

Checklists for Software Engineering Case Study Research

Martin Höst, Per Runeson
Software Engineering Research Group
Lund University, Sweden
martin.host@cs.lth.se, per.runeson@cs.lth.se

Abstract

Case study is an important research methodology for software engineering. We have identified the need for checklists supporting researchers and reviewers in conducting and reviewing case studies. We derived checklists for researchers and reviewers respectively, using systematic qualitative procedures. Based on nine sources on case studies, checklists are derived and validated, and hereby presented for further use and improvement.

1. Introduction

Software engineering is a field of applied research. Research often involves investigating how people work in teams and projects in large organizations aiming to develop software. Case study research methodology is feasible to use when individual, group, organizational and social phenomena are investigated [11]; hence it is often found suitable in software engineering research.

There is a large variety of case study types [2]. A typical case study process consists of the following phases [11], which may be iterated [1]:

1. Case study design: objectives are defined and the case study is planned.
2. Preparation for data collection: procedures and protocols for data collection are defined.
3. Collecting evidence: execution with data collection on the studied case.
4. Analysis of collected data
5. Reporting

There are several examples of case study research initiatives in software engineering, but to our knowledge there are no specific text books on how to perform case study research in this domain. General methodology handbooks exist, which we apply to software engineering, e.g. [8][9][11].

In this paper we report on the development of a set of checklists for case studies in order to improve the case study standards in software engineering. We have identified the need for a checklist, both when conducting case studies and when reviewing case study

reports. In order to derive the checklists, we have applied systematic procedures, based on qualitative research methodology.

2. Methodology

The checklists are derived in six major steps:

1. A literature survey to identify existing checklists.
2. Merging all found checklist items into one list.
3. Classification of checklist items according to case study phase and role.
4. Reduction of the number of items by grouping of related items within each phase and formulating checklist items for each group.
5. Validation of the checklist in a PhD student course.
6. Update of checklists after validation.

The derivation is conducted by the authors of this paper, and the validation was conducted by nine PhD students attending a course on case study methodology (step 5). The first author conducted steps 1 and 2; step 3 was conducted by both authors independently; steps 4 and 6 were conducted by the second author and reviewed by the first author.

3. Results

The survey (step 1) resulted in checklists from nine different sources, see Table 1.

Table 1. Checklist sources

Author	Type	Domain
Corcoran [3]	Journal	Education
Esterhuizen [4]	Handbook	International policy
Kitchenham, Pfleeger [5]	Journal	Software engineering
Kyburz-Graber [6]	Journal	Education
Perry et al [7]	Tutorial	Software engineering
Robson [8]	Textbook	Social science
Stake [9]	Textbook	Social science
Wohlin et al [10]	Textbook	Software engineering
Yin [11]	Textbook	Social science

The merged list of checklists (step 2) comprised 103 items. Each item was classified by both authors independently with respect to phase (see Section 1) and role (step 3). Each item was connected to one or more phases. In case the two authors classified an item differently, a negotiation took place.

In the derivation, we defined three user roles for the checklist (*Researcher*, *Data provider* and *Audience*) and five subroles. However, it was not possible at this level to clearly distinguish between the roles, so the role classification attempt was dropped.

The phase classification resulted in 49 items for the design phase, 22 for preparation, 19 for data collection, 35 for analysis and 52 for the reporting phase.

Within each phase, the checklist items were categorized in three to eight categories (step 4). For example, the design phase items were categorized into either of *author*, *case*, *data*, *method*, *purpose*, *selection* or *theory*. Each subset of items was then reworded into one or more checklist items for our first version of a case study checklist. The resulting checklist had 46 items, grouped per phase (12, 6, 8, 8 and 12 items respectively).

The checklist was validated (step 5) in a PhD student course, where nine students applied the checklist during their review of published case study articles. Each student reported which checklist items they found being useful, and which ones they had trouble understanding. All but one of the students are doing research in the software engineering domain.

A major finding from the validation was that the first version of the checklist was too extensive for the purpose of the review. The reporting section only was not sufficient for this purpose, and reviewing a published paper phase by phase was too extensive. Hence we decided to split the checklist in two, one detailed for researchers conducting a case study, and one broader for those reviewing case study reports. The researcher's checklist is an updated version of the full checklist, see Appendix A. The reviewer's checklist is a shorter one derived from the reporting phase check items, adding quality characteristics to each item, see Appendix B. The items in the reviewer's checklist are linked to the researcher's items for a more detailed description.

During the validation, some redundant items were identified. Those were either removed or integrated with other items, resulting in a researcher's checklist with 38 items, grouped by phase into 10, 5, 6, 6 and 11 items for design, planning, collection, analysis and reporting, respectively.

4. Conclusions

The proposed checklists are derived, using a systematic qualitative approach. An initial validation identified the need for specialization of the checklists for two roles, which also is conducted. We encourage researchers and reviewers to use the checklists and to report their experiences for continuous improvement.

Further work, in addition to evaluation by active case study researchers, possibly includes tailoring of the checklists to software engineering terminology and conditions. The checklists are still defined in general terms, while tailoring or exemplifying in the software engineering domain might make them more useful.

References

- [1] C. Andersson, P. Runeson, "A Spiral Process Model for Case Studies on Software Quality Monitoring - Method and Metrics", *Software Process Improvement and Practice*, 12(2):125-140, 2007.
- [2] J. B. Cunningham, "Case study principles for different types of cases", *Quality & Quantity*, 31:401-423, 1997.
- [3] P.B. Corcoran, "Case Studies, Make-Your-Case Studies, and Case Stories: a Critique of Case-Study Methodology in Sustainability in Higher Education", *Environmental Education Research*, 10(1):7-21, 2004.
- [4] L. Esterhuizen, *Doing Case Studies for the Refugee Sector; A DIY Handbook for Agencies and Practitioners*, The Information Centre about Asylum and Refugees in the UK (ICAR), International Policy Institute, King's College London, 2004.
- [5] B.A. Kitchenham, S.L. Pfleeger, "Case Studies for Method and Tool Evaluation", *IEEE Software*, 12(4):52-62, 1995.
- [6] R. Kyburz-Graber, "Does Case-Study Methodology Lack Rigour? The Need for Quality Criteria for Sound Case-Study Research, as Illustrated by a Recent Case in Secondary and Higher Education", *Environmental Education Research*, 10(1):53-65, 2004.
- [7] D.E. Perry, S. Elliott Sim, S. Easterbrook, "Case Studies for Software Engineers (Tutorial)", *Proceeding of the 28th International Conference on Software Engineering*, pp 1045-1046, 2006.
- [8] C. Robson, *Real World Research*, 2:nd ed., Blackwell, 2002.
- [9] R.E. Stake, *The Art of Case Study Research*, Sage Publications, 1995.
- [10] C. Wohlin, P. Runeson, M. Höst, M.C. Ohlsson, B. Regnell, A. Wesslén, *Experimentation in Software Engineering: An Introduction*, Kluwer Academic Publishers, 2000.
- [11] R.K. Yin, *Case Study Research Design and Methods*, 3:rd ed. Sage, 2003.

Appendix A. Researcher's Checklist

Case Study Design
1. What is the object of study?
2. Is a clear purpose/objective/research question /hypothesis/proposition defined upfront?
3. Is the theoretical basis - relation to existing literature and other cases - defined?
4. Are the authors' intentions with the research made clear?
5. Is the case adequately defined (size, domain, process...)?
6. Is a cause-effect relation under study? If yes, is the cause distinguished from other factors?
7. Will data be collected from multiple sources? Using multiple methods?
8. Is there a rationale behind the selection of roles, artefacts, viewpoints, etc.?
9. Are the case study settings relevant to validly address for the research question?
10. Is the integrity of individuals/organizations taken into account?
Preparation for Data Collection
11. Is a protocol for data collection and analysis derived (what, why, how)?
12. Are multiple data sources and collection methods planned?
13. For quantitative data, are the measurements well defined?
14. Are the planned methods and measurements sufficient to fulfil the objective of the study?
15. Is the study design approved by a review board, and has informed consent obtained from individuals and organizations?
Collecting Evidence
16. Are data collected according to the protocol?
17. Is the observed phenomenon correctly implemented (e.g. to what extent is a design method under study actually used)?
18. Are data recorded to enable further analysis?
19. Are sensitive results identified (for individuals, organization or project)?
20. Are the data collection procedures well traceable?
21. Do the collected data provide ability to address the research question?
Analysis of Collected Data
22. Is the analysis methodology defined, including roles and review procedures?
23. Is a chain of evidence shown with traceable inferences from data to research questions and existing theory?
24. Are alternative perspectives and explanations used in the analysis?
25. Is a cause-effect relation under study? If yes, is the cause distinguished from other factors?
26. Are there clear conclusions from the analysis, including recommendations for practice/further research?
27. Are threats to validity addressed in a systematic way?

Reporting
28. Are the case and its context adequately reported?
29. Are the research questions and corresponding answers reported?
30. Are related theory, hypotheses and propositions clearly reported?
31. Are the data collection procedures presented, with relevant motivation?
32. Are sufficient raw data presented?
33. Are the analysis procedures clearly reported.
34. Are threats to validity analyses reported?
35. Are ethical issues reported openly (personal intentions, integrity issues)
36. Does the report contain conclusions, implications for practice and future research?
37. Does the report give a realistic and credible impression?
38. Is the report suitable for its audience, easy to read and well structured?

Appendix B. Reviewer's Checklist¹

1. Are the research questions, objects of study and case study context well defined? 1, 2, 5, 28, 29
2. Is it motivated that the case is suitable to address the research questions? 8, 9, 14
3. Are the hypotheses and propositions clear and relevant? 2, 30
4. Are the data collection procedures sufficient for the purpose (data sources, collection, storage, validation)? 11, 13, 16, 18, 21, 32
5. Are sufficient raw data presented to provide understanding of the case? 31
6. Are the analysis procedures sufficient for the purpose (repeatable, transparent)? 22, 33
7. Is the case study based on theory and linked to existing literature? 3
8. Is a clear chain of evidence established from observations to conclusions? 6, 17, 20, 23, 25
9. Are threats to validity analyses addressed in a systematic way? 27, 34, 37
10. Are different views taken on the case (multiple collection and analysis methods, multiple authors)? 7, 12, 22, 24
11. Are ethical issues addressed properly (personal intentions, integrity issues, consent, review board approval)? 4, 10, 15, 19, 35
12. Are conclusions, implications for practice and future research, reported suitably for its audience? 26, 29, 36, 37, 38

¹ The numbers after each item refer to corresponding items in the Researcher's checklist.