Quality Assessment and Strength of Evidence in a Mapping Study

Based on Tore Dyba˚ *, Torgeir Dingsøyr (2008). Empirical studies of agile software development: A systematic review. IST, 50.

Research Questions

1. What is currently known about the benefits and limitations of agile software development?
2. What is the strength of the evidence in support of these findings?
3. What are the implications of these studies for the software industry and the research community?

Quality Issues

- ***Rigour***. Has a thorough and appropriate approach been applied to key research methods in the study?
- ***Credibility***. Are the findings well-presented and meaningful?
- ***Relevance***. How useful are the findings to the software industry and the research community?

Screening Questions

1. Is this a research paper? *Consider:* *—Is the paper based on research (or is it merely a "lessons learned" report based on expert opinion?*	☐ Yes ☐ No
2. Is there a clear statement of the aims of the research? *Consider:* *—Is there a rationale for why the study was undertaken?* *—Is the study's focus or main focus on Agile Software Development?* *—Does the study present empirical data?* *—Is there a clear statement of the study's primary outcome (i.e. time-to-market, cost, or product or process quality)?*	☐ Yes ☐ No
3. Is there an adequate description of the context in which the research was carried out? *Consider whether the researcher has identified:* *—The industry in which products are used (e.g. banking, telecommunications, consumer goods, travel, etc)* *—The nature of the software development organization (e.g. in-house department or independent software supplier)* *—The skills and experience of software staff (e.g. with a language, a method, a tool, an application domain)* *—The type of software products used (e.g. a design tool, a compiler)* *—The software processes being used (e.g. a company standard process, the quality assurance procedures, the configuration management process)*	☐ Yes ☐ No

Detailed Questions

Research design 4. Was the research design appropriate to address the aims of the research? *Consider:* – *Has the researcher justified the research design (e.g. have they discussed how they decided which methods to use)?*	☐ Yes ☐ No
Sampling 5. Was the recruitment strategy appropriate to the aims of the research? *Consider:* *—Has the researcher explained how the participants or cases were identified and selected?* *—Are the cases defined and described precisely?* *—Were the cases representative of a d efined population?* *—Have the researchers explained why the participants or cases they selected were the most appropriate to provide access to the type of knowledge sought by the study?* *—Was the sa mple size sufficiently lar ge?*	☐ Yes ☐ No
Control group 6. Was there a control group with which to compare treatments? *Consider:* *—How were the contro ls selected?* *—Were they repres entative of a defined population?* *—Was there anyth ing special about the controls?* *—Was the no n-response high? Could non-respondents be different in any way?*	☐ Yes ☐ No

Detailed Questions

Data collection 7. Was the data collected in a way that addressed the research issue? *Consider:* *—Were all measures clearly defined (e.g. unit and counting rules)?* *—Is it clear how data was collected (e.g. semi-structured interviews, focus group etc.)?* *—Has the researcher justified the methods that were chosen?* *—Has the researcher made the methods explicit (e.g. is there an indication of how interviews were conducted, did they use an interview guide)?* *—If the methods were modified during the study, has the researcher explained how and why?* *—Whether the form of the data is clear (e.g. tape recording, video material, notes etc.)* *—Whether quality control methods were used to ensure completeness and accuracy of data collection*	☐ Yes ☐ No
Data analysis 8. Was the data analysis sufficiently rigorous? *Consider:* *—Was there an in-depth description of the analysis process?* *—If thematic analysis was used, is it clear how the categories/ themes were derived from the data?* *—Has sufficient data been presented to support the findings?* *—To what extent has contradictory data been taken into account?* *—Whether quality control methods were used to verify the results*	☐ Yes ☐ No

Detailed Questions

Reflexivity (research partnership relations/recognition of researcher bias) 9. Has the relationship between researcher and participants been considered adequately? *Consider:* *—Did the researcher critically examine their own role, potential bias and influence during the formulation of research questions, sample recruitment, data collection, and analysis and selection of data for presentation?* *—How the researcher responded to events during the study and whether they considered the implications of any changes in the research design.*	☐ Yes ☐ No
Findings 10. Is there a clear statement of findings? *Consider:* *—Are the findings explicit (e.g. magnitude of effect)?* *—Has an adequate discussion of the evidence, both for and against the researcher's arguments, been demonstrated?* *—Has the researcher discussed the credibility of their findings (e.g. triangulation, respondent validation, more than one analyst)?* *—Are limitations of the study discussed explicitly?* *—Are the findings discussed in relation to the original research questions?* *—Are the conclusions justified by the results?*	☐ Yes ☐ No

Detailed Questions

Value of the research 11. Is the study of value for research or practice? *Consider:* —*Does the researcher discuss the contribution the study makes to existing knowledge or understanding (e.g. do they consider the findings in relation to current practice or relevant research-based literature)?* —*Does the research identify new areas in which research is necessary?* —*Does the researcher discuss whether or how the findings can be transferred to other populations, or consider other ways in which the research can be used?*	☐ Yes ☐ No

Extraction Form

Study description		
1.	Study identifier	Unique id for the study
2.	Date of data extraction	
3.	Bibliographic reference	Author, year, title, source
4.	Type of article	Journal article, conference paper, workshop paper, book section
5.	Study aims	What were the aims of the study?
6.	Objectives	What were the objectives?
7.	Design of study	Qualitative, quantitative (experiment, survey, case study, action research)
8.	Research hypothesis	Statement of hypotheses, if any
9.	Definition of agile software development given in study	Verbatim from the study
10.	Sample description	Size, students, professionals (age, education, experience)
11.	Setting of study	Industry, in-house/supplier, products and processes used
12.	Control group	Yes, no (number of groups, sample size)
13.	Data collection	How was the data obtained? (questionnaires, interviews, forms)
14.	Data analysis	How was the data analyzed? (qualitative, quantitative)
Study findings		
1.	Findings and conclusions	What were the findings and conclusions? (verbatim from the study)
2.	Validity	Limitations, threats to validity
3.	Relevance	Research, practice

Strength of Evidence – GRADE

- Study Design
  - High: Randomized Controlled Experiments
  - Low: Observational Studies
  - Very Low: All others
- Study Quality
  - Methods, issues of bias, validity, reliability, data collection and analysis, etc.
- Consistency
  - differences in the direction of effects and the size of the differences in effects
- Directioness
  - the extent to which the people, interventions, and outcome measures are similar to those of interest

Quality for Quantitative Studies

Question	Quantitative Empirical Studies (no specific type)	Correlation (observational studies)	Surveys	Experiments	Source
Design					
Are the aims clearly stated?	X	X	X	X	[11], [10]
Was the study designed with these questions in mind?			X		[25]
Do the study measures allow the questions to be answered?			X	X	[10], [25]
What population was being studied?			X		[25]
Who was included?			X		[12]
Who was excluded?			X		[12]
How was the sample obtained (e.g. postal, interview, web-based)?			X		[10], [12], [25]
Is the survey method likely to have introduced significant bias?			X		[25]
Is the sample representative of the population to which the results will generalise?			X	X	[10], [25]
Were treatments randomly allocated?				X	[10]
Is there a comparison or control group?	X		X	X	[12]
If there is a control group, are participants similar to the treatment group participants in terms of variables that may affect study outcomes?	X		X	X	[10], [12]
Was the sample size justified	X		X	X	[10], [12]
If the study involves assessment of a technology, is the technology clearly defined?	X	X	X	X	[11]
Could the choice of subjects influence the size of the treatment effect?				X	[10], [11], [19],[25]
Could lack of blinding introduce bias?				X	[10]
Are the variables used in the study adequately measured (i.e. are the variables likely to be valid and reliable)?	X	X	X	X	[10], [11], [19],[25]
Are the measures used in the study fully defined?	X	X	X	X	[11]

Kitchenham, B.; Charters, S. (2007), Guidelines for performing systematic literature reviews in software engineering, Technical Report EBSE-2007-01, School of Computer Science and Mathematics, Keele University.

Quality for Qualitative Studies

Number	Question	Source
1	How credible are the findings?	[12], [25]
1.1	If credible, are they important?	[12]
2	How has knowledge or understanding been extended by the research?	[12], [25]
3	How well does the evaluation address its original aims and purpose?	[25]
4	How well is the scope for drawing wider inference explained?	[25]
5	How clear is the basis of evaluative appraisal?	[25]
6	How defensible is the research design?	[12], [25], [11]
7	How well defined are the sample design/target selection of cases/documents?	[12], [25], [11]
8	How well is the eventual sample composition and coverage described?	[25]
9	How well was data collection carried out?	[12], [25], [11]
10	How well has the approach to, and formulation of, analysis been conveyed?	[12], [25], [11]
11	How well are the contexts and data sources retained and portrayed?	[25]
12	How well has diversity of perspective and context been explored?	[25]
13	How well have detail, depth, and complexity (i.e. richness) of the data been conveyed?	[25]
14	How clear are the links between data, interpretation and conclusions – i.e. how well can the route to any conclusions be seen?	[25]
15	How clear and coherent is the reporting?	[25]
16	How clear are the assumptions/theoretical perspectives/values that have shaped the form and output of the evaluation?	[12], [25], [11]
17	What evidence is there of attention to ethical issues?	[25]
18	How adequately has the research process been documented?	[25]

Kitchenham, B.; Charters, S. (2007), Guidelines for performing systematic literature reviews in software engineering, Technical Report EBSE-2007-01, School of Computer Science and Mathematics, Keele University.