

Using a Protocol Template for Case Study Planning

Pearl Brereton

School of Computing and Mathematics, Keele University, Keele, Staffordshire ST5 5BG, UK
O.P.Brereton@cs.keele.ac.uk

Barbara Kitchenham

School of Computing and Mathematics, Keele University, Keele, Staffordshire ST5 5BG, UK
B.A.Kitchenham@cs.keele.ac.uk

David Budgen

*Department of Computer Science, Durham University, Durham, Science Laboratories,
South Road, Durham City, DH1 3LE, UK.*
david.budgen@durham.ac.uk

Zhi Li

School of Computing and Mathematics, Keele University, Keele, Staffordshire ST5 5BG, UK
z.li@epsam.keele.ac.uk

Abstract. In order to undertake a series of case studies aimed at investigating systematic literature reviews, we have developed a case study protocol template. This paper introduces the template and discusses our experiences of using the template and the resulting case study protocol. We suggest that using our template to prepare a protocol and using the protocol for case study planning can improve the rigour of software engineering case studies.

Keywords: case study, template, protocol.

1.INTRODUCTION

We are currently undertaking a program of research aimed at better understanding the role of systematic literature reviews (SLRs) in software engineering. This is part of the Evidenced-based Practices Informing Computing (EPIC) project which is funded by the UK Engineering and Physical Sciences Research Council. SLRs are a method used to find, critically evaluate and aggregate all relevant research papers (referred to as primary studies) on a specific research question or research topic. The method was developed in medicine and has been adopted by many other disciplines including social sciences, economics, and nursing. The method is intended to ensure that the review is unbiased, rigorous and auditable (see for example, Petticrew and Roberts, 2006; Fink, 2005; Khan et al., 2003).

We are using the participant-observer case study approach as our main research methodology for investigating software engineering SLRs, and plan a series of case studies investigating issues such as:

- The problems facing novice researchers.
- The impact of limited search procedures.
- The repeatability of the SLR process.
- The usefulness of the SLR guidelines (Kitchenham and Charters, 2007).

We believe that case studies are a very useful research method for software engineering research but are a method that is quite difficult to apply rigorously. Zannier et al. (2006) reviewed 63 papers randomly selected from 1227 papers published in 29 years of ICSE proceedings. Looking at the type of evaluation undertaken in the articles, they commented:

"From our perspective, the term case study was frequently used improperly. Case studies lacked hypotheses and/or a real world case."

In addition, we believe the case study methodology has wider applicability than it might at first appear. We agree with Yin (2003) that a proper case study takes place in the field, which in the case of software engineering means in real commercial software projects. However, we also think that validation exercises based on implementing an existing system using a new methodology can provide useful information about the advantages and limitations of new technologies (although **not** information about potential productivity or timescale advantages – since examples are seldom produced under the strict process controls required for commercial products). Such studies are not case studies; they should be referred to as large-scale examples. To improve their rigour, such large-scale examples would benefit from adopting some of the procedures recommended for case studies.

Recently, Höst and Runeson (2007) developed checklists for undertaking case studies and reviewing case studies. These provide help both for case study researchers and readers of case studies. In this paper we present a generic template that may help case study researchers to construct a case study protocol. This relates to the case study design and preparation for data collection sections in Höst and Runeson's checklist for researchers.

In section 2, we explain how we developed our case study protocol and present the basic case study template. In section 3, we discuss our experiences of using the template and how we needed to amend it to suit our specific case study. In section 4, we discuss the benefits of using a case study protocol, and present our final conclusion in section 5.

2. TEMPLATE CONSTRUCTION

As part of our overall planning of our research program (Brereton et al., 2007) we identified the need to plan a series of case studies. To ensure a consistent planning process, we decided to develop a basic case study protocol template. The aim of the template was to provide a common structure for our case study protocols and guidance on how to construct them. We reviewed some of the well-known case study papers and text book but did not find any existing template, although Yin (2003) presented an example of a case study protocol. For this reason we constructed our own template based on basic case study methodologies described by Eisenhardt (1989), Stake (1995), and Yin (2003).

At a high level most of the approaches to case study conduct are quite similar, in spite of the very different philosophical approaches taken by different case study experts. For example Eisenhardt (1989) who is concerned with using case studies to develop theories suggests the following activities:

- **Getting started** by defining the research question and a priori questions but not defining hypotheses or theory.
- **Selecting cases** by considering a particular population and using theoretical concerns to focus on specific cases.
- **Crafting instruments and protocols** using multiple data collection methods, using qualitative and quantitative data, and preferably multiple researchers.
- **Entering the field** i.e. incorporating field notes with data analyses and using flexible and opportunistic data collection methods.
- **Analysing the data** both within case and across cases.
- **Shaping hypotheses** by iterative tabulation of evidence looking for identified constructs, replication logic across cases, and looking for evidence to explain why relationships exist.
- **Enfolding the literature** i.e. comparing with existing similar and conflicting literature.
- **Reaching closure** using the concept of "theoretical saturation" which says researchers stop looking for more cases/data when they believe more data will only give a marginal improvement to the existing results.

Stake (1995), like Eisenhardt, takes an interpretive approach to case study research, but is particularly concerned with "program evaluation" which involves evaluation of social or education policies. He discusses the following case studies issues:

- **Criteria for selection of cases** including whether it is an intrinsic case (i.e. the case itself is the object of interest) or an instrumental case (i.e. the case is used to study something else, for example we intend to use a specific systematic literature review to investigate the SLR methodology).
- **Research questions, issues, and information questions**, where issues may specify possible cause-effect relationships, or potential problems (and are important for intrinsic case studies), and information questions may provide information the case study researcher needs to perform his/her task but are more detailed than the main research question.

- **Data gathering** involving organising data gathering, access and permissions required for field visits, types of data gathering such as observation, interview and document review.
- **Analysis and Interpretation** considering aggregation of data by categories or direct interpretation.
- **Case study researcher roles** considering the approach to be taken in the study (e.g. teacher, advocate, evaluator, biographer, interpreter).
- **Triangulation and validation** including member checking (asking people interviewed or observed to comment on drafts of the case study report).
- **Report Writing** which includes organising the report early on, identifying potential readers, story telling (i.e. the organizational structure of the report), and the use of vignettes (short descriptions that illustrate some aspect of the case).

Yin (2003) takes a slightly more positivist approach to case studies and discusses the following tasks:

- **Case study design** i.e. multiple or single case, holistic or embedded, flexible, or closed; the five components of a research design: questions, propositions, units of analysis, logic linking data to propositions, criteria for interpretation; and validation.
- **Preparing for data collection** including general training (e.g. interviewing techniques), training for the specific case, development of a protocol, selecting the cases(s), conducting a pilot case study.
- **Collecting evidence** which involves considering six sources of evidence and basic principles i.e. using multiple sources, creating a case study database, and maintaining a chain of evidence.
- **Analysing evidence.** Determining the appropriate analysis strategy (i.e. using theoretical propositions, using a framework based on rival explanations, and/or developing case descriptions) and using appropriate analysis techniques (i.e. pattern matching, explanation building, time-series analysis, logic models and/or cross-case synthesis), keeping in mind three principles (attending to all the evidence, presenting the evidence separately from any interpretation, and exploring alternative interpretations).
- **Reporting case studies.** Identify the audience, determine the best structure, have the report reviewed, and start writing early.

These lists demonstrate that there is some general agreement about the individual tasks in a case study, although details vary. In particular, Eisenhardt's goal is to develop theory, so not until the last-but-one stage in her case study process does she consult existing literature.

Our research program aims to understand the SLR process and to evaluate and to improve our SLR guidelines. Thus, we are more concerned with procedures and less with the attitudes of people to the procedures (although we are concerned about problems facing novices). Generally our goals are more consistent with Yin's case study methodology than Eisenhardt's, but like Stake we are concerned with evaluation, although in our case we are concerned with the evaluation of a technology. For this reason, we based our template on Yin's case study methodology while also incorporating some of Stake's ideas.

Yin provides an example of a case study protocol and suggests that a case study protocol should include:

- An overview of the case study project including objectives, issues and relevant findings about the topic being investigated.
- Field procedures, including procedures governing access to case study sites.
- Case study questions, data collection forms and sources of information for answering each question
- A guide for the case study report including an outline of the report and references.

Using Yin's approach, we produced a case study template suitable for our research program. We show a generalised version of the template in Appendix A. The actual template we used for our own research program included elements that were specific to the type of case studies we intended to undertake. We removed these elements to construct the generalised template shown in Appendix A.

In addition to covering the general case study tasks outlined by Yin and Stake, we included several additional sections:

- We included two validation sections, one intended to consider validation of the plan itself, the other intended to specify issues related to the context of the case study that might impact results.
- We included a Divergences section to record changes to the protocol. One of the issues emphasised by case study researchers was that unlike a classic quantitative experiment data collection and analysis are not sequential steps in a case study. Case study researchers need to begin analysis as

soon as data becomes available and to be prepared to change their data collection and analysis plans if necessary. This means a case study protocol is likely to be frequently amended.

- We included a schedule section intended to specify major case study milestones.

3. TEMPLATE USE

The reason we constructed a template was to assist in planning our series of case studies. In this section we will discuss some of the issues that arose when we constructed our protocol, but first we will briefly describe the case study context and goals.

3.1 The First EPIC Case Study

Our case study scoping document (Brereton et al., 2007) identified a list of research questions the EPIC project should address and the basic case study design we would use (i.e. multiple-case embedded design), and some constraints within which the case studies would take place:

- The case study topic should be of interest in its own right and so lead to an SLR which would be of value to software engineering researchers and practitioners
- Whenever appropriate a previous published software engineering review (formal or informal) would be used as a baseline against which our SLR would be evaluated.
- If no previous review was available on the topic, and two independent SLRs were required, the research team members at Durham and Keele would undertake the SLR independently of one another.

The research topics and questions we intended to address in the first case study are shown in Table 1. Other research questions defined in the scoping document were not considered relevant to this case study. (Note the original wordings of the questions in the scoping document were amended slightly in the case study protocol, and we present only the latest wording.)

The software engineering topic we proposed to use in our case study was Unit Testing. This topic was chosen because it is an important software engineering topic that has already been subject to a systematic literature review under taken by Juristo and her colleagues (Juristo et al., 2004; Juristo et al., 2006). However, Juristo et al. restricted their search process to the IEEE and ACM digital libraries. Thus, we would have a case study appropriate to address RQ1 and RQ3. Since this would be the first SLR undertaken by the two research assistants appointed to the EPIC project (one at Durham University, the other at Keele University), it would also address RQ8 which was agreed by the research team to be the main research question. RQ9 and RQ10 relate to issues that affect all SLRs, so we intend to consider these issues in all our case studies. In addition, questions related to search strings (RQ6 and RQ7) would be investigated on an opportunistic basis in all our case studies.

| Case Study Topic | ID | Research Question |
|---|------|---|
| Stability of SLR process | RQ1 | To what extent does the use of the systematic literature review methodology lead to repeatable results? |
| Breadth of Literature Search | RQ3 | To what extent is the adoption of an extended search space vital for answering detailed research questions? |
| Problems facing Novices | RQ8 | Given the existing literature on evidence-based software engineering, including the <i>Guidelines</i> developed by the EBSE project, what are the particular problems facing novice researchers when conducting their first systematic literature review? |
| Manual versus Automated Search Strategies | RQ6 | Are automated search strategies preferable to manual search strategies in the software engineering? |
| Complex versus simple search strings | RQ7 | Are complex focused search strings effective in software engineering? |
| Protocols | RQ9 | To what extent do systematic literature reviews deviate from their protocols? |
| Quality Checklists | RQ10 | How useful are the quality checklists provided in the current version of the <i>Guidelines</i> (Kitchenham & Charters, 2007)? |

Table 1 Research Topics and Questions

3.2 Constructing the Protocol

Constructing a protocol from a template involves not only completing the template sections but also reviewing the relevance of the template section in the specific context. For example, when we used the template for our first case study plan, we omitted some elements of the template:

- We did not include a section related to case study selection because this issue was addressed in a higher level document (Brereton et al., 2007).
- We did not include a section on field procedures, since the object of our study is systematic literature reviews, and members of the EPIC research team (i.e. two research assistants, one at the University of Keele and the other at the University of Durham) would themselves be performing the reviews.

We also found that we needed to revise the basic structure of sections 5 and 6 because we trying to address several very different research questions in the same case study, i.e.

- The problems faced by novice researchers when they use the SLR methodology.
- The extent to which the SLR methodology leads to repeatable results.
- The implications of restricted search procedures.

Each of these research questions was refined into sub-questions or propositions. However, when we attempted to define data collection and analysis for all the questions together sections 5 and 6 became very difficult to understand. We found it necessary to re-write section 5 and 6 for each research question, so we ended up with headings for each top level research question and subsections defining the data, data collection, and analysis procedures that addressed each of the relevant sub-questions or propositions.

One issue that we did not include in the first version of our template was the role of team members. This was a major error, which fortunately was detected during a team review of the protocol. Since we are using a participant-observer design, we needed to define the roles of the research team in terms of their roles as part of the case study team and as part of the SLR team. In particular, two of the senior researchers have two roles:

- They are members of the case study team responsible for planning the case study activities.
- They are members of the SLR team responsible for supervising the research assistants.

Both these roles needed to be properly specified to ensure that the different responsibilities of each role were understood. We also needed to consider how much of the case study plan should be visible to the research assistants. In practice, we provided a separate specification of work for RAs outlining the tasks they were to perform and the data they were to collect without informing them of the specific case study research questions.

One area where we deviated from the advice of all the case study experts is that our protocol does not include a planned report structure in the report section. We take the view that a well-organised and maintained case study protocol should provide much of the information needed in a case study report. Furthermore, we do not have to report formally to external sponsors and other stakeholders, as is assumed by case study researchers in the social sciences and education.

In terms of the development process for the protocol, a major benefit that arose from using the template was that it made it easier to manage and control the process. Writing a protocol is essentially a design task, and a recognised characteristic of design activities is that they are largely opportunistic in nature, rather than following a rigid or sequential set of steps. Here, our experience was that our ideas were able to evolve in an opportunistic manner, but with the use of the template ensuring that we did not lose sight of the overall goals.

4. THE VALUE OF A PROTOCOL

From our experience, one of the most important benefits of our case study protocol came from forcing us to specify in detail how we intended to answer our research questions. Relating the data we intended to collect back to the research questions and recalling the need to triangulate our results (i.e. collect different data to address the same question) caused us to refine our data collection procedures.

With respect to analysis, thinking in detail about the data and how they would be interpreted forced us to consider the different outcomes we might find, and to specify the different causes for the different outcomes. This again led us to refine our data collection procedures so that we could properly distinguish between the different causes.

In practice, we identified the need for data prospectively that it would have been difficult or invalid to attempt to extract at the end of the case study (e.g. de-briefing questionnaires for RAs that can be applied after each main SLR stage, procedures for the supervisors to assist the evaluation of the quality checklists in the SLR guidelines).

Another advantage of a protocol is that when circumstances change, it is easier to understand how to adjust the case study to address those circumstances against a formal case study plan than against an informal understanding. In our case, we found from reviewing the SLR protocols that the RAs were unable to duplicate the

data collection and analysis process in the SLR we were using as a baseline. Reviewing our case study protocol, we could see that stopping the investigation of the repeatability of the SLR process after the data extraction would provide some useful information about repeatability but would avoid problems with data extraction and analysis confounding the investigation of our other research questions.

5. CONCLUSIONS

This paper has introduced a template that can help the production of a case study protocol. We believe that this template will support Höst and Runeson (2007) checklists to improve the rigour of case studies in software engineering. The use of a protocol is an important means of improving the rigour of case studies and other related evaluation methods. We have used the template to produce a protocol for the first of a series of case studies. In our experience, the discipline of producing a protocol has great benefits with respect to ensuring that the data collection and analysis procedures are able to reliably answer the research questions. It also provides a baseline against which a case study plan can be revised if required. In addition, the provision of a written research plan that can be formally reviewed helps to reduce the problem of completing a template in the letter but not in the spirit.

Software engineering researchers often employ examples (or scenarios) that enable them to illustrate the value of their work. Recent instances of such an approach includes the papers by Bleinstein et al. (2006) and Boness et al. (2008). On a smaller scale, examples are often employed in student projects and are particularly valuable for MSc students, who often have to complete their development and evaluation activities within a very constrained timeframe. For both groups, a more formal case study structure would provide an underpinning methodology that could improve the validity of their conclusions -- and the use of a template can make this approach much more accessible.

However, currently we have only limited experience of using the template, although we intend to use it for future EPIC case studies. In addition, it is being trialled by students at the University of Durham. We welcome any feedback from other researchers currently planning case studies on the value of the template.

REFERENCES

- Bleinstein, S.J., Cox, K., Verner, J. and Phalp, K. T. (2006) B-SCP: A requirements analysis framework for validating strategic alignment of organizational IT based on strategy, context, and process," *Information and Software Technology*, vol. 46, pp. 846-868.
- Boness, K., Finklestein, A. and Harrison, R. (2008) A lightweight technique for assessing risks in requirements analysis. *IET Software* 2(1) pp46-57.
- Brereton, O. P., Kitchenham, B.A. (2007) The Scope of EPIC Case Studies. EPIC technical Report EPIC-2007-01.
- Eisenhardt, K.M.(1989) Building Theories from Case Study Research, *Academy of Management Review*, Vol. 14, No. 4, October, pp532-550 (see also rejoinder by Dyer & Wilkins, 1991, and response by Eisenhardt, 1991).
- Fink, A. (2005) *Conducting Research Literature Reviews. From the Internet to Paper*, Sage Publication, Inc.
- Juristo, N., Moreno, A.M. and Vegas,S. (2004) Reviewing 25 years of Testing Technique Experiments, *Empirical Software Engineering Journal*, Issue 1-2, March, pp. 7-44.
- Juristo, N., Moreno, A.M. Vegas,S. and Solari, M. (2006) In Search of What We Experimentally Know about Unit Testing, *IEEE Software*, 23 (6), pp.72-80.
- Höst, M. and Runeson, P. (2007) Checklists for Software Engineering Case Study Research, in *Proceedings of 1st International Symposium on Empirical Software Engineering & Measurement (ESEM)*. IEEE Computer Society Press, pp. 479-482.
- Khan, Khalid, S., Kunz, Regina, Kleijnen, Jos and Antes, Gerd. (2003) *Systematic Reviews to Support Evidence-based Medicine*, The Royal Society of Medicine Press Ltd.
- Kitchenham, B.A. and Charters, S. (2007) Guidelines for performing Systematic Literature Reviews in Software Engineering Technical Report EBSE-2007-01.
- Petticrew, Mark and Helen Roberts. (2005) *Systematic Reviews in the Social Sciences: A Practical Guide*, Blackwell Publishing, ISBN 1405121106
- Stake, R.E. (1995) *The Art of Case Study Research*, Sage, Thousand Oaks, CA.
- Yin, Robert K. (2003) *Case Study Research: Design and Methods*, 3rd Edition, Sage Publications.
- Zannier, C, Melnick, G. and Maurer, F. (2006) On the Success of Empirical Studies in the International Conference on Software Engineering.ICSE06, pp. 341-350.

Acknowledgements

This study was funded by the UK Engineering and Physical Sciences Research Council project EP/E046983/1.

APPENDIX A: CASE STUDY PROTOCOL TEMPLATE

1. Background

- a) identify previous research on the topic
- b) define the main research question being addressed by this study
- c) identify any additional research questions that will be addressed

2. Design

- a) identify whether single-case or multiple-case and embedded or holistic designs will be used, and show the logical links between these and the research questions
- b) describe the object of study (e.g. a new testing procedure; a new feature in a browser)
- c) identify any propositions or sub-questions derived from each research question and the measures to be used to investigate the propositions

3. Case Selection

- a) Criteria for case selection

4. Case Study Procedures and Roles

- a) Procedures governing field procedures
- b) Roles of case study research team members

5. Data Collection

- a) identify the data to be collected
- b) define a data collection plan
- c) define how the data will be stored

6. Analysis

- a) identify the criteria for interpreting case study findings
- b) identify which data elements are used to address which research question/sub question/proposition and how the data elements will be combined to answer the question
- c) consider the range of possible outcomes and identify alternative explanations of the outcomes, and identify any information that is needed to distinguish between these
- d) the analysis should take place as the case study task progresses

7. Plan Validity (see Figure 2.3 and Chapter 5 in Yin (2003))

- a) general: check plan against Höst and Runeson's (2007) checklist items for the design and the data collection plan
- b) construct validity - show that the correct operational measures are planned for the concepts being studied. Tactics for ensuring this include using multiple sources of evidence, establishing chains of evidence, expert reviews of draft protocols and reports
- c) internal validity - show a causal relationship between outcomes and intervention/treatment (for explanatory or causal studies only).
- d) external validity – identify the domain to which study finding can be generalized. Tactics include using theory for single-case studies and using multiple-case studies *to investigate outcomes in different contexts*.

8. Study Limitations

Specify residual validity issues including potential conflicts of interest (i.e. that are inherent in the problem, rather than arising from the plan).

9. Reporting

Identify target audience, relationship to larger studies (Yin, 2003)

10. Schedule

Give time estimates for all of the major steps: Planning, Data Collection, Data Analysis, Reporting. Note Data Collection and Data Analysis are not expected to be sequential stages

11. Appendices

- a) **Validation:** report results of checking plan against Höst and Runeson's (2007) checklist items
- b) **Divergences:** update while conducting the study by noting any divergences from the above steps.