Information and Software Technology 85 (2017) 60–70

Contents lists available at ScienceDirect

Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof

Systematic literature reviews in agile software development: A tertiary study

Rashina Hoda[a,*], Norsaremah Salleh[b], John Grundy[c], Hui Mien Tee[a]

[a] SEPTA Research, Electrical and Computer Engineering, The University of Auckland, Building 903, 368 Khyber Pass, New Market, Auckland 1023, New Zealand
[b] Department of Computer Science, International Islamic University Malaysia, Kuala Lumpur, Malaysia
[c] School of Information Technology, Deakin University, Melbourne, Australia

ARTICLE INFO

Article history:
Received 15 May 2016
Revised 15 January 2017
Accepted 16 January 2017
Available online 17 January 2017

Keywords:
Agile software development
Tertiary study
Systematic literature reviews
Mapping study

ABSTRACT

Context: A number of systematic literature reviews and mapping studies (SLRs) covering numerous primary research studies on various aspects of agile software development (ASD) exist.
Objective: The aim of this paper is to provide an overview of the SLRs on ASD research topics for software engineering researchers and practitioners.
Method: We followed the tertiary study guidelines by Kitchenham et al. to find SLRs published between late 1990s to December 2015.
Results: We found 28 SLRs focusing on ten different ASD research areas: adoption, methods, practices, human and social aspects, CMMI, usability, global software engineering (GSE), organizational agility, embedded systems, and software product line engineering. The number of SLRs on ASD topics, similar to those on software engineering (SE) topics in general, is on the rise. A majority of the SLRs applied standardized guidelines and the quality of these SLRs on ASD topics was found to be slightly higher for journal publications than for conferences. While some individuals and institutions seem to lead this area, the spread of authors and institutions is wide. With respect to prior review recommendations, significant progress was noticed in the area of connecting agile to established domains such as usability, CMMI, and GSE; and considerable progress was observed in focusing on management-oriented approaches as Scrum and sustaining ASD in different contexts such as embedded systems.
Conclusion: SLRs of ASD studies are on the rise and cover a variety of ASD aspects, ranging from early adoption issues to newer applications of ASD such as in product line engineering. ASD research can benefit from further primary and secondary studies on evaluating benefits and challenges of ASD methods, agile hybrids in large-scale setups, sustainability, motivation, teamwork, and project management; as well as a fresh review of empirical studies in ASD to cover the period post 2008.

1. Introduction

Agile software development (ASD) was formally introduced to the software engineering community in 2001 through a set of four core values and twelve principles, laid out in the "Agile Manifesto" [1,42]. Agility, or the ability to rapidly adapt to volatile requirements, is a cornerstone of ASD. This stands in stark contrast to the plan-driven approach prescribed by more traditional models of software development, such as the waterfall model [2]. Other distinguishing features of ASD include enhanced focus on human and social aspects of software engineering [3,7,11–13,15–16,42]; increased collaboration between business customers and software teams [3,21]; and a strong emphasis on frequent delivery of business value [4,5].

Since its inception about two decades ago, ASD has rapidly become a mainstream software development model in use today [30] through industrial adoption of several of its concrete manifestations, such as Scrum [4] and eXtreme Programming [5]. A similar dramatic impact has been witnessed in the research community with publications of copious numbers of primary research studies on ASD topics. There have also been a considerable number of secondary studies in the form of literature reviews and mappings published within this domain over the past decade or so [7,33]. Among these are early ASD reviews from 2003–2004, which focused on categorizing [31] and describing agile methods [34]; studies showing the limitations of agile research [33]; those formulating a theoretical definition of agility [29]; and presenting the

* Corresponding author.
E-mail address: r.hoda@auckland.ac.nz (R. Hoda).

http://dx.doi.org/10.1016/j.infsof.2017.01.007
0950-5849/

current state of agile research at regular intervals: e.g. 2008 [7,33], 2012 [27], and 2014 [26]. A number of other reviews on specific aspects of ASD, such as the role of communication in agile teams [44], test-driven and feature-driven development [41], agile approaches to user centered-design [40], and ASD in global contexts [39], have also been conducted over the past decade.

While these reviews have covered various disparate aspects of ASD over this time, there has been no tertiary study thus far that identifies and catalogues individual SLRs in this significant research field. To achieve this aim, we conducted a tertiary study of ASD research. A tertiary study is a study that involves a review of existing secondary studies (such as systematic literature reviews and systematic mappings) that is expected to answer wider research questions and uses the same methodology as the systematic literature review [9]. An exemplar of a tertiary study is one by Kitchenham et al. [10] who identified and cataloged SLRs in software engineering published between 2004–2008 and compared the results to their previous study to comment on the number and quality of the SLRs in software engineering. We conducted a similar exercise but for the specific area of ASD research, identifying and cataloging SLRs in ASD. In order to maintain high quality and research rigor [7], only those reviews conducting a systematic literature review or systematic mapping of empirical studies (collectively referred to as SLRs) were considered in this tertiary study. As a result, reviews as such [28,30,31] were not included since they were not based on empirical evidence, rather on compiled practitioner comments, industrial surveys and fundamental agile principles respectively.

Applying the guidelines by Kitchenham and Charters [9], we found 28 SLRs focusing on ten different ASD research areas: adoption, methods, practices, human and social aspects, CMMI, usability, global software engineering, organization, embedded systems, and product line engineering. We also analyzed the number and quality of these SLRs.

The rest of the paper is structured as follows: Section 2 presents the review Method including the research questions, search and selection process, inclusion and exclusion criteria, quality criteria, and data extraction. Section 3 presents the Data Extraction Results while the Discussion of the research questions is presented in Section 4. This is followed by the Study Limitations in Section 5 and paper concludes in Section 6. Much of the structure of this paper follows the tertiary study by Kitchenham et al. [10] as we found it to be particularly effective in conveying the process and results from a tertiary study.

2. Method

We conducted a tertiary study using the guidelines provided by Kitchenham and Charters [9]. Some examples of tertiary studies include [8,10,20]. We formulated the research questions based on the guidelines [9] and an exemplar tertiary study by Kitchenham et al. [10].

2.1. The research questions

RQ1 How many SLRs were published since the inception of ASD (late 1990s) to date (2015)?
RQ2 What research areas and topics are being addressed in ASD?
RQ3 Which individuals, organizations, and publication venues are most active in SLR-based research in ASD?
RQ4 What is the quality of the SLRs in ASD?
RQ5 What progress has been achieved with respect to prior recommendations for ASD?

RQs 1–4 follow the research questions addressed by Kitchenham et al. [10] in their tertiary study of SLRs in software engineering in general. Since no prior tertiary studies of ASD exist, RQ5 involved comparing the recommendations from prior (secondary) review studies that we were aware of and that closely related to our topic, described further in Section 4.5.

2.2. The search process

The search process involved the use of the standard online databases that are used for software engineering SLRs i.e. that index the software engineering and computer science related literature. These include the digital libraries IEEE Xplore, ACM, Springerlink, Science Direct, and ISI Web of Science. As previous software engineering SLRs have shown, these provide excellent coverage for the available software engineering literature.

We searched for all relevant SLR papers published up to and including December 2015 from these online databases. The search string format we used was necessarily modified to fit the search requirements of different digital databases e.g. IEEE Xplore, Web of Science and Springer require slightly modified strings due to the query terms they use. The search phrase was developed to answer our research questions and the string we eventually used was:

"agile" AND ("systematic review" OR "systematic literature review" OR "systematic map" OR "systematic mapping" OR "mapping study")

2.3. Study selection

The selection of studies was conducted by applying a set of inclusion and exclusion criteria as we wanted to include as many formal or systematic reviews and mapping on the subject as possible (collectively referred to as SLRs here). The inclusion criteria were:

(a) The SLRs must be directly related to agile topics; and
(b) The studies must be conducted following SLR guidelines.

Our aim to provide an overview of industrial ASD research guided the formulation of the major exclusion criteria, which comprised of:

(a) SLRs that focused on ASD in academic settings; or
(b) SLRs related to Lean practices but not linking this to software development, e.g. application of Lean principles in the construction industry; or
(c) Review works that appeared as abstracts, work-in-progress papers, posters, short papers (less than or equal to 6 pages) and papers not written in English.

The paper selection process is shown in Fig. 1. A total of 1083 papers appeared from our initial searches in the specified databases.

After applying the inclusion and exclusion criteria at the title and abstract level, a total of 125 papers remained. The next filtration was applied at the content level with the inclusion/exclusion criteria being used to remove (i) duplicate studies and (ii) papers that do not relate to agile topics or (iii) did not actually conduct an SLR of agile studies. After these exclusions, this resulted in 31 studies. Finally, the last applied was a quality filter (described in the next section) which resulted in the removal of two (2) papers and left a final set of 29 SLR papers for meta-analysis. Out of 29 studies, we found two SLRs presenting the same data-set (the earlier was published as a mapping study and the later as a journal paper). We included the recent and most comprehensive paper to avoid double counting. This resulted to the inclusion of 28 papers for our analysis.

Table 1
Quality ranking criteria [10].

	Yes (1.0 score)	Partial (0.5 score)	No (0 score)
Q1	Inclusion criteria defined explicitly	Inclusion criteria are implicit	Inclusion criteria not defined and cannot be inferred
Q2	Four or more academic and reputable online databases searched	3–4 online databases searched	Two or fewer online databases searched
Q3	Quality criteria defined explicitly	Quality criteria assessed but not defined	No effort has been made to perform quality assessment
Q4	Information presented clearly and can be traced back to individual studies	Information presented in groups/category, but not linked back to individual studies	Information presented was not referenced

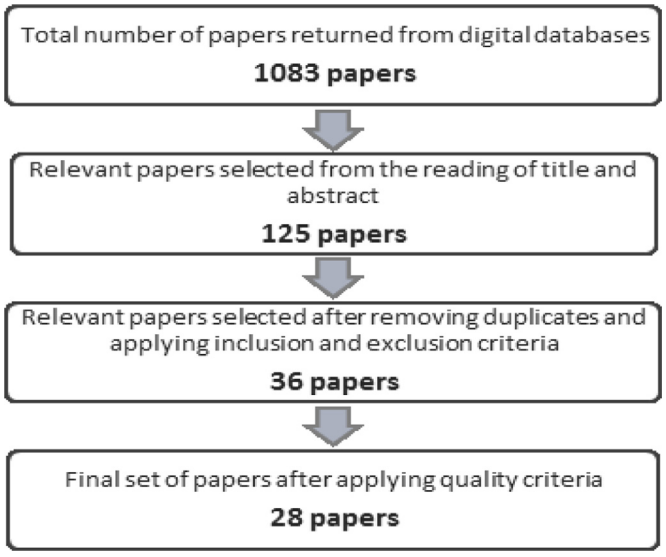

Fig. 1. Flow of the paper selection process.

The paper selection process was jointly discussed and agreed to by all researchers. The data extraction process was primarily carried out by three of the researchers. In order to minimize potential biases introduced into the data, any doubt and uncertainties were put up for discussion between all researchers until a consensus agreement was obtained.

2.4. Quality assessment

The quality of the SLRs was evaluated following the quality criteria defined in the tertiary study conducted by Kitchenham et al. [10]. Four quality criteria based on the DARE criteria [48] in the form of questions were used to assess the quality of each SLR [10]:

- Q1. Are the review's inclusion and exclusion criteria described and appropriate?
- Q2. Is the literature search likely to have covered all relevant studies?
- Q3. Did the reviewers assess the quality/validity of the included studies?
- Q4. Were the basic data/studies adequately described?

All of the SLRs were scored based on how well they satisfied the quality criteria described in Table 1. The following points system was used to determine the individual criteria scores: Yes (Y) = 1 point, Partial (P) = 0.5 point, No (N) = 0 point. The overall quality score was calculated by summing up the four individual criteria scores. Thus, the total quality score for each study ranged between 0 (very poor) and 4 (very good). The quality scores of the SLRs are presented in the next sub-section (Fig. 2) and details of the exclusion based on the quality scores are described in Section 3.

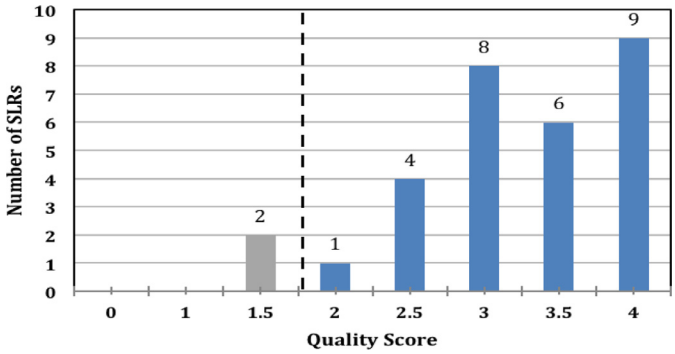

Fig. 2. Quality scores of SLRs (2 excluded studies shown before cut-off line).

2.5. Data extraction and analysis process

The data extraction was conducted using a structured extraction form in Microsoft Excel to capture all information required for further synthesis. The following data were extracted from all included SLRs:

- Bibliographic information (title, abstract, publication year, publication type: conference/journal)
- Keywords
- Number of primary studies
- Research Question(s)
- Search term used
- Online databases searched
- Years covered by the secondary study
- ASD research topic/area(s) discussed
- Agile method(s) discussed
- SLR guidelines used
- Summary of findings
- Quality score

The data was analyzed with a view to answer the research questions using descriptive analysis for the quantitative data and using thematic analysis approach [45] for the qualitative data. For example, answering RQ1 involved descriptive analysis of the number of SLRs. For answering RQ2, a categorization process was carried out whereby the topics covered by the SLRs were grouped into similar themes or areas. Further details of this process are described in Section 4.2.

3. Data extraction results

The 28 SLRs that were published in the time period 2008 to 2015 are shown in Table 2. For each of the SLRs, we identified: the total quality score, year of publication, type of publication (journal/conference), number of primary studies covered by the review, years covered by the SLR, SLR guidelines cited, and review topic. Based on the review topic, the SLRs were grouped into similar categories represented by the last column: research area. These reviews represent primary studies covering a period of 24 years from

Table 2
Agile software development SLRs published between 2008 and 2015.

SLR#	Quality total score	Publication Year	Paper type	#Primary studies covered	#Years Covered	SLR guidelines cited[a]	Review topic	Research area
[S1]	3	2012	Journal	81	1999–2009	KC	Global software engineering	Global software engineering
[S2]	3	2012	Journal	42	2002–2009	KC, Bo	Reconciling software models	Methods
[S3]	3	2012	Conference	12	1980–2011	K	Developers motivation	Human and social aspects
[S4]	3	2012	Conference	23	2002–2010	KC	Embedded systems	Embedded systems
[S5]	4	2011	Journal	39	2003–2010	K	Product line engineering	Product line engineering
[S6]	3	2015	Conference	46	2005–2014	Pe	User-centered design	Usability
[S7]	3.5	2011	Conference	48	2002–2010	KC	Test-driven development	Practices
[S8]	4	2010	Conference	46	1998–2010	KC, HG	Quality of agile practices	Practices
[S9]	4	2009	Conference	20	2003–2009	KC	Global software development	Global software engineering
[S10]	3	2011	Conference	58	2002–2010	KC, Bo	User-centered design	Usability
[S11]	2.5	2014	Conference	34	2001–2013	KC	Maturity model	CMMI
[S12]	3	2013	Journal	333	2000–2013	WW	Communication	Human and social aspects
[S13]	4	2008	Journal	36	2001–2005	KC, HG	Methods evaluation	Methods
[S14]	3.5	2010	Conference	38	2000–2009	KC	Adoption	Adoption
[S15]	4	2015	Journal	83	2002–2012	KC	User-centered design	Usability
[S16]	3.5	2015	Journal	30	2002–2013	PR	Using metrics	Practices
[S17]	4	2015	Journal	21	2006–2014	KC	Geographically distributed communication	Global software engineering
[S18]	2	2008	Conference	24	1991–2007	WW	IT organizations	Organizational agility
[S19]	3.5	2014	Conference	25	2003–2013	KC	Effort estimation	Practices
[S20]	2.5	2015	Journal	56	2004–2014	K	Methods tailoring	Methods
[S21]	4	2014	Conference	71	2001–2012	NE	User-centered design integration	Usability
[S22]	3	2013	Journal	28	2003–2012	KC	Embedded systems	Embedded systems
[S23]	3	2013	Conference	110	2002–2012	Pe	Testing	Practices
[S24]	2.5	2014	Conference	76	2002–2013	Wi	User-centered design integration	Usability
[S25]	4	2014	Journal	21	2002–2013	KC	Requirements engineering	Practices
[S26]	4	2015	Journal	81	1998–2011	KC	CMMI	CMMI
[S27]	3.5	2014	Conference	50	2002–2014	Pe	Information visualization	Miscellaneous
[S28]	2.5	2013	Journal	81	2002–2011	KC	Distributed projects	Global software engineering

[a] KC = Kitchenham & Charters [9]; K = Kitchenham [36]; Bo = Boilchini et al. [18]; HG = Higgins & Green [14]; Pe = Petersen et al. [37]; PR = Petticrew & Roberts [17]; WW = Webster & Watson [38]; Wi = Wieringa et al. [22]; NE = Not Explicitly referring to any guidelines.

as early as 1991 to 2014. Among the 28 SLRs, 13 studies were published in journals and the remaining 15 studies were published in conference proceedings.

A majority of SLRs (approximately 72%, see third last column in Table 2) employed SLR guidelines by Kitchenham [36] or Kitchenham and Charters [9]; while others included Petersen et al. [37] and Webster and Watson [38]. Only one (1) SLR [S21] did not explicitly refer to any specific SLR guidelines but fulfilled the quality criteria.

Fig. 2 shows the distribution of the quality scores assigned to the final set of 28 SLRs. The quality score is dependent on how well the reviews conducted met the quality criteria as described in the previous section. The quality scores were used to exclude two papers with the lowest quality scores (1.5 out of 4 each; shown before the cut-off line in Fig. 2). The main reason for excluding these two papers was that they did not satisfy two or more of the four quality criteria (listed in Table 1): the inclusion/exclusion criteria (Q1); the quality criteria (Q2); and/or did not provide reference or access to all of their primary studies (e.g. through references or supplementary information, Q4).

The majority of the papers achieved a score of 3.0 and above (23 studies, 82%, see second column in Table 2). Hence, most of the studies achieved good quality. The distribution of the quality score gives us an indication of how the quality varied between the SLRs.

Analysis of the data extracted found that Scrum and Extreme Programming (XP) were each investigated by approx. 30% of the reviews, followed by Lean (13.6%), agile software development in general (10.6%), Feature Driven Development (FDD) (6%), Pair Programming (PP) (6%), and Test Driven Development (TDD) (3%). In several of the included reviews, combinations of agile methods were also discussed (e.g. XP + Scrum, Lean + Scrum). In its entirety, XP together with specific practices such as PP and TDD, was the most popular agile method at approx. 40%.

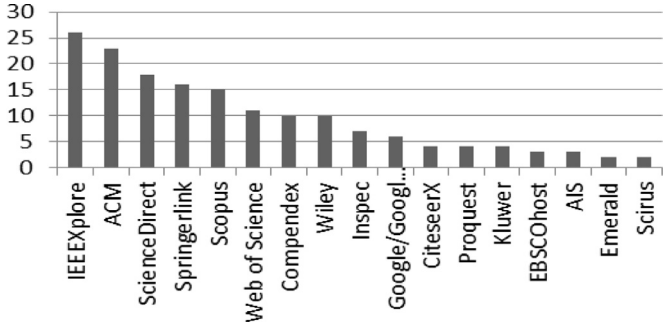

Fig. 3. Online databases used by the SLRs.

Fig. 3 shows the number of SLRs that conducted their searches using each online database. IEEEXplore and ACM emerged as the top two preferred online databases to conduct ASD-related searches. On the other hand, Emerald and EBSCOhost were less used.

Table 3 shows quality scores for each of the four quality question and the total quality score for all 28 SLRs. There were only four mapping studies [S6,S23,S24,S27] while the remaining were systematic literature reviews.

4. Discussion of research questions

4.1. RQ1: how many SLRs were published since the inception of ASD (late 1990s) to date (2015)?

We found that 28 SLRs were published between 2008 and 2015 on various ASD topics. Five of these were mapping studies while the remaining were systematic literature reviews. Fig. 4 shows the number of SLRs that were published each year. The trend indicates

Table 3
Quality scores of the SLRs for each of the quality question and total score.

S#	Study Type*	Q1	Q2	Q3	Q4	Total score
[S1]	SLR	Y	Y	N	Y	3
[S2]	SLR	Y	P	P	Y	3
[S3]	SLR	Y	P	Y	P	3
[S4]	SLR	Y	Y	N	Y	3
[S5]	SLR	Y	Y	Y	Y	4
[S6]	MS	Y	P	Y	P	3
[S7]	SLR	Y	P	Y	Y	3.5
[S8]	SLR	Y	Y	Y	Y	4
[S9]	SLR	Y	Y	Y	Y	4
[S10]	SLR	Y	Y	N	Y	3
[S11]	SLR	Y	P	N	Y	2.5
[S12]	SLR	Y	Y	N	Y	3
[S13]	SLR	Y	Y	Y	Y	4
[S14]	SLR	Y	Y	Y	P	3.5
[S15]	SLR	Y	Y	Y	Y	4
[S16]	SLR	Y	P	Y	Y	3.5
[S17]	SLR	Y	Y	Y	Y	4
[S18]	SLR	P	Y	N	P	2
[S19]	SLR	Y	Y	Y	P	3.5
[S20]	SLR	Y	Y	N	P	2.5
[S21]	SLR	Y	Y	Y	Y	4
[S22]	SLR	Y	Y	N	Y	3
[S23]	MS	Y	P	P	Y	3
[S24]	MS	N	Y	P	Y	2.5
[S25]	SLR	Y	Y	Y	Y	4
[S26]	SLR	Y	Y	Y	Y	4
[S27]	MS	Y	Y	Y	P	3.5
[S28]	SLR	P	Y	N	Y	2.5

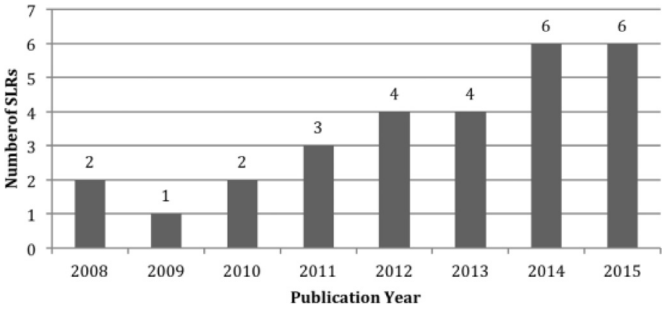

Fig. 4. Number of SLRs published each year (2008–2015).

an increasing interest in the research related to agile methods in various fields since 2008. The number of studies peaked at 2014 and 2015, suggesting that publications related to ASD and interests in consolidating research findings in this field are still on the rise.

Fig. 5a is a graph that visualizes the years covered by the included SLRs in their searches. Fig. 5b is a graph showing the number of primary studies included in the SLRs. It aims to provide a sense of the size and coverage of the individual SLRs and does not take into account any duplication of primary studies across the SLRs. Two SLRs ([S12] and [S23]) provide the highest number of primary studies among all others: 333 and 110 respectively. The reasons for these two SLRs including an unusually large number of primary studies may include:

- The broadness of their research questions. For example, [S23] is a mapping study that covers the broad area of agile testing.
- Hummel et al. [S12] accessed the most popular conferences and journals in the engineering domain as well as in the information system domain, to provide a thorough analysis of the role of communication in ASD.

In Fig. 5b, the bubbles on the right represent the publication venue of the SLRs. Each bubble is labeled based on whether it is a journal (J) or a conference (C) paper. Forty-five (46%) of the SLRs were published through journals while the remaining 54% were published as conference proceedings.

4.2. RQ2: what research areas and topics are being addressed in ASD?

Based on the analysis of the included secondary studies, we grouped the review topics under ten categories. The categorization process was carried out by analyzing the research questions and topics that were covered by the SLR papers, followed by grouping similar review topics together to form a category. The categorization was performed based on a thematic analysis approach which included identifying the research categories used throughout the different secondary studies, defining and naming the key themes, and reviewing them amongst the authors [45]. The first, second and fourth authors conducted independent categorizations and then shared and discussed their recommendations in a review meeting. There were no major conflicts in the categorizations other than the choice of terms (e.g. distributed vs. global software engineering.) The final set of categories was discussed with the third author who further confirmed them. The categories were identified based on their specified topic or key focus of each of the included SLRs as discernible from their title, abstract, overall focus, and the main research area in which the authors seemed to placed their reviews. For example, Chagas et al. [S11] and Silva et al. [S26] were clearly focused on the integrating the CMMI with Agile methods which gave rise to the theme *Agile and CMMI*. Other themes were identified and named in a similar manner. One of the papers [S28] seemed to include aspects of two different categories: adoption and global software development, however, the latter was perceived to be the dominant theme and so the paper was categorized under *Agile and Global Software Engineering* category.

The ten research areas in ASD are shown in Table 4. The first column lists the ten categories (or research areas): adoption, human and social aspects, methods, practices, CMMI, usability, global software engineering, information systems development, embedded systems and miscellaneous applications. The second column provides a brief description of what was covered by the category. The third column lists the references to the SLR that fall under each category, followed by the number of unique SLRs under each category (e.g. agile practices had the most number of SRs, seven); and finally the last column lists the total number of primary studies underlying the SLRs in each category (e.g. Melo et al. [S3] and Hummel et al. [S12] together covered 345 primary studies.) Since analysis was performed at the secondary level, analysis of the primary studies (e.g. number of unique, non-duplicated studies under each category) is not listed.

A brief description of the ten research areas and SLRs covered under each is provided below:

- Agile adoption: Jurca et al. [S14] identified two strategies of agile adoption: wholesale (practices adopted all-at-once) and incremental strategies (practices adopted step-by-step), where incremental strategy was more common in the industry.
- Agile human and social aspects: Melo et al. [S3] identified motivators for agile software developers, e.g. technically challenging work, equity, employee participation with others etc. Hummel et al. [S12] covered 333 primary studies and reported that the role of communication was infrequently presented in scrum practices other than in the daily scrum and was largely focused on the pair programming practice of XP, leaving room for more research in this area.
- Agile methods: Three SLRs focused on different aspects of agile methods, identified as method reconciliation, method analysis, and method tailoring. Magdaleno et al. [S2] identified that large amounts of previous work focused on reconciling agile and

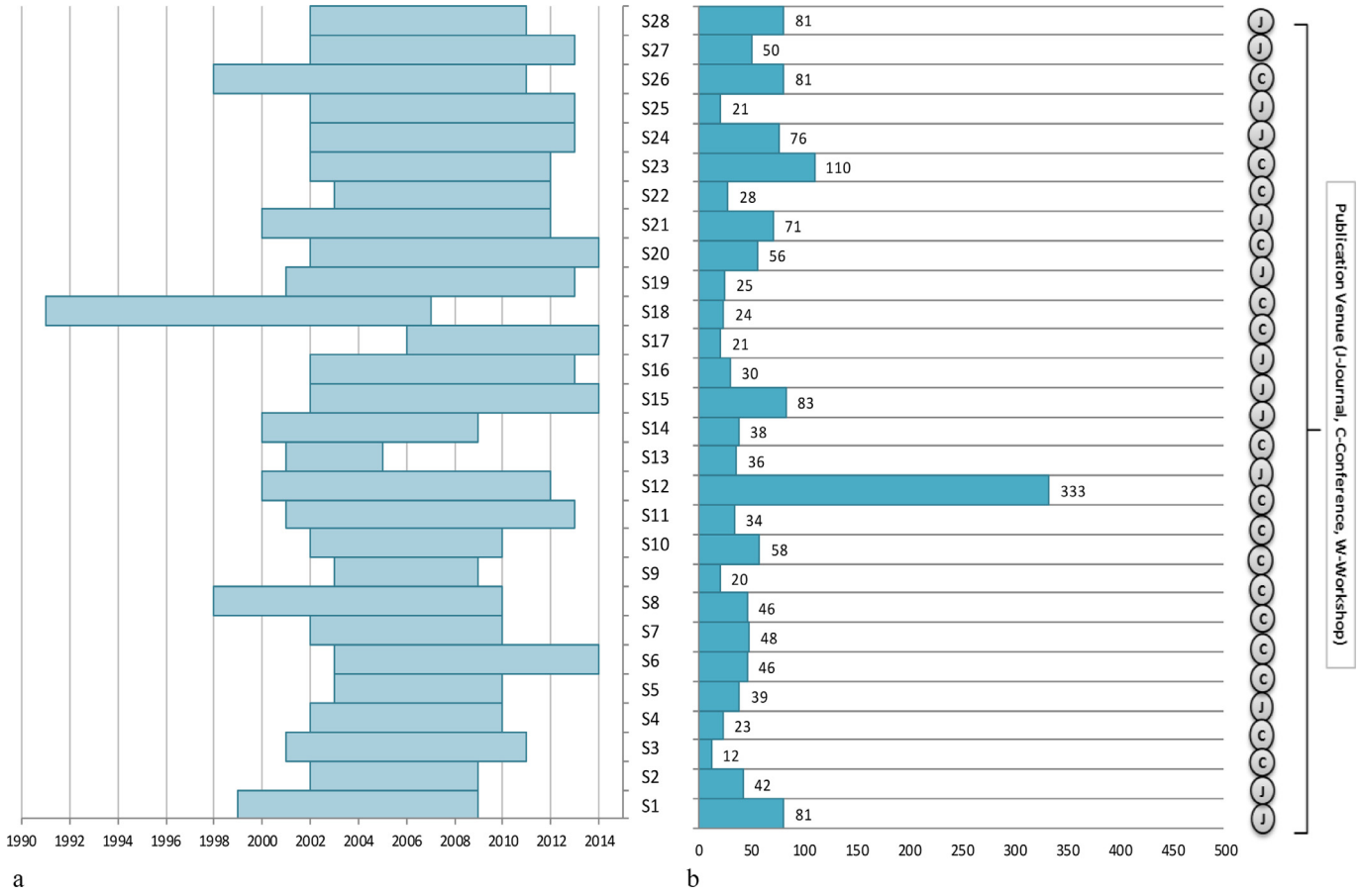

Fig. 5. (a) Years covered by primary studies of the SLRs. (b) Number of primary studies covered by the SLRs.

Table 4
Ten research areas in ASD identified in the tertiary study (#SLR = number of SLRs).

Category	Description	SLRs	#SLR
Agile adoption	Various adoption approaches	[S14]	1
Agile human & social aspects	Human and social aspects in Agile	[S3,S12]	2
Agile methods	Various aspects of agile methods	[S2,S13,S20]	3
Agile practices	Various agile practices	[S7,S8,S16,S19,S23,S25,S27]	7
Agile and CMMI	Integrating capability maturity model with agile	[S11,S26]	2
Agile and usability	Integration Usability with agile methods	[S6,S10,S15,S21,S24]	5
Agile and global software engineering	Use of agile methods in global software development contexts	[S1,S9,S17,S28]	4
Agile and the organization	Use of agile methods in organizations	[S18]	1
Agile and embedded systems	Use of agile methods in embedded systems contexts	[S4,S22]	2
Agile product line engineering	Use of agile methods in product line engineering	[S5]	1
Total			28

plan-driven software development models, while other models such as free and open-source development were less explored. Dybå and Dingsøyr [S13] examined and identified a list of benefits and limitations of agile methods and recommended combining traditional and agile project management for different project contexts. Campanelli, Parreiras [S20] focused on method tailoring and found that method engineering [23] was the overwhelming choice (69.7%) while others were either unclear or used contingency factors [24].

- Agile practices: Seven studies focused on aggregating evidence related to agile practices [S7,S8,S16,S19,S23,S25,S27] making this one of the most popular areas. Research topics that were classified under this category include: test driven development (TDD), agile testing, quality in agile practices, effort estimation in agile practices, use of metrics, requirements engineering and information visualization aspects used in agile practices.
- Agile and CMMI: Silva et al. [S26] presented a state-of-research review on the use of CMMI [25] in ASD and showed that combining CMMI and agile methods allowed companies to achieve levels two and three of CMMI with little effort. Chagas et al. [S11] concluded that the joint approach of agile and maturity models carries potential benefits and the use of tool support can help increase the success rate of such initiatives.
- Agile and usability: da Silva et al. [S6,S21], and [S15] covered various aspects of agile and user-centered design (UCD) such as integration and challenges. da Silva et al. [S10] conducted a state-of-research review on usability research in ASD while [S24] listed recommendations to ensure successful integration between Agile and UX including: a) provide organizational support (e.g. provide support for little design up front- LDUF, sprint zero, and one sprint ahead); b) allocate sufficient UX designers and provide shared workspace between UX designers and

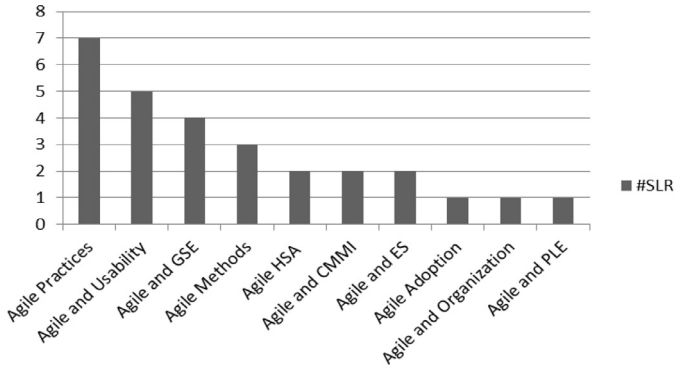

Fig. 6. Agile research areas ranked number of SLRs in that area
(GSE: Global Software Engineering; HSA: Human and Social Aspects; CMMI: Capability Maturity Model Integration; ES: Embedded Systems; PLE: Product Line Engineering.).

developers; and c) use of artifacts such as concept maps, cognitive walkthrough variants, workshops, lo-fi prototypes, interviews, scenarios, and meetings with users.

- Agile and global software engineering: Jalali and Wohlin [S1] presented a state-of-research review covering 81 primary studies. They reported that challenges to GSE and agile integration included: differences in time zones, communication, trust level, knowledge in management, culture and personnel. Hossain et al. [S9] focused exclusively on the use of scrum in GSE contexts; while Matalonga et al. [S28] and Alzoubi et al. [S17] explored adoption and communication challenges in globally distributed agile development respectively.
- Agile and organization: Tapanainen et al. [S18] identified dimensions of organizational agility as: organizational structures, workforce, development process, management and leadership, and infrastructure.
- Agile and embedded systems: Albuquerque et al. [S4] found that applying agile methods has a positive impact on the development of embedded systems. Kaisti et al. [S22] explored the compatibility of agile with embedded systems and suggested strategies such as top-level documentation and some up-front designing.
- Agile product line engineering: Diaz et al. [S5] focused on the use of agile methods in product-line engineering.

Fig. 6 shows the ten research areas ordered by number of SLRs in each area. We found that the category *Agile Practices* had the largest number of SLRs (seven) followed by *Agile and Usability* with five SLRs, and *Agile and Global Software Engineering* with four SLRs. The remaining areas had three, two or one SLR each.

4.3. RQ3: which individuals, organizations, and publication venues are most active in SLR-based research in ASD?

F. Maurer co-authored a maximum number of the SLRs, five: [S6,S10,S23,S24,S27], followed by Hellmann with three SLRs: [S6,S23,S24]. Two other individuals co-authored two SLRs each: de Silva and Silveira [S6,S10]. All other authors were involved with one SLR each. There were a total of 93 unique authors co-authored the 28 SLRs, with the largest number of co-authors on an SLR being seven [S22,S26]; the smallest being two authors [S1,S8,S13,S20]; and the average authors per publication being 3.2.

The 28 SLRs all together accounted for a total of 1892 citations (Google scholar, as of 3rd May 2016) with an average of 67.5 citations each. The SLR titled *"Empirical studies of agile software development: A systematic review"* [S13] by Dybå and Dingsøyr published in the Information and Software Technology journal in 2008 was the single most cited SLR with 1265 citations accounting for approximately 60% of all citations across the 28 SLRs. Table 5 summarizes the five most cited SLRs in ASD.

The SLRs emerged from a total of 31 institutions in 15 countries: Australia, Brazil, Canada, Finland, Germany, Greece, Iran, New Zealand, Netherlands, Norway, Spain, Sweden, Malaysia, Uruguay and the UK. The University of Calgary, Canada led the group with five SLRs across 14 authors; followed by Blekinge Institute of Technology, Sweden with three SLRs across five authors; and Pontifícia Universidade Católica do Rio Grande do Sul, Brazil with 3 SLRs across four authors.

Geographically, Brazil had the most SLRs (eleven) coming from eight of its universities. The popularity and evolution of ASD in the Brazilian IT industry has been documented in [43]. The next highest number of SLRs came from Finland (six SLRs) from four universities; and Germany and Australia with three SLRs each from three universities each.

In terms of the journals and conferences publishing the SLRs, the journal of *Information and Software Technology* had published the largest number of the SLRs, four: [S13,S15,S16,S26]; followed by the *Agile* conference with three [S10,S23,S24]; and the *Journal of Software and Systems* with two SLRs [S2,S20].

4.4. RQ4: what is the quality of the SLRs in ASD?

The quality scores of the individual SLRs included in this review were presented in Fig. 2 earlier. Table 6 below shows the average quality scores of all the SLRs included in this study as well as the distribution by journals and conferences. The average quality score across all SLRs is 3.3 out of 4.0. For journals, the average quality score of the SLRs in ASD is approximately 3.5 out of 4.0. Conferences, by comparison, have a slightly lower average quality score of 3.2.

The contributions of Kitchenham and Charters [9] and Dybå and Dingsøyr [7,35] in standardizing the guidelines for systematic literature reviews and quality in empirical studies, respectively, have been instrumental in achieving good quality in the SLRs on ASD. An overwhelming 72% of all SLRs were conducted using the guidelines by Kitchenham [36] or Kitchenham and Charters [9], while others included Petersen et al. [37] and Webster and Watson [38].

4.5. RQ5: what progress has been achieved with respect to prior recommendations for ASD?

Since no prior tertiary studies of ASD exist, we revisited the recommendations made by some prior reviews that we were aware of and that related closely to the topic, for example: [7,27–33]. We then compared their recommendations with the evidence obtained from our tertiary study and rated the degree of progress on a scale of 1: *little*; 2: *marginal*; 3: *moderate*; 4: *considerable*; and 5: *significant*. We have used our subjective judgment to make these ratings through rounds of independent ratings by first three authors followed by joint discussions to arrive at the final ratings. Table 7 summarizes prior recommendations, ratings, and evidence we found on their progress based on this tertiary study.

Recommendations for better connections between ASD research and established perspectives advocated earlier [27,29] have met with a significant degree of progress as evidenced by three out of the ten research areas based on 12 SLRs that focused on combining ASD with established domains such as usability [S6,S10,S15,S21,S24], CMMI [S11,S26], and GSE [S1,S9,S17,S28].

Focus on management-oriented approaches such as Scrum has been considerable: approximately one third of the SLRs investigating scrum-based empirical studies. ASD was also seen to sustain in different contexts such as global and distributed software development [S1,S9,S17,S28], embedded systems [S4,S22], and product line engineering [S5].

Table 5
Five most cited systematic literature reviews in agile software development.

SLR#	Title	Authors	Venue*	Year	Pages	Citations**
[S13]	Empirical studies of agile software development: a systematic review	Dybå and Dingsøyr	IST	2008	27	1265
[S9]	Using scrum in global software development: a systematic literature review	Hossain, Babar, andPaik	ICGSE	2009	10	182
[S10]	User-centered design and agile methods: a systematic review	da Silva, Martin, Maurer, and Silveira	Agile Conference	2011	10	82
[S1]	Global software engineering and agile practices: a systematic review	Jalali and Wohlin	JSSEP	2011	14	61
[S7]	Factors limiting industrial adoption of test driven development: a systematic review	Causevic, Sundmark, and Punnekkat	ICST	2011	10	42
Total						**1632**

* IST: Information and Software Technology; ICGSE: IEEE International Conference on Global Software Engineering; JSSEP: Journal of Software: Evolution and Process; ICST: International Conference on Software Testing, Verification and Validation
** as of 3rd May 2016

Table 6
SLR quality for journals and conferences.

	SLRs in agile software development
Number of SLRs	**28**
Journal	13
Conference	15
Mean quality score	**3.3**
Journal	3.5
Conference	3.2

Prior reviews called for more and better quality research [7,29,31,33]. In this tertiary study, we have documented that both the quantity and quality of SLRs in ASD are increasing (see Figs. 2 and 4). However, a tertiary study is not able to comment on the quality of the underlying primary studies covered by the SLRs. As such we rate this recommended area as one achieving moderate progress. Similarly, the development of theoretical underpinnings for agile, as advocated often [27,29,32,33], has achieved moderate progress as a number of frameworks [S12,S14,S15] and models [S3] emerge.

A key area where progress appears to have been limited to date is the combination of research rigor and industrial relevance in ASD research [28,32].

4.6. Recommendations for ASD research

Combining research rigor and industrial relevance still remains a 'grand challenge' for ASD research [28,32], arguably true of all new software engineering process innovations. Acknowledging the wide variety of research aims, methods, and outcomes accompanying ASD research, we recommend this is more of a discipline-wide challenge to be addressed by the international ASD research community as a whole rather than by individual ASD research studies as individual studies will likely continue to contribute to theory and practice in varying degrees.

For studies aiming to achieve industrial relevance, approaches such as the collaborative model proposed by [47] and the use of a 'challenge wall' [28] to elicit industrial challenges to guide research topics and questions could be useful. Such a 'pull' based research model – much like its counterpart in Lean development [46] – promises to focus on industrial relevance from the outset. Research methods such as case studies, grounded theory, and ethnography complement such an approach as they enable the collection and analysis of industrial data through rigorous research procedures to achieve a research-industry balance.

ASD research areas and topics that can benefit from further primary and secondary studies in the near future include: evaluating benefits and challenges of ASD methods, agile hybrids in large-scale setups, sustainability, motivation, teamwork, and project management. Of the ones represented in this tertiary study, some areas and topics can benefit from a fresh review, e.g. ASD and embedded systems as [S4] and [S22] were published in 2012 and 2013 respectively; scrum practices in global contexts as [S9] was the only SLR on this topic, published in 2009; and organizational aspects of agile as Tapanainen et al. [S18] was the only SLR on this topic, published in 2008. Finally, the last overarching SLR of empir-

Table 7
Prior recommendations for ASD research and evidence of progress to date; *1 = little progress; 2 = marginal progress; 3 = moderate progress; 4 = considerable progress; 5 = significant progress.

Prior recommendations	Degree of progress*	Evidence from the SLRs in this tertiary review
Better connection with established perspectives [27,29]	5	Three of the ten research areas based on 12 SLRs representing 44% of primary studies on agile and usability [S6,S10,S15,S21,S24]; agile and CMMI [S11,S26]; and agile and GSE [S1,S9,S17,S28]
Good coverage of management-oriented approaches.	4	Significant focus on scrum (Section 3.)
Sustaining agile in different contexts [28,29]	4	Agile and GSE [S1,S9,S17,S28]; agile and embedded systems [S4,S22]; product line engineering [S5]
More and better quality research [7,29,31,33]	3	Increasing number and quality of SLRs representing large numbers of primary studies (Figs. 2 and 4.)
Stronger theoretical foundations [27,29,32,33]	3	[S14] (theoretical agile adoption frameworks), [S3] (theoretical model in analysis of developers' motivation in agile teams), [S12] (theoretical framework developed based on Unified model of SD success), and [S15] (analysis framework derived to identify generic principles of user-centered agile software development).
Grand challenge in ASD		
Combining research rigor with industrial relevance [28,32]	2	Formal efforts by agile research network [28], collaborative model [47]

ical studies in ASD was in 2008 and another one can be conducted to cover the period post this SLR ([S13]).

In addition to the above recommendations for ASD research, this review also contributes to the area of SLR studies as follows. We recommend that the standard quality criterion "*information presented was not referenced*" [10] should in fact be an exclusion criterion. In other words, SLRs that do not list or provide access to their sources (i.e. primary studies) should not be included. We have set a precedent for this in this tertiary study by using this quality criterion to filter low quality papers. This will ensure stricter quality control in tertiary studies and encourage researchers conducting SLRs to be transparent about the evidence they used. It will also enable future studies to analyze the included SLRs on the level of primary studies.

5. Limitations

Since this is a tertiary review only secondary studies performing either systematic literature reviews or systematic mapping studies were included. As such, if a popular agile research topic had several primary studies devoted to it but did not have a systematic review published on the topic, we did not include it. Also, we only included *systematic* literature reviews and mappings as compared to informal literature reviews in order to maintain a high quality of results.

In order to ensure validity of the conclusions, the search process was carried out in five reputable digital databases without year restrictions to include as many relevant papers as possible. The search string used was also not constrictive, so that the best possible number of agile related SLRs could be picked up by the search. Since we did not perform a manual search process, we referred to a set of online databases that were often used in other tertiary study (e.g. [10]) to conduct our searches, allowing us to minimize the probability of missing evidence.

One of the fundamental search terms used in this study was 'agile' which was then searched in combination with other terms. A result of this approach was that a known SLR on specific agile practice (e.g. pair programming [19]) was omitted during the initial screening as the paper did not include the word 'agile' in the title, abstract, or keywords. To include such a paper required changing our search terms to include all individual practices from all agile methods individually with and without the term 'agile' which was not possible. We also excluded any SLRs on academic practice of agile as our aim was to present an overview of industrial ASD research. This meant that SLRs such as [15] were not included.

Since the researchers independently performed the data extraction, a related threat is the potential personal/author bias in the extraction process which can influence the accuracy of the extracted data. To overcome this potential issue, we discussed and resolved conflicts or disagreement in a joint meeting.

The categorization of the research areas (RQ2) was performed on the basis of the most dominant theme of the paper as discernible by its title, abstract, overall focus, and the main research area that the authors seemed to place their study in. However, the categorization did not cover other, less dominant themes that may have been present in the studies. For example, Matalonga et al. [S28] was primarily focused on distributed or global software engineering but included some aspects of adoption. A majority of the SLRs, however, were clearly focused on a single research area.

In answering RQ4, we have used mean values of the quality scores as it is common practice [10]. However, it should be noted that doing so assumes an interval or ratio scale while the quality scoring is an ordinal scale. In other words, the quality scoring allows us to order or rank one quality aspect as lower or higher than the other but the interval between two ranks is neither easily quantifiable nor equal. As such, the quality scores and inferences therefrom are limited due to this divergence of fundamental concepts and popular practice.

As a baseline for answering RQ5, we included prior reviews closely related to our topic that we were aware of (e.g. [78, 27, 28]) so to compare their recommendations with the evidence from our findings. This raises the threat of availability bias, where the rating made to the degree of progress was limited to the set of recommendations available from those studies. We included at least seven (7) prior reviews (published between 2003 and 2015) which we hope could minimize this bias.

Being a tertiary study, the descriptive and qualitative analysis was performed at the secondary studies level [6] and an analysis of the overlap between the sets of primary studies was not performed. This has particular reference to RQ5 since a potential high level of overlap of primary studies between SLRs in the same research area can provide a skewed view of the progress achieved. However, we have measured progress based on not only multiple SLRs in the same research area but also across multiple research areas to minimize this threat to validity. Furthermore, since all included SLRs provide clear reference or access to their primary studies such analysis is possible in the future. A tertiary study can be conducted to assess the quality of the included SLRs, however, it is not able to assess the quality of the primary studies underlying the SLRs [6].

6. Conclusion

We conducted a tertiary study of systematic literature reviews in agile software development based on the tertiary study by Kitchenham et al. [10]. In response to the research questions *RQ1. how many SLRs were published since the inception of ASD (late* 1990s*) to date?* and *RQ2. what research areas and topics are being addressed in ASD?* We found 28 SLRs published between 2008 and 2015 focusing on ten different areas of ASD research: adoption, methods, practices, human and social aspects, CMMI, usability, global software engineering, organizational agility, embedded systems, and product line engineering.

In response to the research question *RQ3. which individuals, organizations, and publication venues are most active in SLR-based research in ASD?* We found that F. Maurer and T. Hellman had authored a majority of the SLRs: five and three each; while the University of Calgary, Canada was home to the most SLRs followed by Blenkinge Institute of Technology, Sweden and PUCRS, Brazil with five, three, and three SLRs each. Information and software technology published a majority of the SLRs (four), followed by the Agile conference with three, journal of systems and software with two, and product-focused software improvement, computational science and its application, and global software engineering conferences with two SLRs each. The SLR titled *"Empirical studies of agile software development: A systematic review"* [S13] by T. Dybå and T. Dingsøyr was the single most cited SLR with 1265 citations accounting for approximately 60% of all citations across the 29 SLRs.

In response to the research question *RQ4. What is the quality of the SLRs in ASD?* The average quality score of all included SLRs was 3.3 out of 4.0. Of these, SLRs published in journals had a slightly higher average quality score (3.5 out of 4.0) than those published in conferences (3.2 out of 4.0). The SLR guidelines by Kitchenham and Charters [9] and the exemplar SLR by Dybå and Dingsøyr have been instrumental in standardizing and improving the quality of SLRs in ASD.

In response to the last research question *RQ5. What progress has been achieved with respect to prior recommendations for ASD?* We found that significant progress has been made with regards to better connecting ASD with established domains such as usability, CMMI, and global software engineering. Considerable progress has been achieved in focusing on management-oriented approaches as

Scrum and sustaining ASD in different contexts such as embedded systems. Moderate progress has been made in conducting more and better quality ASD research, where assessing the quality of primary studies was not within the scope of this tertiary study. Similarly, some progress seems to have been made towards building stronger theoretical foundations of ASD.

We did not find much evidence to support a significant progress toward resolving the 'grand challenge' of ASD: combining research rigor with industrial relevance, as a topic. Since individual studies tend to contribute to theory and practice in varying degrees, a balance needs to be maintained by the international ASD research community as a whole. For studies focusing on industrial relevance, we recommend the use of an industrial 'pull' model to enable industry needs to guide research topics and questions and achieve research rigour through established methods such as case studies, grounded theory, and ethnography.

Some areas and topics that can benefit from further primary and secondary studies in the near future include: evaluating benefits and challenges of ASD methods, agile hybrids in large-scale setups, sustainability, motivation, teamwork, and project management. Whereas topics and areas such as ASD and embedded systems, scrum practices in global contexts, and agile and the organization can benefit from fresh SLRs to cover the recent research period. Finally, another SLR of empirical studies in ASD can be conducted to cover the period post the last one published eight years ago.

Based on our experiences of conducting this review, we also recommend a stricter quality control in tertiary studies through the exclusion of SLRs that do not list or provide access to their primary studies.

This tertiary study focused on collating secondary studies with underlying empirical evidence from industrial use of ASD. Future work can focus on a similar approach for secondary studies in academic settings.

Acknowledgments

This study was supported by a FRDF-NS Fund (#3703888) from The University of Auckland.

Reviews included in the tertiary study

[S1] S. Jalali, C. Wohlin, Global software engineering and agile practices: a systematic review, J. Softw. Evol. Process 24 (6) (2011) 643–659.
[S2] M. Magdaleno, C.M.L. Werner, R.M. de Araujo, Reconciling software development models: a quasi-systematic review, J. Syst. Softw. 85 (2) (2012) 351–369.
[S3] C. de, O. Melo, C. Santana, F. Kon, Developers motivation in agile teams, in: Proceedings of the 38th Euromicro Conference on Software Engineering and Advanced Applications, 2012.
[S4] O. Albuquerque, P.O. Antonino, E.Y. Nakagawa, An investigation into agile methods in embedded systems development, Comput. Sci. Appl. 2012 (2012) 576–591.
[S5] J. Díaz, J. Pérez, P. Alarcón, J. Garbajosa, Agile product line engineering-a systematic literature review, Softw. Pract. Exp. 41 (8) (2011) 921–941.
[S6] T.S. da Silva, F.F. Silveira, M.S. Silveira, T. Hellmann, F. Maurer, A systematic mapping on agile UCD across the major agile and HCI conferences, in: Proceedings of the International Conference on Computational Science and its Application (ICCSA2015), 2015, pp. 86–100.
[S7] A. Causevic, D. Sundmark, S. Punnekkat, Factors limiting industrial adoption of test driven development: a systematic review, in: Proceedings of the Fourth IEEE International Conference on Software Testing, Verification and Validation, 2011, pp. 337–346.
[S8] P. Sfetsos, I. Stamelos, Empirical studies on quality in agile practices: a systematic literature review, in: Proceedings of the Seventh International Conference on the Quality of Information and Communications Technology (QUATIC), 2010, pp. 44–53.
[S9] E. Hossain, M.A. Babar, H. Paik, Using scrum in global software development: a systematic literature review, in: Proceedings of the Fourth IEEE International Conference on Global Software Engineering (ICGSE), 2009, pp. 175–184.
[S10] T.S. da Silva, A. Martin, F. Maurer, M. Silveira, User-centered design and agile methods: a systematic review, in: Proceedings of the AGILE Conference, 2011, pp. 77–86.
[S11] L.F. Chagas, D.D. Carvalho, A.M. Lima, C.A.L. Reis, Systematic literature review on the characteristics of agile project management in the context of maturity models, Softw. Process Improv. Capab. Determ. (2014) 177–189.
[S12] M. Hummel, C. Rosenkranz, R. Holten, The role of communication in agile systems development, Bus. Inf. Syst. Eng. 5 (5) (2013) 343–355.
[S13] T. Dybå, T. Dingsøyr, Empirical studies of agile software development: a systematic review, Inf. Softw. Technol. 50 (9–10) (2008) 833–859.
[S14] A. Rohunen, P. Rodriguez, P. Kuvaja, L. Krzanik, J. Markkula, Approaches to agile adoption in large settings: a comparison of the results from a literature analysis and an industrial inventory, Product-Focused Softw. Process Improv. (2010) 77–91.
[S15] M. Brhel, H. Meth, A. Maedche, K. Werder, Exploring principles of user-centered Agile software development, Inf. Softw. Technol. 61 (2015) 163–181.
[S16] E. Kupiainen, M.V. Mäntylä, J. Itkonen, Using metrics in agile and lean software development, Inf. Softw. Technol. 62 (2015) 143–163.
[S17] Y.I. Alzoubi, A.Q. Gill, A. Al-Ani, Empirical studies of geographically distributed agile development communication challenges: a systematic review, Inf. Manag. 53 (1) (2016) 22–37.
[S18] T. Tapanainen, M. Hallanoro, J. Päivärinta, H. Salmela, Towards an agile it organisation: a review of prior literature, in: Proceedings of the 2nd European Conference on Information Management and Evaluation (ECIME), 2008, pp. 425–432.
[S19] M. Usman, E. Mendes, F. Weidt, R. Britto, Effort estimation in agile software development: a systematic literature review, in: Proceedings of the 10th International Conference on Predictive Models in Software Engineering - PROMISE '14, 2014, pp. 82–91.
[S20] A.S. Campanelli, F.S. Parreiras, Agile methods tailoring – a systematic literature review, J. Syst. Softw. 110 (2015) 85–100.
[S21] D. Salah, R. Paige, P. Cairns, A systematic literature review for agile development processes and user centred design integration, in: Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering - EASE '14, 2014.
[S22] M. Kaisti, V. Rantala, T. Mujunen, S. Hyrynsalmi, K. Könnölä, T. Mäkilä, T. Lehtonen, Agile methods for embedded systems development - a literature review and a mapping study, EURASIP J. Embed. Syst. 2013 (1) (2013) 1–16.
[S23] T.D. Hellmann, A. Chokshi, Z.S.H. Abad, S. Pratte, F. Maurer, Agile testing: a systematic mapping across three conferences – understanding agile testing in the XP/Agile Universe, agile and XP conferences, in: Agile Conference, 2013, pp. 32–41.
[S24] G. Jurca, T.D. Hellmann, F. Maurer, Integrating agile and user-centered design. a systematic mapping and review of evaluation and validation studies of agile-UX, in: Agile Conference, 2014, pp. 24–32.
[S25] I. Inayat, S.S. Salim, S. Marczak, M. Daneva, S. Shamshirband, A systematic literature review on agile requirements engineering practices and challenges, Comput. Hum. Behav. (2014).
[S26] F.S. Silva, F.S.F. Soares, A.L. Peres, I.M. de Azevedo, A.P.L.F. Vasconcelos, F.K. Kamei, S.R. de Lemos Meira, Using CMMI together with agile software development: a systematic review, Inf. Softw. Technol. 58 (2015) 20–43.
[S27] J. Paredes, C. Anslow, F. Maurer, Information visualization for agile software development teams, in: Proceedings of the 2nd IEEE Working Conference on Software Visualization, 2014, pp. 157–166.
[S28] S. Matalonga, M. Solari, G. Matturro, Factors affecting distributed agile projects: a systematic review, Int. J. Softw. Eng. Knowl. Eng. 23 (09) (2013) 1289–1301.

References

[1] J. Highsmith, M. Fowler, The Agile Manifesto, Softw. Dev. Mag. 9 (8) (2001) 29–30.
[2] W.W. Royce, Managing the development of large software systems, in: proceedings of IEEE WESCON, 26(8), 1970, pp. 328–338.
[3] A. Cockburn, J. Highsmith, Agile software development: the people factor, Computer 34 (11) (2001) 131–133 Nov.
[4] K. Schwaber, M. Beedle, Agile Software Development with Scrum, Prentice-Hall, 2002.
[5] K. Beck, Extreme Programming Explained: Embrace Change, first ed., Addison-Wesley Professional, 1999.
[6] B. Kitchenham, O.P. Brereton, D. Budgen, M. Turner, J. Bailey, S. Linkman, Systematic literature reviews in software engineering – a systematic literature review, Inf. Softw. Technol. 51 (2009) 7–15.
[7] T. Dybå, T. Dingsøyr, Empirical studies of agile software development: a systematic review, Inf. Softw. Technol. 50 (9–10) (2008) 833–859.
[8] G. Hanssen, D. Šmite, N.B. Moe, Signs of agile trends in global software engineering research: a tertiary study, in: Proceedings of the IEEE Sixth International Conference on Global Software Engineering Workshop, 2011, pp. 17–23.
[9] B.A. Kitchenham and S. Charters, "Procedures for Performing Systematic Literature Review in Software Engineering," EBSE Technical Report version 2.3, EBSE-2007-01, Software Eng. Group.
[10] B.A. Kitchenham, R. Pretorius, D. Budgen, P. Brereton, M. Turner, M. Niazi, S. Linkman, Systematic literature reviews in software engineering – a tertiary study, Inf. Softw. Technol. 52 (8) (2010) 792–805.
[11] H. Sharp, H. Robinson, An ethnographic study of XP practice, Emp. Softw. Eng. 9 (4) (2004) 353–375.
[12] E. Whitworth, R. Biddle, The social nature of agile teams, in: Proceedings of the Agile '07, 2007, pp. 26–36.

[13] L. Williams, R.R. Kessler, W. Cunningham, R. Jeffries, Strengthening the case for pair programming, IEEE Softw. 17 (4) (2000) 19–25 July/Aug.
[14] J.P.T. Higgins, S. Green (Eds.), Cochrane Handbook for Systematic Reviews of Interventions, The Cochrane Collaboration, 2008 Version 5.0.0 (updated February 2008).
[15] N. Salleh, E. Mendes, J. Grundy, Empirical studies of pair programming for CS/SE teaching in higher education: a systematic literature review, IEEE Trans. Softw. Eng. 37 (4) (2011) 509–525.
[16] R. Hoda, J. Noble, S. Marshall, Self-organizing roles on agile software development teams, IEEE Trans. Softw. Eng. 39 (3) (2013) 422–444.
[17] M. Petticrew, H. Roberts, Systematic Reviews in the Social Sciences: A Practical Guide, Blackwell Publishing, Oxford, United Kingdom, 2006.
[18] J. Biolchini, P.G. Mian, C.C.C. Natali, G.H. Travassos, Systematic Review in Software Engineering, COPPE /UFRJ, 2005.
[19] J.E. Hannay, T. Dybå, E. Arisholm, D.I.K. Sjøberg, The effectiveness of pair programming: a meta-analysis, Inf. Softw. Technol. 51 (7) (2009) 1110–1122.
[20] D.S. Cruzes, T. Dybå, Research synthesis in software engineering: a tertiary study, Inf. Softw. Technology 53 (5) (2011) 440–455.
[21] S. Augustine, Managing Agile Projects, Prentice Hall PTR, 2005.
[22] R. Wieringa, N. Maiden, N. Mead, Requirements engineering paper classification and evaluation criteria: a proposal and a discussion, Requir. Eng. 11 (1) (2006) 102–107.
[23] K. Conboy, B. Fitzgerald, Method and developer characteristics for effective agile method tailoring: a study of XP expert opinion, ACM Trans. Softw. Eng. Methodol. 20 (1) (2010) 2.
[24] B. Fitzgerald, N. Russo, T. O'Kane, An empirical study of system development method tailoring in practice. In: Proceedings of the ECIS 2000
[25] C. Mellon, 'Brief History of CMMI'. [Online]. Available: http://resources.sei.cmu.edu/asset_files/Brochure/2009_015_001_28416.pdf. [Accessed: 24-Mar-2015].
[26] S.-W. Chuang, T. Luor, H.-P. Lu, Assessment of institutions, scholars, and contributions on agile software development (2001–2012), J.Syst. Softw. 93 (2014) 84–101.
[27] T. Dingsøyr, S. Nerur, V. Balijepally, N.B. Moe, A decade of agile methodologies: towards explaining agile software development, Jo. Syst. Softw. 85 (2012) 1213–1221.
[28] P. Gregory, L. Barroca, K. Taylor, D. Salah, H. Sharp, Agile challenges in practice: a thematic analysis, in: C. Lassenius, T. Dingsøyr, M. Paasivaara (Eds.), Agile Processes, in Software Engineering, and Extreme Programming, 212, ed: Springer International Publishing, 2015, pp. 64–80.
[29] K. Conboy, Agility from first principles: reconstructing the concept of agility in information systems development, Inf. Syst. Res. 20 (3) (2009) 329–354.
[30] S. Stavru, A critical examination of recent industrial surveys on agile method usage, J. Syst. Softw. 94 (2014) 87–97.
[31] P. Abrahamsson, J. Warsta, M.T. Siponen, J. Ronkainen, New directions on agile methods: a comparative analysis, in: Proceedings of the International Conference on Software Engineering, IEEE, Portland, Oregon USA, 2003.
[32] T. Dingsøyr, T. Dybå, P. Abrahamsson, A preliminary roadmap for empirical research on agile software development, in: Proceedings of AGILE, IEEE, Toronto, Canada, 2008, pp. 83–94.
[33] P. Abrahamsson, K. Conboy, X. Wang, "Lots done, more to do": the current state of agile systems development research, Eur. J. Inf. Syst. 18 (2009) 281–284.
[34] D. Cohen, M. Lindvall, P. Costa, An introduction to agile methods, in: M.V. Zelkowitz (Ed.), Advances in Computers, Advances in Software Engineering, 62, Elsevier, Amsterdam, 2004.
[35] T. Dybå, 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, ACM, 2008, pp. 178–187.
[36] B. Kitchenham, Procedures for Performing Systematic Reviews, TR/SE-0401, Software Engineering Group, Department of Computer Science, Keele University, 2004.
[37] K. Petersen, R. Feldt, S. Mujtaba, M. Mattsson, "Systematic mapping studies in software engineering", Proceedings of the 12th International Conference on Evaluation 71–80.
[38] J. Webster, R.T. Watson, Analysing the past to prepare for the future: writing a literature review, MIS Q. 26 (2) (2002) xiii–xxiii.
[39] S. Matalonga, M. Solari, G. Matturro, Factors affecting distributed agile projects: a systematic review, Int. J. Softw. Eng. Knowl. Eng. 23 (09) (2013) 1289–1301.
[40] T.S. da Silva, F.F. Silveira, M.S. Silveira, T. Hellmann, F. Maurer, A systematic mapping on agile UCD across the major agile and HCI conferences, in: Proceedings of the International Conference on Computational Science and its Application (ICCSA2015), 2015, pp. 86–100.
[41] A. Causevic, D. Sundmark, S. Punnekkat, Factors limiting industrial adoption of test driven development: a systematic review, in: Proceedings of the Fourth IEEE International Conference on Software Testing, Verification and Validation, 2011, pp. 337–346.
[42] A. Alliance, Agile manifesto, 2001 http://www.agilemanifesto.org, Online at (accessed 19.01.17).
[43] C.D.O. Melo, V. Santos, E. Katayama, H. Corbucci, R. Prikladnicki, A. Goldman, F. Kon, The evolution of agile software development in Brazil, J. Braz. Comput. Soc. 19 (4) (2013) 523–552.
[44] M. Hummel, C. Rosenkranz, R. Holten, The role of communication in agile systems development, Bus. Inf. Syst. Eng. 5 (5) (2013) 343–355.
[45] V. Braun and V. Clarke. Using thematic analysis in psychology. 3 pp. 77–101. 2006.
[46] M. Poppendieck, T. Poppendieck, Lean Software Development: An Agile Toolkit: An Agile Toolkit, Addison-Wesley, 2003.
[47] C. Wohlin, Empirical software engineering research with industry: top 10 challenges, in: Proceedings of the 1st International Workshop on Conducting Empirical Studies in Industry (CESI), IEEE, 2013, pp. 43–46.
[48] Database of Attributes of Effects, Centre for reviews and dissemination (CRD), 2016. Last accessed: 5th December Last accessed: 5th December http://www.york.ac.uk/crd/about/.