Citation for the published Conference paper:

Title:
Action research as a model for industry-academia collaboration in the software engineering context

Author:
Kai Petersen, Cigdem Gencel, Negin Asghari, Dejan Baca, Stefanie Betz

Conference Name:
International Workshop on Long-term Industrial Collaboration on Software Engineering

Conference Year:
2014

Conference Location:
Västerås

Action Research as a Model for Industry-Academia Collaboration in the Software Engineering Context

Kai Petersen
Blekinge Institute of Technology
Karlskrona, Sweden
kai.petersen@bth.se

Cigdem Gencel
Free University of Bolzano
Faculty of Computer Science
Bolzano, Italy
cigdem.gencel@unibz.it

Negin Asghari
Ericsson AB
Kista, Sweden
negin.asghari@ericsson.com

Dejan Baca
Ericsson AB
Karlskrona, Sweden
dejan.baca@ericsson.com

Stefanie Betz
Karlsruhe Institute of Technology
Karlsruhe, Germany
Stefanie.Betz@kit.edu

ABSTRACT

Background: Action research is a well-established research methodology. It is following a post-positivist research philosophy grounded in critical thinking. The methodology is driven by practical problems, emphasis participatory research, and develops practically useful solutions in an iterative manner.
Objective: Two objectives are to be achieved: (1) Understanding the state of the art with respect to action research usage in the software engineering literature, and (2) reflecting and providing recommendations of how to foster industry-academia collaboration through action research.
Method: Based on our experience with two action research studies in close collaboration with Ericsson lessons learned and guidelines are presented.
Results: In both cases presented action research led to multiple refinements in the interventions implemented. Furthermore, the close collaboration and co-production with the industry was essential to identify and describe the required refinements to provide an in-depth understanding. In comparison with previous studies, we required multiple iterations while previous software engineering studies reported mostly one iteration, or were not explicit regarding the number of iterations studied.
Conclusion: We conclude that action research is a powerful tool for industry-academia collaboration. The success of the method highly depends on the researchers and practitioners working in a team. Future studies need to improve the reporting with respect to describing the type of action research used, the iterations, the model of collaboration, and the rationales for changes in each iteration.

WISE'14, September 16, 2014, Vasteras, Sweden.

http://dx.doi.org/10.1145/2647648.2647656.

Categories and Subject Descriptors

D.2.0 [**Software Engineering**]: General

General Terms

Practice

Keywords

Action Research; Software Engineering; Industry-Academia Collaboration

1. INTRODUCTION

In the information system (IS) community a concern was raised that there is a lack of relevance of IS research from a practical point of view (cf. [19, 35]). Baskerville [4] points out that the issue raised in [19, 35] was the reason for the increasing popularity of action research in IS. A similar concern is frequently raised in secondary studies synthesizing primary studies. Ivarsson and Gorschek [16] defined a rubric to assess the scientific rigor and practical relevance of primary studies. Many studies have been using the rubric in their synthesis and found that only few studies achieved high scores on practical relevance (cf. [11, 15, 20, 25]).

Consequently, our proposition is that action research, as was the case in IS, is very promising in increasing the relevance of software engineering research by providing a model of co-production between software engineering researchers and practitioners. There exist no single definition of what an action research is, though different characteristics have been defined in the literature (see e.g. [6]). All sources share the view that action research aims at understanding (diagnosing stage [6]) a situation in a practical context and aims at improving it by changing the situation (therapeutic [6] or action stage [33]). The action stage has been refined, e.g. in action planning, action taking, evaluation, and specifying learning [33].

In software engineering only few studies followed an action research methodology. The systematic review by Santos and Travassos [9] summarized existing action research studies conducted in the software engineering context. Key findings were that genuine action research is rare, and that reporting is incomplete.

In order to facilitate the planning, execution, and reporting of action research in the software engineering context, we report on findings from the software engineering literature. Furthermore, we reflect on two cases in which we utilized action research to introduce different changes at Ericsson AB. The first change is the introduction of a method for software security risk analysis [3]. The second change is the introduction of an elicitation instrument to operationalize GQM+Strategies in the organization [27].

The remainder of the paper is structured as follows: Section 2 provides background information on the action research methodology. Section 3 presents action research conducted in the software engineering context. Section 4 describes the approaches and results obtained utilizing action research at Ericsson AB. Section 5 discusses the lessons learned from the application. Section 6 concludes the paper.

2. BACKGROUND

In action research the research problem studied is emerging from the context of investigation (e.g. an organization). Action research is characterized as:

- Intervention focused: An action (intervention) is introduced by the researcher with the goal of understanding the consequence of the action in the studied context (f. ex. [4]).
- Iterative: The action taken (e.g. a method or tool) is iteratively improved in multiple action research cycles. Different models have been proposed of how to structure the cycles (cf. [4, 23, 33]).
- Participatory: The researcher takes an active part in implementing the action and making observations in the studied context [4].

Other views on action research exist as well. For example, French and Bell [13] distinguish between different types of action research studies, namely diagnostic, participatory, empirical, and experimental.

- Diagnostic: The researcher only helps the organization in systematically identifying and describing the problem, and makes a solution proposal (action) without being further involved.
- Participatory: The researcher is involved in the further steps, as the participating organization shares experiences of applying the action.
- Empirical: In empirical action research members of the organization systematically collect the data capturing the effect of the action.
- Experimental: Multiple cycles of action taking and evaluation are conducted with one or more organizations to improve the action.

The definition by French and Bell [13] would allow to classify studies of only identifying challenges and proposing solutions based on literature as action research, which would not be in-line with the description of action research processes in [4]. The experimental version with participatory elements would best reflect the action research cycle described in the literature (e.g. [33]).

2.1 School of Thought

When classifying action research with respect to the schools of thought (see Table 1) action research is following the post-positivist, interpretivist, and critical thinking schools of thought. That is, one is not aiming to find a generally applicable truth (positivist), but rather a practically useful solution in a context (pragmatist/post-positivist). As Karl Popper [34] pointed out, there is no absolute truth, rather we can only empirically falsify and not verify. The interpretivist view states that a subjective truth is captured, by which people construct their reality. Whenever the action researcher makes an interpretation, it is often qualitative and hence subjective in nature. Though, objective measures independent from the interpretation could also be collected. Critical thinking is interruptive, which is reflected in the action taken. Which school of thought is followed is essential, as the school of thought determines which requirements are posed on the research in terms of validity.

2.2 The Action Research Cycle

The previously discussed traits of action research are reflected in the action research process. The very basic action research process described by [6] distinguishes two stages: namely diagnosing and therapeutic. In the diagnosing stage the problem is understood, which is then addressed in the therapeutic stage by taking an action. Further refinements have been made. A well known process is the action research cycle [33] consisting of five steps presented in Figure 1.

1. *Diagnosing*: This step is focused on understanding and describing the problem. The problem to be solved is described and agreed upon.
2. *Action planning and design:* Alternatives of how to solve the problem are identified and a choice is made of how to solve the problem.

Table 1: Schools of Thought

World views	Truth
Positivist	Objective (independent from participants)
Post-positivist	Objective (explicitly considering possible biases)
Constructivist/ Interpretivist	Subjective (constructed by participants)
Advocacy/ Participatory	Subjective (constructed by participants and the observer)
Pragmatist	Depends on what works at the time

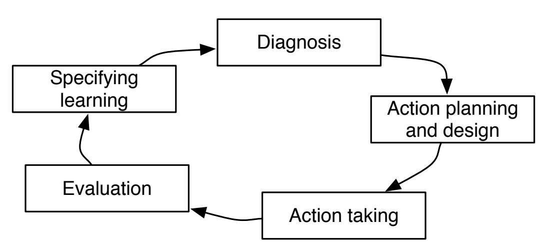

Figure 1: Action Research Cycle

3. *Action taking:* The action plan is implemented.
4. *Evaluation:* The effects of the action are captured. This can be done using different data collection methods, such as measurements, observation, questionnaires, interviews, or focus groups.
5. *Specifying learning:* The general learnings are specified based on the evaluation. These are used to decide how to proceed (e.g. whether further cycles are needed in order to solve the problem identified in diagnosis).

Further proposals of how to structure the action research process have been made. Checkland [8] adds the steps of establishing roles before starting the action planning and taking steps, which may represent a model of collaboration. Furthermore, an exit step is added after which a reflection over all iterations is taking place. The same step has been added by McKay [23] who makes the exist step conditional depending on whether the problem identified in the diagnosis stage has been fulfilled. Furthermore, McKay [23] adds that research literature should be studied to identify potential solutions. Action research processes (cf. [33, 8, 23]) reflect continuous learning and at each cycle we may falsify by identifying limitations in our actions during the evaluation and specifying learning steps and making adjustments accordingly.

2.3 Role of the Researcher

Research methods such as case studies and controlled experiments make a clear division of labor between the researcher and the researched [10]. For instance, in case study research the researcher designs the study, interviews participants, or acts as a passive observer to understand the real world context [30]. In controlled experiments [36] the division is even more pronounced when blinding with respect to the subjects under investigation. Action research emphasizes that the division of labor should be removed [10], i.e. the researcher and the researched work as a team. For example, when researchers introduce a new inspection technique in an organization then they actively take part in conducting the inspection process.

2.4 Research Validity

Three main threats to validity are to be highlighted that are associated with action research.

Action research is highly context dependent. The action is implemented in a specific social setting, which hinders the transferability of results (external validity). Though, the transferability within the setting studied (e.g. groups within an organization) may be high if the context is similar (internal validity).

Furthermore, action research is subjective as the results are highly dependent on the reflection of the action researcher. Multiple biases may occur:

- The researcher is part of the organization and does not provide an objective/external view of the situation after spending too much time in the organization.
- The researcher as an inventor of the action may interpret the results positively (selective bias) when reporting the results.

With every new action research cycle improvements may not be associated with the changes in the action, as other confounding factors come into play, such as learning effects or contextual changes.

In Section 5 we will further discuss the validity of action research in software engineering, elaborating on actions that could be taken to increase the validity of action research studies in that context.

3. RELATED WORK

Researchers in software engineering have investigated the extent of action research usage in the software engineering context.

Glass et al. [14] reviewed 369 papers in the time period 1995 to 1999 from six software engineering journals. One of their research questions was related to the use of research methods. The highest number of studies were using conceptual analysis, case study, and laboratory experiments. None of the studies used action research.

Santos and Travassos [9] conducted a systematic literature review to identify action research studies in the software engineering context. In total they identified 16 action research studies, distinguishing between inspired by and based on action research ideas, and genuine action research. Eight studies (cf. [32, 31, 5, 29, 17, 18, 21, 5]) could be classified as genuine action research (cf. [9]). All genuine action research studies were published in 1999 or later. We summarize the action research processes presented in these studies, followed by the conclusions of Santos and Travassos [9].

Staron et al. [32] conducted a case study on a framework for measurement program planning they developed. The action research process was integrated in the case study by iteratively developing the measurement systems used. The researcher became an active actor in the team driving measurement program planning. An example of iteratively developing a measurement system (in this case predictor system) is presented in Staron and Meding [31] who developed a prediction model for defect inflow. The iterations and versions of the prediction model are clearly articulated; in particular the initial model showed a clear need for redesign given that there was a large prediction error. Staron and Meding point out the benefit of being part of the organization when conducting the research, stating that the ability of directly observing the needs led to better models and thus results.

Fernandez-Medina and Piattini [12] applied action research to redesign a database with the researchers taking an active part in the redesign to improve security. Multiple benefits from a solution perspective have been identified after continuously iterating and reflecting. Specifically, security and privacy awareness have been increased and system security has improved. However, from the report the number of iterations and learnings for each iteration were not explicit.

Polo et al. [29] present an participatory action research study building a software maintenance process for a software organization. Paolo et al. classify their action research as participatory (cf. [13]). The diagnosis step to understand the problem was done by having meetings with representatives of the organization, and an agreement of the problem to be solved was reached. Thereafter, potential solutions were reviewed from standards and experience reports to make a proposal for an action. The reference group of the organization reviewed the proposals and intended to choose a particular standard. The researcher advised for a different one providing a rational; the recommendation was followed by the organization. In the subsequent steps the maintenance process model was selected and the process was modeled in dialog with the researcher. Beyond the reference group, further practitioners (e.g. developers) were involved to provide their input to the process. The process was then applied and evaluated, refined, and extended. Polo et al. conclude that the action research process made the practitioners more active participants in the research, and led to a tangible improvement in the organization. What was missing was a more systematic approach to capture the data as information got lost in the research process.

Kauppinen et al. [17] utilized an action research process inspired by the IDEAL [22] model. The process was iterative and consisted of the steps: assessing current situation (diagnosing), developing, piloting and implementing new processes (action), and monitoring new processes (evaluation and learning). Kauppinen et al. highlighted that activities may be interleaved, such as piloting and implementation. The main role of the researchers was to act as facilitators, and the data collection was qualitative (interviews). The action research was conducted with three case organizations, strengthening external validity.

The IDEAL model utilized by Kauppinen et al. has been evaluated through an action research by Kautz et al. [18] to conduct process improvement. The action research process was divided into the initiating, diagnosing, establishing, acting, and leveraging phase. During the initiating phase the researchers were contracted as consultants and expectations and goals were clarified. Thereafter, in the diagnosing phase the current process was assessed and diagnosed using questionnaires, interviews, and document reviews. The results were then presented to the organization. Based on the problem understanding the establishing phase resulted in an improvement and action plan. The plan was further refined in meetings between researchers and industry practitioners, and resulted in well defined routines that could be piloted. After finalizing the overall project, the leveraging phase was concerned with reviewing the whole project considering the actions as being part of the whole. One of the improvements (structured meetings) was successfully followed and considered beneficial. Overall, it was concluded that having a structured model to conduct the action research was important. Similar to Kauppinen et al. interleaving activities were identified. Another important success factor was the openness of the participating organization.

Maiden and Robertson [21] present an action research of the RESCUE process for use case and scenario development during requirements elicitation. The process consisted of four stages (use case development based on context models, discovery of new requirements based on use cases, generate use case scenarios, and scenario walk-through with tool support). Each stage is presented in the paper from a retrospective view.

Bengtsson and Bosch [5] presented an action research where the researchers actively participated in the design process of the architecture. The goal was to learn from applying the architecture design process suggested by the researchers. A clear problem (poor maintainability and challenges in certification) has been identified, and detailed context in relation to the system studied has been provided. Aggregated lessons learned were presented across all design iterations conducted.

Santos and Travassos [9], in their synthesis, raise completeness of reporting as the main concern in the studies. They point out that papers do not report on the number of action research cycles and their construction, research context, and data collection techniques to the extent desired.

4. ACTION RESEARCH CASES

It is widely acknowledged that software engineering is a complex social activity and the success or failure of processes, methods, and tools highly depends on the application context [28]. Two cases are presented where we used action research at Ericsson AB. Both cases are related to social activities where a team of people uses the methods introduced. The first method is a security risk analysis method (countermeasure graphs) [3], and the second an elicitation instrument to gather goals, information needs, and strategies from stakeholders to operationalize GQM+Strategies [27]. In this section we describe the action research processes and models of collaboration, while further details on the research context and actions taken can be found in [3, 27].

4.1 Case 1: Development of Countermeasure Graphs

Action research process: The first action research contained six cycles (see Figure 2). Each cycle consisted of the steps action planning, action taking, and evaluation. Each cycle was later labeled during the learning phase to reflect what additions have been made.

In the first cycle (Cycle 1) the diagnosis took place. Peltier as a risk analysis method was used on two products initially (A and B), which were closely related and later split into two separate ones (see Cycle 3). Practitioners have conducted the risk analysis using Peltier as a method. Looking at the risk analysis results, it was evident that the method identified security threats, but not those with the highest risk levels (in terms of damage and likelihood of occurrence). To get further confidence that this observation is prevalent in other parts of the organization, the Peltier analysis of another product (D) was evaluated. A similar pattern was

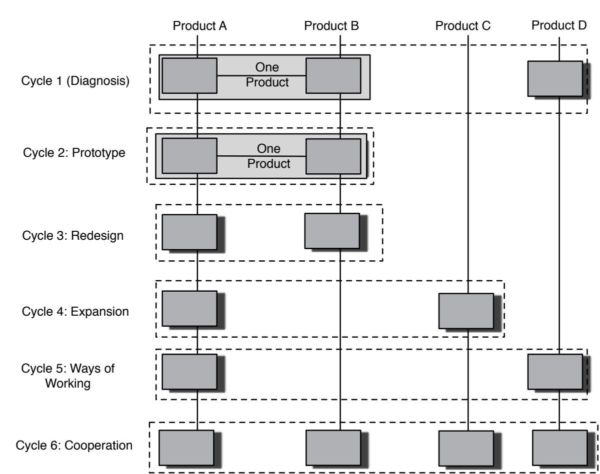

Figure 2: Action Research Cycle of Case 1 (from [3])

observed. As a means to understand the reasons a workshop was held with the participants to figure out why they believe the Peltier method was ineffective. The main reason was that the focus is on threats and no holistic picture of the risk situation could be obtained using Peltier.

In the second cycle (Cycle 2) the prototype method (countermeasure graphs) was designed and tested. The initial idea of the method has been presented in [2]. Countermeasure graphs provide a holistic risk model considering attackers goals, agents (actors or malicious users), attacks they can do, and countermeasures to mitigate attacks. The first risk analysis using countermeasure graphs was conducted with products A and B, which provided the motivation for taking an action. During the use of countermeasure graphs two needs for changes became evident; the need for tool support and a minor change in the structure of countermeasure graphs to improve maintainability.

During the third cycle (Cycle 3) we restructured countermeasure graphs, the restructuring being minor, but important. The focus was still on products A and B. Furthermore, a tool has been developed where the practitioners could enter attackers goals, agents, attacks and countermeasures, and could connect them. Prioritization has also been implemented, which was an integral part of the initial countermeasure method. We observed that it was easy to maintain the information entered, and that the method appeared stable from a conceptual view.

Given the conceptual stability of the method, another product (Product C) was added in Cycle 4. Furthermore, we shifted the focus on further improving the tool. That is, In this cycle the main emphasis was on interface design as countermeasure graphs documented in the tool grew, i.e. mechanisms for grouping and filtering needed to be implemented for the coming cycles.

Cycles 5 and 6 did not lead to further changes in the method itself. Only minor changes were made (e.g. the need for clarifying the terminology was needed). From an evaluation perspective another product was added that, until this stage, only used Peltier as a risk analysis method. The tool was further enhanced based on suggestions, e.g. by adding network support and screen sharing to support distributed development, and a result repository to apply configuration management to the risk models.

Data collection: Throughout the action research process the main means of data collection were field notes, where we collected feedback from the practitioners as well as any observations we made. Furthermore, quantitative and tangible data was obtained through the risk analyses conducted with the teams.

Model of collaboration: Both authors already working in the company, and were employed at the time the study was conducted. Earlier, the authors have already been involved in conducting security research with the company (cf. [1]). During Cycles 1 to 4 the first researcher's responsibility was to facilitate the risk analysis, while in Cycles 5 and 6 this responsibility was with practitioners, and the researcher acted as an active participant in the risk analysis. This allowed to concentrate on contributing content to the risk analyses and recording observations. The second researcher was providing design input and reflections based on the outcomes noted down by the first researcher.

Outcome: The initial version was conceptually relatively close to the final version. Only small structural changes were needed. However, for the action to become really usable the following iterations were very important to develop the tool and clarify the terms. Countermeasure graphs also appeared to be robust with respect to the context as they were applicable on multiple products, even though they were developed in the same company.

4.2 Case 2: Elicitation Instrument for GQM+ Strategies

Action research process: In the second action research study we required three iterations to arrive at a stable release version of our method (see Figure 3). After having arrived at a stable version, a case study has been conducted to test the developed solution on a larger scale.

The diagnosis started with understanding the current processes and terminology used in the parts of the organization for which measurement program planning should take place. The need of how to capture stakeholder information needs for software measurement emerged from ongoing work of defining overall processes for measurement program planning (cf. [7]). Elicitation was the first step in that process and there was a need for a systematic approach. At first, the organization utilized Goal Question Metric templates as suggested in literature, though the terms were not well understood, and no overall structure was visible in the templates aiding the practitioners' understanding. The challenge of practitioners filling in the templates was also observed in the literature.

In the first iteration we defined a notation to model the relationships between goals, strategies, measurement context and metrics. Furthermore, refinements of goals and strategies could be modeled together with the practitioners. Thus, the model became a means for gathering information from the practitioners. The very first interview was conducted face to face. The outcome was that the notation failed as the practitioner's focus was on understanding the notation and negotiating/remembering the concepts used rather than providing information of his information needs. Overall, the interviewee stopped talking about what we were interested in from a measurement point of view (even though the information we received was valuable to understand the usefulness of our solution). A second interview was held using the notation to gain further confidence, the findings were the same. Both interviews were conducted in the same improvement program (program A).

During the second iteration a semi-structured interview guide was developed. We also did not utilize the terminology used in the notation anymore,. For example, rather than asking for the goals directly, we asked when the practitioners believe their improvement program is successful, and how the success could be achieved. The details and ratio-

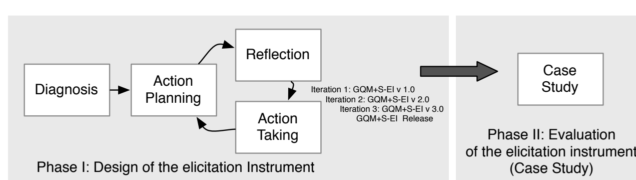

Figure 3: Action Research Cycle of Case 2 (from [27])

nales for formulating the questions in the interview guide are presented in [27]. Only one interview was conducted for improvement program A. We found that the interviewee talked freely and fluently, did not require clarifications, and focused on actual content. This was a major improvement in comparison to the previous iteration, and we found to be on the right track. We captured the information gathered on the white-board so that it is visible for the interviewee, though this was challenging due to the speed in which information had to be captured.

Consequently, in the third iteration we utilized a mind-mapping tool on the computer and projected the documentation on a screen. Three interviews have been conducted in this iteration, one in program A, and for two other programs (B and C) one interview was added, respectively. Some interviews were distributed given that the company is operating with distributed development sites; thus the mind-mapping tool made communication easier. In the subsequent interviews further questions have been added based on relevant topics discussed by the practitioners that could not yet be elicited with the existing questions. After the interviews the interview guide was ready for release and its ability to accurately and completely capture information was evaluated in a case study.

The case study as such could be considered as a fourth iteration. Eight interviews in one improvement program were conducted, and the completeness and accuracy of eliciting the required information was assessed. Overall, we found that the elicitation instrument was accurate and led to complete GQM+Strategy models.

Data collection: For data collection we used documentation and interviews in the diagnosis stage. In each iteration we gathered all information provided by the practitioners, and hence could see how well we elicited the information, and whether the questions have been answered. One actor had the designated role of being the scribe to document everything on the white-board or the mind-maps. We also took field-notes during the interview sessions to keep track of noteworthy events related to the action introduced (here elicitation instrument). Right after each interview session we discussed what went well and what needs to be improved, and reflected on the reasons. Thereby, we still remembered all important events. Furthermore, we recorded these discussion sessions for further analysis and transcribed them, avoiding information loss.

Model of collaboration: Three researchers were involved in all activities of the action research, one researcher being employed during the action research study, and one having an office on-site while conducting her master thesis research. The third researcher was employed at the university. All researchers had access cards and could always visit the company. In addition, two practitioners responsible for measurement program planning and supporting the improvement programs were involved. They participated in the interviews, in the beginning as observers and later as facilitators. Furthermore, a steering group was established where we reported our progress and discussed intermediate results. This further helped to develop ideas and get early feedback before implementing them in the iterations.

Outcome: The initial version of the solution developed by the researchers (notation) was not successful. A complete redesign of the solution was required. The new version worked well, as was demonstrated in the final case study. Multiple improvement programs have been used to develop the solution to make sure relevant questions asked are not dependent on one particular program.

4.3 Comparison

When comparing the two action research studies, both would be classified as experimental by French and Bell. All other important characteristics (intervention focused, iterative, and participatory) were also implemented.

Both action research studies also utilized similar approaches for collaboration between practitioners and researchers, who were working as a team. With additional iterations more responsibility was shifted to the practitioners, who in the beginning were mainly observers. The researchers were mainly observers in the later iterations.

Furthermore, both studies utilized multiple units of analysis (products and improvement programs) from different parts of the organization. Given the size of the organization, the context between them differed. Switching context allows to test the sensitivity of the action with respect to the context, and allows to reflect on the sensitivity as well.

Two differences are highlighted. First, the initial solution proposed by the researchers was conceptually relatively close to the final version (countermeasure graphs) in Case 1, while a complete redesign was needed in Case 2 (elicitation instrument). A similar experience was reported by Staron and Meding [31] where they had to redesign their initial model to forecast defect inflow.

Second, a steering group was established in Case 2 (elicitation instrument) which was found useful to inform about progress and keep commitment from several parts of the organization.

5. DISCUSSION

Contribution: Based on the experiences made we believe that action research may lead to evolutionary or revolutionary results, which cannot be known beforehand. A complete failure in a research cycle (such as in Case 1) may lead to a complete rethinking and design of a conceptually novel idea. We also found that in each case multiple cycles were needed to develop and make new ideas usable. Thus, it was surprising that most action research studies only report a single cycle. Stability of the solution across contexts is a good indicator for the practical applicability of the solution; this is when no new cycle is required. Besides the practical contributions, the software engineering community also recognizes action research as a sound methodology. The studies surveyed in the related work were published in international journals, and multiple papers were published at the International Conference on Software Engineering (ICSE).

Industry-Academia Collaboration: *Process improvement and action research:* Action research is similar to software process improvement in many ways (e.g. when compared to the Plan-Do-Check-Act cycle), while adding scientific guidelines and emphasizing collaboration between researchers and practitioners. Given that there are similarities, and most companies are familiar with conducting improvement programs, we believe that action research may ease the collaboration between academia and research. The presence of the researcher and following scientific guidelines, and also bringing scientific knowledge from literature and research propositions into the process is an advantage over regular improvement activities.

Collaboration strategy: In action research practitioners and researchers have to work in a team. This requires long-term commitment and intensive collaboration. When conducting the research we found that in the beginning the researchers should be the facilitators of implementing the action, and the practitioners should act as the observers. With additional iterations the roles should be switched, the practitioners becoming the main actors and the researchers the observers. Furthermore, it was very important to be present and have frequent access to the company; depending on the research conducted the researchers should have an access card, office space, and computer to access the information they need. Being physically present regularly also helps to become part of the team and pick up information "on-the-fly" during informal discussions.

Funding: To realize such a collaboration funding is necessary. Funding agencies in Sweden (e.g. the KK stiftelssen) require in-kind funding from companies, where they invest their time to conduct research. The funding application should describe the means of co-production with the companies. Action research, given that it is a truly collaborative and iterative process, may be very useful to plan and document the co-production. Another good option is direct industry funding. In Case 1 the researchers were employed as industrial PhD students, one PhD student being fully funded by the company. Additional researchers could be brought in through part-time temporary position to quickly and efficiently assemble a research team to solve a problem.

Research Design and Reporting: *Data collection:* It was recognized that data collection may be incomplete as no systematic approach for data collection was followed (cf. [29]). We recommend to utilize multiple researchers in the process to note down as much information as possible, and after each interactive session/action cycle the observations should be discussed immediately after that, given that the memories of relevant events are fresh. The discussions and reflections should be recorded and transcribed, as this provides all the design rationales and changes in each iteration and makes them traceable (see Case 2 in Section 4). The work artifacts produced during the action cycles also provided valuable insights in the studied cases. In addition, well known methods shall be used (interviews, questionnaires, observations, and study of documentation).

Validity concerns: In Section 2.4 we highlighted three validity concerns of action research. The first one is that the research is highly context dependent. This threat can never be mitigated, though it can be reduced. We added further units of analyses (products and improvement programs) to always evaluate the sensitivity of our solutions to different contexts. The benefit is that the context switch may reveal limitations, but at the same time may provide explanations of why something works in one context, but not in another. Hence, an awareness of the context is a key element of study validity in action research. Lessons learned may then be transferred between similar contexts. The second threat is the bias of the researcher. This can be best achieved by involving multiple researchers, and having a steering group to present and discuss the reflections. The collaboration with different practitioner groups also aids in reducing the bias. The third threat is the time factor (learning and changes in the context). This threat can also be reduced by involving different people over time, and being aware of major changes. However, given that action research is not conducted in a controlled environment this threat will stay relevant. Though, it is important to note that action research is not measured by the same set of validity threats as research methods inspired by positivist science (e.g. experiments). Gencel and Petersen [26] discuss how to identify and consider the threats of validity depending on different schools of thought and research methods.

Santos and Travassos [9] highlighted incompleteness of reporting in their results. With respect to reporting it is important to (1) describe the context in detail (cf. [28]), (2) describe every cycle focusing on action planning, taking, and evaluation/learning, (3) provide a detailed reflection on validity with relevant threats for the action research methodology (cf. [26]). As highlighted by [24] action research is often written as a case study, which should be avoided. The learning in each iteration, and why changes are made, should be transparent as this is one of the strengths of action research: Providing deep insights and detailed explanations why an action does not work, or why one version of an action is less flawed than another.

6. CONCLUSION

In this paper we presented action research as a method, reviewed literature on action research conducted in software engineering, and reported our experience from two action research studies. Our experience, and experience from literature indicates that action research is a powerful tool for industry-academia collaboration and for transferring research results into practice. Multiple guidelines and recommendations have been identified to make the collaboration through action research a success. In future work, action research studies reporting multiple iterations and action cycles are of particular interest.

7. REFERENCES

[1] D. Baca, B. Carlsson, K. Petersen, and L. Lundberg. Improving software security with static automated code analysis in an industry setting. *Softw., Pract. Exper.*, 43(3):259–279, 2013.

[2] D. Baca and K. Petersen. Prioritizing countermeasures through the countermeasure method for software security (cm-sec). In *Product-Focused Software Process Improvement (PROFES 2010)*, pages 176–190, 2010.

[3] D. Baca and K. Petersen. Countermeasure graphs for software security risk assessment: An action research. *Journal of Systems and Software*, 86(9):2411–2428, 2013.

[4] R. L. Baskerville. Investigating information systems with action research. *Communications of the AIS*, 2(3es):4, 1999.

[5] P. Bengtsson and J. Bosch. Haemo dialysis software architecture design experiences. In *Int. Conf. on Software Engineering (ICSE 99)*, pages 516–525, 1999.

[6] F. H. Blum. Action research–a scientific approach? *Philosophy of science*, pages 1–7, 1955.

[7] Çigdem Gencel, K. Petersen, A. A. Mughal, and M. I. Iqbal. A decision support framework for metrics selection in goal-based measurement programs: Gqm-dsfms. *Journal of Systems and Software*, 86(12):3091–3108, 2013.

[8] P. Checkland. From framework through experience to learning: the essential nature of action research.

Information systems research: Contemporary approaches and emergent traditions, pages 397–403, 1991.

[9] P. S. M. dos Santos and G. H. Travassos. Action research use in software engineering: An initial survey. In *Int. Symp. on Empirical Software Engineering and Measurement (ESEM 2009)*, pages 414–417, 2009.

[10] O. Eikeland. The validity of action research-validity in action research. 2006.

[11] F. Elberzhager, A. Rosbach, J. Münch, and R. Eschbach. Reducing test effort: A systematic mapping study on existing approaches. *Information and Software Technology*, 54(10):1092–1106, 2012.

[12] E. Fernández-Medina and M. Piattini. Designing secure databases. *Information and Software Technology*, 47(7):463–477, 2005.

[13] W. L. French and C. H. Bell. *Organization development: Behavioral science interventions for organization improvement.* Prentice-Hall Englewood Cliffs, NJ, 1973.

[14] R. L. Glass, I. Vessey, and V. Ramesh. Research in software engineering: an analysis of the literature. *Information and Software technology*, 44(8):491–506, 2002.

[15] M. Ivarsson and T. Gorschek. Technology transfer decision support in requirements engineering research: a systematic review of rej. *Requir. Eng.*, 14(3):155–175, 2009.

[16] M. Ivarsson and T. Gorschek. A method for evaluating rigor and industrial relevance of technology evaluations. *Empirical Software Engineering*, 16(3):365–395, 2011.

[17] M. Kauppinen, M. Vartiainen, J. Kontio, S. Kujala, and R. Sulonen. Implementing requirements engineering processes throughout organizations: success factors and challenges. *Information and Software Technology*, 46(14):937–953, 2004.

[18] K. Kautz, H. W. Hansen, and K. Thaysen. Applying and adjusting a software process improvement model in practice: the use of the ideal model in a small software enterprise. In *Int. Conf. on Software engineering (ICSE 2000)*, pages 626–633. ACM, 2000.

[19] P. G. Keen. Relevance and rigor in information systems research: improving quality, confidence, cohesion and impact. *Information systems research: Contemporary approaches and emergent traditions*, pages 27–49, 1991.

[20] S. Mahdavi-Hezavehi, M. Galster, and P. Avgeriou. Variability in quality attributes of service-based software systems: A systematic literature review. *Information and Software Technology*, 55(2):320–343, 2013.

[21] N. Maiden and S. Robertson. Developing use cases and scenarios in the requirements process. In *Int. Conf. on Software engineering (ICSE 2005)*, pages 561–570. ACM, 2005.

[22] B. McFeeley. Ideal: A user's guide for software process improvement. Technical report, DTIC Document, 1996.

[23] J. McKay and P. Marshall. The dual imperatives of action research. *Information Technology & People*, 14(1):46–59, 2001.

[24] J. Meyer. Evaluating action research. *Age and ageing*, 29(suppl 2):8–10, 2000.

[25] H. Munir, M. Moayyed, and K. Petersen. Considering rigor and relevance when evaluating test driven development: A systematic review. *Information and Software Technology*, 56(4):375–394, 2014.

[26] K. Petersen and Çigdem Gencel. Worldviews, research methods, and their relationship to validity in empirical software engineering research. In *IWSM/Mensura (2013)*, pages 81–89, 2013.

[27] K. Petersen, C. Gencel, N. Asghari, and S. Betz. An elicitation instrument for operationalising gqm+ strategies (gqm+ s-ei). *Empirical Software Engineering*, pages 1–38, 2014.

[28] K. Petersen and C. Wohlin. Context in industrial software engineering research. In *Int. Symp. on Empirical Software Engineering and Measurement*, pages 401–404. IEEE Computer Society, 2009.

[29] M. Polo, M. Piattini, and F. Ruiz. Using a qualitative research method for building a software maintenance methodology. *Software: Practice and Experience*, 32(13):1239–1260, 2002.

[30] P. Runeson and M. Höst. Guidelines for conducting and reporting case study research in software engineering. *Empirical software engineering*, 14(2):131–164, 2009.

[31] M. Staron and W. Meding. Predicting weekly defect inflow in large software projects based on project planning and test status. *Information and Software Technology*, 50(7):782–796, 2008.

[32] M. Staron, W. Meding, and C. Nilsson. A framework for developing measurement systems and its industrial evaluation. *Information and Software Technology*, 51(4):721–737, 2009.

[33] G. I. Susman and R. D. Evered. An assessment of the scientific merits of action research. *Administrative science quarterly*, pages 582–603, 1978.

[34] P. Tichỳ. On popper's definitions of verisimilitude. *British Journal for the Philosophy of Science*, pages 155–160, 1974.

[35] R. Westfall. An is research relevancy manifesto. *Communications of the AIS*, 2(2es):5, 1999.

[36] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén. *Experimentation in software engineering.* Springer, 2012.