematic reuse in software projects. Studies indicate that 80% of testing cost is regression testing and more than 50% of software maintenance cost is related to testing [2].

There is a gap between research and practices of regression testing. Research on regression testing mainly focuses on selection and prioritization of test cases. Several techniques for regression test selection are proposed and evaluated. Engström *et al.* reviewed the literature in the field recently [3] and highlights the importance of the test context to the outcome of regression testing techniques. Only few empirical evaluations of regression test selection techniques are carried out in a real industrial context [4], [5], [6].

However industry practice on regression testing is mostly based on experience alone, and not on systematic approaches. There is a need for researchers to better understand the needs and practices in industry. Rooksby *et al.* [7] argue for the need for investigation and characterization of real world work. They conclude

M. Ali Babar, M. Vierimaa, and M. Oivo (Eds.): PROFES 2010, LNCS 6156, pp. 3–16, 2010.

that improvements of current testing practices are meaningful in its specific local context and "cannot be brought about purely through technically driven innovation". In their paper they highlight, based on experiences from testing in four real projects, that improvements in industry are not always sophisticated and accurate as is often pursued in research.

In order to retrieve a better understanding of real world needs and practices, a qualitative survey [8, p. 61-78] of industry practice of regression testing is conducted, by means of focus group discussions in a software process improvement network (SPIN) and a questionnaire to validate the results. Issues discussed in the focus group were definitions and practices of regression testing in industry as well as challenges and improvement suggestions. A total of 46 software engineers from 38 different organizations participated in the focus group and questionnaire survey. Results are qualitative and of great value in that they highlight relevant and possible directions for future research.

To the extent of our knowledge no industrial surveys on regression testing practices have been reported on. However experience reports on regression testing in industrial software development projects can be found [9]. Onoma *et al.* conclude that regression testing is used extensively and that several companies develop in-house regression testing tools to automate the process. Re-test all is a common approach and the selection of test cases is not a critical issue.

When it comes to testing practices in general a couple of industrial surveys have been undertaken [10], [11], [12], [13], concluding that test automation is a key improvement issue [13] and that test case selection for continuous regression testing is a hard task. No systematic approach for test case selection was used by the companies but instead they relied on the developers expertise and judgment [12].

This paper is organized as follows: Section 2 describes how the survey is conducted and discusses validity issues. In section 3 results are presented and analyzed. Finally conclusions are provided in section 4.

2 Method Description

The study's overall goal is to characterize current regression testing practices in industry for the sake of research. It also aims at identifying good practices for spreading across different companies as well as areas in need for improvement within the companies and possibly identification of future research topics. Hence, a qualitative survey is found appropriate [8, p. 61-78]. The research questions for the survey are:

RQ1 What is meant by *regression testing* in industry?
RQ2 Which *problems* or *challenges* related to regression testing exist?
RQ3 Which *good practices* on regression testing exist?

The survey is conducted using two different research methods, one focus group discussion [14, p. 284-289] in a SPIN group, and one questionnaire in a testing interest network. The focus group was used to identify concepts and issues related to regression testing, while the questionnaire was used to validate the findings

in a different setting. A similar approach was used for a unit testing survey in 2006 [12].

2.1 Focus Group

The focus group meeting was arranged at one of the monthly meetings of SPIN-syd, a software process improvement network in Southern Sweden [15]. The members of the network were invited to a 2.5 hour session on regression testing in May 2009. 15 industry participants accepted the invitation, which is about the normal size for a SPIN-syd monthly meeting, and the same as for our previous unit testing survey [12]. The focus group meeting was moderated by two academics and one industry participant, and observed by a third academic. An overview of the focus group participants is shown in Table 1.

Table 1. Participants in focus group meeting. Number of developers in the surveyed company: extra small is 1, small is $2-19$, medium is $20-99$, and large $100-999$.

Company	Domain	Size	Role
A	Automation	Medium	Participant
A	Automation	Medium	Participant
A	Automation	Medium	Participant
G	Medical devices	Medium	Participant
G	Medical devices	Medium	Participant
I	Information systems	Large	Moderator
I	Information systems	Large	Participant
S	Telecom	Large	Participant
S	Telecom	Large	Participant
E	Telecom	Large	Participant
X	Consultant	Extra small	Participant
C	Consultant	Extra small	Participant
Q	Consultant	Medium	Participant
K	Consultant	Medium	Participant
O	Consultant	Large	Participant
L	Academics	N/A	Researcher
L	Academics	N/A	Researcher
L	Academics	N/A	Observer

The industry participants represented automation, medical devices, information systems (IS), and telecom domains. Consultants also participated which were working with testing for their clients. The product companies all produce embedded software and were both of medium and large size, while consultancy firms of all sizes were represented.

The session was organized around five questions:

- What is regression testing?
- When do the participants regression test?
- How do the participants regression test?

- What are the participants' problems regarding regression testing?
- What are the participants' strengths regarding regression testing?

For each of the questions, the moderator asked the participants to write their answers on post-it charts. Then each participant presented his or her view of the question and the responses were documented on white boards.

After the session, key findings were identified using qualitative analysis methods. Statements were grouped into themes, primarily structured by the five questions, and secondary according to keywords in the statements. Further, the results were restructured and turned into questions for use in the questionnaire.

2.2 Questionnaire

The resulting questionnaire consists of 45 questions on what regression testing is, with five-level Likert-scale response alternatives: *Strongly disagree, Disagree, Neutral, Agree, Strongly Agree* and an additional *Not Applicable* option (see Fig 1). One question on automation vs manual used five scale alternatives from *Automated* to *Manual* (see Fig 2). Further, 29 questions on satisfaction with regression testing practices in the respondents' organizations had the response alternatives *Very Satisfied, Satisfied, Neutral, Dissatisfied, Very Dissatisfied* and *Not Applicable* (see Fig 3). The questionnaire was defined in the SurveyGizmo questionnaire tool for on line data collection [16].

Respondents were invited through the SAST network (Swedish Association for Software Testing) through their quarterly newsletter, which is distributed to some 2.000 testers in Sweden, representing a wide range of company sizes and application domains. Respondents were promised an individual benchmarking report if more than three participants from one company responded, and a chance for everybody to win a half-day seminar on testing given by the second author. Thirty-two respondents answered the complete questionnaire, which are presented in Table 2.

The respondents cover the range of company sizes and domains. Out of the 32 respondents, 9 were developing embedded systems in particular within the telecom domain, 12 developed information systems in particular within the domains of business intelligence and finance, and 11 were consultants. Out of 21 product companies, 3 represent small development organizations, 9 represent medium sized organizations and 8 represent large organizations. The size of the consultancy organizations are not specifically relevant, but is reported to indicate the variation.

2.3 Threats to Validity

The study does not aim at providing a statistically valid view of a certain population of companies, as intended with general surveys [8]. The research questions are focused on existence and not on frequencies of responses. Hence, we consider the survey having more character of multiple case studies on a certain aspect of several cases and consequently we discuss threats to validity from a case study perspective [17].

Question	Item	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
What is regression testing?	1. Repetitive tests	0	3	3	14	11
	2. Retest of functionality	0	0	1	8	22
	3. Reexecution of testcases	0	1	7	12	11
	4. Same as system testing	11	12	4	2	1
Why is regression testing applied?	5. To assess if the system has the desired properties	1	12	6	8	5
	6. To find defects	0	4	4	15	9
	7. To ensure that nothing has been affected or destroyed	0	0	0	2	30
	8. To guide further priorities in the project	1	9	9	11	0
What kinds of changes generate regression testing?	9. New versions	0	0	3	11	18
	10. New configurations	0	1	7	11	12
	11. Fixes	0	0	6	9	16
	12. Changed solutions	0	1	3	8	18
	13. New hardware	1	5	5	8	11
	14. New platforms	1	3	4	6	17
	15. New designs	0	1	7	9	15
	16. New interfaces	0	2	4	10	15
	17. RT is applied regardless of changes	3	12	6	5	5
At which levels are RT carried out?	18. Single components	3	4	7	11	4
	19. Single modules	1	3	5	15	5
	20. Whole system	0	0	1	9	22
When in the development process is RT applied?	21. As early as possible	3	6	6	7	7
	22. Continuosly during the whole process	1	4	5	11	9
	23. At the end	3	1	3	9	16
	24. Daily	8	11	6	2	1
	25. At each software integration	3	4	9	9	5
	26. At each milestone	1	6	6	9	6
	27. Before each release	0	0	1	9	22
	28. As often as we have resources	6	9	5	6	3
What determines the amount and frequency of RT?	29. The assessed risk	0	0	3	14	13
	30. The amount of new functionality	0	2	2	15	12
	31. The amount of fixes	0	2	3	16	9
	32. The amount of available resources	4	6	6	8	6
Which tests are used in regression testing?	33. A selection of developer's tests	5	10	6	6	2
	34. A selection of tester's tests	0	2	1	21	8
	35.A selection from a specific regression test suite	0	4	1	8	18
	36. New test cases are designed	1	9	8	12	2
How are regression test cases selected?	37. The same tests are run each time	1	4	13	8	6
	38. Selection depends on the situation	0	4	6	13	9
	39. We do a complete retest each time	2	8	12	3	6
	40. We do a complete retest of safety critical parts	0	2	9	10	8
	41. Test cases on changes and possible side effects	1	2	5	14	10
	42. A selection is made ad hoc	9	12	7	3	0
	43. Run as many as possible from a prioritized list	4	9	7	6	4
	44. We focus on functional test cases	0	2	8	14	8
	45. We execute smoke test	1	4	11	10	5

Fig. 1. Number of responses for each questionnaire alternative on regression test practices

Construct validity concerns the underlying constructs of the research, i.e. terms and concepts under study. We mitigated construct validity threats by having the first question of the focus group related to terminology and concepts. Thereby, we ensured a common understanding for the rest of the group meeting. In the

Question		Manual		Equal		Automated
46. How do you execute regression tests?		8	2	15	4	3

Fig. 2. Number of responses for each questionnaire alternative on automated vs. manual regression testing

Question	Item	Very dissatisfied	Dissatisfied	Neutral	Satisfied	Very satisfied
How satisfied are you with following in your organization?	47. Processes/practices for change impact analysis	4	7	9	10	1
	48. Assessment of the extent of test coverage	2	6	12	8	2
	49. Assessment of the amount of required tests	1	4	11	14	1
	50. Prioritization of test cases wrt product risks	1	3	9	16	3
	51. Prioritization of test cases wrt fault detection ability	0	4	13	10	2
	52. Time to design good test cases	4	6	9	10	1
	53. Methods/tool support to design good test cases	3	12	9	4	4
	54. Assessment of cost/benifit of automating RT	3	11	12	2	4
	55. Time to RT	3	12	7	7	2
	56. Balance between manual and automated RT	3	14	4	4	4
	57. Execution of automated RT	2	12	8	2	3
	58. Environment for automated RT	2	14	7	3	2
	59. Execution of manual RT	0	2	9	19	0
	60. RT in real target environment	0	4	8	15	4
	61. RT in simulated target environment	0	5	9	13	3
	62. RT of GUI	2	5	7	15	2
	63. RT of data base applications	1	4	10	13	0
	64. RT of third party products	1	5	15	6	0
	65. Consistentcy of verdict reporting	0	3	14	11	1
	66. Time to analyze results	1	3	15	12	1
	67. Processes/practices for analyzing results	1	5	14	10	1
	68. Presentation of results from automated tests	2	8	6	5	4
	69. Maintenance of tests in case of changes in products	3	7	13	8	1
	70. Methods/tool for traceability between TC and reqs	6	9	6	7	3
	71. Minimization of redundant tests (wrt test coverage)	2	10	8	11	0
	72. Coordination between designers and testers	1	1	8	21	1
	73. Minimization of dependencies in the system	1	12	12	6	0
	74. Modularization of the system	0	9	15	7	0
	75. Testability issues in design guidelines	2	9	14	5	0

Fig. 3. Number of responses for each questionnaire alternative on satisfaction with regression test practices

survey, however, the terms may be interpreted differently and this is out of control of the researchers.

Internal validity relates to identification of casual relationships. We do not study any casual relationships in the study, and we just briefly touch upon correlations between factors. Patterns in the data that might indicate correlations are interpreted conservatively in order not to over interpret the data.

Table 2. Respondents to the questionnaire. Number of developers in the surveyed company: extra small is 1, small is $2-19$, medium is $20-99$, and large $100-999$.

Company	Size	Domain
Me	Small	Automation
Te	Medium	Automation
V	Large	Automotive
Tc	Small	Business intelligence
Ql	Medium	Business intelligence
Ti	Medium	Business intelligence
C	Large	Consultant
Ha	Large	Consultant
H	Large	Consultant
H	Large	Consultant
Q	Medium	Consultant
R	Small	Consultant
K	Medium	Consultant
Si	Large	Consultant
So	Large	Consultant
T	Small	Consultant
Tp	Medium	Consultant
Eu	Medium	Finance
Sk	Large	Finance
A	Medium	Finance
U	Medium	Information systems
Sm	Medium	Information systems
W	Small	Information systems
B	Large	Information systems
L	Large	Insurance
Mu	Large	Insurance
Ma	Large	Medical devices
E	Large	Telecom
Hi	Medium	Telecom
M	Medium	Telecom
S	Large	Telecom
S	Large	Telecom

External validity relates to generalization from the findings. We do not attempt to generalize in a statistical sense; any generalization possible is analytical generalization [17]. In order to help such generalization, we report characteristics of the focus group members and questionnaire respondents in Tables 1 and 2.

3 Analysis of the Results

The focus group and survey results were analyzed using the Zachman framework, which originally was presented for analysis of information systems architectures [18]. The framework has six categories, *what, how, where, who, when* and *why*,

although these terms were not originally used. For each category, questions are defined and tailored to the domain under investigation. Originally intended for IS development, Zachman proposed that it might be used for developing new approaches to system development [18]. We use it similar to Runeson [12], i.e. to structure the outcome of the focus group meetings and to define the validation questionnaire, although we primarily focus on *what, how* and *when*.

An overview of the questionnaire results is shown in Figures 1, 2 and 3. Questions are referred to in the text as [Qx] for question *x*. The analysis is then presented according to the framework questions and identified strengths and weaknesses in subsections 3.1 to 3.4.

3.1 What?

There is good agreement in the focus group and among the survey respondents regarding what regression testing is. Regression testing involves repetitive tests and aims to verify that previously working software still works after changes to other parts. Focus can be either re-execution of test cases or retest of functionality. As for testing in general the goal of the regression testing may differ between different organizations or parts of an organization. The goal may be either to find defects or to obtain a measure of its quality. Regression testing shall ensure that nothing has been affected or destroyed, and give an answer to whether the software has achieved the desired functionality, quality and stability etc. In the focus group discussion, an additional goal of regression testing was mentioned as well; to obtain a guide for further priorities in the project. Regression testing offers a menu of what can be prioritized in the project, such as bug fixes. This additional goal was only confirmed to some extent by 35% of the respondents [Q8].

Different kinds of changes to the system generate regression testing. Mentioned in the focus group discussion and confirmed by the majority of the respondents were: new versions, new configurations, fixes, changed solutions, new hardware, new platforms, new designs and new interfaces [Q9-16]. One third of the respondents, mostly small and medium sized organizations, indicated that regression testing is applied regardless of changes, while in larger organizations, regression testing was tighter connected to changes [Q17]. The amount and frequency of regression testing is determined by the assessed risk, the amount of new functionality, the amount of fixes and the amount of available resources. The first three factors are confirmed by the majority of the respondents [Q29-31] while the agreement on the dependency on resources availability varies to a greater extent among the respondents [Q32].

3.2 When?

Regression testing is carried out at different levels (e.g. module level, component level and system level [Q18-20]) and at different stages of the development process. From focus group discussions it was found that that some organizations regression test as early as possible while other regression test as late as possible in the process, and some claimed that regression testing is continuously carried

out throughout the whole development process. The purpose may be slightly different for the three options; early regression test to enable early detection of defects, and late regression testing for certification or type approval purposes.

How often regression testing is carried out differed as well; some organizations regression test daily while others regression test at each software integration, at each milestone, or before releases [Q24-26]. In some cases the availability of resources is determinant. Among the questionnaire responses, there were large variations on how often regression testing is applied. The most common approach is to regression test before releases (indicated by 95% of the respondents) [Q27]. Only 10% of the respondents regression test daily [Q24].

3.3 How?

From the focus group discussions it was identified that tests used for regression testing may be a selection of developer's tests, a selection of tester's tests, a selection of tests from a specific regression test suite, or new test cases are designed. According to questionnaire responses, the most common is to reuse test cases designed by testers. Strategies for regression test selection mentioned in the focus group were: complete retest, combine static and dynamic selection, complete retest of safety critical parts, select test cases concentrating on changes and possible side effects, ad-hoc selection, smoke test, prioritize and run as many as possible, and focus on functional test cases. Questionnaire results confirm that it is common to run a set of specified regression test cases every time, together with a set of situation dependent test cases. Ad-hoc selection seems not to be a common approach; only 10% of the respondents indicate that approach [Q42]. 70% of the respondents confirm the focus on functional test cases [Q44] and 50% confirm the usage of smoke tests [Q45].

A project may include several different regression testing activities. Both manual and automatic regression testing are applied. 50% of the respondents indicate an equal amount of manual and automatic regression testing while 30% perform regression testing exclusively manually [Q46].

3.4 Weaknesses and Strengths

The focus group had an open discussion about both weaknesses and strengths in their regression testing practices, and it showed that in several cases representatives from one organization had solution proposals where others had problems. Some problems were common to most of the participants (e.g. lack of time and resources to regression test and insufficient tool support) while others were more specific. The outcome of the discussion was a list of 29 possible problem areas which were validated in the questionnaire.

Test case selection. Several problems related to test case selection were discussed in the focus group. It was mentioned that it is hard to assess the impact of changes on existing code and to make a good selection. It is hard to prioritize test cases with respect to product risks and fault detection ability, and to be

confident in not missing safety critical faults. Determining the required amount of tests was also considered a problem, and it is hard to assess the test coverage.

Participants wished for a regression test suite with standard test cases and for regression testing guidelines at different stages of a project with respect to quality aspects. Some participants were satisfied with their impact analysis and with their test management systems. As a response to the test selection problem, exploratory testing was recommended and also to have a static test set used for each release. No specific test selection technique was referred to, such as the ones reviewed by Engström *et al.* [3].

The results from the questionnaire responses are in this respect not conclusive. The responses are divided evenly across the whole spectrum, with a slight shift towards satisfaction. However, in terms of processes for impact analysis and assessment of test coverage the challenges identified in the focus group where confirmed by a third of the respondents even though as many were satisfied. [Q47-51].

Test case design. Lack of time and resources for regression testing was a recurring complaint in the discussions. So also in the case for test case design. Among respondents to the survey were as many satisfied as dissatisfied in this matter [Q52]. One proposal mentioned in the focus group was to focus on test driven development and thus make developers take test responsibility, hence building test automation into the development process, which may be reused for regression testing purposes as well.

Automated and manual regression testing. Automating regression testing causes problems and manual testing is time and resource consuming. Both problems and proposals were discussed in the focus group. Within the focus group, participants were satisfied and dissatisfied with automation as well as with their manual testing. Most participants wanted a better balance between automated and manual testing and support in determining cost benefit of automating regression testing.

It is not only costs for implementing the automated tests that need to be considered, but also costs for maintaining the test suites and in many cases manual analysis of results. It was proposed to define interfaces for automation below the user interface level in order to avoid frequent changes of the test scripts, due to user interface changes. Use of manual testing was recommended for testing of user experience and for exploratory testing.

The problems of automation was confirmed by questionnaire responses. 60% of the respondents were dissatisfied with the balance between manual and automated regression testing [Q56], the assessment of cost/benefit, execution of automated regression tests as well as the environment for automated regression testing. In contrast, as many were satisfied with their manual testing, 60% [Q59].

Regression testing problem areas. Specific problem areas for regression testing, mentioned in the discussion forum were: regression tests in real target environment and in simulated target environment, regression testing of third party

products and of GUI's. For each problem mentioned, were among the participants both those who had problems and those who were satisfied with their solutions. None of the problem areas was confirmed by a majority of negative answers in the questionnaire even though between 10-25% were dissatisfied in each case [Q60-64]. As testing of databases is subject to regression testing research, this area was added to the questionnaire, although not mentioned in the focus group.

Test results. Several of the participants in the focus group were unsatisfied with how test results were presented and analyzed. In many cases verdict reporting is inconsistent and often there is no time to do a thorough analysis. Some participants said that their reporting of results and analysis works well and gave examples of good factors, such as having an independent quality department and having software quality attributes connected to each test case, which is good not only for for reporting results but also for prioritization and selection of test cases.

The questionnaire responses were generally neutral regarding consistency of verdict reporting and processes and practices for analyzing results, but agreed that practices for presentation of results from automated tests were not good enough [Q68].

Test suite maintenance. The focus group named maintenance of test suites and test cases as a problem. Participants stated that much of the regression testing is redundant with respect to test coverage and that there is a lack of traceability from tests to requirements. Some of the participants were satisfied with their tools and processes for traceability and claimed that they are good at maintenance of test cases in case of changes in the product. A recommendation was to have independent review teams reviewing the test protocols.

Questionnaire responses confirmed the lack of good tools for documenting traceability between test cases and requirements but otherwise the variation in the responses to the questions regarding maintenance was great [Q69-71].

Testability. An issue brought up in the focus group were the amount of dependencies in the software and its relation to testability. Participants expressed a wish for a test friendly design where the structure enables a simple delimitation of relevant tests. There is a need for design guidelines considering testability, modularization of the software and clearer dependencies in order to make it easier to set test scopes.

Questionnaire responses indicate satisfaction with coordination/communication between designers and testers [Q72] and neutrality to modularization of the system [Q74]. Further they confirmed the need for minimization of dependencies in the system [Q73] as well as for testability issues in design guidelines [Q75].

Test planning. Finally some needs and recommendations regarding the test planning was given. Again a cost model was asked for: *It would be nice to have a cost model for environments and technical infrastructure covering; automated*

testing, test data, test rigs, unit tests, functional tests, performance tests, target/simulator and test coverage.

Everyone in the focus group agreed that it is better to test continuously than in large batches. A rule of thumb is to plan for as much test time as development time even when the project is delayed. It is also good to have a process with a flexible scope for weekly regression tests, e.g. core automated scope, user scenarios, main regression scope, dynamic scope, dynamic exploratory scope etc. In order to broaden the coverage, it was proposed to vary the test focus between different test rounds.

4 Conclusions

Regression testing increases in software projects as software becomes more and more complex with increasing emphasis on systematic reuse and shorter development cycles. Many of the challenges, highlighted in the study, are *not* specific to regression testing but are general to all testing. However, they have a significant impact on how effective the regression testing becomes. Questions involving automated testing is of course particularly important for regression testing, as the same tests are repeated many times. Similarly, a test-friendly design is of great importance when one wants to do a selective retesting. Literature on regression testing tends to focus on the selection of test cases based on changes in the code, but for practitioners it does not seem to be the most important issue.

Regression testing definitions (RQ1) are very much the same across all surveyed companies and in line with formal definitions [1] although the regression testing practices differ. Regression testing is applied differently in different organizations, at different stages of a project, at different levels and with varying frequency. Regression testing is not an isolated one-off activity, but rather an activity of varying scope and preconditions, strongly dependent on the context in which it is applied. In most development organizations, regression testing is applied continuously and at several levels with varying goals. This further underlines the need for industrial evaluations of regression testing strategies, where context information is clearly reported, as was previously noted [3].

Regression testing challenges (RQ2) relate to test case selection, trade-offs between automated and manual testing and design for testability. Issues related to test automation are:

- Assessment of cost/benefit of test automation
- Environment for automated testing and the presentation of test results.

Design issues affect regression testing since there is a strong relation between the effort needed for regression testing and the software design. Design for testability, including modularization with well defined and observable interfaces, helps verifying modules and their impact on the system. This could be addressed by including testability in design guidelines. Except for the design issues, coordination and communication between designers and testers work well.

Good practices (RQ3) were also reported on:

- Run automated daily tests on module level.
- Focus automation below user interface.
- Visualize progress monitoring.

These practices are not specific to regression testing. The latter item is not specific testing at all, but is a management practice that becomes critical to regression testing as it constitutes a key part of the development project progress. This indicates that regression testing should not be addressed nor researched in isolation; rather it should be an important aspect of software testing practice and research to take into account.

Acknowledgment

The authors would like to thank Per Beremark for moderating the focus group meeting and to all participants in the focus group and questionnaire. The work is partly funded by The Swedish Governmental Agency for Innovation Systems (VINNOVA) in the UPPREPA project under grant 2005-02483, and partly by the Swedish Research Council under grant 622-2004-552 for a senior researcher position in software engineering.

References

1. IEEE: IEEE standard for software test documentation. IEEE Std(829-1983, Revision) (1998)
2. Chittimalli, P.K., Harrold, M.J.: Recomputing coverage information to assist regression testing. IEEE Transactions on Software Engineering 35(4), 452–469 (2009)
3. Engström, E., Runeson, P., Skoglund, M.: A systematic review on regression test selection techniques. Information and Software Technology 52(1), 14–30 (2010)
4. Engström, E., Runeson, P., Wikstrand, G.: An empirical evaluation of regression testing based on fix-cache recommendations. In: Proceedings of the 3rd International Conference on Software Testing Verification and Validation, pp. 75–78 (2010)
5. Skoglund, M., Runeson, P.: A case study of the class firewall regression test selection technique on a large scale distributed software system. In: International Symposium on Empirical Software Engineering., pp. 72–81 (2005)
6. White, L., Robinson, B.: Industrial real-time regression testing and analysis using firewalls. In: Proceedings 20th IEEE International Conference on Software Maintenance, pp. 18–27 (2004)
7. Rooksby, J., Rouncefield, M., Sommerville, I.: Testing in the wild: The social and organisational dimensions of real world practice. Computer Supported Cooperative Work (CSCW) 18(5), 559–580 (2009)
8. Flink, A.: The survey handbook, 2nd edn. SAGE Publications, Thousand Oaks (2003)
9. Onoma, A.K., Tsai, W.T., Poonawala, M.H., Suganuma, H.: Regression testing in an industrial environment: Progress is attained by looking backward. Association for Computing Machinery. Communications of the ACM 41(5), 81–86 (1998)

10. Causevic, A., Sundmark, D., Punnekkat, S.: An industrial survey on contemporary aspects of software testing. In: Proceedings of the 3rd International Conference on Software Testing Verification and Validation, pp. 393–401 (2010)
11. Grindal, M., Offutt, J., Mellin, J.: On the testing maturity of software producing organizations. In: Testing: Academia & Industry Conference-Practice And Research Techniques, TAIC/PART (2006)
12. Runeson, P.: A survey of unit testing practices. IEEE Software 23(4), 22 (2006)
13. Runeson, P., Andersson, C., Höst, M.: Test processes in software product evolution - a qualitative survey on the state of practice. Journal of Software Maintenance and Evolution: Research and Practice 15, 41–59 (2003)
14. Robson, C.: Real World Research, 2nd edn. Blackwell Publishing, Malden (2002)
15. Runeson, P., Beremark, P., Larsson, B., Lundh, E.: SPIN-syd - a non-profit exchange network. In: 1st International Workshop on Software Engineering Networking Experiences, Joensuu, Finland (2006)
16. Surveygizmo (December 2009) a web tool for questionnaires and polls, `http://www.surveygizmo.com`
17. Runeson, P., Höst, M.: Guidelines for conducting and reporting case study research in software engineering. Empirical Software Engineering 14(2), 131–164 (2009)
18. Zachman, J.A.: A framework for information systems architecture. IBM Systems Journal 26(3), 276–293 (1987)