Value-Based Empirical Research Plan Evaluation

Stefan Biffl Dietmar Winkler
Vienna University of Technology, Institute of Software Technology and Interactive Systems
{biffl, winkler}@qse.ifs.tuwien.ac.at

ABSTRACT
Empirical studies are crucial to gain evidence on the effects of software engineering methods and tools in defined contexts. However, empirical studies can be costly and thus need to be planned well to achieve best benefits for study stakeholders, e.g., researchers and industry partners, with limited resources and reasonable project risks (e.g., insufficient validity, unaddressed stakeholder win conditions). The design of empirical studies typically covers issues of empirical methodology, but seldom discusses tradeoffs between conflicting study goals explicitly. Unaddressed conflicts during study planning can put the study success at risk and study planners may overlook study variations that could offer better value or lower cost/risk.

This paper proposes a value-based empirical research planning framework for eliciting and reconciling research stakeholder win conditions in order to compare the benefits and risks of potential empirical study variants. We reports findings from an initial feasibility study in a session at a recent ISERN meeting of empirical research experts, where teams of empirical researchers applied the process to their studies. Main results of the feasibility study were: a) empirical researchers found the process useful and easy to use; b) several stakeholders, win conditions, and conflicts were found that were not addressed in the original study designs but could be accommodated with reasonable changes; c) an investment of some 60 minutes allowed in several cases to improve potential value contribution or mitigate risks in a study design.

Categories and Subject Descriptors
D2.9 Management

General Terms
Experimentation, Economics, Measurement.

Keywords
Empirical study design evaluation, value-based empirical research, risk management.

1. Introduction
In recent software engineering research two trends can be observed: empirical evidence has become an explicit evaluation criterion in many important publication channels, which fosters the need for appropriate study planning for a growing variety of research contexts [1][3]; and value-based software engineering (VBSE) [6][7][8] has gained momentum in linking technical research and development to project stakeholders and their win conditions, which may conflict and introduce new demands and risks on project planning. Empirical studies can take considerable effort and thus should be designed well to contribute the expected value. The plan of an empirical study often has to trade-off expected stakeholder value with reasonable effort, cost, and risk.

While a growing number of empirical studies are reported [12][14], researchers who are new to empirical work are often surprised by the complexity of issues to take into account for conducting an empirical study with reasonable validity [10][13]. On the other hand practitioners complain about empirical studies that seem focused on academic needs and seem to provide less value to industry stakeholders [2]. These symptoms reflect the need to support researchers in planning an empirical study with an approach, which helps to elicit and reconcile research stakeholder win conditions. Such an approach should help to elicit stakeholder value propositions and link study deliverables to stakeholder win conditions in order to explicitly support negotiating the tradeoffs between study variants. In addition to strengthening the value of an individual study, the consideration of potential industry partners and the EMSE community as explicit stakeholders can help to ease conducting empirical studies with industry participants and building a empirical body of knowledge in EMSE areas.

The remainder of the paper is structured as follows: Section 2 summarizes recent trends in empirical study planning and the application of concepts from value-based software engineering. Section 3 introduces the process for value-based empirical research study plan evaluation. Section 4 reports on an initial feasibility study at the session on “value-based empirical research” at the ISERN 2006 meeting; and Section 5 summarizes lessons learned and next research steps.

2. Empirical Study Planning Issues
In the last few years researchers interested in building a body of SE evidence from several related empirical studies in a particular area of SE found a number of issues in comparing the results and contributions of different empirical studies [14]. As a consequence they have proposed standardized ways to report empirical studies more comprehensively [12][13]. However, these reporting guidelines tend to increase the data collection effort of individual researchers and are applicable only to new empirical research projects as it has been found to be very hard to obtain valid additional data on past empirical work.

Further, with the growing number of empirical researchers the need for understanding the interests of the stakeholders involved in proposing, conducting, reporting, and publishing empirical studies gained importance [2]. Typical potential stakeholders in an empirical study are the researchers who conduct the study and/or data analysis; study participants; funding agencies or industrial sponsors; and potential users of the study results in academia and practice. These stakeholders usually have diverse and often conflicting interests, which are only partly represented in study planning as study planners are often focused on dealing

with the multitude of technical and organizational issues to smoothly run the study. In such a setting the interests of the study planners (often academic researchers) is likely to receive highest priority and shape the strengths and limitations of the study.

General examples for (industry) stakeholder value from sessions at ISERN meetings[1] in 2004 and 2005 were: a) to shorten development cycle time; b) to identify and prioritize improvement opportunities for process and/or product quality; c) to provide more effective risk management; d) to quantify impacts, predictions, and tradeoffs; and e) to provide a safe environment to try out new approaches in local industry context. Many sessions at recent ISERN meetings raised the question on how to improve the value of empirical studies for the EMSE research community and/or industrial stakeholders; and identified the following challenges for value of empirical research in SE [6]:

- Stakeholder value may differ between stakeholder situations;
- Stakeholder value may change over time;
- SE phenomena tend to change faster than in other empirical disciplines;
- Data is often competitively sensitive.

Thus, it seems that empirical research study planning can benefit from a value-based approach similar to planning a software development project: define the study scope with explicit stakeholder win conditions and then consider the benefits, costs, and risks of study plan alternatives. While senior empirical researchers may do this already intuitively, we see a need for a repeatable and transitionable approach that can be taught to researchers who are new to the EMSE field and that supports empirical study planning and evaluation:

- Focus on the most important EMSE topics first (long-term vs. short term benefits)
- Consideration of key stakeholders (industry / academia).
- Definition of deliverables / stakeholder value propositions.
- Address conflicts between stakeholder interests
- Dependencies of stakeholder interest realization

3. Value-Based Empirical Research

This paper proposes a value-based empirical research (VBER) planning framework for eliciting and reconciling research stakeholder win conditions in order to compare the benefits and risks of potential empirical study variants for planning and monitoring empirical studies.

The VBER framework builds on the VBSE benefits realization approach (BRA) [9] and consists of a) research study deliverables (the outcome of the study), b) stakeholders and their value propositions (benefits of the study and the derived results for a stakeholder), and c) the relationship between study deliverables and stakeholder value propositions.

The process of the evaluation of value considerations in a research study consists of the following steps (see Figure 3 and an illustrative example in Figure A1):

Step 1: Characterize *empirical study* and its *key deliverables* (e.g., using the GQM template [4]) and identify potential dependency dependencies between key deliverables.

Step 2: Elicit empirical study *principal stakeholders* (industry and academia) and their *key value propositions* expected from conducting the empirical study, e.g., improvement of process characteristics; more accurate/flexible SE results; faster time-to-market, lower cost, improved customer satisfaction, larger body of empirical evidence.

Step 3: Mark whether stakeholder *values support each other* (+) or *clash* (-) identifying potential conflicts that indicate project risks.

Step 4: *Link* study deliverables to stakeholder value propositions (using benefits realization analysis [9]); a link between a study deliverable and stakeholder value propositions may be direct (if the deliverable *per se* contributes to the stakeholder value); where needed, *add success-critical initiatives and assumptions* between study deliverables and stakeholder benefits, e.g., risk countermeasures such as training, extra documentation, or determination of study validity for the local context [5]. Unaddressed activities on indirect deliverable-stakeholder links can put the actual value of the study at risk.

Based on this model one can check whether there are major risk symptoms in the study plan, e.g., study deliverables that seem not to provide value to any stakeholder; value propositions of main stakeholders that seem to be unaddressed by study deliverables; important risks on links between study deliverables and stakeholder value propositions that seem hard or costly to mitigate in the scope of the planned study. The model further allows to explore ways to make the study less risky and/or add extra benefits; to compare effects of proposed empirical study approach to alternative approaches (e.g., industry best practice); and to assess the adequacy of the study (i.e., whether is it worthwhile to conduct a study or even not).

4. Feasibility Study at ISERN 2006

We conducted an initial feasibility study in a session on "Value-Based Empirical Research"[2] at the ISERN 2006 meeting of empirical research experts, where teams of experienced empirical researchers (ISERN members) applied the process to their studies. The goal of the initial feasibility study was, by exploring stakeholder value in empirical studies in the ISERN community, to find out whether the VBER framework described in section 3 could effectively support experienced researchers in a) identifying how well study deliverables can support stakeholder value propositions and b) finding ways to improve the study value by identifying and mitigating potential risks.

[1] http://isern.iese.de/network/ISERN/pub/isern-meetings.html

[2] Related sessions at past ISERN meetings were ISERN 2003 "(Decision-oriented) Empirical Studies in SE and PM"; ISERN 2004 "Best practice for industry cooperation"; and ISERN 2005 "Value-Based Empirical Methods". We currently prepare a technical report with more in-depth data analysis from the study results that we expect to provide online for ISERN/ESEM 2007.

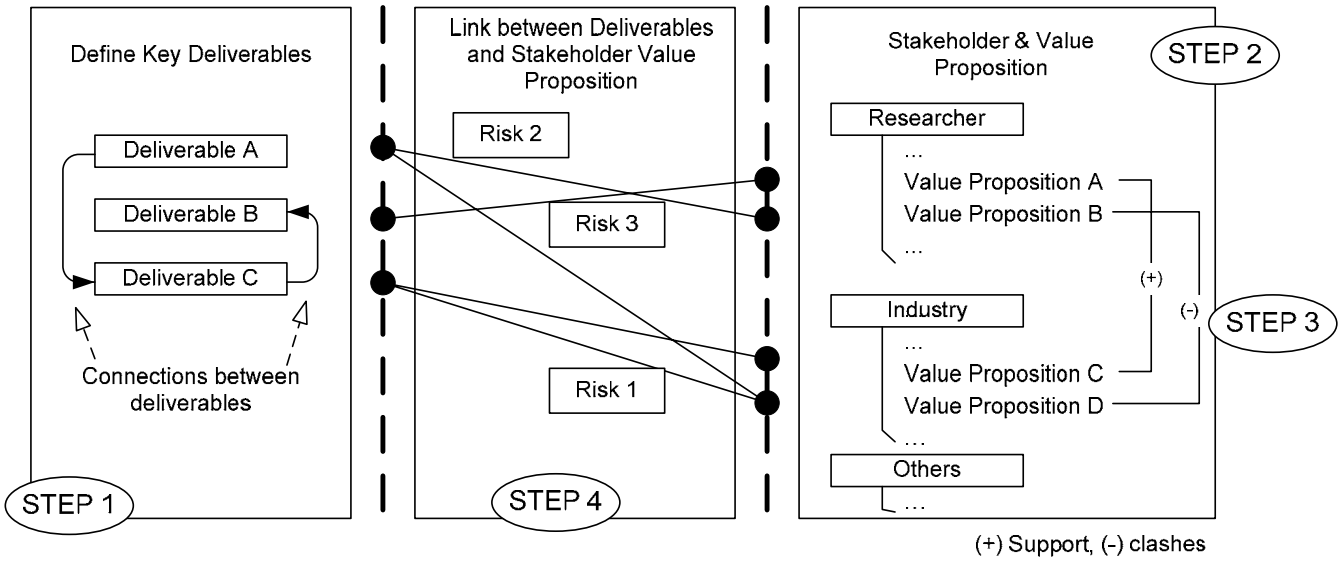

Figure 3. Value-Based Empirical Research (VBER) evaluation framework.

The study participants were asked to brainstorm a list of recent empirical studies they were involved in and then focus on one empirical study in this list. Then the participants followed the steps 1 to 4 listed in section 3. Further, the participants were asked to *consider the time horizon of stakeholder value* in empirical study planning, i.e. to clarify in the group whether stakeholder benefits were rather "short term" or "long term" to find out whether there was a correlation between the time horizon of stakeholder value and the study interest to industry stakeholders.

The group work in 3-person teams worked well and brought lively discussion; participants found the group work interesting and were willing to extend the timeframe for more in-depth analysis of the study plan implications.

Main results of the feasibility study were:

a) The participants found the VBER process useful and easy to use for providing a big picture on the key deliverables of the study, main stakeholders, their value propositions, and relationships between these elements. This overview picture easily allowed spotting missing relationships between key elements that indicate potential omissions and risks in the study plan.

b) While many study deliverables provided value to some stakeholders, such as the academic researchers, often stakeholders and win conditions came up in the brainstorming and analysis that were not addressed in the original study designs but could be accommodated with reasonable changes. However, in some studies clashes were identified not only between the value propositions of industry and research stakeholders, but even among researchers.

c) We could observe a considerable correlation between strong interests of industry stakeholders and short-term focus of study goals.

d) The effort for conducting an initial analysis of a typical empirical study took up to 60 minutes, which seemed like a small investment compared to the study plan improvement due to added potential value contribution or mitigated risks.

5. CONCLUSION

Appropriate plans are the key to conducting valuable empirical studies and need to be evaluated accordingly. In this paper we proposed a value-based approach for evaluating the plan for an empirical research study in order to help elicit success-critical stakeholders (that may have been missing in the plan), identify potential clashes between stakeholder value propositions, deliverables that do not address stakeholder value propositions or unaddressed stakeholder value propositions, and additional initiatives to improve value of deliverables (to industry).

Important lessons learned were: Explicit model of stakeholder value links allows reasoning about risks in empirical study planning; often stakeholder values clash. Interestingly, clashes occur not only between industry stakeholders and researchers, but even among researchers, e.g., if experimenters could provide comprehensive data that would benefit meta-analysis but would have to take considerable extra effort.

Further work will be to refine the framework and ensure that important stakeholders such as potential industry partners and the EMSE community are not forgotten. We plan to analyze the value considerations in EMSE success stories (e.g., CoCoMo, process improvement, the inspection body of knowledge; successful funding of EMSE research organizations and projects in order to support planning empirical studies that make the winners of all involved stakeholders and can strengthen the appeal for industry stakeholders to participate in empirical studies.

Acknowledgements

We want to thank all participants of the ISERN session on "Value-Based Empirical Research" in 2006 for their discussion contributions and for providing valuable input to this work.

REFERENCES

[1] Basili V., Elbaum S.: "*Better Empirical Science for Software Engineering*", Invited Presentation, ICSE 2006,

[2] Basili, V.: "*Is there a future for Empirical Software Engineering*", presentation at ISESE, 2006.

[3] Basili V.: The Role of Experimentation in Software Engineering: Past, Current, and Future, Proc. ICSE, 1996.

[4] Basili V., Caldiera G., Rombach H.D.: "The Experience Factory," in Encyclopedia of Software Engineering, vol 1, J.J. Marciniak, John Wiley & Sons, 1994, pp. 469-476.

[5] Biffl S., Aurum A., Boehm B., Erdogmus H., Grünbacher P.: "*Value-Based Software Engineering*", Springer, 2005.

[6] Biffl S.: "*Value-Based Empirical Research – Application of VBSE stakeholder value propositions to Empirical Research with industry focus*", ISERN session reading material, 2006.

[7] Boehm B.: "Software Engineering is a value-based Contact Sport", IEEE Software, 2002.

[8] Boehm B.: "*Value-Based Software Engineering*", Software Engineering Notes, Vol 28, 2, 2003.

[9] Boehm B.; "Value-Based Software Engineering: Seven Key Elements and Ethical Considerations", in: Biffl et al. (eds.) "*Value-Based Software Engineering*", Springer, 2005.

[10] Freimut B., Punter T., Biffl S., Ciolkowksi M.: „*State-of-the-Art in Empirical Studies*", ViSEK Report, Fraunhofer IESE, 2002.

[11] Jedlitschka A., Ciolkowski M.: "*Towards Evidence in Software Engineering*", presentation at ISESE, 2004.

[12] Jedlitschka A., Pfahl D.: "*Reporting Guidelines for Controlled Experiments in Software Engineering*", ISESE 2005.

[13] Kitchenham B., Pfleeger, S.L., Pickard L.M., Jones P.W., Hoaglin D.C., El Emam, K., Rosenberg J.: "*Preliminary Guidelines for Empirical Research in Software Engineering*", IEEE Transactions on Software Engineering Vol 28, 2002.

[14] Kitchenham B., Dyba T., Jorgensen M.: "*Evidence-Based Software Engineering*", Proc. ICSE 2004.

APPENDIX

Figure A1 shows an illustrative example from the feasibility study in the ISERN session.

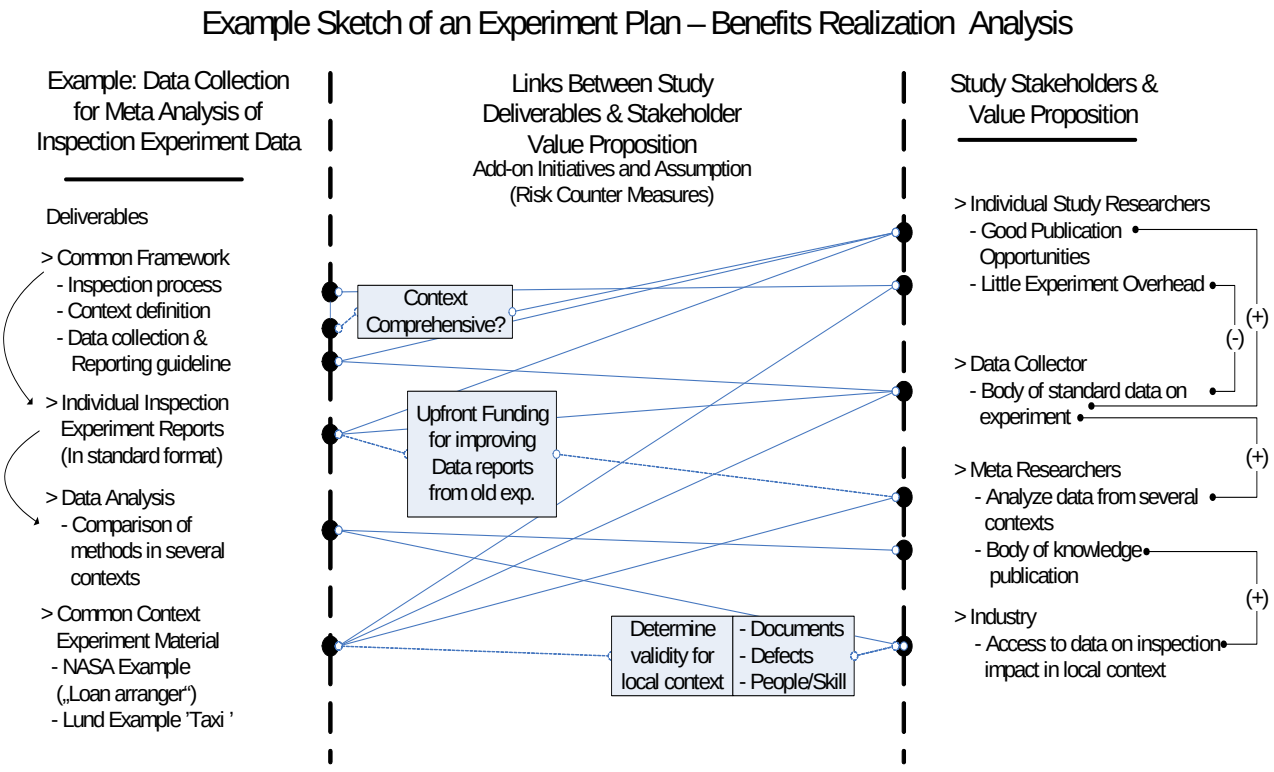

Figure A1: Value-based research planning sketch.