Journal of Theoretical and Applied Information Technology
30th September 2018. Vol.96. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

QUALITATIVE ASSESSMENT OF SYSTEMATIC LITERATURES IN SOFTWARE ENGINEERING

J.O. OKESOLA[1], K.O. OKOKPUJIE[2], A.A. ADEBIYI[3], C.K. AYO[4]

[1]Department of Computational Sciences, Technical University, Ibadan, Nigeria
[2]Department of Electrical and Information Engineering, Covenant University, Ota, Nigeria
[3,4]Department of Computer & Information Sciences, Covenant University, Ota, Nigeria

[1]Olatunjiokesola@gmail.com, [2]okokpujie@covenantuniversity.edu.ng,
[3]ayo.adebiyi@covenantuniversity.edu.ng, [4]Charles.ayo@covenantuniversity.edu.ng

ABSTRACT

Several systematic literature reviews (SLRs) have been published on many aspects of Software engineering (SE) in the last two decades. However, researchers are yet to evaluate the quality of those studies in order to determine the reliability of their findings. This work employed SLR method and performed automated search of studies published between 2012 and 2017 aiming at evaluating the quality of the recent SLRs published in SE. This paper adapted Dybå and Dingsøyr quality criteria using dichotomous scale of grading to assess the quality of the primary studies in SLRs. A total of 15 of 53 primary studies have suitable recruitment strategy for their research aims, and 19 mentioned the control group (s) with wish their methods were compared. All the 53 papers passed all the standard quality conditions. The quality of the SLRs are high with only very small percentage failing in three out of 11 quality criteria. The research methodologies applied in their primary studies are comprehensive and were based on clear description of the context, thereby making their findings valid and reliable. The current SLRs in SE are of good quality but adequate consideration should be given to the relationship between the researchers and the participants.

Keywords—*Quality assessment, Requirement Engineering, Software Engineering, Structured Review, Systematic Literature Review.*

1. INTRODUCTION

The use of Evidenced-based research and practice started in Medicine because expert opinion based on results from scientific experiments appear to be more reliable than those based on medical advice [1]. This is a Systematic Literature review (SLR), a practice that is primarily aimed at "providing the means by which current best evidence from research is being integrated with practical experience and human values in the decision making process as regard software development and maintenance" [1]. It is a recommended methodology [2] because it brings and combines together findings from different sources. Hence, its application to Software engineering - Evidenced-based Software Engineering (EBSE) - has been on the increase since 2004 [1]. However, it is pertinent to determine the level of confidence to be placed on the findings and recommendations arising from an SLR. This can be achieved either by qualitative or quantity assessement. We are restricted to the former as the latter is out of this research scope.

Rigorous quality assessment of primary studies in an SLR is a notable mechanism of raising the confidence in the findings and recommendations from SLRs [3]. It reduces bias in systematic review process and gives a clearer understanding of potential comparison to guide interpretation of results [4],[5],[6]. Even though, quality assessment of primary studies is complicated as "quality" does not have a generally accepted definition [7], a number of guidelines and tools have been developed specifically for SLRs in Software Engineering (SE). Many of these, including Critical Appraisal Skills Programme - CASP [8] and others developed by Jadad et. al. [9], Dybå and Dingsøyr [7], Kitchenham et al [10], and Sjoberg et al [11] are all influential scales used to evaluate qualitative studies considering the following three broad questions: (1) are the results of the study valid?, (2) what are the results?, and (3)

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

will the results help locally? Methodologies employed in the primary studies of SLRs are hardly considered.

In this work, the quality of SLRs published between 2012 and 2017 is assessed following the comprehensive review of the past related works [12]. We admit that *quality* is an extent to which systematic errors or bias may be prevented through the review and analysis of primary studies [6], However, since "quality of a study is closely linked to the research methods used and the validity of the findings generated by the study"[11], it becomes imperative that the methods of all primary studies of SLRs are critically appraised in this paper and the validity of results clearly established.

2 RELATED WORKS

A good number of existing tools, guidelines and checklist for quality assessment of primary studies in SE are being built for the treat to the validity of the research findings while others are particular about the methodological characteristics. CASP [8] and Jadad et al. [9] are most popular and widely adapted in Medicine, sociology, and software engineering [3]. While the former came up with 10 questions and which not all are applicable to software engineering[3], the latter developed a scale that reflects the validity/bias issues and concerns only on three important items - whether the study is described as randomized, as double blind, or there is a description of withdrawals and dropouts.

CONSORT statement is another popular tool widely used to improve the quality of reports of parallel-group randomized trials. The statement contains flow diagram, and 22 checklist that cut across the content of the title, abstract, introduction, methods, results, and discussion sections of a paper. The flow diagram represents details from the four stages of a trial (intervention allocation, enrollment, follow-up, and analysis).

Sjoberg et al. [11] discussed measures to increase the quality of empirical studies in SE in general, while Kitchenham et al. [10] have proposed a set of more concrete guidelines to assist researchers in performing empirical studies. Based on this and CASP guidelines, Dybå and Dingsøyr [7] subsequently developed eleven criteria of which the revised ten are universal questions and could easily be applied in SE. Quite a number of other systematic reviewers now revise questions from original questions to formulate their own quality assessment checklists [3].

The Grade of Recommendation Assessment, Development and Evaluation (GRADE) Working Group has developed a system for grading the quality of evidence and strength of recommendations. This system has either been adopted or modified by several organizations including the British Medical Journal (BMJ), the World Health Organization (WHO), and Cochrane Collaboration for quality evaluating of reports in systematic reviews [6]. The GRADE approach classifies four grades of evidence as high, moderate, low, and very low. It defines the quality in a way to give an assurance that an estimate of effect or association is correct. When compared with CASP [8] and Dybå and Dingsøyr [7], the approach assesses the evidence more precisely even though, the process may be complicated and therefore better handled by experienced SLR researchers [3].

3 METHODOLOGY

Following the raising of research questions and comprehensive literatures search process, this study adapted the 11 criteria specified by Dybå and Dingsøyr [7] to assess the quality of the primary studies in the SLRs and employed a dichotomous scale (Yes=1; No=0) to grade each of the criteria. These criteria as listed in section 4.0 cover the following four main issues to be considered when assessing the quality of primary studies:

- **Reporting**: Criteria (1-3) are connected to the reporting quality of a study's rationale, aims, and context.
- **Rigor**: Criteria (4-8) have to do with the objectivity of the research methods used to establish the validity of data collection tools and the analysis methods, and hence the trustworthiness of the findings.
- **Credibility**: Criteria (9-10) are concerned with the assessment of the credibility of the study methods for ensuring that the findings were valid and meaningful.
- **Relevance**: The last criterion (11) is linked to the assessment of the relevance of the study for the software industry at large and the research community.

3.1 The Study Questions

The primary Research Question (RQ) is *"How reliable are the SLRs in software engineering domain"?* This question is broken down into the following secondary questions to address the study objectives

- RQ1: Are the SLRs research based on clear aims and description of the context?
- RQ2: How comprehensive are the research methodologies applied in the primary study?

- RQ3: Are the research findings valid?

3.2 The Search Process

The following digital libraries and broad indexes were auto-searched: Citeseer, IEE Computer society digital library, ACM, Web of Science, SpringerLink, EBSCO, Science Direct and Scopus. The search in [2] was limited to the first five but they are relevant to this study to validate our search process. Following the same process, the literature search were independently handled by two separate researchers and their notes were compared for reliability and completeness. All searches were limited to the document title and keywords as quality papers hardly keep silent on their main research subject using the following set of *search strings* (SS):

1. "software engineering" AND "review of studies"
2. "software engineering" AND "structured review"
3. "software engineering" AND "systematic review"
4. "software engineering" AND "literature review"
5. "software engineering" AND "literature analysis"
6. "software engineering" AND "in-depth survey"
7. "software engineering" AND "literature survey"
8. "software engineering" AND "meta-analysis"
9. "software engineering" AND "past studies"
10. "software engineering" AND "subject matter expert"
11. "software engineering" AND "Analysis of research"
12. "software engineering" AND "empirical body of knowledge"
13. "software engineering" AND "overview of existing knowledge"
14. "software engineering" AND "body of published research"
15. "Evidence-based software-engineering" OR "evidence-based software engineering"
16. "software engineering" AND "review"
17. "software engineering" AND "literature analysis"
18. "software engineering" AND "literature listing"

The first search was conducted for 2004-2007 using search string -SS1 to SS15 over the broad indexes excluding EBSCO and Science Direct as applied in [1], and SS16 as included in [2]. The output here help to validate this search process when compared with papers found in[1]. The search process found 16 papers as against 18 found in [1]. However, since "the two missed papers were border line and discuss more of computer science rather than software engineering" [2], we conclude that our search process is good and accurate.

The second search was conducted for 2012 to date (2017) using search string (SS) 1 to 18 over all the eight indexes. Searches on Citeseen, IEE and ACM were complicated and done over 18 times using the set of simple search strings as the databases do not allow easy construction of complex searches. However, searches over SCOPUS, SpringerLink, and Science Direct were straight forward and the following Complex searches were easily applied over one time period.

TITLE-KEY("software engineering" AND ("review of studies" OR "structured review" OR "systematic review" OR "literature review" OR "literature analysis" OR "in-depth survey" OR "literature survey" OR "meta-analysis" OR "past studies" OR "literature listing")) or TITLE--KEY("software engineering" AND ("literature analysis" OR "review" OR "evidence-based software engineering" OR "body of published research" OR "overview of existing knowledge" OR "empirical body of knowledge" OR "Analysis of research" OR" subject matter expert"))[All Sources(Computer Science, Engineering)

3.3 Study Selection

We integrated the result of the different searches and subject the 4,201 papers found to initial screening paying specific attention to abstract, title and keywords. Exporting the total list to Microsoft excel and sorting on the document title, the total papers was reduced to 2,319 as many have multiple indexes. Searching through the abstract column to *reject* papers that did not have at all [13] or include *literature review* or any other *search strings*(SS), the total were screened to 74. Our assumption is that any paper that fails to have sentence(s) on any of the search strings in its abstract submission cannot be a true representative of SLRs and will be improper to be recognised as such. Meanwhile, we removed[14] and[15]as they both have publication date of 2018 which is out of our date scope. Hence, we were left with 72 SLRs which were further subjected to a more detailed assessment:

Electronic copies of each SLRs were downloaded and two (two-member) teams of researchers were formed to independently screen each paper for possible inclusion. To avoid false rejection of relevant papers, notes were compared between the teams and discrepancies were resolved going by the following exclusion criteria:

- The paper is not a full flesh research submission

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- The topics are more of Computer sciences or Information systems and *not related* to Software engineering or did not discuss *literature review*
- The papers did not follow a defined search process.

This process rejected four papers that were not a full fleshed research submission but a mere abstract extension, template or PowerPoint presentation. A total number of 13 SLRs are irrelevant as they were either silent on literature review or discussed the subject but on unrelated domains like Computer sciences or Information systems. Of the refined total, four papers ([16],[17],[18],[19]) were dropped because they failed to follow a defined search process leaving us with a total of **53** SLRs subjected to data extraction process (see Fig. 1).

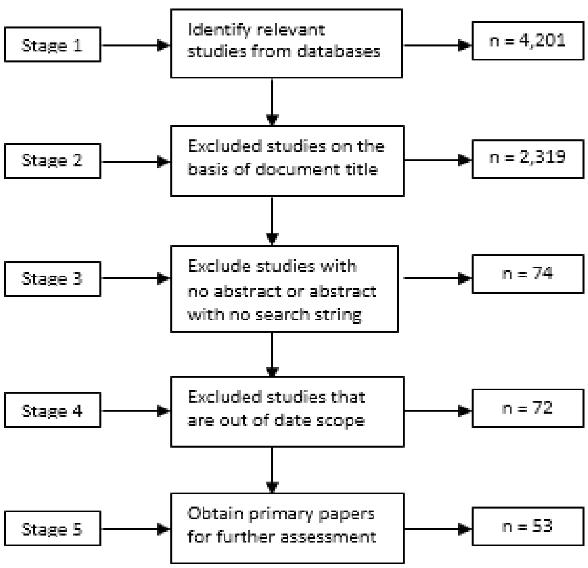

Fig. 1. Flow diagram of the study selection

3.4 Data Extraction Process

Data extraction follows the same procedure used in data selection. The 53 SLRs published between 2012 and 2017 are shown in Table 1. For each paper, we could identify:

- The year of publication (to show how current the SLRs)
- The document type (journal, conference, workshop, book chapter).
- The number of primary studies in software engineering (as explicitly stated or tabulated by the author).This is a quality assurance measure.

4 QALITY ASSESSMENT

The criteria postulated by Dybå and Dingsøyr [7]were used to evaluate the SLRs quality in this study. Unlike some other SLRs with five or fewer criteria and pay attention to only one or two aspects, the 11 criteria checklist covers the four main aspects of study quality – Reporting, Rigor, credibility and relevance. When adapted to this study, those citations identified as SLRs are independently assessed using the following criteria:

Q1 Is the SLR a research based or a mere "lessons learned" report?
Q2: Is there a clear statement of the aims of the research?
Q3: Is the context in which the research was carried out adequately described?
Q4: Was the research design appropriate to address the aims of the research?
Q5: Was the recruitment strategy appropriate to the aims of the research?
Q6: Was there a control group with which to compare treatments?
Q7: Was the data collected in a way that addressed the research issue?
Q8: Was the data analysis sufficiently rigorous?
Q9: Has the relationship between researcher and participants been adequately considered?
Q10: Is there a clear statement of findings?
Q11: Is the study of value for research or practice?

The Dichotomous scale of grading was used to score each question where

Yes (Y) = 1; FALSE (F) = 0.

The assessment follows a similar procedure used for data selection (fig. 1) where:

1. Two (2-member) teams of researchers are formed and the SLRs were equally shared amongst the teams.
2. Each member of each team independently answered the quality questions on each study giving adequate justification.
3. Each pair of researchers compared their results and form a position
4. The two teams compared their results, resolved issues and formed a consensus.
5. The quality result is considered and presented in Table 1.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1: Quality scores for the SLRs

Study No	Year	Document Type	No of 1^0 studies	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
[20]	2012	Review	-	1	1	1	1	1	1	1	1	0	1	1
[21]	2012	Conference	96	1	1	1	1	0	0	1	1	0	1	1
[22]	2012	Article	132	1	1	1	1	0	0	1	1	0	1	1
[23]	2012	Review	9	1	1	1	1	1	0	1	1	0	1	1
[24]	2012	Conference	-	1	1	1	1	0	1	1	1	0	1	1
[25]	2013	Conference	25	1	1	1	1	0	0	1	1	0	1	1
[26]	2013	Review	330	1	1	1	1	1	0	1	1	0	1	1
[27]	2013	Article	68	1	1	1	1	0	0	1	1	1	1	1
[28]	2012	Review	208	1	1	1	1	1	0	1	1	0	1	1
[29]	2013	Article	26	1	1	1	1	0	1	1	1	0	1	1
[30]	2012	Review	-	1	1	1	1	0	1	1	1	0	1	1
[31]	2013	Conference	-	1	1	1	1	0	1	1	1	0	1	1
[32]	2013	Article	-	1	1	1	1	1	0	1	1	0	1	1
[33]	2013	Conference	116	1	1	1	1	1	0	1	1	0	1	1
[34]	2013	Conference	23	1	1	1	1	0	1	1	1	0	1	1
[35]	2013	Conference	14	1	1	1	1	0	0	1	1	0	1	1
[36]	2013	Conference	-	1	1	1	1	0	0	1	1	0	1	1
[37]	2014	Article	693	1	1	1	1	1	1	1	1	0	1	1
[38]	2014	Conference	100	1	1	1	1	0	0	1	1	0	1	1
[39]	2015	Article	-	1	1	1	1	0	0	1	1	0	1	1
[40]	2015	Article	-	1	1	1	1	0	1	1	1	0	1	1
[41]	2015	Conference	-	1	1	1	1	0	1	1	1	0	1	1
[3]	2015	Conference	127	1	1	1	1	0	0	1	1	0	1	1
[42]	2015	Conference	43	1	1	1	1	0	0	1	1	0	1	1
[43]	2015	Conference	21	1	1	1	1	1	1	1	1	0	1	1
[44]	2015	Conference	36	1	1	1	1	0	0	1	1	1	1	1
[45]	2015	Conference	44	1	1	1	1	1	1	1	1	0	1	1
[46]	2016	Review	-	1	1	1	1	0	0	1	1	0	1	1
[47]	2015	Conference	-	1	1	1	1	0	0	1	1	0	1	1
[48]	2015	Review	12	1	1	1	1	0	0	1	1	0	1	1
[49]	2015	Conference	15	1	1	1	1	0	0	1	1	0	1	1
[50]	2016	Conference	-	1	1	1	1	0	1	1	1	0	1	1
[51]	2016	Review	350	1	1	1	1	1	0	1	1	0	1	1
[52]	2015	Conference	82	1	1	1	1	0	1	1	1	0	1	1
[53]	2016	Conference	-	1	1	1	1	0	0	1	1	0	1	1
[54]	2016	Conference	-	1	1	1	1	0	1	1	1	0	1	1
[55]	2016	Conference	-	1	1	1	1	1	0	1	1	0	1	1
[56]	2016	Conference	-	1	1	1	1	0	0	1	1	0	1	1
[57]	2016	Conference	15	1	1	1	1	0	0	1	1	0	1	1
[58]	2016	Review	-	1	1	1	1	0	1	1	1	0	1	1
[59]	2016	Conference	52	1	1	1	1	0	1	1	1	0	1	1
[60]	2017	Conference	33	1	1	1	1	1	0	1	1	0	1	1
[61]	2017	Conference	-	1	1	1	1	0	0	1	1	0	1	1
[62]	2017	Article	78	1	1	1	1	0	0	1	1	0	1	1
[63]	2017	Conference	-	1	1	1	1	0	0	1	1	0	1	1
[64]	2017	Article	-	1	1	1	1	0	0	1	1	1	1	1
[65]	2017	Conference	316	1	1	1	1	1	1	1	1	0	1	1
[66]	2017	Conference	-	1	1	1	1	0	0	1	1	0	1	1
[67]	2017	Conference	1270	1	1	1	1	1	0	1	1	0	1	1
[5]	2017	Conference	2	1	1	1	1	0	1	1	1	0	1	1
[68]	2017	Conference	-	1	1	1	1	0	0	1	1	0	1	1
[69]	2017	Editorial	-	1	1	1	1	1	0	1	1	0	1	1
[70]	2017	Conference	-	1	1	1	1	1	1	1	1	0	1	1
TOTAL				53	53	49	53	15	19	50	52	3	52	50

Journal of Theoretical and Applied Information Technology
30th September 2018. Vol.96. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5 DQISCUSSION OF RESEARCH QUESTIONS

This section addresses our specific study questions to ascertain that the primary question has been answered and the quality of the SLRs has been established.The summary results of the quality assessment for the 53 SLRs are depicted in Table 2. We fix the pass mark at 50%.

Table 2: Qualitative assessement of 53 SLRs on SE

SL Rs	Quality scores (QS)	% Scores (QS/53)	RQ Addressed	Related Quality Conditions (QC)	QC [7] Met?
Q1	53	100			
Q2	53	100	RQ1	Reporting	Yes, by 97.50%
Q3	49	92.5			
Q4	53	100			
Q5	15	28.3			
Q6	19	35.9	RQ2	Rigour	Yes, by 71.32%
Q7	50	94.3			
Q8	52	98.1			
Q9	3	5.7	RQ3	Credibility	Yes, by 51.90%
Q10	52	98.1			
Q11	50	94.3	Nil	Relevance	

5.1: RQ1: Are the SLRs Research Based on Clear Aims and Description of the Context?

Table 2 confirms this question to be positive as RQ1 addressed by Q1, Q2 and Q3 is average scored 97.5%. All studies were rated OK except four that failed on Q3 because they did not justify their research design as the authors failed to discuss how they arrived at the choice of their research method.The 49 representing 97.5% of the selected samples clarified the goal of their research and their relevance. The studies (attempted to) exemplify or elucidate subjective actions of the research participants, employed qualitative approaches which was the right methodology to address the research goals.

5.2: RQ2: How Comprehensive are the Research Methodologies Applied in the Primary Study?

Yes, the research methodologies in the various primary studies were comprehensive considering the pass mark of 71.32% scored by "Rigour" on Table 2. RQ2 is being assessed by quality scores of Q4 – Q8 and confirmed by average status of "Rigour".

Obviously, 38 of the 53 primary studies did not have suitable recruitment strategy for their research aims, and 34 are silent on the control group (s) with wish their methods are (can be) compared. Notwithstanding, over 94% gave extensive description of their analysis processes and present sufficient data to support their findings (Q7 and Q8); they were comprehensive on their data collection and analysis procedures.

5.3: RQ3: Are the Research Findings Valid?

This question is affirmative going by the research "credibility" status of 51% even though, the performance on Q9 was poor. The researchers failed to examine their roles and influence in the formulation of their research questions, data collection, sample selection and choice of location. This is true for 50 primary studies as the possibility of researcher bias was mentioned in only three studies. Hence, the relationship between the researcher and the participants were not adequately considered leading to a very low quality score of three in Q9 (Table 2).

However, all the studies except one, have clear statements of findings as the findings are unambiguous. The credibility of findings and researchers arguments were also adequately discussed in relation to the study questions.

6 STUDY LIMITATION

Since identifying relevant studies is a major challenge in SLRs, we extended our search to seven quality sources in this study. Our search was based on our search strings (SS) but limited to document title and keywords. Notwithstanding, we may have missed some qualified papers that may have been cited outside our selected databases. There is also a probability that some studies that used other terminologies to describe their literature review have been missed out.

In our attempts to reduce our sample size to manageable number, we categorised papers whose topics are more of Computer sciences or Information systems as irrelevant and excluded them from our list. It is however possible that we have missed some SLRs that fall in-between information technology, software engineering and computer sciences.

Our search also make assumptions that quality SLRs will either be in journal or conference proceedings. Hence, we did not search for technical reports, books, study thesis and dissertations because of publication bias where authors fail to publish negative results. However, this may not necessarily be an issue in software engineering as some negative studies

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

are already being published [2] [12].

7 CONCLUSION

Although the SLRs failed in three quality questions, they have quality scores above 92% in the rest eight questions and passed all the four quality conditions (Table 2) thereby confirming the good quality of the primary studies. However, the researchers should ensure the recruitment strategies are properly aligned to research aim and objectives by giving clear explanations and discussion of how and why the participants were selected.

We acknowledge that ensuring completeness of auto search remains an important issue. We hereby recommend for future studies that as many quality sources as possible should be searched for qualified SLRs, and authors should expand the study search strings to include other terminologies such as "research aggregation" and "study/research synthesis" which are now being used by some studies.

ACKNOWLEDGMENT

This study was supported by Covenant University, Ota, Nigeria

REFERENCES

[1] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering - A systematic literature review," *Inf. Softw. Technol.*, vol. 51, no. 1, 2009.

[2] B. Kitchenham *et al.*, "Systematic literature reviews in software engineering-A tertiary study," *Inf. Softw. Technol.*, vol. 52, no. 8, 2010.

[3] Y. Zhou, H. Zhang, X. Huang, S. Yang, M. A. Babar, and H. Tang, "Quality assessment of systematic reviews in software engineering: A tertiary study," in *ACM International Conference Proceeding Series*, 2015.

[4] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," *Engineering*, vol. 2, p. 1051, 2007.

[5] V. Garousi and M. Felderer, "Experience-based guidelines for effective and efficient data extraction in systematic reviews in software engineering," in *ACM International Conference Proceeding Series*, 2017.

[6] J. Higgins and S. Green, "Chapter 22: Overview of reviews. Cochrane handbook for systematic reviews of interventions," *Cochrane Database Syst. Rev.*, pp. 187–235, 2008.

[7] T. Dybå and T. Dingsøyr, "Strength of Evidence in Systematic Reviews in Software Engineering," in *Proceedings of the Second ACM-IEEE international symposium on Empirical software engineering and measurement*, 2008, no. 7465, pp. 178–187.

[8] P.H.R.U., "Critical Appraisal skills programme (CASP)," no. 2017, pp. 1–5, 1994.

[9] A. R. Jadad *et al.*, "Assessing the quality of reports of randomized clinical trials: Is blinding necessary?," *Control. Clin. Trials*, vol. 17, no. 1, pp. 1–12, 1996.

[10] B. A. Kitchenham *et al.*, "Preliminary Guidelines for Empirical Research in Software Engineering," *Main*, vol. 28, no. 8, pp. 721–734, 2002.

[11] D. I. K. Sjoberg, T. Dyba, and M. Jorgensen, "The Future of Empirical Methods in Software Engineering Research," *Futur. Softw. Eng. (FOSE '07)*, no. JUNE 2007, pp. 358–378, 2007.

[12] O. J. Okesola, K. Okokpuji, P. R. O. Oyom, O. Kalesanwo, and O. Awodele, "Structuring Challenges in Requirement Engineering Techniques," in *International MultiConference of Engineers and Computer Scientists, IMECS2018, London*, 2018.

[13] S. Anderson, P. Sagar, B. Smith, and K. Wallace, "What we have learnt adopting evidence-based software engineering for industrial practice," in *ACM International Conference Proceeding Series*, 2016, vol. 01–03–June.

[14] A. Amin, S. Basri, M. F. Hassan, and M. Rehman, *A snapshot of 26 years of research on creativity in software engineering - A systematic literature review*, vol. 425. 2018.

[15] J. C. Guzmán, G. López, and A. Pacheco, *Defining "Architecture" for software engineering – a review of terminology*, vol. 598. 2018.

[16] L. Kumar, S. Sripada, and A. Sureka, "A review of six years of Asia-pacific software engineering conference," in *Proceedings - Asia-Pacific Software Engineering Conference, APSEC*, 2017.

[17] T. E. Colanzi, S. R. Vergilio, W. K. Guez Assunção, and A. Pozo, "Search Based Software Engineering: Review and analysis of the field in Brazil," *J. Syst. Softw.*, vol. 86, no. 4, 2013.

[18] Y. Zhang, J. Dong, C. Li, Z. Ji, and J. Wu, *Review on innovative practice teaching and its quality evaluation system for software*

engineering specialty, vol. 163 AISC. 2012.
[19] S. Koushik and R. Selvarani, *Review on cost effective software engineering using program slicing techniques*, vol. 132 AISC. 2012.
[20] C. Pons, R. Giandini, and G. Arévalo, "A systematic review of applying modern software engineering techniques to developing robotic systems | Revisión sistemática de la aplicación de técnicas modernas de ingeniería de software al desarrollo de sistemas robóticos," *Ing. e Investig.*, vol. 32, no. 1, 2012.
[21] B. Penzenstadler, V. Bauer, C. Calero, and X. Franch, "Sustainability in software engineering: A systematic literature review," in *IET Seminar Digest*, 2012, vol. 2012, no. 1.
[22] J. Portillo-Rodríguez, A. Vizcaíno, M. Piattini, and S. Beecham, "Tools used in Global Software Engineering: A systematic mapping review," *Inf. Softw. Technol.*, vol. 54, no. 7, 2012.
[23] L. Rodríguez-Martínez, M. Mora, F. Álvarez, L. Garza, H. Durán, and J. Muñoz, "Review of relevant system development life cycles (SDLCs) in service-oriented software engineering (SoSE)," *J. Appl. Res. Technol.*, vol. 10, no. 2, 2012.
[24] R. Matinnejad and R. Ramsin, "An analytical review of process-centered software engineering environments," in *Proceedings - 2012 IEEE 19th International Conference and Workshops on Engineering of Computer-Based Systems, ECBS 2012*, 2012.
[25] M. Freire *et al.*, "Automated support for controlled experiments in software engineering: A systematic review," in *Proceedings of the International Conference on Software Engineering and Knowledge Engineering, SEKE*, 2013, vol. 2013–Janua, no. January.
[26] S. Schneider, R. Torkar, and T. Gorschek, "Solutions in global software engineering: A systematic literature review," *Int. J. Inf. Manage.*, vol. 33, no. 1, 2013.
[27] B. Kitchenham and P. Brereton, "A systematic review of systematic review process research in software engineering," *Inf. Softw. Technol.*, vol. 55, no. 12, 2013.
[28] T. Hall, S. Beecham, D. Bowes, D. Gray, and S. Counsell, "A systematic literature review on fault prediction performance in software engineering," *IEEE Trans. Softw. Eng.*, vol. 38, no. 6, 2012.
[29] A. Kasoju, K. Petersen, and M. V. Mäntylä, "Analyzing an automotive testing process with evidence-based software engineering," *Inf. Softw. Technol.*, vol. 55, no. 7, 2013.
[30] S. Jalali and C. Wohlin, "Global software engineering and agile practices: A systematic review," *J. Softw. Evol. Process*, vol. 24, no. 6, 2012.
[31] H. Zhang and M. Ali Babar, "Systematic reviews in software engineering: An empirical investigation," *Inf. Softw. Technol.*, vol. 55, no. 7, 2013.
[32] A. Sharma and F. Maurer, *A roadmap for software engineering for the cloud: Results of a systematic review*, vol. 1–4. 2013.
[33] S. Imtiaz, M. Bano, N. Ikram, and M. Niazi, "A tertiary study: Experiences of conducting systematic literature reviews in software engineering," in *ACM International Conference Proceeding Series*, 2013.
[34] M. F. Bosu and S. G. MacDonell, "Data quality in empirical software engineering: A targeted review," in *ACM International Conference Proceeding Series*, 2013.
[35] C. Marshall and P. Brereton, "Tools to support systematic literature reviews in software engineering: A mapping study," in *International Symposium on Empirical Software Engineering and Measurement*, 2013.
[36] A. S. Sayyad and H. Ammar, "Pareto-optimal search-based software engineering (POSBSE): A literature survey," in *2013 2nd International Workshop on Realizing Artificial Intelligence Synergies in Software Engineering, RAISE 2013 - Proceedings*, 2013.
[37] M. M. Ahmed and S. Letchmunan, "A systematic literature review on challenges in service oriented software engineering," *Int. J. Softw. Eng. its Appl.*, vol. 9, no. 6, 2015.
[38] J. R. M. Viana, N. P. Viana, F. A. M. Trinta, and W. V. De Carvalho, "A Systematic Review on Software Engineering in Pervasive Games Development," in *Brazilian Symposium on Games and Digital Entertainment, SBGAMES*, 2014, vol. 2014–Decem, no. December.
[39] B. Rizvi, E. Bagheri, and D. Gasevic, "A systematic review of distributed Agile software engineering," *J. Softw. Evol. Process*, vol. 27, no. 10, 2015.
[40] P. Lenberg, R. Feldt, and L. G. Wallgren, "Behavioral software engineering: A definition and systematic literature review," *J. Syst. Softw.*, vol. 107, 2015.
[41] F. J. Ekaputra, M. Sabou, E. Serral, and S. Biffl, "Collaborative exchange of systematic literature review results: The case of empirical software engineering," in *WWW 2015 Companion - Proceedings of the 24th*

International Conference on World Wide Web, 2015.

[42] D. Heaton and J. C. Carver, "Claims about the use of software engineering practices in science: A systematic literature review," *Inf. Softw. Technol.*, vol. 67, 2015.

[43] C. Marshall, P. Brereton, and B. Kitchenham, "Tools to support systematic reviews in software engineering: A cross-domain survey using semi-structured interviews," in *ACM International Conference Proceeding Series*, 2015, vol. 27–29–Apri.

[44] Z. Sharafi, Z. Soh, and Y.-G. Guéhéneuc, "A systematic literature review on the usage of eye-tracking in software engineering," *Inf. Softw. Technol.*, vol. 67, 2015.

[45] G. Liu, G. Rong, H. Zhang, and Q. Shan, "The adoption of capture-recapture in software engineering: A systematic literature review," in *ACM International Conference Proceeding Series*, 2015, vol. 27–29–Apri.

[46] M. Gasparic and A. Janes, "What recommendation systems for software engineering recommend: A systematic literature review," *J. Syst. Softw.*, vol. 113, 2016.

[47] T. Clear, "'Follow the moon' development: Writing a systematic literature review on Global Software Engineering Education," in *ACM International Conference Proceeding Series*, 2015, vol. 19-22-Nov-.

[48] S. Sepúlveda Cuevas and A. C. Leal, "Protocol adaptations to conduct systematic literature reviews in software engineering: A chronological study | Adaptaciones al protocolo para realizar revisiones sistemáticas de literatura en ingeniería de software: Un estudio cronológico," *Ing. e Investig.*, vol. 35, no. 3, 2015.

[49] A. C. S. Dutra, R. Prikladnicki, and C. Franca, "What Do We Know about High Performance Teams in Software Engineering? Results from a Systematic Literature Review," in *Proceedings - 41st Euromicro Conference on Software Engineering and Advanced Applications, SEAA 2015*, 2015.

[50] R. Colomo-Palacios, L. O. Colombo-Mendoza, and R. Valencia-García, *Towards supporting international standard-based software engineering approaches using semantic web technologies: A systematic literature review*, vol. 658. 2016.

[51] M. Kosa, M. Yilmaz, R. V. O'Connor, and P. M. Clarke, "Software engineering education and games: A systematic literature review," *J. Univers. Comput. Sci.*, vol. 22, no. 12, 2016.

[52] T. Clear *et al.*, "Challenges and recommendations for the design and conduct of global software engineering courses: A systematic review," in *ITiCSE-WGP 2015 - Proceedings of the 2015 ITiCSE Conference on Working Group Reports*, 2015.

[53] P. Gupta, I. Arora, and A. Saha, "A review of applications of search based software engineering techniques in last decade," in *2016 5th International Conference on Reliability, Infocom Technologies and Optimization, ICRITO 2016: Trends and Future Directions*, 2016.

[54] K. R. Felizardo, E. Mendes, M. Kalinowski, E. F. Souza, and N. L. Vijaykumar, "Using Forward Snowballing to update Systematic Reviews in Software Engineering," in *International Symposium on Empirical Software Engineering and Measurement*, 2016, vol. 08–09–Sept.

[55] B. Cartaxo, "Integrating evidence from systematic reviews with software engineering practice through evidence briefings," in *ACM International Conference Proceeding Series*, 2016, vol. 01–03–June.

[56] V. Garousi, M. Felderer, and M. V. Mäntylä, "The need for multivocal literature reviews in software engineering: Complementing systematic literature reviews with grey literature," in *ACM International Conference Proceeding Series*, 2016, vol. 01–03–June.

[57] J. S. Molléri, K. Petersen, and E. Mendes, "Survey Guidelines in Software Engineering: An Annotated Review," in *International Symposium on Empirical Software Engineering and Measurement*, 2016, vol. 08–09–Sept.

[58] V. Garousi, K. Petersen, and B. Ozkan, "Challenges and best practices in industry-academia collaborations in software engineering: A systematic literature review," *Inf. Softw. Technol.*, vol. 79, 2016.

[59] K.-J. Stol, P. Ralph, and B. Fitzgerald, "Grounded theory in software engineering research: A critical review and guidelines," in *Proceedings - International Conference on Software Engineering*, 2016, vol. 14-22-May-.

[60] D. Pflüger *et al.*, "The scalability-efficiency/maintainability-portability trade-off in simulation software engineering: Examples and a preliminary systematic literature review," in *Proceedings of SE-HPCCSE 2016: 4th International Workshop on Software Engineering or High Performance Computing*

in Computational Science and Engineering - Held in conjunction with SC 2016: The International Conference for High Performance Computing, Netwo, 2017.

[61] B. Cartaxo, "Supporting researchers to transfer knowledge from systematic reviews to software engineering practice," in *CIbSE 2017 - XX Ibero-American Conference on Software Engineering*, 2017.

[62] R. Malhotra, M. Khanna, and R. R. Raje, "On the application of search-based techniques for software engineering predictive modeling: A systematic review and future directions," *Swarm Evol. Comput.*, vol. 32, 2017.

[63] A. S. Barroso, J. S. M. Da Silva, M. S. Soares, and R. P. C. Do Nascimento, "Influence of human Personality in software engineering a systematic literature review," in *ICEIS 2017 - Proceedings of the 19th International Conference on Enterprise Information Systems*, 2017, vol. 3.

[64] L. Hernández, M. Muñoz, J. Mejía, A. Peña, N. Rangel, and C. Torres, "A systematic literature review focused on the use of gamification in software engineering teamworks | Una revisión sistemática de la literatura enfocada en el uso de gamificación en equipos de trabajo en la ingeniería de software," *RISTI - Rev. Iber. Sist. e Tecnol. Inf.*, no. 21, 2017.

[65] X. Zhou, Y. Jin, H. Zhang, S. Li, and X. Huang, "A map of threats to validity of systematic literature reviews in software engineering," in *Proceedings - Asia-Pacific Software Engineering Conference, APSEC*, 2017.

[66] A. Bacchelli and M. Beller, "Double-blind review in software engineering venues: The community's perspective," in *Proceedings - 2017 IEEE/ACM 39th International Conference on Software Engineering*

[67] M. Kuutila, M. V. Mantyla, M. Claes, and M. Elovainio, "Reviewing literature on time pressure in software engineering and related professions: Computer assisted interdisciplinary literature review," in *Proceedings - 2017 IEEE/ACM 2nd International Workshop on Emotion Awareness in Software Engineering, SEmotion 2017*, 2017.

[68] S. Z. Qasim and M. A. Ismail, "Research problems in Search-Based Software Engineering for many-objective optimization: A literature survey," in *ICIEECT 2017 - International Conference on Innovations in Electrical Engineering and Computational Technologies 2017, Proceedings*, 2017.

[69] S. Beecham, D. Bowes, and K.-J. Stol, "Introduction to the EASE 2016 special section: Evidence-based software engineering: Past, present, and future," *Inf. Softw. Technol.*, vol. 89, 2017.

[70] R. Manjula, A. Patil, and R. Shingade, "A review on Service-Oriented Software Engineering (SOSE)," in *Proceedings of 2016 Online International Conference on Green Engineering and Technologies, IC-GET 2016*, 2017.