Mutual knowledge transfer between industry and academia to improve testing with defect taxonomies

Michael Felderer[1], Armin Beer[2]

[1]University of Innsbruck, Innsbruck, Austria
michael.felderer@uibk.ac.at

[2]Beer Test Consulting, Baden, Austria
info@arminbeer.at

Abstract: Software engineering is an applied research area preferably conducted jointly by academia and industry to enable transfer of knowledge in both directions and at the end improvement software engineering in industry. In this paper we present how the required mutual knowledge transfer via empirical evaluation was performed to improve testing with defect taxonomies in industry.

1 Introduction

Empirical studies have a value stand-alone, but they can also be part of a knowledge exchange and improvement endeavor jointly between academia and industry [WRH+12]. Software engineering is an applied research area [WRH+12], from which research on industrially relevant problems is expected. It is in many cases insufficient to just do academic research on, for instance, requirements engineering or software testing with the motivation that these areas are challenging in industry [WRH+12]. Industry-relevant research involves more than merely producing research results and delivering them in publications and technical reports. It demands close cooperation and collaboration between industry and academia throughout the entire research process [GWC+06]. Thus, software engineering is preferably conducted jointly by academia and industry to enable transfer of knowledge in both directions and at the end transfer of new methods, technologies or tools from academia to industry [WRH+12]. Joint research with mutual knowledge transfer provides an excellent opportunity to improve industrial software development based on concrete empirical evidence, and hence being a good example of evidence-based software engineering [KDJ04].

In this paper we present how the required mutual knowledge transfer via empirical evaluation was performed to improve (requirements-based) testing with defect taxonomies in industry. Testing with defect taxonomies exploits defect taxonomies to support all phases of the system test process, i.e., test planning, design, execution and evaluation [FB14a].

The remainder of this paper is structured as follows. In Section 2 we provide an overview of testing with defect taxonomies. In Section 3 we present the process of mutual knowledge transfer between industry and academia to improve testing with defect taxonomies. Finally, in Section 4 we conclude and present future work.

2 Testing with defect taxonomies

In practice, most defect taxonomies are only used for the a-posteriori allocation of testing resources to prioritize failures for debugging purposes. But the full potential of these taxonomies to control and improve all phases of the overall system test process has so far remained unexploited. To overcome this limitation, in previous publications [FB12], [FB13a], [FB13b], [FB14a], [FB14b], [FBP14] we defined and evaluated *requirements-based testing with defect taxonomies* (RTDT) which supports all phases of testing requirements, i.e., test planning, design, execution and evaluation. In RTDT, the defect taxonomy is created on basis of the requirements specification. Then the requirements are linked to defect categories and specific requirements quality attributes like completeness, traceability or comprehensibility are validated [FB13a]. Afterwards tests are planned taking defect taxonomies into account which results in a test strategy that assigns test design techniques to combinations of defect categories and requirements. From the test strategy abstract and executable test cases are designed and then executed. Finally, the test results are evaluated and failures are assigned to defect categories. In empirical case studies [FB12], [FB13b] performed in a public health insurance institution in Austria it has been shown that RTDT has the potential to improve the effectiveness [FB13b] and efficiency [FB12] of system testing. Furthermore, in a student experiment we showed that RTDT is independent of a specific type of defect taxonomy [FBP14].

3 Mutual knowledge transfer between industry and academia

In this section we present the process of mutual knowledge transfer between industry and academia to improve testing with defect taxonomies. The knowledge exchange was mainly performed between a researcher and a test manager who are both authors of this paper. The steps of the knowledge transfer process between industry and academia are shown in Figure 1 and presented in the following.

Initially, *defect taxonomies were informally applied for testing* in projects of a public health insurance institution in Austria by a test manager. In this institution a standard test process based on an iterative development process is applied. In several discussions with a researcher from the University of Innsbruck, knowledge on how defect taxonomies are actually applied was transferred to the researcher and the potential of defect taxonomies to improve testing in practice but also potential for research was jointly recognized.

Then, the researcher performed a *state-of-the-art analysis* on defect taxonomies as well as their application for software testing and requirements engineering. In several telephone conferences and face-to-face discussions, the knowledge on state-of-the art was transferred from the researcher to the test manager. The resulting state-of-the-art analysis is presented in [FB13a] and [FB13b].

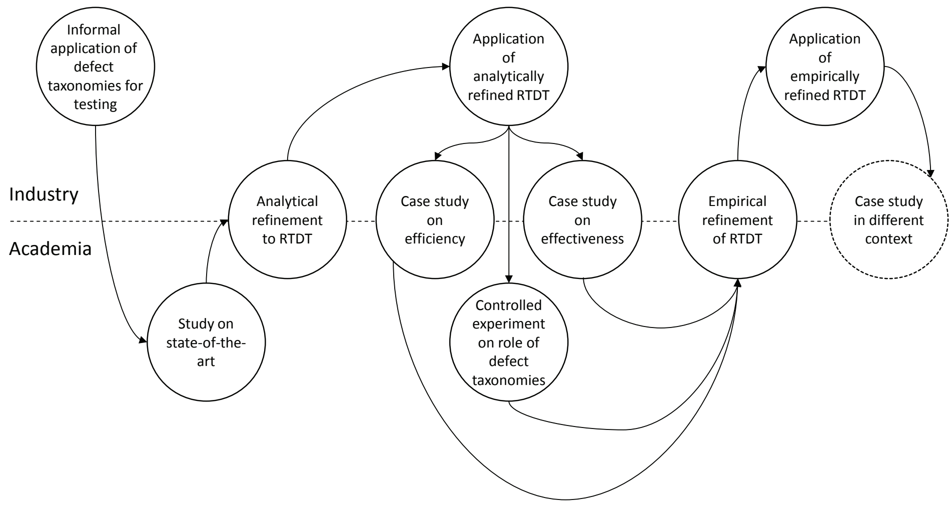

Figure 1: Steps of mutual knowledge transfer between industry and academia

On the basis of knowledge of state-of-the-art, the application of defect taxonomies for testing was *analytically refined* by the researcher and the test manager to RTDT. In course of this refinement, the process steps of RTDT, the underlying test design techniques, and the applied taxonomy itself were formalized. In addition, the potential of RTDT to also validate requirements quality was recognized and elaborated. The refined approach of RTDT is presented in [FB13a] and [FB13b].

Afterwards RTDT was *applied in industry*. The informal feedback of this application was positive but also raised two main issues. First, empirical evidence whether the approach improved effectiveness and efficiency of testing was required to confirm the informal feedback and to derive profound guidelines for successful application of RTDT. Second, as the application showed that the applied defect taxonomy is a critical artifact of RTDT, the role of defect taxonomies required further investigation.

To show improved effectiveness and efficiency of RTDT, *empirical case studies on effectiveness and efficiency* were performed. A case study in software engineering is an empirical enquiry that draws on multiple sources of evidence to investigate one instance (or a small number of instances) of a contemporary software engineering phenomenon in its real-life context, especially when the boundary between phenomenon and context cannot be clearly specified [RHR+12]. The case studies were jointly performed by the researcher and the test manager in context of the standard test process of the investigated public health insurance institution and showed improved effectiveness (based on the observed defects) and efficiency (based on the time needed for testing with defect taxonomies compared to testing without defect taxonomies). For this purpose, case studies were selected as empirical strategy as RTDT can best be investigated in its industrial context. The case studies are based on test projects of the public health insurance institution where RTDT was applied on the basis of the established standard test process. The case study on effectiveness is presented in [FB13b] and on efficiency in [FB12].

To investigate the role of defect taxonomies, a *controlled experiment on the role of defect taxonomies* was performed. In an industrial context the role of defect taxonomies could not be investigated in isolation. We therefore performed a controlled experiment with student participants at two universities in Austria, i.e., Graz University of Technology and University of Innsbruck. For sure, external validity can be threatened when experiments are performed with students, and the representativeness of these subjects may be doubtful in comparison to that of test professionals. However, our experimental setting was similar to system-level testing in industry, where it is common to recruit test personnel with some domain knowledge, but only little experience in systematic testing. The required skills in test design are then provided in short trainings (i.e., similar to preparation units in our experiment). Similar to students, also testers in industry have to be concerned with the requirements when a new project starts. Therefore, we think the student experiment could be considered as appropriate, because students and test professionals are in a comparable situation with respect to investigate the role of defect taxonomies. The experiment on the role of defect taxonomies is presented in [FBP14].

Based on the case studies and the student experiment, an *empirical refinement of RTDT* was performed. Extraction of knowledge from the empirical investigations by the researcher and the test manager resulted in a refined process for RTDT and concrete practical guidelines for the application of RTDT. Guidelines comprise the management of the underlying defect taxonomy, the required tool support, the quality of requirements, preconditions of the development and test process as well as the cost-benefit analysis of the approach. The empirical refinements of RTDT are presented in [FB14a] and [FB14b].

The *empirically refined process of RTDT* and the guidelines are now successfully applied in a public health insurance institution in Austria. The knowledge gathered from the case studies and the experiments helped to improve RTDT in practice.

RTDT has so far only been applied in a public health insurance institution. But we expect similar benefits in other domains. We therefore plan *further empirical case studies* on the effectiveness and efficiency of RTDT in different contexts.

4 Conclusion

In this paper we presented how mutual knowledge transfer via empirical evaluation was performed to improve requirements-based testing with defect taxonomies in industry. Requirements-based testing with defect taxonomies exploits defect taxonomies to support all phases of the system test process, i.e., test planning, design, execution and evaluation, and thereby improves its effectiveness and efficiency. The mutual knowledge transfer between industry and academia was jointly performed in several steps by a test manager from a public health insurance institution and a researcher from the University of Innsbruck. The steps for mutual knowledge transfer are (1) informal application of defect taxonomies for testing, (2) study on state-of-the-art, (3) analytical development of RTDT, (4) application of analytical refinement, (5) empirical evaluation in case studies and a controlled experiment, (5) empirical refinement, as well as (6) application of empirical refinement. Appropriate size and complexity of the projects selected for the case studies, support by the management, as well as established procedures to quantify process aspects turned out to be key prerequisites in industry to enable knowledge transfer with academia in case of RTDT. Main success factors for the mutual knowledge transfer were the personal commitment of a researcher and an industrial test manager, the stepwise refinement of the approach by analytical and empirical methods, as well as the applied mix of empirical methods which resulted in concrete applicable practical guidelines. In future, we plan further industrial case studies in different contexts. This is accompanied and supported by disseminating the approach via publications in magazines relevant to industry [FB14a], [FB14b], [FB14c] as well as presentations on leading industrial conferences like Eurostar[1] and expo:QA[2].

Acknowledgement. Parts of the work described in this paper were supported by the project QE LaB – Living Models for Open Systems (FFG 822740).

References

[FB12] M. Felderer and A. Beer, Estimating the Return on Investment of Defect Taxonomy-Supported System Testing in Industrial Projects, 38th Euromicro Conference on Software Engineering and Advanced Applications, IEEE, 2012.

[FB13a] M. Felderer and A. Beer, Defect Taxonomies for Requirements Validation in Industrial Projects, 21st International Conference on Requirements Engineering (RE 2013), 2013.

[FB13b] M. Felderer and A. Beer, Using Defect Taxonomies to Improve the Maturity of the System Test Process: Results from an Industrial Case Study, Software Quality Days 2013, Springer LNBIP, 2013.

[FB14a] M. Felderer and A. Beer, Using Defect Taxonomies for Testing Requirements, IEEE Software, 2014, doi: doi.ieeecomputersociety.org/10.1109/MS.2014.56 (online first).

[FB14b] M. Felderer and A. Beer, Practical Guidelines for System Testing with Defect Taxonomies, ASQT 2013, OCG, 2014.

[FB14c] M. Felderer and A. Beer, , Softwaretechnik Trends, Eine industriell erprobte Methode für den Review und Test von Anforderungen mit Hilfe von Fehlertaxonomien. Softwaretechnik-Trends, 34(1), 2014 (in German)

[FBP14] M. Felderer, A. Beer, and B. Peischl, On the role of defect taxonomy types for testing requirements: Results of a controlled experiment, 40th EUROMICRO Conference on Software Engineering and Advanced Applications (SEAA 2014), 2014.

[GWC+06] T. Gorschek, C. Wohlin, P. Carre, and S. Larsson, A model for technology transfer in practice, IEEE Software, 23(6), 88-95, 2006.

[KDJ04] B. A. Kitchenham, T. Dyba, M. Jorgensen, Evidence-based software engineering, 26th International Conference on Software Engineering (ICSE 2004), IEEE, 2004.

[RHR+12] P. Runeson, M. Host, A. Rainer, A., and B. Regnell,, Case study research in software engineering: Guidelines and examples. John Wiley & Sons, 2012.

[WRH+12] C. Wohlin, P. Runeson, M. C. Höst, B. Regnell, and A. Wesslén, Experimentation in Software Engineering, Springer, 2012.

[1] http://conference.eurostarsoftwaretesting.com/event/2012/leveraging-defect-taxonomies-for-testing-results-from-a-case-study/ (last accessed January 11, 2015)

[2] http://www.expoqa.com/en/conference-programa-detallado.php#e217 (last accessed January 11, 2015)