Structuring Software Engineering Case Studies to Cover Multiple Perspectives

Emil Börjesson and Robert Feldt
Software Engineering and Technology
Chalmers University
Gothenburg, Sweden
Emil.Borjesson@Chalmers.se

Abstract—Case studies are used in software engineering (SE) research for detailed study of phenomena in their real-world context. There are guidelines listing important factors to consider when designing case studies, but there is a lack of advice on how to structure the collected information and ensure its breadth. Without considering multiple perspectives, such as business and organization, there is a risk that too few perspectives are covered.

The objective of this paper is to develop a framework to give structure and ensure breadth of a SE case study.

For an analysis of the verification and validation practices of a Swedish software company we developed an analytical framework based on two dimensions. The matrix spanned by the dimensions (perspective and time) helped structure data collection and connect different findings. A six-step process was defined to adapt and execute the framework at the company and we exemplify its use and describe its perceived advantages and disadvantages.

The framework simplified the analysis and gave a broader understanding of the studied practices but there was a trade-off with the depth of the results, making the framework more suitable for explorative, open-ended studies.

Index Terms—Case Study, Multi-perspective, Framework, Empirical

Version 1, May 5, 2011.

I. INTRODUCTION

The case study is an observational research method used in many different fields of research due to its flexibility and its ability to investigate a phenomenon in its context [1]. Case studies are also applicable when there is no clear distinction between the phenomena and its context. This is particularly true in empirical software engineering research where there are many factors that impact the phenomenon, such as the type and organization of the company, the development processes used etc. To understand a contemporary phenomenon we also need to understand its history and how the different factors have evolved over time.

Existing advice for empirical research in software engineering focuses on experiments and systematic reviews while guidelines for case studies was only recently published by Runeson and Höst [1]. There, a high quality case study is defined as a study that produces valid information of academic or industrial significance, either generic or practitioner oriented [1]. A key criteria for achieving this is that the data is collected in a planned and consistent manner and that conclusions are based on a clear chain of evidence. This can be a challenge in practice since case studies are typically flexible research designs with multiple sources of evidence; it is not clear-cut how to find the right balance between a flexible research design that allows multiple factors and causes to be taken into account while providing enough structure and support for planning and analysis.

Software engineering is different from computer science in that it takes more perspectives than only the technical into account. For example the personality or motivation of the engineers [2], [3] can affect the quality of their work and organizational and career considerations can affect important activities such as effort estimation [4], [5]. In general, the characteristics not only of the people but of the organization, business and processes used in software development are all important.

In recent empirical research in industry we were faced with designing a case study to describe and understand the verification and validation practices at a Swedish company developing safety-critical software systems. From initial talks at the company we understood that the practices could not be studied in isolation; they were heavily tied to the whole context of the company as well as how they had evolved over time. To structure our data collection and understanding we based our case study on an analytical matrix combining these two main aspects. On one dimension we wanted to cover at least the main perspectives of the BAPO framework with its four aspects [6]: Business, Architecture (technical aspects), Processes and Organization. The other dimension was time detailed in three steps as Past, Current and Future. Together this created a matrix of 12 different sub areas to be considered during the case study. We also defined a general six-step process to design, collect and analyze the findings of the study and adapted it to the company. This paper describes this analytical framework, called the BAPO/PCF framework for software engineering case studies, covering both the matrix and the process to adapt and use it. We exemplify and evaluate the framework based on our application of it at the studied company. In particular this paper addresses the following research questions:

1) What are the perceived advantages of a multi-perspective approach to conducting a case study?
2) What are the perceived disadvantages of a multi-perspective approach to conducting a case study?
3) How does the multi-perspective structure of the design affect research validity?

The definition of a research design is in this paper a methodology used to conduct a research project, such as a case study, an experiment, a longitudinal research project, etc. A research method is defined as a way of eliciting information within a research project; hence a research design can include the use of one or several research methods.

To help the reader differentiate between what was done in the state of practice analysis and what are general case study practices in the paper, we refer to the latter as the ‘case study’ and the former as either ‘company study’ or ‘study at the company’.

The paper is divided as follows, section II will present the research context in which the research design was developed and executed. Section III will present how the framework was developed and executed. In Section IV the advantages and disadvantages of the design will be discussed, and finally some conclusions will be presented in Section V.

II. RESEARCH CONTEXT AND THE STUDIED COMPANY

The project, for which the research design was developed, was a state-of-practice analysis of a small company, less than 50 employees, which develop safety critical software applications. The state-of-practice analysis constituted the first part of a larger project, with the goal of improving the company’s verification and validation practices. The analysis was conducted in order to understand the company’s needs, regarding the company’s processes, organization, etc, to narrow the scope of the process improvement effort.

The company conducts software development in bespoke projects to single customers, but the end applications are developed from a set of core products. These products are maintained according to market demand, making the company’s business strategy a mix of bespoke and market driven engineering. Because of the nature of the software the company is developing, the company’s business is governed by different quality standards and frameworks. These standards affect the development, but primarily the company’s verification and validation practices. Standards and frameworks are also imposed on the company by the company’s customers, who have different needs in their own domains. These demands require the company to be flexible in their development, which has resulted in the adoption of iterative as well as an agile software development based on Scrum [7]. The architectural granularity of the company’s systems is on a sub-system level, hence coarse grained, but efforts are made to increase the granularity by refactoring the systems with reusable components based on services, so called service oriented architectures (SOA). The decision to change the system architectures was taken internally at the company but driven by external business factors related to a large European project that will change much of the company’s market domain. The migration from sub-systems to SOA has required the company to acquire both new knowledge as well as new practices. Organizational changes have also been made, including changing the responsibilities of different roles and how communication is handled internally at the company. The company study was conducted over approximately 6 months, during which 7 structured interviews, 2 surveys, 1 structured observation, and a considerable amount of hours were spent on document analysis and watercooler discussion.

III. THE BAPO/PCF FRAMEWORK

The term research process is in this paper defined as the six-step methodology that was used to develop the case study design used at the company as well as the execution of the design. Figure 1 presents the process steps and how they incrementally relate to one another.

Get Knowledge about the Domain. Very little was known about the company when the project started and therefore the first step of the process was to gain a deeper understanding about the company and the company context. This information was acquired through a combination of interviews with people at the company and through literature review of internal documentation as well as documentation about the company’s domain. The context in which the company operates was important to understand the goals and needs of the company, in order to align the improvement effort with these goals, such as process changes or introduction of new tools, etc. The initial interviews also served as a way of finding new sources by using snowballing, sources that could be used later during the company study. Snowballing is conducted by asking questions at the end of interviews and looking at document references to locate further artefacts to study [8]. The majority of the domain documentation that was studied was market quality standards and frameworks. Development and quality standards were important to study because the results of the company study would be used as a base for a larger project that focused on the company’s verification and validation (V&V) practices, hence changes done to the V&V practices would, other than to comply with company needs, also have to comply with such standards. Domain specific documentation is often hard for someone from outside the domain to understand, so therefore all interpretations and question marks related to the analyzed domain documentation had to be verified with people at the company.

The collected data from step 1 showed that a broader view would be required to capture the state-of-practice at the company in its context. To meet this requirement and structure the research a multi-perspective framework was developed based around an analytical matrix with two dimensions. The first dimension was chosen to represent the perspectives of the company that would be analyzed, defined as company Business, Software Architectures, Processes and Organization (BAPO). These perspectives were chosen because they are used within academic and industrial frameworks such as the Family evaluation framework, which is used to evaluate Software Product Lines [6]. The second dimension of the matrix was chosen to represent time, defined as the Past, Current and
Future (PCF), since a chronological dimension of the elicited information was perceived to enable deeper and more precise conclusions to be drawn from the research findings. Hence the matrix starts with past Business, moving on to current business and so on, until finally reaching future organization. The BAPO/PCF framework was also chosen because it was perceived to be a good fit in this particular study, which might not be the case in the context of another company.

**Develop focus Questions/Areas.** The second step of the process was to use the information that was gathered in the first step, combined with the BAPO/PCF framework, to narrow down the focus of the company study by developing concrete research questions. The research questions were split up among the 12 sub areas/cells of the analytical matrix, which gave an initial understanding of where within the company to elicit information to answer the research questions. For instance it gave a coarse-grained view of which roles within the company that should be interviewed and what documentation to analyze. Information about existing company roles and documentation was provided by the first step of the process. The research questions were split among the 12 matrix cells based on the researchers own opinion where they fit best, with the company context taken into consideration. For instance questions regarding the company’s agile development processes were mostly constricted to the process row of the matrix, only sorted in time, whilst questions about the organization were split up between business, processes and organization in this case. The reason for the spread of organizational questions came from the need to understand how the organization had been affected by the other company aspects, for instance how the introduction of agile processes had affected the organization, or what affect business changes had resulted in, how roles had changed, etc. Some questions could not be confined to just one cell, such questions were mostly of higher level, but they helped to find logical connections between cells within the matrix. The options in these cases were to add the question in several of the cells, or to redefine them to make each of the questions unique and only fit in a specific cell.

**Choice of Detailed Research Methods.** The third step of the process was to choose what research methods that would be used in the company study. Primarily qualitative methods were chosen, such as interviews, visual inspection, structured observation and water-cooler discussions. Water-cooler discussions are conducted during coffee and lunch-breaks and make use of the fact that most information sharing within development companies are conducted during breaks [9]. This allows the researcher to get qualitative data that was current and truthful, but since the information often included the persons own opinions the information had to be verified to remove bias. Surveys were also used to provide quantitative results that would give depth to the research results. The surveys were deductive, hence based on previously elicited information, to verify the collected results and therefore conducted later in the company study. Different research methods are more appropriate to use depending on what sources are available, and therefore it is good to have a set of methods to choose from [1]. Creating a plan of what methods to use in what cells is suitable, but it should be considered a guideline that can be changed depending on the conditions of the study rather than a strict plan.

**Data collection and Data analysis and Alignment.** The first three steps of the process constitute the development of the design, whilst the following three steps of the process discuss how to execute the design. The fourth and fifth steps are defined as data collection and data analysis and alignment, and were executed incrementally, meaning that collected data was continuously analysed and aligned with previously collected information within the BAPO/PCF matrix, before new information was elicited and analyzed. Such analysis and alignment became essential since the data collection governed how the research would proceed, meaning that the collected information was used to evaluate if a certain research question had been answered or not. If the question had been answered
the research could move on to a new cell, otherwise further time was spent to find more information.

Case studies require flexibility to be able to respond to different events that may occur during the course of the study, such as for instance sudden research opportunities. For instance during the company study an opportunity arose to do observation of the company’s system testing practices that would not arise again during the course of the planned study period. Because events, such as opportunities, have to be considered the research has to be flexible and can not follow a strict research plan. The BAPO/PCF framework does however support such flexibility, since the research questions connected to the BAPO/PCF matrix cells can be answered out of order, which allows the researcher to jump between cells and work with the question that is the most urgent at the moment. The progress of the research can also be measured through what research questions within the matrix have been answered. Each cell that gets filled with information, and its research question answered, constitutes a measurable part of the study progress, and once all questions in all cells have been answered the study is complete. The time spent with a certain matrix cell can either be planned according to a fixed time budget, or in a more ad hoc manner where the collected results drive the time spent working with a cell as well as the order which the matrix cells are traversed. There are academic papers that state that research using the BAPO perspectives should start with the business aspect because this has the largest impact on the company, followed by software architectures that has the second largest impact and so on until the organizational aspect has finally been analyzed [6]. No evidence could be found to support this claim during the company study, which used an approach where intermediate results decided what cell to jump to next in the study. The study at the company did not follow the matrix row by row, but instead company processes were first investigated, followed by software architectures, followed by company business and finally company organization. The BAPO/PCF matrix traversal was not chosen at random, it was chosen based on information that became available in the initial part of the forth process step and proved throughout the study to be valuable asset to keep the research focused. Even though several deviations were made from the research plan, because opportunities arose, it proved to be valuable to have a plan to fall back on, which is also supported by Runeson and Høst that state that a plan is crucial for a case study to succeed [1].

To strengthen the analysis of the collected information, in the company study, cause-effect chains (CEC) were combined with the BAPO/PCF matrix. These event chains link the research results together to provide a broader contextual view and also help to find the causes behind a result. Two different types of CEC’s were used during the company study, where the first type focuses on one path through the matrix, the simplest being from a certain perspective’s past to the same perspective’s current or from the current to the future, hence a chronological view of to the result. The first type of a CEC can also span between different company perspectives

to provide a cross-perspective view, but always form a single path through the BAPO/PCF matrix that shows how an event in one perspective in time has affected another perspective in time and so on. The second CEC type differs from the first type in the sense that it starts in one matrix cell and spans out to several cells both in time and between perspectives, which shows the broader impact an event has had within the company. In the company study the CEC’s were developed post-elicitation of the research information by using inspection to find the internal connections between the results, which made it possible to draw clearer conclusions, develop predictions about the future of each company perspective, and also raise the result validity. Predictions for the future perspective of the matrix were developed using induction, with qualitative research data as input, based on the concept that if a specific action A in context B has outcome C in the past, which is repeated in the current in the same context B with result C, it would be likely that the same pattern would reoccur also in the future. A visual representation of the CEC’s was also developed that uses the research results, described in bullet-point lists in the matrix cells, which are connected by drawing arrows from the origin matrix cell to the cell(s) that have been impacted by the event in the origin cell. It should be noted that a CEC can never find a connection between the current to the past, or from the future to the current, but if there seems to be a connection backwards in time it might be possible to create another CEC that was previously overlooked. An example of how to visualize the second type of a CEC is presented in Figure 4, which shows how the introduction of unit-testing impacted the studied company in a larger context, which could have been overlooked if the scope of the research design had been narrower. The general reasoning behind CEC’s has also been visualized in Figure 5.

The BAPO/PCF matrix can also be used as a tool to visualize the research result quality. Several techniques were used to ensure data validity during the company study, but the primary technique was triangulation [1], [10], which states
that in order for information to be valid it must be verified by at least three sources. These sources can be either artefacts or roles, for instance documentation, different roles at the company, etc, or results gathered by using several research methods, such as interviews, literature review, etc. The matrix can be used to show how triangulation was performed during a study by visualizing which sources were used to answer what research question within the BAPO/PCF matrix. This provides a graphical overview of the data validity that is easy to understand, also by non-researchers like for instance the managers of the company under study. The general triangulation technique is presented in Figure 2, and the visualization of what sources that were used for triangulation, in the company study, is presented in Figure 3. Another technique that was used to ensure data validity during this step of the process, in the company study, was to send transcripts of interviews, preliminary documentation analysis reports, etc, to different roles within the studied company for validation. This technique provides the researcher with fast feedback if the elicited information has been correctly interpreted or not, which is valuable in cases where the information includes contextual jargon unknown to the researcher.

**Validation Discussion.** The final step of the process was to validate the research results with the studied company. This was partially done in step four and five since those steps deal with information validation through the means of company review, but those reviews were on a information level, whilst the final validation was on a conclusion level to ensure that the conclusions drawn from the research information were reasonable and correct. The final validation can be done in different ways, like for instance writing a report that the company can review, or the results can be presented orally. In the company study the results were presented orally during a power-point presentation where the analytical matrix was once more used. All significant results from the study were broken down into bullet-points in the matrix cells and added to a power-point slide. The graphical representation of the results gave the audience a clear overview of the results, which could afterwards be discussed further in more detail. Errors or discrepancies in the results were finally analyzed further and rewritten in a final study report.

**IV. Discussion**

The perceived advantages of applying the BAPO/PCF framework for case studies are related to the structure and the breadth of the multi-perspective view of the framework. By applying the multi-perspective approach to a case study it becomes possible to collect results specific to the different perspectives of the company, as well as in time. This provides a contextual structure that makes it clear where and how to acquire information to answer the research questions, but also how to find information that can link the results together. For instance by pointing out a role within the studied company that has knowledge about several perspectives, the information gained from this knowledge can be used to draw cross-perspective conclusions about the cause of a particular result. By looking at the entire matrix, including these cross-perspective results, longer chains of cause and effect can also be drawn that provides a deeper as well as broader understanding of results and of the company. Internal validity is also improved with this approach since as described by Wohlin et al. 2000 [11], a factor that is investigated because of its effects on another factor may itself be affected by a third factor, and if the third factor is overlooked there may be a threat to internal validity. Hence the BAPO/PCF framework improves the internal validity by giving the researcher the means of finding these connections between different factors that could have otherwise been overlooked with a narrower scope. It is however important to recognize that this framework, like other case studies, does not solve the issue regarding to what extent a factor within a given perspective affects another factor in another given perspective. A change within a company is seldom localized to a certain perspective of the company but rather has ripple effects to several different perspectives. An example from the company study is how changes to a company’s business goals resulted in the introduction of new processes, new architectural development methods as well as

---

**Fig. 4. Cause-effect chain example**

**Fig. 5. General Cause-effect chain structure**
organizational change.

Other advantages that are side effects of the BAPO/PCF framework’s structure include support for project planning, visualization of the validation through triangulation, as well as to visualize the research results. The framework supports research plans that are flexible and allow events, such as research opportunities, to be taken advantage of as they arise during the study. Visualization of the triangulation provides an overview of the validation effort and also makes it easy for non-researchers to see how the results have been validated, for instance making it easier to trust results and conclusions that are previously unknown to the studied company. Result visualization has similar advantages in terms of providing an overview of the collected results, and help to find connections between the collected information.

As for general case studies this design does not limit what research methods that are applicable for acquiring information, but unlike many other case study designs that rely on purely qualitative methods for data elicitation and analysis, this design allows longitudinal design concepts to be used as well, which are quantitative. This is made possible by the chronological axis of the matrix that links the past to the future and therefore allows quantitative metrics to be developed. An example of such a metric could be organizational change over time, i.e. the company’s growth rate. Such information can also be used deductively to draw more plausible predictions about the future, for instance how the growth of the company will continue or decrease.

The design does provide context to the research results but by broadening the research it becomes necessary to sacrifice information depth in projects with fixed budgets. Hence one of the BAPO/PCF framework’s greatest strengths is also its primary weakness. There is a trade-off that must be made when using the framework, which states that in order to gain information depth in one perspective, information depth will have to be sacrificed in another perspective. The consequences of this trade-off has not been investigated in this study, and since the design has only been used during one empirical case study this subject is still open to speculation. The broadness of the design does however make it more suitable for research with open ended questions where very little or nothing is known about the phenomenon under study, rather than research with narrow research questions of a more deductive nature.

The results and conclusions that were developed during the company study proved to have high validity, however since the design has only been used in one study it is uncertain if the same results would be achieved in another study. The design is flexible in many ways, including the core concept of the analytical matrix, but once again, since only the BAPO/PCF configuration has been tested nothing can be said about another configuration even though it can be speculated that another configuration could be more beneficial in another company context.

V. Conclusions

This paper introduced the BAPO/PCF framework for structuring case studies in software engineering and ensuring they cover multiple perspectives. The framework was evaluated in a case study of verification and validation practices in a Swedish software company.

The combination of the BAPO (Business, Architecture/technical, Process and Organization) and PCF (Past, Current and Future) dimensions resulted in 12 sub areas to consider in designing and executing the case study. The matrix allowed the research questions to be connected to the research effort, as well as providing the researcher with tools for result visualization, project planning and result validation. The structure also allowed an analysis of cause-effect chains across perspectives and in time providing a broader understanding and increased validity. Most importantly the framework helped uncover issues and connections the company themselves were not aware off.

The largest perceived disadvantage is that the approach can become too broad and therefore require considerable effort to cover all 12 sub areas. This can be addressed by sacrificing depth of analysis within less prioritized sub areas but this needs further research.

In summary, a flexible yet powerful case study research design can be created by adding structure through the use of an analytical matrix and a simple process to adapt it to the context being studied. The analytical matrix used, based on BAPO and PCF, can be of general value for such software engineering research.

REFERENCES