What Do We Know about Alignment of Requirements Engineering and Software Testing?

Pertti Karhapää Alireza Haghighatkhah Markku Oivo
Empirical Software Engineering in Software, Systems and Services (M3S) Research Unit, ITEE, University of Oulu. 2017.
P.O.Box 8000
FI-90014 University of Oulu

+358 50 3295783 pertti.karhapaa@oulu.fi
+358 46 9204711 alireza.haghighatkhah@oulu.fi
+358 50 4092332 markku.oivo@oulu.fi

ABSTRACT

Context: The alignment of different software engineering activities for coordinated functioning and optimized product development is of great importance, particularly in industrial-scale development. The link between intermediate activities has been researched extensively, but the link between requirements engineering (RE) and software testing (ST) is a relatively less explored area.

Objective: The objective of this study is to aggregate, structure, and classify all existing research regarding alignment of RE and ST published by the end of 2015.

Method: We conducted a systematic mapping study (SMS) and aggregated all studies relevant to our scope. The primary studies are analyzed in terms of publication trend, focus area, i.e., how alignment is supported, the application domain and benefits and challenges, methodological data, and scientific rigor and industrial relevance.

Results: There is a growing interest towards the topic. Several different techniques have been identified to improve RE and ST alignment. Test generation from requirements specification has received most attention. Alignment of RE and ST is particularly important for large safety-critical domains. While many challenges have been reported, the supporting evidence for benefits is scarce. Frameworks/methods/techniques is the most frequent contribution type. Solution proposal and evaluation research were the most frequently applied research type. Case study research was the most frequently applied research method, however, almost half of the studies did not clearly report any research method.

Conclusion: Despite the numerous approaches that are proposed, it is not clear what approach is suitable in what context and why. To support industry in RE and ST alignment, guidelines and tool support are needed. The supporting evidence for claimed benefits is very limited. Overall, the research area is in its early stages and an increase in both the number and rigor of empirical studies are required.

KEYWORDS

Software engineering; alignment; requirements engineering; software testing; verification; validation

1. INTRODUCTION

Software development, regardless of whether plan driven or agile, follows the same concept: from initial conceptualization to realization through different analysis, design, implementation and testing activities. In all development, it is equally important to verify and validate that the final product operates as intended. The link between intermediate activities, e.g., analysis and design, or design and implementation, has been researched extensively, but the link between requirements engineering (RE) and software testing (ST) is less explored [1]. Since the requirements of a software product dictates how the product should operate and testing should show that the product does what it is supposed to do [2], it would be beneficial to explore this link.

Modern software development in industrial context can be very complex, which makes it difficult to maintain the link between RE and ST. The product to be developed can be a system of systems with tens of millions of lines of code and thousands of requirements and all these systems must function properly together [3]. This need is highlighted in safety critical domains, such as the automotive, clinical, or avionics domains, where a system that function incorrectly may lead to loss of lives. It is very important that the systems in these domains function as required and that makes the activities and alignment of RE and ST very important.

Another characteristic of modern big organizations is that the intra organizational distances starts to grow [4] and, like in the automotive industry, different activities of the development process might be executed outside the original equipment manufacturer's (OEM) organization [3]. This makes it even more difficult to maintain the link between RE and ST and makes coordination even more important. This coordination is referred to as alignment of RE and ST, that Unterkalmsteiner et al. [1] define as "the adjustment of RE and ST efforts for coordinated functioning and optimized product development".

The importance of linking RE and ST has been recognized in industry. Uusitalo et al. [5] reported that a better link between RE and ST ensures better flow of information about requirements to the testing process. Kukkanen et al. [6] found that this improves the control and estimation of project cost and schedule leading to positive effects on project work, product quality, and ultimately leads to more satisfied customers. Barmi et al. [7] conducted a systematic mapping

study (SMS) regarding the alignment of non-functional requirements (NFR) and testing, but expanded later to include also functional requirements due to lack of focus on NFRs. Their study discussed challenges related to methods but did not classify the benefits and challenges. They presented that evaluation of proposals is lacking and validity discussion is scares, however, more detailed analysis of rigor and relevance could be done. Unterkalmsteiner et al. [1] found later additional relevant studies that should have been included in the previous mapping study. A dedicated workshop has also been held on the topic in 2014 [8] and 2015 [9]. The research area has evolved significantly and many articles has been published recently. Thus, there is a need for an up-to-date overview of the research on alignment of RE and ST.

The aim of this SMS is to attempt to gather all the scientific literature on alignment of RE and ST, to see how the field has evolved after the study by Barmi et al., and to structure the body of knowledge. The main research question (RQ) thus is: *What is the current state of research on the alignment of RE and ST?* The body-of-knowledge is characterized and classified in terms of publication year and venues, approaches to support alignment, in what context alignment is supported and the benefits and challenges, contribution and research type and research method, and scientific rigor and industrial relevance.

The next chapter gives more information about the background. Chapter 3 present the applied research method and how it was conducted in this study. Chapter 4 presents the findings of the SMS. Discussion is presented in Chapter 5, and Chapter 6 gives concluding remarks.

2. BACKGROUND AND RELATED WORK

The waterfall model of the development process can help development of large complex software systems [10]. However, this model puts the activities of RE and ST very far from each other. In the V-model view of the development process different analysis and design activities are linked to corresponding testing activities [11]. Thus, the V-model helps to preserve the connection between RE and ST. There is, in many agile methods too, a closer relationship between requirements and tests, however, there is not much explicit focus on the alignment of RE and testing activities. The favoring of more face-to-face communication and less documentation in agile methods [12] might be a concern since it has been found that the linking of people does not make documentation redundant [5].

The importance of the flow of information was identified by Uusitalo et al. [5] stating that the most important function of the link between RE and ST is to ensure the flow of information. Bjarnason et al. [4] further stated that not only the information itself is important but the distances between activities, artefacts, and people are also important, and constructed a theory of distance that explains how certain practices affect this distance. Unterkalmsteiner et al. [1] studied and constructed a taxonomy for alignment of RE and ST based on the premise that information and the flow of information between RE and ST is a key ingredient in better coordination of the two activities and proposed a framework called REST-bench for the assessment of alignment in development organizations. The REST-bench framework was later validated successfully in industrial context in five different case studies of varying size in both plan driven and agile development [13].

The previous mapping study on the topic by Barmi et al. [7] focused on the classification of different methods and techniques that support alignment that had been published between 2000 and 2010. Their mapping study classified research into 6 categories: formal approaches, traceability, code centric approaches, model centric approaches, test cases, and alignment problems and good practices. The authors found that most studies focused on model centric approaches and traceability.

3. RESEARCH METHOD

The research method of this study was SMS as explained and introduced by Petersen et al. in [14] for software engineering research. The main difference between SMS and a systematic literature review (SLR) is that, while SLRs aim to identify best practice with respect to specific procedures, technologies, methods or tools by aggregating information from comparative studies, SMSs focus on classification and thematic analysis of literature on a software engineering topic [15].

3.1 Objective and Research Questions

The aim of this SMS was to structure the body of knowledge on alignment of RE and ST. Several sub questions help answer the main question: *What is the current state of research on the alignment of RE and ST?*

RQ1 – What is the intensity of research on the alignment of RE and ST? The aim of this question is to get an overview of publication trends with respect to number of publications per year and type of venues.

RQ2 – What are the foci of research regarding alignment of RE and ST and what is proposed to support alignment in the research? The purpose of this RQ is to find out how alignment of RE and ST can be achieved and to find and classify all approaches and solutions presented in the primary studies.

RQ3 – In what context has alignment of RE and ST been applied and what effects in terms of benefits and challenges have been reported? The purpose here is to give an overview of the application domains where alignment have been applied and synthezise reported benefits and challenges.

RQ4 – What are the most frequently applied research types and research methods, and what kind of contributions are provided by the studies? The purpose is to understand what kind of research has been done in the studies, using what research method and to investigate the contribution types.

RQ5 – What is the scientific rigor and industrial relevance of the research? The purpose of this question is to investigate the scientific rigor of existing research and their relevance for industrial practitioners.

3.2 Search Strategy

Three search strategies were used in this study: database search, snowballing, and manual search. The key-words for the search phrase in database search were formulated by inspecting an initial set of studies. These studies were obtained from the previous SMS [7], the taxonomy paper by Unterkalmsteiner et al. [1], and through an initial manual search for relevant studies. The search phrase was constructed and improved to cover almost all the studies in the initial set without creating too much noise in the search result. The different categories of the final search phrase are presented in Table 1 and were used in the form C1 AND C2 AND C3 AND C4. The exact search phrases can be found in [17].

Five different databases were used in search. Figure 1 shows the used databases together with search results and the complete study selection process. Search was conducted in December 22nd 2015. The snowballing, as proposed in [16] was conducted before full text review to catch any missing relevant studies. Thereafter, we consulted an expert in the area, as proposed in [14], and it was decided that a manual search would be needed. The manual search was conducted by applying the inclusion and exclusion criteria on papers published in the dedicated requirements engineering and testing workshops [8] [9].

The terms "trac*", "gap", and "coverage" in search phrase category 4 (C4) are not synonyms of "alignment". The initial set of studies described the concept of alignment in many ways. Thus, it was deemed necessary to include these terms to capture these studies of the initial set and potential other studies referring to the concept in similar terms.

3.3 Selection Criteria and Selection Process

The inclusion and exclusion criteria applied in the selection process are presented in Table 2. The selection criteria were first piloted by two researchers to improve the criteria and to ensure that both researchers had a common understanding of the concept. Conflicting decisions regarding inclusion or exclusion of studies were resolved in meetings through discussion and argumentation.

Table 1. Categories of the search phrase

Category	Search terms
Software (C1)	Software
Requirements (C2)	Requirement
Testing (C3)	(test* OR valid* OR verif*)
Alignment (C4)	(align* OR link* OR trac* OR bridg* OR gap OR coordinat* OR coverage)

The selection of studies was conducted in steps of the protocol and the results of each step of the selection process can be seen in Figure 1. To avoid exclusion of relevant studies during abstract screening, the step of extended review was included for unclear cases of abstract screening. Extended review consisted mainly of the inspection of introduction and conclusion of studies. In the end, 80 studies were deemed relevant for the topic of the mapping study.

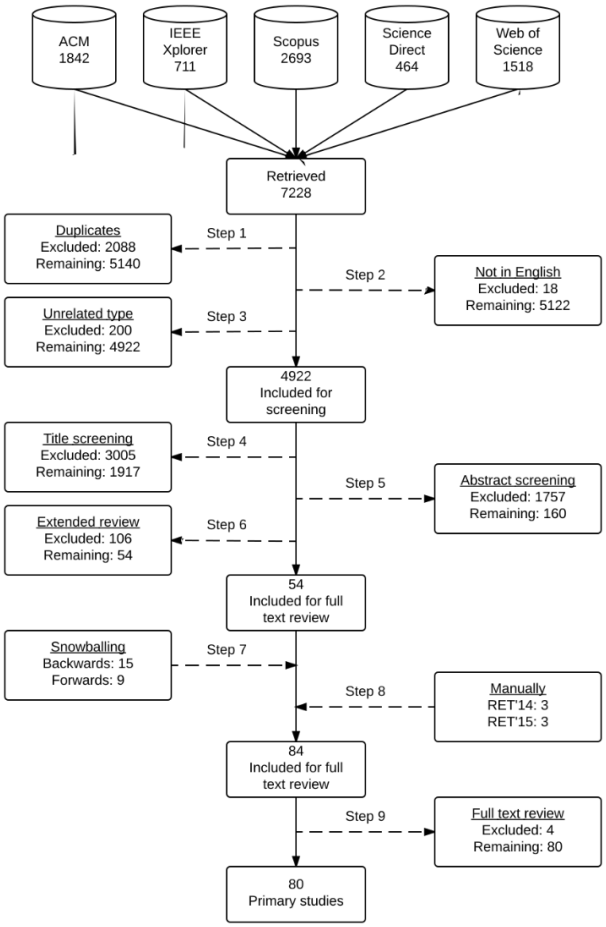

Figure 1. Inclusion and exclusion of studies

Table 2. Exclusion and inclusion criteria

Inclusion criteria
The study is a peer-reviewed study
The study focuses on development of software-intensive products, systems or services
The study includes any software development activity with the intention of alignment of requirements engineering with verification, validation or testing
Exclusion criteria
Duplicate, non-English studies, short papers, non-peer-reviewed studies
The study is not related to the software domain
The study does not clearly discuss the alignment of requirements engineering and validation, verification or testing of software intensive products, systems or services

3.4 Data Properties and Data Extraction

Data properties needed to answer the research questions are presented in Table 3. A more detailed definition of these data properties can be found in Appendix B in [17].

The method by Ivarsson and Gorschek [18] was used to evaluate the scientific rigor of research in the studies and their relevance for industrial practitioners.

Data extraction was first piloted with a second researcher. Any differences in the piloting results were discussed until agreement could be achieved before the data extraction was done on the full set of primary studies by both researchers. Results of the data extraction were once more compared before moving to analysis and synthesis of data.

Table 3. Data properties

	Data property	Research question
General data	Publication year	RQ1, RQ4
	Reference type	RQ1
Content data	Research focus	RQ2
	Reported domains, benefits and challenges	RQ3
	Application domain	RQ3
Methodological data	Research type (definition from [19])	RQ4
	Contribution type (adapted from [20])	RQ4
	Research method	RQ4
	Rigor and relevance (used method from [18])	RQ5

3.5 Data Analysis and Interpretation

When the data had been extracted, the data was analyzed and synthesized to answer the research questions and to structure the body of knowledge. Descriptive statistics were used when investigating frequencies and qualitative synthesis [21] regarding benefits, challenges, and how the studies support alignment.

3.6 Threats to Validity

The four concerns of validity [16], that is, construct validity, internal validity, external validity, and reliability, have been considered.

Construct validity refers to the identification and selection of primary studies. The threats to construct validity were minimized by using three different search strategies and including only peer reviewed scientific publications. All publications published up to the date of search were considered, and an expert was consulted regarding the list of primary studies. Same expert was consulted regarding the concept of alignment when constructing the inclusion criteria. The criteria were also subjected to peer review. Abstract screening was done in a conservative way to minimize the risk of excluding relevant studies. An abstract can be of poor quality, misleading or miss important information [9], thus, the extended review was deemed necessary.

Threats to internal validity concern whether wrong conclusions can be made from the data. The data extraction protocol vas reviewed by and piloted with a second researcher. Through the piloting, the researchers could get a common understanding and gave the opportunity to improve the data extraction protocol. In overall three researchers were involved during the study, thus reducing researcher bias.

External validity refers to the extent to which findings are generalizable within and beyond the scope of the study. The purpose is not to generalize the results beyond software development or alignment of RE and ST.

Reliability refers to the repeatability of the study. The search process must be repeatable by strict implementation of research protocol. However, classification and synthesis of primary studies is a more subjective process and relates to the research creativity and innovation. A protocol to follow in search, inclusion and exclusion of studies, and in data extraction, was devised and reviewed by a second researcher before the conducting of the mapping study. The synthesis of data and classification of studies, was also subjected to peer review, however, another researcher might device completely different classification just by taking another perspective.

4. FINDINGS

This section presents the findings regarding each research question. Due to space restrictions, only the studies mentioned in this paper are listed in the reference list. The full list of

primary studies can be found in the Appendix A of [17] and the full version of data extraction result is available over the web[1].

4.1 Intensity of Research (RQ1)

The publication trend is presented in Figure 2. Starting from 1998, there has been great variance in number of publications from year to year.

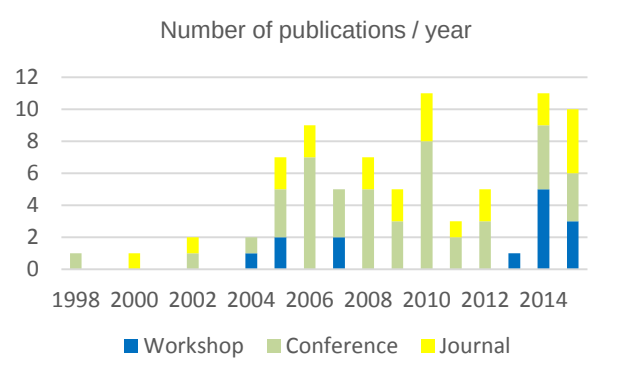

Figure 2. Publication trend

The most prominent publication type is conference proceeding which constitutes 72.5% (58/80) of all publications. The intensity of conference proceedings is highest in the period 2006 – 2010 with 26 publications. The number of workshop and journal publications is much lower during this period. The most recent 5-year period shows a more even number of conference, workshop, and journal publications (12, 12 and 9 publications respectively). Despite the slight decrees of publications in the recent years, overall, the research interest on the topic is on the increase. The recent dedicated RET workshops [8] [9] are an indication of this.

The distribution of publications in different venues is 22 journal publications, 44 conference proceedings, and 14 studies have been published through different workshops. Inspecting the different venues shows that the 22 journal publications have been published in 14 different venues. All the conference proceedings come out of 36 different conferences and the 14 workshop publications comes from 6 different workshops. The publications are very scattered in different venues.

4.2 Focus of Research (RQ2)

A high-level overview of the primary studies shows that the greatest number of studies (84%) were technical in nature presenting techniques and methods that development organizations can apply for better alignment of RE and ST. 10 studies (12.5%) focus on practices to apply in the development process of which six also list challenges and two benefits. Two of the 80 studies focus on the assessment of alignment and one focus on theory construction. The primary studies were also grouped based on how alignment is proposed to be supported. The resulting eight categories are shown in Table 4.

[1] https://figshare.com/articles/New_draft_item/4962758

Table 4. Categorization of techniques and approaches to support alignment of RE and ST

Group	Description	#	The studies
V&V in RE	These studies support alignment by presenting techniques or approaches that focus more on verification and validation of requirements in the early phases of the development process	12	S5, S12, S27, S28, S31, S36, S37, S42, S43, S60, S78, S79
Test generation from requirements specification	These studies support alignment by presenting techniques or approaches that focus on generation of tests, test models, or test cases from the requirements specification.	25	S6, S8, S11, S17, S18, S22, S24 – S26, S33, S38, S41, S44, S49, S53, S56, S61, S62, S66, S67, S70, S71, S73 – S75
Improved traceability	These studies support alignment through techniques or approaches to improve traceability between tests and requirements or traceability between models on different abstraction level	12	S7, S9, S13, S15, S23, S46, S47, S52, S57, S58, S64, S76
Improved testing	These studies support alignment by presenting ways of improving testing but they do not address test generation	3	S14, S65, S69
Formalization of the requirements	These studies support alignment by presenting techniques or approaches for the formalization of the requirements, which can then facilitate generation of tests	10	S10, S16, S30, S32, S39, S40, S45, S48, S55, S59
Practices to support alignment	These studies support alignment by presenting practices that can be applied in the way of working in the development project	4	S3, S4, S34, S54
Assessment of alignment	These studies support alignment by presenting frameworks for the	2	S1, S51

	assessment of alignment efforts in practice		
Other approaches	This group contains studies that support alignment through other approaches that could not be included in any of the other groups	5	S20, S21, S29, S77, S80

V&V in RE: This group (12 studies) emphasize early verification and validation (V&V). It is recommended that testing activities should be performed as soon as there is something to test [22] [23]. Leaving V&V activities to the end of RE would mean that V&V would have to consider the whole set of requirements after a time, which requires stakeholders to revisit the requirements to recollect their context and rationale [24].

Test generation from requirements specification: This group contains the greatest number of studies (25 studies). These studies focus on supporting alignment through test generation from requirements specification. Model-based testing (MBT) is a technique in which test cases are generated from behavioral models [25]. For example, [25], [26], and [27] propose to support alignment through test case generation from requirements specified as models. Not all studies start, however, with already specified requirements. [28] and [29] present methods to generate model-based test cases from natural language (NL) requirements to bring RE and ST closer. Use cases can also be used to deal with ambiguities of NL requirements. Techniques to generate test cases from use cases to support alignment has been presented in e.g., [30] and [31]. Goal orientation cannot only be used for understanding the domain and organizational setting, but can also be used for the derivation of test cases. [32] and [33] support alignment through this approach.

Improved testing: This group (3 studies) supports alignment by focusing more on the improvement of testing, but do not discuss test generation. [34] proposes to annotate source code with goals and events that are emitted when source code is run and can be compared to plans of how goals are achieved. [35] presents an idea for minimizing the gap between RE and ST through better information gathering in the RE phase, improving requirements quality, and better usage of this information in the ST phase reducing the number of test scenarios and test cases. [36] examines possible strategies of combining use cases and usage based testing through transformation of UCs or extension of UCs to facilitate testing.

Formalization of the requirements: Formalization of the requirements is important for test generation, especially when automation of test generation is pursued. The studies of this group (10 studies) supports alignment by facilitating test generation through formalization of the requirements. A formal specification gives a more precise and unambiguous description of the informal requirements [37]. This allows for formal verification of the requirements that can guarantee that the implementation conform to specification [38]. This is of great importance in safety critical domains, such as the automotive industry, that must meet safety standards [39].

Improved traceability: Traceability is important to understand the impact of change during the development [40]. Thus, the improvement of traceability is one way to support alignment and 12 studies focus more on this aspect. Traceability does not concern only test traceability to requirements or traceability between items in a model, but also traceability between items in different models and on different abstraction levels [40].

Practices to support alignment: Four studies focus on practices, i.e., practices in the way of working, to support alignment. The most important themes throughout these studies are the linking of people and linking of processes through cooperation and communication. Involvement of testers in RE is one such practice [5] [22].

Assessment of alignment: These studies, [1] and [13], provide a method for the assessment of RE and ST alignment (REST-bench) and thus provide improvement opportunity for better alignment. The assessment framework was introduced in [1] and in the later study [13] the assessment framework was successfully validated in real industrial settings. The results also showed that large organizations with plan driven development processes can benefit more from alignment than small agile projects.

Other approaches: This group contains studies that cannot be grouped in the previous categories. Example of such studies are, [41] that address the issue of alignment with a model for defining and validating RE quality in a precise and systematic way, and [42] presenting a tool to model NFRs to support alignment.

4.3 Domains, Benefits and Challenges (RQ3)

Different ways of improving alignment of RE and ST have been proposed in a variety of domains with varying size of application or organization. However, the three most frequently reported domains are large industrial domains: automotive industry (15 studies), telecommunication and mobile devices (11 studies), and avionics (6 studies).

Five studies, [6], [5], [43], [44], and [45], reported challenges out of which 62 individual challenges were extracted. The challenges could be grouped into eight categories for which the categories of [43] were used. The categories are: organization and process related challenges, people related challenges, tool related challenges, RE related challenges, ST related challenges, change management related challenges, traceability related challenges, and metrics related challenges. Short description of categories is presented below.

Organization and process related challenges: The biggest challenge is posed by the separation of activities which, especially in big organizations, translates to increased distance between RE and ST, e.g., through outsourcing. This calls for better communication. The applied development process might not be well suited for closer collaboration of requirements engineers and testers, but still the change of processes or

process improvement initiatives can make it difficult to maintain alignment.

People related challenges: These manifest themselves in people's skills or willingness to cooperate with other units. Also, the missing knowledge of the work of others was mentioned as a challenge.

Tool related challenges: Tools may not be appropriate for realizing alignment and many organizations use a diversity of different tools for different activities. It may be difficult to make all these heterogeneous tools fit together. The use of tools was also mentioned as a challenge.

RE related challenges: The defining of good, verifiable, and complete requirements is a concern and especially regarding quality requirements, the NFR. The large number of requirements also pose a challenge. Cooperation with other units was mentioned as a challenge in RE. Requirements engineers may not consider testing sufficiently when defining requirements and testers' involvement in RE may be minimal.

ST related challenges: The NFRs are a challenge in ST also since they are many times difficult to verify. Like for RE, the cooperation with other units has been reported as a challenge. Large amount of test cases and test coverage was seen as a challenge. The definition of a good testing process has also been mentioned as a n issue.

Change management related challenges: The evolving and changing requirements are a challenge especially in the cases where there is a lack of change management strategy. It was also reported that the updating of information can pose a challenge for alignment since in many cases this updating of information may require additional resources and is left undone. Additionally, it was reported that sometimes it may be hard to find responsible people.

Traceability related challenges: Despite existing tools and practices, maintaining traceability in practice was seen as a challenge. Links between tests and requirement are sometimes missing and large legacies may imply that many test cases do not have requirements linked. The different abstraction levels pose a challenge and poor quality of requirements have a negative effect on traceability.

Metrics related challenges: There is a lack of metrics for alignment, lack of experience regarding metrics, and there is a need for appropriate performance metrics that includes both operative and top management levels.

Many studies report on perceived benefits and give convincing argumentation for the benefits of alignment, others reference the reported benefits, but only two studies report on benefits, [6] and [5]. 16 individual benefits were reported that are categorized in the following four categories: benefits for project quality, benefits for project quality, benefits for RE, and benefits for ST.

Benefits for project quality: It was found that better alignment led to improved control and estimation of project cost and schedule. Better traceability improved the efficiency of change management and the practice of linking testers to requirements owners made it possible to progress despite lower quality requirements. Letting testers do requirements suggestions was another proposed practice that showed that it was more likely that testing will be completed both in time and planned scope.

Benefits for product quality: It was found that improving alignment increased the reliability of test results and thus also improved the likelihood of products satisfying customer needs. In addition, better test traceability to requirements resulted in more efficient error removal.

Benefits for RE: The practice of early tester participation improved the quality of the requirements and helped surface deficiencies and omissions in requirements. Tester participation in requirements reviews helped to identify requirements that would be difficult to validate and thus pay more attention to those early on.

Benefits for testing: The practice of early tester participation was beneficial also for testing. It led to that testing activities are properly considered in planning and the domain and system knowledge is improved. Letting testers make requirements suggestions improved testability and possibilities for automation, and reduced the testing effort. Test coverage was improved by improved test traceability to requirements and the amount of assumptions made by testers were reduced by linking testers to requirements owners.

From the amount of reported challenges, it could be said that it is well understood what needs to be in place for better alignment of RE and ST, however, very little evidence of the benefits of improved alignment exists in the literature.

4.4 Methodological data (RQ4)

Two research types, solution proposal and evaluation research, are most frequently used in the studies. These constitute over 83.7% (67/80) of all research types. The number of solution proposals is 29, which is 36.2% of all studies, and the number of evaluation research is 38, which is 47.5% of the studies. There are just a few cases of validation research (4), experience reports (4), and opinion papers (3). Two studies conducted philosophical argumentation.

The most frequently applied research method is case study, 42.5% (34/80), which is in accordance with the high number of evaluation research. However, what is even more noticeable is that the number of studies not reporting any research method is even higher, 45% (36/80) of studies. Four studies used a mixed method and three of the studies conducted experimentation. Two studies used survey as research method and one study reported doing action research.

The number of extracted contributions (98) is higher than the number of studies (80), since some studies provide more than one kind of contribution. The by far most frequent contribution type is framework/method/technique, at 60% (59/98) of all extracted contributions. The second highest contribution type is tool and the third highest lessons learned, representing 16.3% (16/98) and 13.3% (13/98) of all contributions respectively. The contribution types advice/implication and model are provided in three studies both, two studies contribute with metrics, and there is one study providing a theory and one providing guidelines.

As most studies present technical descriptions, the most frequent contributions type is framework/method/technique. This indicates that researchers and practitioners have a good idea about what is needed for better alignment. These contributions are distributed equally amongst research types solution proposal and evaluation research. However, even though there are some studies reporting lessons learned, concrete guidelines which support practitioners in practice and studies presenting advice/implication are very scares. Even though research has been done in industrial context, it is not clear what method or techniques is suitable in what context and why.

4.5 Rigor and Relevance (RQ5)

The bubble chart of Figure 3, visualizing the result of the evaluation of rigor and relevance, shows that there is an approximately equal distribution of studies with both low and high relevance, whilst most studies exercise low scientific rigor. This is also reflected by the finding that most studies do not report using any research method (RQ4). The highest number of studies can be found in the sector of both low rigor and low relevance and a high number of studies also in the sector of low rigor and high relevance. As stated in RQ4, distribution of contribution type framework/method/technique is equal over evaluation research and solution proposals. A small portion of the studies exercise both high scientific rigor and have high relevance for industry.

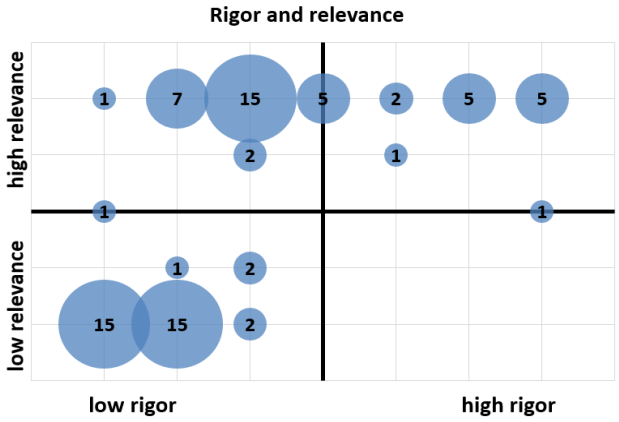

Figure 3. Scientific rigor and industrial relevance

Low relevance and low rigor: There are 35 studies in this category. The studies consist mostly of solution proposals that have not yet been evaluated in practices. They don't report using any specific research method. Thus, their scientific rigor and applicability is limited.

High relevance and low rigor: 25 studies can be found in this group. These studies consist mostly of case studies in which different approaches to link RE and ST are evaluated in industrial context.

High relevance and high rigor: There are 13 studies in this category. Most of these studies also consist of evaluations done in industrial case studies, however, these studies exercise higher scientific rigor. These studies are also more recent. Eight of the studies of this group have been published in 2014 and 2015.

5. DISCUSSION

What is the intensity of research on the alignment of RE and ST? (RQ1): Results showed that the greatest number of studies have been published between the period of 2006 – 2010 with a slight decrease in the following 5-year period. Overall, the publication trend shows that there is growing interest in the area among both practitioners and researchers. Furthermore, the fact that there is an annual dedicated workshop, starting in 2014, shows that there is an interest towards the topic.

What are the foci of research regarding alignment of RE and ST and what is proposed to support alignment in the research? (RQ2): Studying how the publications support alignment showed that most of the publications contain more technical presentations of different techniques to improve alignment through modeling of requirements, formalization of requirements, earlier verification and validation, improved traceability and generation of tests based on requirements. All these different proposals of how to support alignment show that there is no single way of how alignment can be achieved. Some studies presented practices that can be applied in the development process to support alignment. This also shows that practitioners and researchers know how to change or improve practices in place for better alignment. A few of these studies investigated also the challenges regarding these practices and alignment in general together with some benefits. An assessment framework for assessment of alignment was introduced and validated [1] [13]. This framework could also be used to elicit improvement opportunities. If this evaluation framework proves to be effective in both assessment of alignment and in eliciting improvement opportunities, it could be expected that we will get more data regarding alignment of RE and ST in the near future. Even though many approaches are proposed for RE and ST alignment, there is a lack of comparative studies. Meaning that it is not clear what approach is suitable in what context and why.

In what context has alignment of RE and ST been applied and what effects in terms of benefits and challenges have been reported? (RQ3): Results showed that alignment techniques and practices had been applied in a variety of domains but the automotive industry was the most frequently reported domain. The automotive industry deals with large and complex systems with a great number of requirements and, in addition, the automotive software is safety critical. The third most frequently reported domain was avionics, which is also safety critical. Thus it would seem that alignment of RE and ST is particularly important for big safety critical domains.

There are many diverse challenges reported regarding different aspects and different phases of the development process. Some studies of challenges focus on the whole of the development process while others report challenges regarding different practices. The supporting evidence for benefits are presented

only in two studies. The number of challenges also indicate that it is well known what is needed for alignment but the low number of reported benefits show that the effects of alignment have not yet been studied to greater lengths.

What are the most frequently applied research types and research methods, and what kind of contributions are provided by the studies? (RQ4): Methodological data showed that almost half of the studies had research type evaluation research and almost an equal amount presented solution proposals. Only a few studies were of type validation research and experience report. Almost half of the studies did not report clearly on any research method and from those that did, the majority reported doing case studies. In that light, almost half of the studies proposed solutions and evaluated those solution proposals in industrial settings. Also this indicates that the concept of alignment is understood and it is known what kind of techniques could be used for better support of alignment. However, even if there are some studies presenting lessons learned, the number of contribution types advice/implications and guidelines are very low. Putting all this together implies that alignment efforts have been tried out in industrial setting and evaluation of the different proposals have been done, but not yet to the extent that it would be possible to report what works in what context. With more comparative and rigorous evaluation of techniques and practices, the number of lessons learned might rise, and eventually also the number of guidelines.

What is the scientific rigor and industrial relevance of the research? (RQ5): The results showed that most research exercise low scientific rigor. The distribution of studies between high and low industrial relevance was approximately equal but the biggest number of studies could be found in the sector of both low industrial relevance and low scientific rigor. This finding supports the discussion in the previous research question. With more rigorous research it could be expected that we gain a better understanding of what works for alignment, why, and in what context.

Implications: The studies of alignment have been published in many different journals, conferences, and workshops. As such this study provides an up to date overview of all existing knowledge on the area and points to sources where future studies can be found. For researchers, it provides a basis for future studies. Practitioners can use this study to better understand approaches, benefits and challenges of alignment of RE and ST.

Limitation: Since this is a mapping study, the aim is to cover the whole of a certain research area. As such, the biggest limitation regards the finding of all relevant studies. Three search strategies were used in this study: database search snowballing, and manual search. Even though snowballing was used to complement database search and not miss any relevant studies, a manual search of the dedicated workshop was required to find additional known relevant studies. Another limitation regarding the finding of all relevant studies comes from the search phrase. The key-words in the alignment category of the search phrase (C4 in Table 1) were elicited from the ways alignment was referred to in the known studies. There might be additional ways to refer to the same concept. Thus, it cannot be guaranteed that all relevant studies have been found. Categorization of how alignment is supported is based on an existing categorization. This categorization can be done in a different way taking a different perspective.

Future directions: The effects of alignment in terms of benefits have been reported to lesser extent. Many studies present convincing arguments for the benefits of alignment but only two studies report found benefits. There are many proposed methods and techniques, however, it is not clear what technique is suitable in what development context and why. Thus it would seem that the research area is in its early stages. Studying the effects of alignment and evaluating approaches and practices in real industrial context with more scientific rigor could reveal what works in which type of organizations, in what context, and why. This could provide more guidelines in the future. Overall, concrete metrics and guidelines to support practitioners in industry are required.

6. CONCLUSIONS

This paper presents the results of a SMS that structures and classifies the research regarding alignment of RE and ST providing up to date information. 80 studies contributing to the topic were found through search of databases, snowballing and manual search up till the end of 2015 that could answer the question: What is the current state of research on the topic of alignment of RE and ST?

This alignment of RE and ST activities can be supported through many different techniques including early V&V, formalization of the requirements facilitating test generation, test generation form requirements specification, and improved traceability. Ways of supporting closer collaboration of requirements engineers and testers have been identified as practices that can improve the alignment of RE and ST. Better communication between people, organizational units, and organizations is of outmost importance since the flow of information is the key for better alignment.

Challenges have been identified on all levels of the development ranging from organizational issues to people related challenges to challenges regarding tools and applied practices both in RE and ST. Amongst the benefits are improved project and product quality with greater likelihood of meeting schedule and budget demands, and a greater likelihood of more satisfied customers. However, although the many challenges have been identified, the effects in terms of benefits have been less explored.

The primary studies of this mapping study have been published in many different journals or magazines, conferences and workshops and thus this SMS provides an inventory of relevant studies. This up to date information of the topic provides a basis for future studies for researchers. Practitioners can use this study to better understand approaches, benefits and challenges of alignment of RE and ST. Examining methodological data and the rigor and relevance of the primary studies shows that more rigorous research is needed to

examine how and why the different approaches to support alignment works and in what context.

7. REFERENCES

[1] M. Untrkalmsteiner, R. Feldt and T. Gorschek, "A taxonomy for requirements engineering and software test alignment," Acm Transactions on Software Engineering and Methodology, vol. 23, no. 2, p. 16, 2014. S1.
[2] R. Pressman, Software Engineering: A Practitioner's Approach, 7th ed., Bookman, 2011.
[3] M. Broy, I. H. Kruger, A. Pretschner and C. Salzmann, "Engineering automotive software," Proceedings of the IEEE, vol. 95, no. 2, pp. 356-373, 2007.
[4] E. Bjarnason, K. Smolander, E. Engström and P. Runeson, "A Theory of Distances in Software Engineering," Information and Software Technology, 2015. S2.
[5] E. J. Uusitalo, M. Komssi, M. Kauppinen and A. M. Davis, "Linking requirements and testing in practice," in International Requirements Engineering, 2008. RE'08. 16th IEEE, 2008. S4.
[6] J. Kukkanen, K. Väkeväinen, M. Kauppinen and E. Uusitalo, "Applying a systematic approach to link requirements and testing: A case study," in Software Engineering Conference, 2009. APSEC'09. Asia-Pacific, 2009. S3.
[7] Z. A. Barmi, A. H. Ebrahim and R. Feldt, "Alignment of requirements specification and testing: A systematic mapping study," in Software Testing, Verification and Validation Workshops (ICSTW), 2011 IEEE Fourth International Conference on, 2011.
[8] M. Felderer, E. Bjarnason, M. Staats, M. Morandini, M. Borg and M. Unterklamsteiner, in 1st IEEE International Workshop on Requirements Engineering and Testing, RET 2014, Karlskrona, Sweden, 2014.
[9] E. Bjarnason, M. Morandini, M. Borg, M. Unterkalmsteiner, M. Felderer and M. Staat, in 2nd International Workshop on Requirements Engineering and Testing (RET 2015), 2015.
[10] W. W. Royce, "Managing the development of large software systems," Proceedings of IEEE WESCON, vol. 26, no. 8, pp. 328-338, 1970.
[11] S. L. Pfleeger and J. M. Atlee, Student Study Guide for Software Engineering: Theory and Practice, Prntice Hall, 2009.
[12] M. Fowler and J. Highsmith, "The agile manifesto," Software Development, vol. 9, no. 8, pp. 28-35, 2001.
[13] M. Unterkalmsteiner, T. Gorschek, R. Feldt and E. Klotins, "Assessing requirements engineering and software test alignment—Five case studies," Journal of Systems and Software, vol. 109, pp. 62-77, 2015. S51.
[14] K. Petersen, S. Vakkalanka and L. Kuzinarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," Information and Software Technology, vol. 64, pp. 1-18, 2015.
[15] B. A. Kitchenham, D. Budgen and O. Pearl Brereton, "Using mapping studies as the basis for further research–a participant-observer case study," Information and Software Technology, vol. 53, no. 6, pp. 638-651, 2011.
[16] C. Wohlin, "Guidelines for Snowballing in Systematic Literature Studies and a Replication in Software Engineering," in 18th International Conference on Evaluation and Assessment in Software Engineering (EASE 2014), 2014.
[17] P. Karhapää, "Alignment of Requirements Engineering and Software Testing (Master's thesis). Retrieved from http://jultika.oulu.fi/Record/nbnfioulu-201611103000".
[18] M. Ivarsson and T. Gorschek, "A method for evaluating rigor and industrial relevance of technology evaluations," Empirical Software Engineering, vol. 16, no. 3, pp. 365-395, 2011.
[19] R. Wieringa, N. Maiden, N. Mead and C. Rolland, "Requirements engineering paper classification and evaluation criteria: a proposal and a discussion," Requirements Engineering, vol. 11, no. 1, pp. 102-107, 2006.
[20] M. Shaw, "Writing good software engineering research papers: minitutorial," in Proceedings of the 25th international conference on software engineering, 2003.
[21] D. S. Cruzes and T. Dybå, "Recommended steps for thematic synthesis in software engineering," in 2011 International Symposium on Empirical Software Engineering and Measurement, 2011.
[22] D. Graham, "Requirements and testing: Seven missing-link myths," IEEE Software, vol. 19, no. 5, pp. 15-17, 2002. S34.
[23] R. C. Martin and G. Melnik, "Tests and requirements, requirements and tests: A möbius strip," IEEE software, vol. 25, no. 1, pp. 54-59, 2008. S43.
[24] L. O. Lobo and J. D. Arthur, "Local and global analysis: Complementary activities for increasing the effectiveness of requirements verification and validation," Proceedings of the 43rd annual Southeast regional conference, vol. 2, pp. 256-261, 2005. S28.
[25] F. Abbors, D. Truscan and J. Lilius, "Tracing requirements in a model-based testing approach," in Advances in System Testing and Validation Lifecycle, 2009. VALID'09. First International Conference on, 2009. S46.
[26] M. F. Granda, N. Condori-Fernández, T. E. Vos and O. Pastor, "Towards the automated generation of abstract test cases from requirements models," in Requirements Engineering and Testing (RET), 2014 IEEE 1st International Workshop on, 2014. S66.
[27] M. Felderer, P. Zech, F. Fiedler and R. Breu, "A tool-based methodology for system testing of service-oriented systems," in Advances in System Testing and Validation Lifecycle (VALID), 2010 Second International Conference on, 2010. S11.
[28] E. Sarmiento, J. C. S. do Padro Leite and E. Almentero, "C&L: Generating model based test cases from natural language requirements descriptions," in Requirements Engineering and Testing (RET), 2014 IEEE 1st International Workshop on, 2015. S75.
[29] V. A. de Santiago Júnior and N. L. Vijaykumar, "Generating model-based test cases from natural language requirements for space application software," Software Quality Journal, vol. 20, no. 1, pp. 77-143, 2012. S25.
[30] C. Nebut, F. Fleurey, Y. Le Traon and J. M. Jézéquel, "A requirement-based approach to test product families," in International Workshop on Software Product-Family Engineering, 2003. S8.
[31] M. El-Attar and J. Miller, "Developing comprehensive acceptance tests from use cases and robustness diagrams," Requirements engineering, vol. 15, no. 3, pp. 285-306, 2010. S22.
[32] D. C. Ngyuyen, A. Perini and P. Tonella, "A goal-oriented software testing methodology," in International Workshop on Agent-Oriented Software Engineering, 2007. S6.
[33] R. Hassan, M. Eltoweissy, S. Bohner and S. El-Kassas, "Formal analysis and design for engineering security automated derivation of formal software security specifications from goal-oriented security requirements," IET software, vol. 4, no. 2, pp. 149-160, 2010. S24.
[34] K. Winbladh, T. A. Alspaugh, H. Ziv and D. J. Richardson, "An automated approach for goal-driven, specification-based testing," in 21st IEEE/ACM International Conference on Automated Software Engineering (ASE'06), 2006. S14.
[35] K. Smilgyte and R. Butleris, "Towards Software Testing Process Improvement from Requirements," in International Conference on Information and Software Technologies, 2012. S65.
[36] B. Rgnell and P. Runeson, "Combining Scenario-based Requirements with Static Verification and Dynamic Testing," in REFSQ, 1998. S69.
[37] T. Miller and P. Strooper, "A case study in model-based testing of specifications and implementations," . Software Testing, Verification and Reliability, vol. 22, no. 1, pp. 33-63, 2012. S5.
[38] H. Post, C. Sinz, F. Merz, T. Gorges and T. Kropf, "Linking functional requirements and software verification," in 2009 17th IEEE International Requirements Engineering Conference, 2009. S27.
[39] S. Siegl, K. S. Hielscher and R. German, "Model based requirements analysis and testing of automotive systems with timed usage models," in 2010 18th IEEE International Requirements Engineering Conference, 2010. S30.
[40] S. Ibrahim, N. B. Idris, M. Munro and A. Deraman, "A Software Traceability Validation for Change Impact Analysis of Object Oriented Software," Software Engineering Research and Practice, pp. 453-459, 2006. S9.
[41] H. Femmer, J. Mund and D. M. Fernández, "It's the activities, stupid!: a new perspective on RE quality," in Proceedings of the Second International Workshop on Requirements Engineering and Testing, 2015. S77.
[42] J. Zou and C. J. Pavlovski, "Control cases during the software development life-cycle," in 2008 IEEE Congress on Services-Part I, 2008. S20.
[43] G. Sabaliauskaite, A. Loconsole, E. Engström, M. Unterkalmsteiner, B. Regnell, P. Runeson, T. Gorschek and R. Feldt, "Challenges in aligning requirements engineering and verification in a large-scale industrial context," in International Working Conference on Requirements Engineering: Foundation for Software Quality, 2010. S19.
[44] E. Bjarnason, M. Unterkalmsteiner, E. Engsström and M. Borg, "An industrial case study on test cases as requirements," in International Conference on Agile Software Development, 2015. S50.
[45] E. Bjarnason, P. Runeson, M. Borg, M. Unterkalmsteiner, E. Engström, B. Regnell and R. Feldt, "Challenges and practices in aligning requirements with verification and validation: a case study of six companies," Empirical Software Engineering, vol. 19, no. 6, pp. 1809-1855, 2014. S54.