Pós-Graduação em Ciência da Computação

Bruno Falcão de Souza Cartaxo

A Model to Transfer Knowledge from Research to Software Engineering Practice Based on Rapid Reviews and Evidence Briefings

Recife
2018

Bruno Falcão de Souza Cartaxo

A Model to Transfer Knowledge from Research to Software Engineering Practice Based on Rapid Reviews and Evidence Briefings

Work presented to the Computer Science Postgraduate Program at the Center of Informatics of the Federal University of Pernambuco as partial requirement to obtain the degrre of Doctor degree in Computer Science.

Advisor: Sérgio Castelo Branco Soares
Gustavo Henrique de Lima Pinto

Recife
2018

Catalogação na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

C322m Cartaxo, Bruno Falcão de Souza
A model to transfer knowledge from research to software engineering practice based on rapid reviews and evidence briefings / Bruno Falcão de Souza Cartaxo. – 2018.
174 f.: il., fig., tab.

Orientador: Sérgio Castelo Branco Soares.
Tese (Doutorado) – Universidade Federal de Pernambuco. CIn, Ciência da Computação, Recife, 2018.
Inclui referências e apêndices.

1. Engenharia de software. 2. Revisões sistemáticas. I. Soares, Sérgio Castelo Branco (orientador). II. Título.

005.1 CDD (23. ed.) UFPE- MEI 2018-114

Bruno Falcão de Souza Cartaxo

“A Model to Transfer Knowledge from Research to Software Engineering Practice Based on Rapid Reviews and Evidence Briefings”

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutora em Ciência da Computação.

Aprovado em: 11/05/2018.

_____

Orientador: Prof. Sérgio Castelo Branco Soares

BANCA EXAMINADORA

_____

Prof. Fabio Queda Bueno da Silva
Centro de Informática /UFPE

_____

Prof. Fernando José Castor de Lima Filho
Centro de Informática / UFPE

_____

Prof. Guilherme Horta Travassos
Programa de Engenharia de Sistemas / UFRJ

_____

Profa. Maria Teresa Baldassarre
Departamento de Ciência da Computação / UNIBARI

_____

Profa. Tayana Uchôa Conte
Departamento de Ciência da Computação / UFAM

This thesis is dedicated to every person I once shared a moment with.

ACKNOWLEDGEMENTS

To master Sérgio Soares, for the space.

To Gustavo Pinto, for the braveness.

To everybody from INES laboratory, for the presence.

To Rafaella Queiroga, for the love.

To Aucélia and Mauro, for the life.

To Marina, Nara, Carlinhos and Vitinho, for the comfort.

To everybody else, for the opportunity.

ABSTRACT

Inspired by the promising results of Evidence Based Medicine (EBM), Evidence Based Software Engineering (EBSE) was introduced more than one decade ago. One of EBSE's main goal is to integrate the best research evidence with Software Engineering (SE) practice. However, some claim Secondary Studies (SSs) — one of the main products of EBSE — lack connection with SE practice. To better understand this claim, the first contribution of this thesis is an analysis on how SSs are related to problems faced in SE practice. We introduced a coverage technique to compare topics investigated by SSs and issues practitioners report in Stack Exchange Q&A websites. We could observe a low coverage (15.6%), which corroborates with the claims about lack of connection between SSs and SE practice. Motivated by the findings of our first contribution, and inspired by recent advances in EBM, we propose a model to transfer knowledge from SSs to SE practice. The model adhere to two of the most prominent approaches proposed in EBM: Rapid Reviews (a lightweight research method that focus on understanding practitioners problems, and mitigating them with evidence delivered in a timely manner), and Evidence Briefings (one-page documents summarizing the main findings of an empirical study). To evaluate the EBs, we created 12 of them based on one SS each, and conducted two surveys. One with the authors of the SSs, and other with practitioners that reported issues considered as covered by the SSs. Both practitioners (32) and researchers (7), positively evaluated the EBs. We also invited authors of papers published in four SE conferences to create EBs of their papers. We conducted another survey to evaluate their perceptions. The positive results show that EBs can play a role as a transfer medium transferring knowledge from research to SE practice. To evaluate the applicability of RRs in a real context, we conducted two RRs in two software companies. The problems reported by the practitioners were due to low customer collaboration (Company 1) and low team motivation (Company 2). Following RRs strategies, we searched for evidence that might help to mitigate the problems, synthesized the results, and introduced the findings to practitioners through EBs. Through a series of interviews, we evaluated the practitioners perceptions about the introduction of RRs to support decision-making. Our results show that practitioners are rather positive about RRs. In summary, and based on the positive results we obtained with the six empirical studies we conducted, we believe the model we proposed to transfer knowledge based on RRs and EBs has a promising future as an alternative to make research closer to SE practice.

Key-words: Rapid Reviews. Knowledge Transfer. Evidence Briefings. Systematic Reviews. Empirical Software Engineering. Evidence Based Software Engineering.

RESUMO

Há mais de uma década a Engenharia de Software Baseada em Evidência (ESBE) foi introduzida, inspirada nos resultados promissores da Medicina Baseada em Evidência (MBE). Um dos principais objetivos da ESBE é integrar as melhores evidências com a prática da Engenharia de Software (ES). No entanto, alguns pesquisadores afirmam que estudos secundários – um dos principais produtos da ESBE – apresentam uma falta de conexão com a prática da ES. A primeira contribuição dessa tese é analisar como estudos secundários se relacionam com os problemas enfrentados na prática. Definimos uma técnica de análise de cobertura que permite comparar os tópicos investigados pelos estudos secundários com os problemas reportados pelos profissionais da prática em *websites* da plataforma Stack Exchange. Podemos observar que a baixa taxa cobertura (15.6%) corrobora com argumentos no sentido da falta de conexão dos estudos secundários com a prática de ES. Inspirados em avanços recentes da MBE, propusemos um modelo para transferir conhecimento de estudos secundários para a prática. O modelo adere a dois conceitos que tem ganhado notoriedade na MBE: as *Rapid Reviews* (RRs), que são métodos que focam em entender e mitigar os problemas dos profissionais da prática fornecendo evidências científicas em tempo hábil; bem como os *Evidence Briefings* (EBs), que são documentos de uma página que resumem as principais evidências de um estudo empírico. Para avaliar os EBs, criamos 12 deles baseados em um estudo secundário cada e conduzimos dois *surveys*. Um survey com os autores (7) dos estudos secundários, e outro com profissionais (32) que reportaram problemas cobertos pelos estudos secundários. Também convidamos os autores de artigos publicados em quatro conferências para criar EBs dos seus artigos e através de um terceiro *survey* avaliamos suas percepções. Os resultados positivos mostraram que os EBs podem ter um papel importante como meio de transferência de conhecimento. Para avaliar a aplicabilidade de RRs, conduzimos duas delas em duas empresas de software. Os problemas reportados estavam relacionados à baixa colaboração do cliente (Empresa 1) e a baixa motivação da equipe (Empresa 2). Buscamos por evidências que pudessem mitigar os problemas, sintetizamos essas evidências e apresentamos os resultados aos profissionais por meio de EBs. Em uma série de entrevistas avaliamos as percepções dos profissionais sobre o uso de RRs. A percepção dos profissionais foi majoritariamente positiva. Visto os resultados positivos, acreditamos que o modelo de transferência de conhecimento que propusemos, baseado em RRs e EBs, pode aproximar a pesquisa da prática da ES.

Palavras-chaves: *Rapid Reviews*. *Evidence Briefings*. Revisões Sistemáticas. Transferência de Conhecimento. Engenharia de Software Empírica. Engenharia de Software Baseada em Evidência.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

Evidence Based Practice (EBP) aims at investigating the best research evidence in a given domain of expertise and integrating it to practice (MCKIBBON, 1998). Medicine was one of the pioneers establishing the Evidence Based Medicine (EBM). Following the promising results of EBM, many other research fields have also been adopting EBP, such as: psychology (ANDERSON, 2006), nursing (DICENSO; CULLUM; CILISKA, 1998), crime prevention (FARRINGTON et al., 2003), social work (WEBB, 2001), and education (DAVIES, 1999). In 2004, the Software Engineering (SE) community acknowledged the importance of EBP (KITCHENHAM; DYBA; JORGENSEN, 2004). According to Kitchenham et al., the goal of Evidence Based Software Engineering (EBSE) is:

> *"to provide the means by which current best evidence from research can be* ***integrated with practical experience*** *and human values in the decision-making process regarding the development and maintenance of software."* (KITCHENHAM; DYBA; JORGENSEN, 2004) (bold emphasis added)

Considering this goal, it is no coincidence that Secondary Studies (SSs) are the main products of EBSE. It is argued that the knowledge aggregated in SSs is the most appropriate to be transferred to practice because years of EBP has taught that individual studies can often lead to very different conclusions compared to SSs (LAVIS et al., 2003). For instance, a single study comparing the mortality rates of for-profit and non-profit hospitals found a lower risk of death in for-profit hospitals, while a SS considering data from studies that summed up 26,000 hospitals and 38 millions of patients, found a higher risk of death in for-profit hospitals (DEVEREAUX et al., 2002).

After more than a decade of contributions, EBSE has new studies being conducted on regular basis (SILVA et al., 2011; BORGES et al., 2014; BORGES et al., 2015). However, despite its evolution, many claim that EBSE lacks connection with SE practice (HASSLER et al., 2014; SANTOS; SILVA, 2013; SILVA et al., 2011). To illustrate, an investigation with researchers specialized in EBSE revealed that lack of connection with industry is the sixth top barrier to SSs, from a total of 37 barriers listed by them (HASSLER et al., 2014). In the same direction, the study of Santos et al. (SANTOS; SILVA, 2013) presents a survey with 44 authors of 120 SSs, and only six of them affirmed their SSs had direct impact on industrial practice. In addition, a tertiary study identified that only 32 out of 120 SSs provide guidelines to practitioners. This situation is critical because it may indicate that EBSE has not been accomplishing its stated goal.

The EBM community also faced similar problem in its early days, and it is still facing it to some extent today (BEST et al., 1997; TRICCO et al., 2017; TRICCO et al., 2015). One of the most successful initiatives to mitigate the lack of connection between EBM

and practice is what has been called as Rapid Reviews (RRs) (TRICCO et al., 2015). They are a kind of SSs aiming to provide research evidence to support decision-making in practice, and in consequence, must be conducted taking into account the constraints inherent to practice. So, RRs usually deliver evidence in a more timely manner, with lower costs, reporting results through more appealing mediums, and more connected to practice, when compared to Full Secondary Studies (FSSs). By FSSs, we mean the traditional and more rigorous SSs, like: systematic mappings, meta-analyzes, and traditional systematic literature reviews (KITCHENHAM; CHARTERS, 2007). To achieve these goals, RRs omit or simplify some steps of FSSs, for instance: limiting literature search to a specific period or search engine; using just one person to screen primary studies; not conducting quality appraisal of primary studies; or presenting results with no formal synthesis (TRICCO et al., 2015).

Some studies were conducted to evaluate the impact of RRs methodological adaptations by comparing them with FSSs (ABOU-SETTA et al., 2016; CORABIAN; HARSTALL, 2002; BEST et al., 1997; TAYLOR-PHILLIPS et al., 2017; VELDE et al., 2011). Although there are more evidence reporting the similarity of results obtained with RRs and FSSs, there are also evidence on the opposite side. Thus, further investigations are needed to draw more conclusive results. In consequence, instead of considering RRs as replacements for FSSs, we believe RRs should be understood as a complementary EBP product. While FSSs are important to provide in-depth insights, RRs are important to easily and quickly transfer established knowledge to practice.

In parallel to RRs, and closely related to them, there is another movement in EBM to approximate research evidence to practice, the emergence of alternative mediums to report research results. For instance, the Contextual Summaries of Young et al., which are mediums limited to a one-page document (YOUNG et al., 2014); or the Briefings proposed by Chambers and Wilson, that summarize the main findings of empirical studies in one section on a cover page (CHAMBERS; WILSON, 2012); or even the Evidence Summaries of Khangura et al., which use an informative box separated from the main text to highlight the audience and nature of the medium (KHANGURA et al., 2012). Inspired by the characteristics of many alternative mediums in medicine, and together with an information design specialist, we proposed a medium we call Evidence Briefing (EB), which is one of the main contributions of this thesis as it is going to be shown later (CARTAXO et al., 2016).

1.1 RESEARCH GOAL

Due to many claims on the lack of EBSE connection with SE practice, and emerging strategies in EBM to deal with this problem, the goal of this thesis is to *propose*, *evaluate*, and *discuss* a model to transfer knowledge from secondary studies to software engineering

practice, based on Rapid Reviews and Evidence Briefings.

1.2 RESEARCH QUESTIONS

To achieve the stated goal, this thesis aims to answer the following General Research Questions (GRQs):

GRQ1: How secondary studies in software engineering cover the issues practitioners face in practice?

With this question, we investigate whether the SSs being conducted in SE can provide evidence to support decision-making towards the solution, or at least attenuation, of issues practitioners face in practice. Answering this question, we intend to provide evidence that may substantiate the claims that SSs lack connection with practice, corroborating the importance of the model we are proposing in this thesis.

GRQ2: What are the perceptions of researchers and practitioners on Evidence Briefings as transfer mediums?

With this question, we investigate the perceptions of two groups of people highly involved in a possible initiative to use EBs as transfer medium: researchers and practitioners. By EBs, we mean one-page documents with the main findings of a SS. Answering this question is particularly important to understand whether the EBs have chance to be accepted as a knowledge transfer medium, since, according to Rogers (ROGERS, 2003), the perception of individuals about an initiative is one of the main predictors of its adoption.

GRQ3: What are the perceptions of practitioners on the knowledge produced with Rapid Reviews as transfer objects?

With this question, similarly to GRQ2, we investigate the applicability of RRs in SE practice. The intention is, together with practitioners, to diagnose their practical problems, to conduct RRs based on scientific evidence, to report the results through EBs, and to evaluate the perception of practitioners about this entire process.

1.3 PHILOSOFICAL STANCE

To answer research questions, a researcher has to define and report the method s/he will follow. An important step, however, is to present her/his philosophical stance since it will influence which methods will be chosen, under which world-view the results will be interpreted, and what will be accepted as valid answers to the research questions (BOUCHER,

2014; HOLDEN; LYNCH, 2004; EASTERBROOK et al., 2008). Whatever is the researcher's philosophical stance, her/his view of ontology effects her/his view of epistemology which, in turn, effects her/his choice of methodology (HOLDEN; LYNCH, 2004). Regarding **ontology**, one should understand it as the field that investigates the nature of reality, that is, what things, if any, have existence (Objectivism), or whether reality is the product of one's mind or a social construction (Subjectivism) (BURRELL; MORGAN, 2017). In other words, what constitutes reality (RADDON, 2010). The next step is the **epistemology** field, which explores the nature of knowledge, that is, how is it possible, if it is, to gain knowledge of the world? (HUGHES; SHARROCK, 1997) It is concerned with what constitutes valid knowledge and how we can obtain it (RADDON, 2010).

The researcher who conducted this thesis is more aligned with *pragmatism* as philosophical stance. His main concern is to provide a knowledge transfer model that, being put in practice, shows its usefulness from researchers and practitioners point of view.

Pragmatism is – according to Willian James, one of its founding fathers – a stance that tries to settle metaphysical discussions, that otherwise might have no end. A researcher who is aligned with this stance interprets each notion by its practical consequences. In this context, James poses a question, *"What difference would it practically make to anyone if one notion rather than another is true?"* And he himself answers the question affirming that, *"If no practical difference can be traced, then the alternatives mean practically the same thing, and all dispute is idle"* (JAMES, 1907). Easterbrook complements affirming that pragmatists believe knowledge is always approximate and incomplete (EASTERBROOK et al., 2008). Pragmatism's ontology supports that the nature of reality is complex, rich, and reality is the practical consequences of ideas (SAUNDERS; LEWIS; THORNHILL, 2009; YEN, 2011). The pragmatism's epistemology supports that knowledge is valid when it has practical meaning in specific contexts, and "true" knowledge enables successful action. Pragmatists tend to apply methods that better suit research question/problem, instead of follow methods just because they fit in a particular philosophical stance (HOLDEN; LYNCH, 2004). For this reason, there is a common understanding that pragmatists usually adopt **mixed methods**, and acknowledge both quantitative and qualitative data as long as it is believed they are the best way to approach a problem (EASTERBROOK et al., 2008; HOLDEN; LYNCH, 2004; SAUNDERS; LEWIS; THORNHILL, 2009; YEN, 2011; CRESWELL, 2009; SHANNON-BAKER, 2016; MERTENS, 2012). In the end, pragmatists think of ontology, epistemology, and methodology as a continuum. One does not need to dichotomically embrace subjectivity or objectivity, qualitative or quantitative (YEN, 2011; SAUNDERS; LEWIS; THORNHILL, 2009; RADDON, 2010; NEWMAN; BENZ, 1998). Finally, it is important to highlight that the pragmatic view is not synonymous with absence of rigor or criteria. On the contrary, pragmatists do acknowledge the limitations of each kind of method, as well as researchers aligned with different philosophical stances also do. Nevertheless, instead of invalidating evidence produced with specific

methods – just because the methods are not aligned with her/his philosophical stance, or have some known limitations – pragmatists consider as valid knowledge, being cautious to make sure all the known caveats inherent to the methods are explicitly reported. Thus, the ones consuming knowledge can consciously assume the risks inherent to a decision-making supported by that knowledge.

1.4 RESEARCH METHODS OVERVIEW

To answer the GRQs, we conducted various empirical studies, adopting a mixed methods strategy. In accordance with the pragmatic philosophical stance, the research methods were defined taking into account the nature of the questions/problems they intend to answer/solve. Figure 1 summaries this thesis at a glance. Blue boxes denote the GRQs, the green ones represent the empirical studies we conducted to answer the GRQs, the yellow boxes are the main products of this thesis, and the red one represents the evidence of lack of connection between SSs and SE practice.

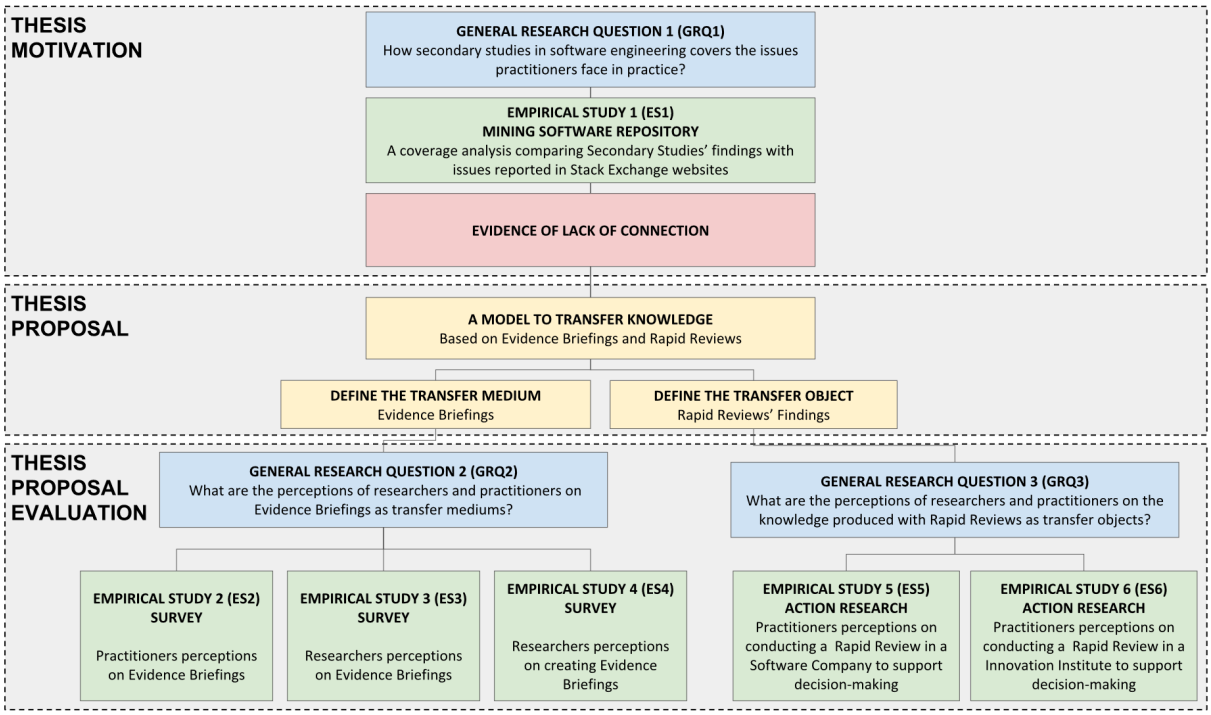


Figure 1 – Overview of research methods.

Aiming to answer GRQ1, we conducted the Empirical Study 1 (ES1) adopting a mining software repository approach. We mined issues reported by practitioners in Stack Exchange websites — a platform widely used by practitioners to discuss the problem they face in practice — and analyzed if SSs could offer knowledge that helps to solve or at least attenuate these issues. With ES1 we could investigate whether there is a lack of

connection between SSs and practice, as shown in the large gray area labeled as "Thesis Motivation" in Figure 1.

The results of ES1 provided evidence that supports the existence of the lack of connection. This motivated us to define a model to transfer knowledge from SSs to SE practice. The model is a instantiation of a well-known generic knowledge transfer model, but applying and adapting the concepts of EBs and RRs, which have been extensively demonstrated to be effective in medicine and other health-care related fields. The large gray area labeled as "Thesis Proposal" in Figure 1 depicts what we proposed in this thesis.

We conducted a series of empirical studies to evaluate the proposed model, in particular studies to evaluate the use of EBs and RRs in SE practice, as shown in the large gray area labeled as "Thesis Proposal Evaluation" in Figure 1. The Empirical Study 2 (ES2), the Empirical Study 3 (ES3), and the Empirical Study 4 (ES4) are surveys to investigate the perceptions of researchers and practitioners on EBs. Due to the positive reception of EBs, as revealed on these empirical studies, we conducted two empirical studies, the Empirical Study 5 (ES5) and the Empirical Study 6 (ES6), to evaluate the perceptions of practitioners on participating in the conduction of RRs — which results were reported in EBs — to support decision-making in practice.

1.5 RESEARCH RESULTS OVERVIEW

In ES1, we introduced a coverage technique to compare findings of SSs and issues practitioners report in five Stack Exchange websites. From 26,000+ issues reported on these websites, just 1.75% are related to the same SE areas of the 24 selected SSs. From that set of issues related the same SE areas as the selected SSs, only 15.6% are in fact covered by those SSs – when findings of a SS offer knowledge that can solve or at least attenuate the a practitioner issue. This low percentage corroborates with the claims related to the lack of connection of SSs with SE practice. It motivated us to take a step further, and propose a model to transfer knowledge to practice as a way to approximate SSs to SE practice.

After defining the knowledge transfer model, ES2, ES3, and ES4 were conducted to evaluate the perceptions of practitioners and researchers on EBs as a knowledge transfer medium. In ES2 and ES3 we found both practitioners (32) and researchers (7) are generally positive about EBs. Most of the practitioners and researchers believe that it is easy to find information on EBs (53% and 71% respectively); that EBs are clear and understandable (82% and 71% respectively); and that EBs look reliable (62% and 56% respectively). Then, we invited authors of papers in four SE conferences (EASE, SBES, SBCARS, and SAST) to create EBs of their papers. We conducted another survey to evaluate their perceptions. Authors of 44 out of 115 (38%) papers published in these four conferences created EBs. Our results revealed that although it is time consuming to create EBs (88%

of researchers took up to three hours), researchers still see benefits creating them, since 88% of the researchers affirmed they Likely or Very Likely are willing to create EBs for their papers.

To evaluate RRs, we conducted two of them in two software companies. The problems they were facing relate to low customer collaboration (Company 1) and low team motivation (Company 2). Following some RRs strategies, we delimited the search for evidence that might help to mitigate the problems, synthesized the results using thematic analysis, and introduced the findings to practitioners through EBs. In a series of interviews, we evaluated practitioners perceptions about the introduction of RRs to support decision-making. Our results show that practitioners are rather positive about RRs. They reported to have learned new concepts, reduced time and cost of decision-making, improved their understanding about the problem they were facing, among other benefits. Additionally, two months after the introduction of the first RRs (in Company 1), in a follow up visit, we perceived that the practitioners have indeed adopted some of the evidence provided by the RR to deal with their problems. They also affirmed to have experienced enhancements regarding the low customer collaboration.

These results provide evidence on the need to make EBSE more connected to SE practice, and specifically show that the model we proposed based on EBs and RRs is a promising way, among others, to reduce the gap between research and practice in SE.

1.6 MAIN CONTRIBUTIONS

This thesis makes the following contributions to SE community:

- It *defines* a systematic coverage method that establishes a relationship between SSs and the issues practitioners face in SE practice;
- It applies the defined method to *investigate* how SSs cover practitioners issues in SE, through a study based on Stack Exchange – ES1;
- It *proposes* a model to transfer knowledge from SSs to SE practice.
- It *proposes* EBs as mediums to transfer knowledge to SE practice, based on mediums from medicine;
- It *evaluates* the perceptions of practitioners, regarding the format and content of EBs – ES2;
- It *evaluates* the perceptions of researchers, regarding the format and content of EBs – ES3;
- It *evaluates* the perceptions of researchers on creating EBs – ES4;

- It *proposes* RRs as alternative to FSSs, in order to adhere to practitioners constraints;
- It *evaluates* the perceptions of practitioners of an applied-research institute on using RRs to support decision-making in practice – ES5.
- It *evaluates* the perceptions of researchers of a software company on using RRs to support decision-making in practice – ES6.

1.7 MAIN ACHIEVEMENTS

Following there is list of research papers published with results of this thesis and the ones submitted/under review; a list of research papers published but out of the scope of this thesis; a list of fruitful collaborations established during the entire Ph.D.; and a list of research already emerged derived from this thesis.

1.7.1 Publications Related to this Thesis

Some of the results of this thesis are published in four papers. Three conference papers (CARTAXO et al., 2017; CARTAXO et al., 2016; CARTAXO; PINTO; SOARES, 2018a) and one journal paper (CARTAXO; PINTO; SOARES, 2018b). Additionally, two journal papers were submitted, and are under review.

1. **MSR 2017 (Qualis A1)**: We published a paper in the 14th International Conference on Mining Software Repositories reporting part of the results of ES1. The coverage analysis with SSs and practitioners issues posted in Stack Exchange websites. It is entitled "Using Q&A Websites as a Method for Assessing Systematic Reviews" (CARTAXO et al., 2017).
2. **IST 2018 (Qualis A2)**: We published a paper in the Information and Software Technology journal presenting our knowledge transfer model. It is entitled "Towards a model to transfer knowledge from software engineering research to practice" (CARTAXO; PINTO; SOARES, 2018b).
3. **ESEM 2016 (Qualis A2)**: We published a paper in the 14th International Conference on Empirical Software Engineering and Measurement reporting the results of ES2 and ES3, which evaluate the perceptions of practitioners and researcher on EBs. It is entitled "Evidence Briefings: Towards a Medium to Transfer Knowledge from Systematic Reviews to Practitioners" (CARTAXO et al., 2016).
4. **EASE 2018 (Qualis B1)**: We published a paper in the 22th International Conference on Evaluation and Assessment in Software Engineering reporting the results of

ES5, which evaluate the perceptions of practitioners on RRs. It is entitled "The Role of Rapid Reviews Supporting Decision-Making in Software Engineering Practice." (CARTAXO; PINTO; SOARES, 2018a).

We also published three Doctoral Symposium papers in international conferences on empirical software engineering throughout the entire research (CARTAXO, 2017a; CARTAXO, 2016; CARTAXO, 2017b). This was crucial to get feedbacks from experienced researchers and sharpen all aspects of this thesis.

1. **ESEM 2017 (Qualis A2)**: 11th International Symposium on Empirical Software Engineering and Measurement. The paper is entitled "Supporting Knowledge Transfer From Secondary Studies To Software Engineering Practice" (CARTAXO, 2017a).

2. **EASE 2016 (Qualis B1)**: 20th International Conference on Evaluation and Assessment in Software Engineering. The paper is entitled "Integrating evidence from systematic reviews with software engineering practice through evidence briefings" (CARTAXO, 2016).

3. **ESELAW 2017 (Qualis B4)**: Experimental Software Engineering Latin America Workshop. The paper is entitled "Supporting Researchers to Transfer Knowledge from Systematic Reviews to Software Engineering Practice" (CARTAXO, 2017b).

A paper was submitted and is under review in the Journal of Brazilian Computer Society **(JBCS - Qualis B1)** reporting the extended results of ES1, which is the coverage analysis with SSs and practitioners issues posted in Stack Exchange websites. It is entitled "How Systematic Reviews Cover Practitioners' Issues: A Study on Stack Exchange Communities."

Another paper as submitted and is under review in the Empirical Software Engineering journal **(EMSE - Qualis B1)** reporting the results of ES4, which evaluates the perception of researchers on creating EBs. It is entitled "Evidence Briefings as Mediums to Transfer Knowledge to Software Engineering Practice."

1.7.2 Publications Out of the Scope of this Thesis

Two papers were published fruit of research we conducted in parallel to this thesis.

1. **ESEM 2015 - Qualis A2**: A paper was published reporting an empirical study we conducted in a development project in partnership with Samsung Brazil (CARTAXO et al., 2015b). The objective of this study was to diagnose performance problems in a static/dynamic analysis tool for digital TVs, as well as propose a solution and empirically evaluate it.

2. **ESELAW 2015 - Qualis B4**: We published a paper about context characterization of empirical studies in software engineering. This research emerged when we were prospecting the Ph.D. theme (CARTAXO et al., 2015a). However, we further decided to change the goal of this research to the one we presented in Section 1.1.

Two other papers were published fruit of collaborations with other researchers.

1. **EASE 2017 - Qualis B1**: A paper was published at the 21th International Conference on Evaluation and Assessment in Software Engineering. We collaborated in a research to define a model to support experiments in SE research. It is entitled "A Comparative Study of Model-Driven Approaches For Scoping and Planning Experiments" (FERREIRA et al., 2017).

2. **EASE 2017 - Qualis B1**: Another paper was published at the 21th International Conference on Evaluation and Assessment in Software Engineering reporting results of an interview in Brazilian software companies about the benefits and limitations of agile. It is entitled "On the Benefits/Limitations of Agile Software Development" (KAMEI et al., 2017).

1.7.3 Partnerships with Researchers for this Thesis

Many partnerships were established to conduct, submit, and publish the above-mentioned papers.

- **Dr. Gustavo Pinto, professor at UFPA**, specialist in mining software repositories, participated in many of the papers we published. In fact, he became the co-advisor of this thesis (CARTAXO et al., 2016; CARTAXO et al., 2017; CARTAXO; PINTO; SOARES, 2018b; CARTAXO; PINTO; SOARES, 2018a).

- **Dr. Marcio Ribeiro, professor at UFAL**, is specialist in software engineering, and participated in ES4, and in the paper submitted to EMSE, which is under review.

- **MSc. Fernando Kamei, professor at IFAL**, specialist in agile methods, participated in ES1 (CARTAXO et al., 2017), and in the paper submitted to JBCS, which is under review.

- **MSc. Elton Vieira, professor at IFPE**, specialist in information design, participated in ES2 and ES3 defining the template of the EBs (CARTAXO et al., 2016).

- **MSc. Danilo Ribeiro, Ph.D. candidate at UFPE**, is specialist in empirical software engineering and human aspects in software engineering, participated in ES1 (CARTAXO et al., 2017), and in the paper submitted to JBCS.

- **MSc. Ronnie Santos, Ph.D. candidate at UFPE**, is also specialist in empirical software engineering and human aspects in software engineering, participated in ES1 (CARTAXO et al., 2017).

1.7.4 Research Derived from this Thesis

- **A master dissertation was successfully concluded** aiming to update the tertiary study of Da Silva (SILVA et al., 2011), which maps systematic reviews in SE. This work, conducted by MSc. Kenelly, was derived from a need emerged from this thesis to have more up to date informations about SSs in SE.
- **A research is on going with Dr. Maria Tereza Baldasarre, professor at UNIBARI in Italy**. The aim of this research is to evaluate the perceptions of researchers about RRs in SE.

1.8 DOCUMENT STRUCTURE

The remainder of this document is structured as follows:

- **Chapter 2** presents the methodological aspects, and the results of ES1, revealing how SSs cover the issues practitioners face in practice. It also provide evidence that motivate this thesis.
- **Chapter 3** shows the model we proposed to transfer knowledge from research to SE practice.
- **Chapter 4** exposes the details about the method as well as the results of ES2, ES3, and ES4. The ones we evaluate the perceptions of practitioners and researchers about EBs as a knowledge transfer medium.
- **Chapter 5** present the details on the method and the results of ES5 and ES6. The ones we evaluate the applicability of RRs in SE practice.
- **Chapter 6** presents the the main conclusions of this thesis, and future directions for research that might derive from it.

2 COVERING PRACTITIONERS ISSUES

Many researchers have been claiming that Evidence Based Software Engineering (EBSE) in general, and Secondary Studies (SSs) in particular, lack connection with Software Engineering (SE) practice (HASSLER et al., 2014; SANTOS; SILVA, 2013; SILVA et al., 2011). This can be an obstacle to EBSE achieving one of its main goals, which is to integrate the best research evidence with practice (KITCHENHAM; DYBA; JORGENSEN, 2004). To investigate if those claims are supported by evidence, we conducted a study aiming to answer the following research question:

GRQ1: How secondary studies in software engineering cover the issues practitioners face in practice?

To provide answers to this important but overlooked question, we selected a set of SSs, previously identified in Da Silva's et al. (SILVA et al., 2011) tertiary study. Although other tertiary studies have been published in the last years, they are focused on specific areas of SE, not in SE as a whole, as we carefully discuss in Section 2.1. The selected SSs vary from several topics, such as agile methods (DYBÅ; DINGSØYR, 2008), usability evaluation (INSFRAN; FERNANDEZ, 2008), knowledge management (BJØRNSON; DINGSØYR, 2008), and others.

Next, we identified issues practitioners face in practice, and analyzed how the selected SSs cover these issues. By **coverage** we mean when at least one finding of a SS offers knowledge that helps to solve or at least attenuate a practical issue. Nevertheless, it is important to highlight we are NOT suggesting SSs should provide definitive evidence to solve practitioners issues. Instead, as Booth et al. discussed, we believe research evidence can help practitioners during decision-making, helping them to solve a practical issue (BOOTH; COLOMB; WILLIAMS, 2003). Additionally, we used Stack Exchange websites as data source for practitioners issues. Stack Exchange is a platform with over 100 high-quality, professional Q&A websites. In the scope of this study, by **practitioner issue** we mean, a question asked on one of the Stack Exchange websites. Generally speaking, questions in Stack Exchange are composed by a title — summarizing the question — and a body — introducing further details.

Part of the results of this coverage analysis were published in the 14th International Conference on Mining Software Repositories (MSR 2017) (CARTAXO et al., 2017). The extended version presented in this chapter was recently submitted to the Journal of Brazilian Computer Society (JBCS).

2.1 BACKGROUND AND RELATED WORK

2.1.1 On Identifying Practitioners Issues

Many studies have been conducted trying to establish a link between research and SE practice. For instance, the work of Begel et al., which identified 145 practitioners issues asking them directly (BEGEL; ZIMMERMANN, 2014). Another example is Carver's et al. research, that shares a very brief paragraph summarizing studies in SE, and asks practitioners how relevant are them (CARVER et al., 2016). Nevertheless, these kinds of research have at least one of the following limitations:

1. Relying on human memory, which is not effective in remembering ordinary or long term events (BERNARD et al., 1984);

2. Difficulties to obtain a diverse and heterogeneous sample (BERNARD et al., 1984; GOLDER; MACY, 2012).

The first limitation happens when one asks practitioners what are the issues they usually face, instead of observing directly the issues occurring in practice. The latter happens when one, trying to overcome the first limitation, immerses in one or a few software development environments to observe practitioners issues in detail, but fails obtaining a diverse and heterogeneous sample due to the nature of immersive methods, which are inherently context bounded, and hardly generalizable. Interestingly, these two limitations have been vastly reported in other research fields, such as social sciences (GOLDER; MACY, 2012).

Over the last years, researchers have been arguing SE has many similarities with social sciences, and consequently should be studied through their methods (RALPH; CHIASSON; KELLEY, 2016). Nonetheless, pure and applied social sciences are facing an important paradigm shift since the appearance of social media (RALPH; CHIASSON; KELLEY, 2016; GOLDER; MACY, 2012; DRURY, 2008). Hence, if SE wants to follow social sciences footsteps, it is important to observe this recent disruptive movement. To illustrate, the paper published by Murthy on the *Sociology Journal* mentions the importance of Twitter to understand the recent social movements in Lybia, Egypt, Tunisia, and Algeria (MURTHY, 2012). Other example is Drury's paper published in *Journal of Direct, Data and Digital Marketing Practice*, about the importance of social media to understand and reach a target audience (DRURY, 2008).

Another critical point in favor of exploring social media to mitigate the two aforementioned limitations, as explained by Golder and Macy (GOLDER; MACY, 2012) in a publication of *American Sociological Association* is:

> *"As a discipline devoted to explaining patterns of human behavior and social interaction, sociologists often have to choose whether to rely on direct real-time observation of very small numbers of non-representative individuals (e.g.,*

> *in field observation or in the laboratory) or to rely on indirect retrospective accounts obtained through survey responses from large representative samples.* ***Social media offers us the opportunity for the first time to both observe human behavior and interaction in real time and on a global scale.*** *"* (GOLDER; MACY, 2012) (bold emphasis added)

In this context, we decided to investigate how knowledge from SSs cover practitioners issues using Stack Exchange websites as data source. Stack Exchange is a platform with over 100 high quality, professional Question & Answer (Q&A) websites. It covers topics as diverse as Mathematics, Home Improvement, Statistics, and English Language. Software development, which is a knowledge intensive activity, is particularly well-supported (ROBILLARD, 1999). Stack Overflow is certainly the most well known Stack Exchange website. It focuses on very technical coding questions. However, there are many other Stack Exchange websites focused on different areas of SE, like software testing, quality, reverse engineering, project management, and more. To illustrate, following are some snippets of questions made on those websites:

> *"[...] How to facilitate communication and peer reviews on a distributed scrum team? [...]"*

> *"[...] Pair programming when driver and observer have different skill level and experience [... this] strategy still work [...] if they have a very different programming skill level? If one never experience in the problem domain while another have? Is it still OK if they have low programming skill level?"*

> *"[..] how to apply agile methods in large complex embedded system software (100+ engineers). Firmware development has some unique characteristics that make it difficult to do agile (ie. Hardware is not available until late in the dev cycle; Once product is released, can't easily update firmware;) [...] I would like to hear any tips or guidelines on how to adopt agile methodology for firmware development projects."*

Stack Exchange is a good example of a social media platform that exhibits characteristics that mitigate the two limitations highlighted by social sciences. First, the possibility to identify practitioners issues when they occur, full of context, without the need to rely on human memory. This can be observed not only looking at questions posted on Stack Exchange websites but also by the official description of the platform when they instruct practitioners to *"focus on questions about* ***an actual problem you have faced****. Include details about what you have tried and exactly what you are trying to do. [...] Not all questions work well in our format. Avoid questions that are primarily opinion-based [...]"* (STACKEXCHANGE..., g). Second, Stack Exchange websites have diverse users distributed all around the world, with a wide variate of age, gender and professional

experience (STACKEXCHANGE..., f; VASILESCU; CAPILUPPI; SEREBRENIK, 2012; MORRISON; MURPHY-HILL, 2013). Stack Exchange websites are changing the way software engineers communicate and coordinate software projects. For instance, Stack Exchange users reported that these websites are replacing web search and traditional forums as their primary source of knowledge (MAMYKINA et al., 2011). Some researchers go further and suggest that answers in these websites often become a substitute for the official documentation if it is not available or incomplete (TREUDE; BARZILAY; STOREY, 2011). This strong interest happens partially due to a community of millions of virtually real-time practitioners willing and eager to help their peers. As an example, a question asked on Stack Overflow has a median answer time of 11 minutes (MAMYKINA et al., 2011). Many studies also use Stack Exchange websites to explore the state of practice in SE. To cite a few, Wang and Godfrey detected iOS and Android API usage obstacles based on questions posted in Stack Exchange websites (WANG; GODFREY, 2013); Pinto and Kamei discovered flaws and desirable features of refactoring tools from questions posted in Stack Exchange websites (PINTO; KAMEI, 2013); or Vasilescu et al., which investigated the interplay between Stack Exchange activities (VASILESCU; FILKOV; SEREBRENIK, 2013). Many other SE studies have been conducted in a steady pace using Stack Exchange websites (PINTO; CASTOR; LIU, 2014; TREUDE; PROLO; FILHO, 2015; REBOU et al., 2016; POSNETT et al., 2012).

2.1.2 On Identifying Secondary Studies

One who wants to analyze how SSs cover practitioners issues must also define how to establish a sample of SSs, so they can be compared to the practitioners issues. A possible way to identify SSs is through tertiary studies. According to Kitchenham and Charters, tertiary studies are reviews of SSs related to the same research question (KITCHENHAM; CHARTERS, 2007).

In 2009, Kitchenham and et al. published a tertiary study and identified 20 SSs in SE (KITCHENHAM et al., 2009). One year later, they updated the first tertiary study and identified 35 new SSs, which clearly shows the growing interest in this kind of method (KITCHENHAM et al., 2010). In 2011, just one year ahead, Da Silva et al. published another update of that tertiary study, this time identifying 65 new SSs, consolidating the view that SSs are becoming increasingly important to the SE community (SILVA et al., 2011).

Some tertiary studies were published after 2011, but none of them focuses on SSs in SE as a whole. They either investigate methodological aspects of SSs or explore a narrower specific SE area. Examples of the former are: the tertiary study conducted by Cruzes and Dybå investigating the types of research synthesis methods used in SSs (CRUZES; DYBå, 2011); the tertiary study of Da Silva et al. critically appraising SSs from the perspective of their research questions (SILVA et al., 2010); the tertiary study of Ali and Petersen investigating which strategies are used in SSs to select primary studies (ALI; PETERSEN,

2014); and finally the tertiary study conducted by Zhou et al., which investigated how SSs assess quality of the primary studies they include (ZHOU et al., 2015). Examples of the latter - tertiary studies that explore a a narrower specific SE area - are: Marques et al.'s tertiary study investigating SSs on distributed software development (MARQUES; RODRIGUES; CONTE, 2012); the tertiary study of Santos et al. also on distributed software development (SANTOS et al., 2012); Verner's et al. tertiary study exploring SSs on global software development (VERNER et al., 2014); Bano's et al. tertiary study identifying SSs on software requirements (BANO; ZOWGHI; IKRAM, 2014); Goulão's et al. tertiary study on model-driven engineering (GOULÃO; AMARAL; MERNIK, 2016); Garousi's et al. tertiary study on software testing (GAROUSI; MäNTYLä, 2016); and the tertiary study of Hoda et al. on agile software development (HODA et al., 2017).

In this thesis, to properly understand the SSs of SE research area as a whole, we chose to focus on tertiary studies that investigates the broader area of SE, and the most recent – at the time we conducted ES1 – is the one published by Dal Silva et al. in 2011 (SILVA et al., 2011).

2.2 RESEARCH METHOD

In this section, we present the steps required to conduct this research, as depicted in Figure 2. The numbers denote the order of each step. Looking at major activities (large gray areas), we started selecting SSs (Section 2.2.1). Then we extracted the search string of these SSs (Section 2.2.2). In parallel, we selected Stack Exchange websites related to SE (Section 2.2.3). After that, we used the search strings we extracted to find practitioners issues in the selected Stack Exchange websites (Section 2.2.4). We then excluded false positives issues (Section 2.2.5). Finally, we conducted a coverage analysis, matching each practitioner issue with SSs findings aiming to calculate coverage rate, and to identify recurrent issues occurring in practice (Section 2.2.6).

2.2.1 Secondary Studies Selection

We relied on Da Silva's et al. tertiary study to select our initial set of SSs (SILVA et al., 2011). Figure 3 shows the procedure we followed. We excluded SSs that do not fit in the purpose of this research. Firstly, we excluded those that, according to Da Silva, do not present guidelines to practitioners, resulting in 32 SSs. They were excluded because their lack of guidelines to practitioners might leave practitioners with no concrete actionable items, which should be avoided when the intention is to provide evidence to practitioners (LAVIS et al., 2003). Otherwise, practitioners under time pressure have little chance to read and search immediate implications on often extensive SSs. Secondly, we excluded SSs that do not report their search strings, since we need them to search for practitioners issues on Stack Exchange websites, resulting in the final set of 24 selected SSs.

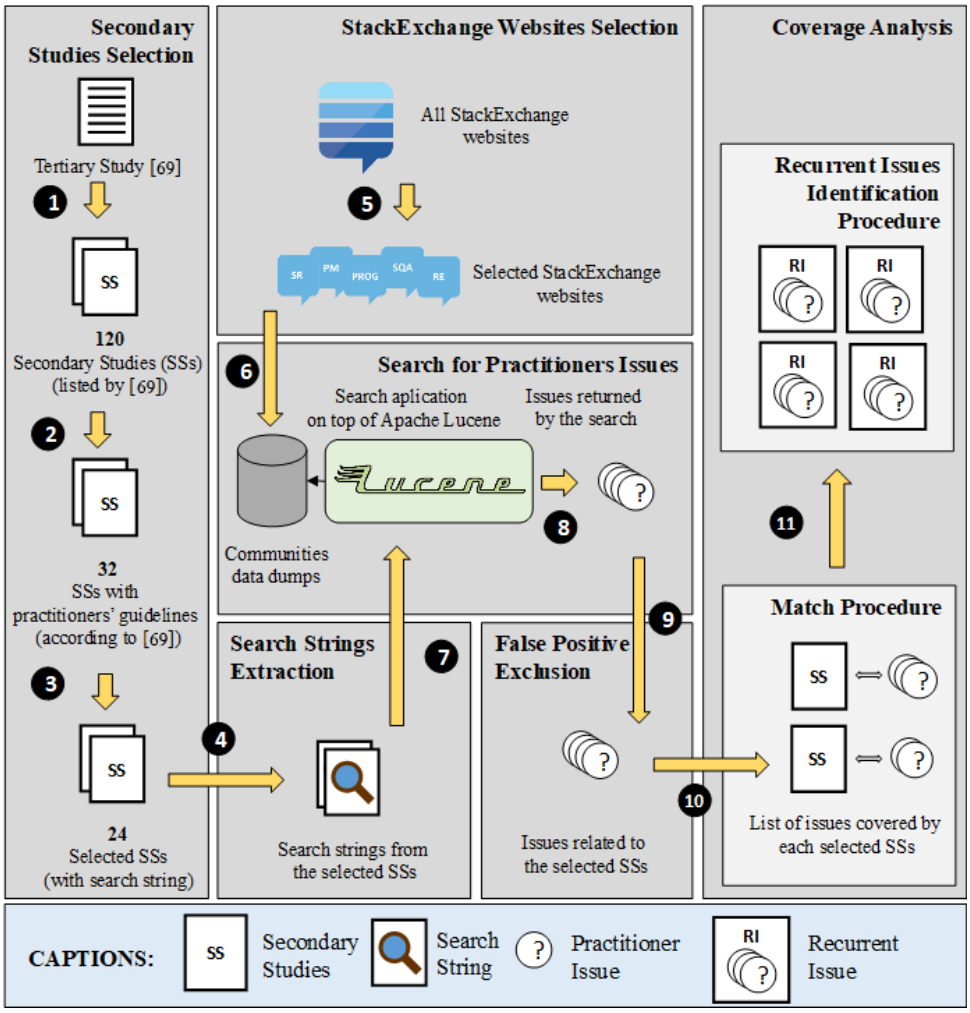


Figure 2 – Research steps (ES1).

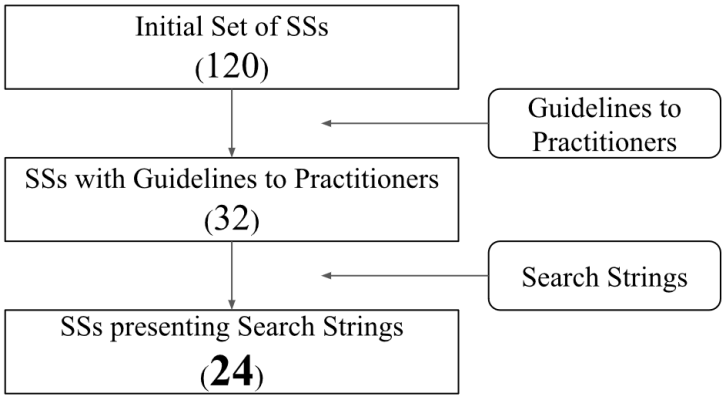


Figure 3 – Secondary studies selection procedure.

We also classified each selected SS according to the 15 SE Areas as defined in SWEBOK (BOURQUE; FAIRLEY, 2014) using the Open Card Sort Technique (SPENCER, 2009). This was necessary to define whether the practitioners issues posted in Stack Exchange websites are indeed related to the SSs, or are false positives (details in Section 2.2.5). Table 1 presents the 24 selected SSs and their respective SE areas. Figure 4 depicts the number

of SSs per SE area giving a quantitative notion of which SE areas have more and less attention from the research community.

Table 1 – Selected secondary studies.

CODE	TITLE	SE AREA
SS1	A systematic review of cross- vs. within- company cost estimation studies	Software Engineering Management
SS2	A systematic review of software maintainability prediction and metrics	Software Maintenance
SS3	A systematic review of usability evaluation in web development	Software Design
SS4	A systematic literature review to identify and classify software requirement errors	Software Requirements
SS5	Automated acceptance testing: A literature review and an industrial case study	Software Testing
SS6	Challenges and improvements in distributed software development: A systematic review	Software Engineering Management
SS7	Critical Barriers for Offshore Software Development Outsourcing Vendors: A Systematic Literature Review	Software Engineering Economics
SS8	Critical success factors for offshore software development outsourcing vendors: A systematic literature review	Software Engineering Economics
SS9	Definitions and approaches to model quality in model-based software development – a review of literature	Software Quality
SS10	Effectiveness of requirements elicitation techniques: Empirical results derived from a systematic review	Software Requirements
SS11	Empirical studies of agile software development: A systematic review	Software Engineering Models and Methods
SS12	Evidence-based guidelines for assessment of software development cost uncertainty	Software Engineering Management
SS13	Factors influencing software development productivity-state-of-the-art and industrial experiences	Software Engineering Economics
SS14	Forecasting of software development work effort: Evidence on expert judgement and formal models	Software Engineering Management
SS15	Harmfulness of code duplication: A structured review of the evidence	Software Construction
SS16	Knowledge management in software engineering: A systematic review of studied concepts, findings and research methods used	Software Engineering Professional Practice
SS17	Model-based testing approaches selection for software projects	Software Testing
SS18	On the generation of requirements specifications from software engineering models: A systematic literature review	Software Requirements
SS19	Risks and safeguards for the requirements engineering process in global software development	Software Engineering Management
SS20	Software process improvement in small and medium software enterprises: A systematic review	Software Engineering Process
SS21	Technology transfer decision support in requirements engineering research: a systematic review of rej	Software Requirements
SS22	The effectiveness of pair programming: A meta-analysis	Software Engineering Models and Methods
SS23	Towards a defect prevention based process improvement approach	Software Quality
SS24	Using scrum in global software development: A systematic literature review	Software Engineering Models and Methods

2.2.2 Search Strings Extraction

After selecting the 32 SSs presenting guidelines to practitioners, we extracted their search strings to search for practitioners issues on Stack Exchange. As previously mentioned, eight of these SSs were excluded because they do not present search strings, resulting in the 24 selected SSs. We are aware that some SSs might employ manual search (e.g. by using snowballing (WOHLIN, 2014)), which does not need search strings. However, we found that seven out of the eight SSs that do not present search strings, explicitly declared to use search engines. The remaining one do not clearly explain its search strategy.

Still, we found that nine out of the 24 selected SSs do not present their search strings properly. For instance, some studies present only a list of search terms but do not mention

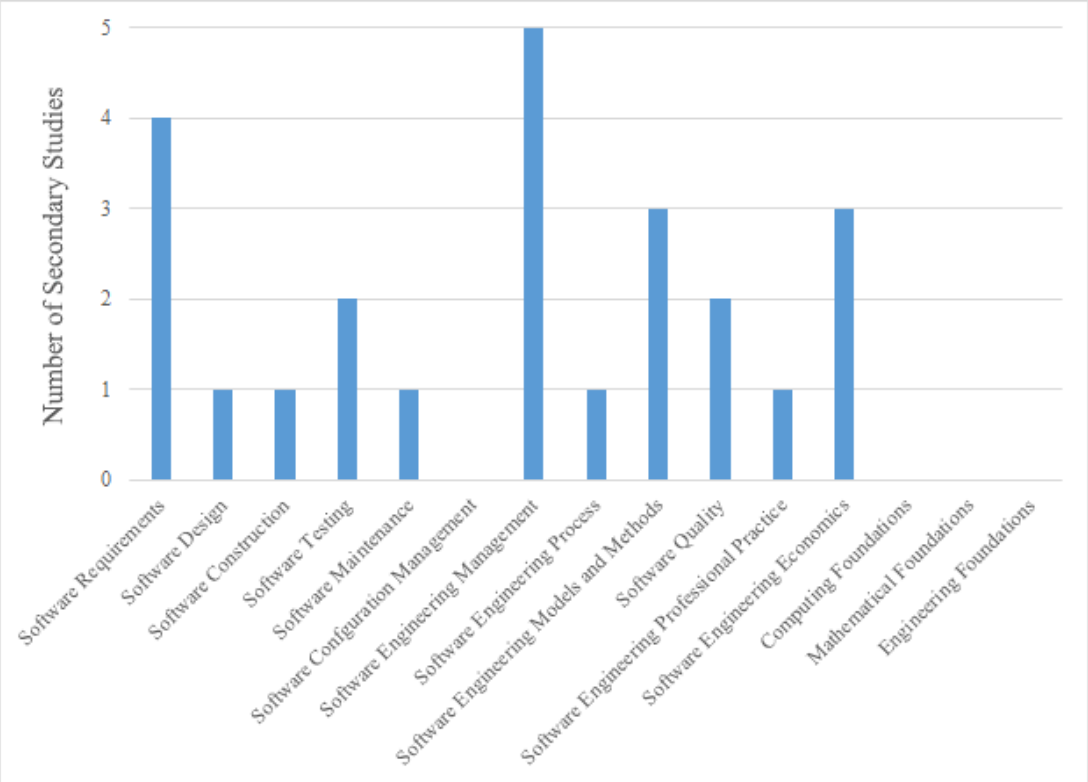


Figure 4 – Distribution of the selected secondary studies per software engineering area.

which logical operator (AND or OR) they used to connect each term. Others connect terms with ambiguous operators. For example, the following search string using "+" (plus) and "," (comma), which we considered as "OR" and "AND", respectively.

"quality + model", "quality + model driven", "model driven + experience"

Or the following search string using the operator "WITH", which we considered as AND.

software AND ((cost OR effort OR productivity) WITH (factors OR indicators OR drivers OR measure))

2.2.3 Stack Exchange Websites Selection

In this step, we selected Stack Exchange websites to search for practitioners issues. We restricted our search to websites related to at least one of the 15 SWEBOK SE areas (BOURQUE; FAIRLEY, 2014). To establish the relationship, we compared the official websites' descriptions with each SWEBOK's SE area. For this research, we selected five Stack Exchange websites from more than 100 active websites.[1] Table 2 shows the relationship between the selected Stack Exchange websites and the SWEBOK's SE areas.

[1] <stackexchange.com/sites>

Table 2 – Selected Stack Exchange websites.

WEBSITE	DESCRIPTION	SWEBOK'S SE AREAS
Programmers (PROG)	Q&A for professional programmers who are interested in getting expert answers on conceptual questions about software development	Software Design Software Construction
Project Management (PM)	Q&A for project managers	Software Management
Software Quality Assurance (SQA)	Q&A for software quality control experts, automation engineers, and software testers	Software Testing Software Quality
Reverse Engineering (RE)	Q&A for researchers and developers who explore the principles of a system through analysis of its structure, function, and operation	Software Maintenance
Software Recommendations (SREC)	Q&A for people seeking specific software recommendations	Software Tools

Stack Exchange maintains a staging zone, named *Area 51*[2], intended to receive requests to create new websites, as well as monitoring a set of metrics to assess how well existing websites are. According to *Area 51*, websites with answers/question ratio above three are considered *good*, and above one are *okay but need improvement*. Table 3 shows some numerical data about the selected websites.

Table 3 – Selected websites numerical data.

COMMUNITY	# QUESTIONS	# ANSWERS	A/Q
PROG	35,560	128,199	3.6
PM	2,362	8,420	3.56
SQA	2,642	6,333	2.39
RE	1,745	2,751	1.57
SREC	4,434	4,894	1.1
TOTAL	46,743	150,597	3.22

As we can see, three selected websites can be classified as *good*, and the remaining ones as *okay but need improvement*. *Area 51* monitors other metrics, such as *number of visits per day* and *number of avid users*. For instance, RE and SREC are considered excellent in these metrics, with 199 avid users and 1,905 visits per day, and 394 avid users and 4,384 visits per day, respectively. We did not use the most well-known Stack Exchange website, Stack Overflow, because it focuses mainly on technical and coding issues, which is often out of the scope of SE SSs.

[2] <area51.stackexchange.com>

2.2.4 Search for Practitioners Issues

We searched for practitioners issues in the selected Stack Exchange websites using the search strings extracted from the selected SSs. A simple way to find issues would be using the Stack Exchange's online search. However, despite providing rich ways to perform advanced online searches in their database, it is not possible to use logical operators (e.g. OR and AND), like the ones in SSs' search strings. Therefore, we implemented a search application on top of Apache Lucene,[3] and used the search strings as input to find issues on the last Stack Exchange data dump at the time we conduced this research (August 18, 2015). Apache Lucene is a highly scalable Java search library widely recognized and used. Stack Exchange itself implements its online search on top of Lucene as reported in their official blog (STACK..., 2011).

To select high-quality issues, we filtered them based on their score. In Stack Exchange, a user can "up-vote" an issue if s/he thinks it is relevant, or "down-vote" otherwise. The score is the resulting value of this voting process. The score is also commonly used as a metric for choosing relevant and high-quality issues (e.g. (TREUDE; PROLO; FILHO, 2015)). Considering the median score as the threshold, we selected only issues above or equal to their respective websites median score. We decided to adopt the median score instead of, for instance, the mean score, because scores have no upper or lower limits. Thus, outliers affecting the mean are common. Table 4 shows descriptive statistics about the issues' scores for each selected Stack Exchange website. Hereafter, we will use the term **more relevant issues** for issues with score above or equal to the median score of their websites.

Table 4 – Issues scores descriptive statistics.

WEBSITE	MEDIAN	MEAN	S.D.	MAX	MIN
PROG	3	7.2	22.1	2,189	-11
PM	3	4.3	5.3	80	-5
SQA	1	2.5	4.1	72	-8
RE	2	3.6	5.2	76	-8
SREC	2	3.4	4.4	75	-4

As we can see, the standard deviation (S.D.) is larger than the mean in all cases, which corroborates the decision of not choosing the mean as a measure for selecting practitioners issues on Stack Exchange websites.

The five selected Stack Exchange websites have a total of 46,743 issues. Among them, 26,687 issues have scores above the median of their websites. From these issues, 1,860 (7%) were found using the search strings of our 24 selected SSs. Table 5 depicts the number of returned issues for each selected SSs, per Stack Exchange website.

[3] <https://lucene.apache.org/core>

Table 5 – Number of returned issues for each selected secondary study.

SS	PROG	PM	RE	SREC	SQA	TOTAL
SS14	471	78	0	8	19	576
SS11	257	84	0	30	24	395
SS18	161	25	0	54	16	256
SS15	147	0	1	6	3	157
SS5	106	12	0	18	20	156
SS21	87	14	0	3	11	115
SS22	75	6	0	1	1	83
SS9	17	4	0	1	7	29
SS7	20	5	0	0	1	26
SS13	17	2	0	1	2	22
SS16	14	5	0	1	0	20
SS24	7	2	0	0	2	11
SS8	5	2	0	0	0	7
SS10	3	0	0	0	0	3
SS3	2	0	0	0	0	2
SS20	1	0	0	0	0	1
SS1	1	0	0	0	0	1
SS4	0	0	0	0	0	0
SS2	0	0	0	0	0	0
SS6	0	0	0	0	0	0
SS17	0	0	0	0	0	0
SS19	0	0	0	0	0	0
SS23	0	0	0	0	0	0
SS12	0	0	0	0	0	0
TOTAL	1,391	239	1	123	106	1,860

Secondary studies with no issues returned: Among the 24 selected SSs, the search string of seven (29%) did not return any issue. We identified three reasons that might explain this fact:

- **The search string has too many key terms:** This makes the search too specific, making it difficult to find issues with all key terms in its title or body since they are connected with the restrictive AND operator. For instance, there is a SS with 13 terms connected by AND operators in its search string. The search strings of SS2 and SS4 have this problem.
- **The search string has key terms with no synonyms:** This kind of search string leaves no room for issues using different words. For instance, consider the following

search string: *"uncertainty assessment" AND motivation*. This search string does not admit a synonym for any of its two terms, which prevent it from finding issues using different words. The search strings of SS12 and SS19 have this problem.

- **The search string has key terms with composed synonyms only:** A composed synonym comprises more than one word, and the order of each word matters.To illustrate a composed synonym for a key term like "requirements" could be "software requirements". Thus, when all the synonyms are composed it forces the issues' title or body to have the same words of the composed synonym in the same order. For instance, lets take the following synonyms of a key term from a search string of a SS: *"model based test" OR "model based testing" OR "model driven test" OR "model driven testing" OR "specification based test" OR "specification based testing" OR "specification driven test" OR "specification driven testing" OR "use case based test" OR "use case based testing" OR "use case driven test" OR "use case driven testing" OR "UML based test" OR "UML based testing" OR "UML driven test" OR "UML driven testing" OR "requirement based test" OR "requirement based testing" OR "requirement driven test" OR "requirement driven testing" OR "finite state machine based test" OR "finite state machine based testing" OR "finite state machine driven test" OR "finite state machine driven testing"*. It has many synonyms, but they are all composed and very specific, reducing the probability of finding an issue that has, at least, one of them. The search strings of SS6, SS7, and SS23 have this problem.

Stack Exchange websites characteristics: Based on the search, we observed the following characteristics about the selected Stack Exchange websites:

- **PROG**: This website returned the highest number of issues (74.7%). It was not surprising since this website is the one with higher number of issues posted by practitioners. Despite its name – Programmers – it is clearly focused on conceptual design and programming issues such as good practices, design patterns, and architectural trade-offs.

- **PM**: This website returned the second highest number of issues (12.8%) with the search strings of the selected SSs. Additionally, in this website, we could find many issues about diverse SE areas, beyond the scope of software project management, such as software requirements, software testing, and others.

- **SQA**: In this website, issues are indeed focused on quality assurance and testing, and it is not common to find issues from other topics, as in PM website.

- **SREC**: This website returned few issues, even though it has the second highest number of issues reported by practitioners, as shown in Table 2. Additionally, we

could observe that many of those few issues were, in fact, false positives. In the end, the majority of the issues posted in this website are recommendation requests about software applications in general, such as applications to burn DVDs, or to remotely access a PC, to mention a few. Requests for tools to support SE practices, such as IDEs, test automation tools, or bug trackers are usually posted on the other websites focused on SE areas, such as in PROG, PM, SQA, and RE. Therefore, we believe this website should not be considered for further studies aiming to investigate SE issues.

- **RE**: This website returned only one issue, and was the one with the lowest amount of issues returned by the search. On the other side, only one of the selected SSs ((RIAZ; MENDES; TEMPERO, 2009)) was considered as belonging to *Software Maintenance* area, which is the broader area that comprises reverse engineering according to SWEBOK (BOURQUE; FAIRLEY, 2014). Thus, we believe further investigations are important to understand whether this website is a good choice to investigate SE issues.

2.2.5 False Positive Exclusion

Some studies have shown a high-rate of false positives when searching for issues in Stack Exchange websites (KAVALER et al., 2013). For instance, in Pinto's et al. work, the authors observed about 50% of the initially selected questions were, in fact, false positives (PINTO; CASTOR; LIU, 2014). To remove them, we classified each practitioners issues as **Related** or **Not Related** to the SE area of the selected SSs, as defined in Table 1. This explains why we classified each selected SS based on the SWEBOK SE areas. To avoid misclassification, this procedure was conducted in pairs, followed by conflict resolution meetings. We performed an agreement analysis using the Kappa statistic. The Kappa value was 0.85, which means an "Excellent Agreement" level according to the Kappa reference table (VIERA; GARRETT, 2005). After the classification, we analyzed Not Related issues to understand the reasons why they were returned by the search.

Our search returned a total of 1,860 issues using the search strings of the selected SSs. Table 6 shows the result of the false positives exclusion. The seven SSs which did not return any issue by the search – SS4, SS2, SS6, SS17, SS19, SS23, SS12 – were omitted from the table.

As we can see, 1,392 issues were discarded due to being considered as Not Related to the selected SSs (i.e. false positives). After discarding false positives, we ended up with 468 practitioners issues. This set represents 1.75% of the 26,687 more relevant issues of the five Stack Exchange websites. This result might indicate a gap between topics explored with SSs and ones demanded by practitioners.

Table 6 – Number of issues Related and Not Related to the selected secondary studies.

SS	**RELATED**		**NOT RELATED**	**TOTAL**
	#	%	#	#
SS11	217	54.9%	178	395
SS22	45	54.2%	38	83
SS18	41	16%	215	256
SS15	34	21.6%	123	157
SS21	31	26.9%	84	115
SS5	24	15.3%	132	156
SS14	23	4%	553	576
SS16	15	75%	5	20
SS7	14	53.8%	12	26
SS24	7	63.6%	4	11
SS13	6	27.2%	16	22
SS9	5	17.2%	24	29
SS8	5	71.4%	2	7
SS1	1	100%	0	1
SS10	0	0%	3	3
SS3	0	0%	2	2
SS20	0	0%	1	1
TOTAL	468	25.1%	1,392	1,860

Moreover, ten out of the 17 SSs presented a high rate of false positives – less than 50% of the Related issues. For instance, no issue was related to SS3, SS10, or SS20. We found a scenario that may explain this situation:

- **Secondary studies using rather common terms in their search strings:** For instance, SS18 (NICOLáS; TOVAL, 2009) uses terms such as "from", "documents", and "features" as part of the search string. These terms are likely to appear in other contexts beyond requirements specifications, which is the SS focus. In consequence, 215 (84%) out of the 256 issues, returned by the search, were classified as Not Related to this particular SS. A similar situation was observed in Kavalers' et al. study, where they looked for issues that reported Java classes with the term "security" in their name. Many issues were found because the term "security" is often mentioned in posts not related to classes with security in their names, rising the number of false positives.

2.2.6 Coverage Analysis

After excluding false positives, we conducted the coverage analysis based on qualitative techniques (STOL; RALPH; FITZGERALD, 2016). The analysis is divided in two parts: *match procedure* and *recurrent issues identification procedure*. To avoid bias, the entire coverage analysis was conducted by one researcher, and revised by another. With this analysis, we could identify which issue is covered, and which is not, mapping gaps between SSs and issues asked by practitioners. Figure 5 depicts the entire coverage analysis.

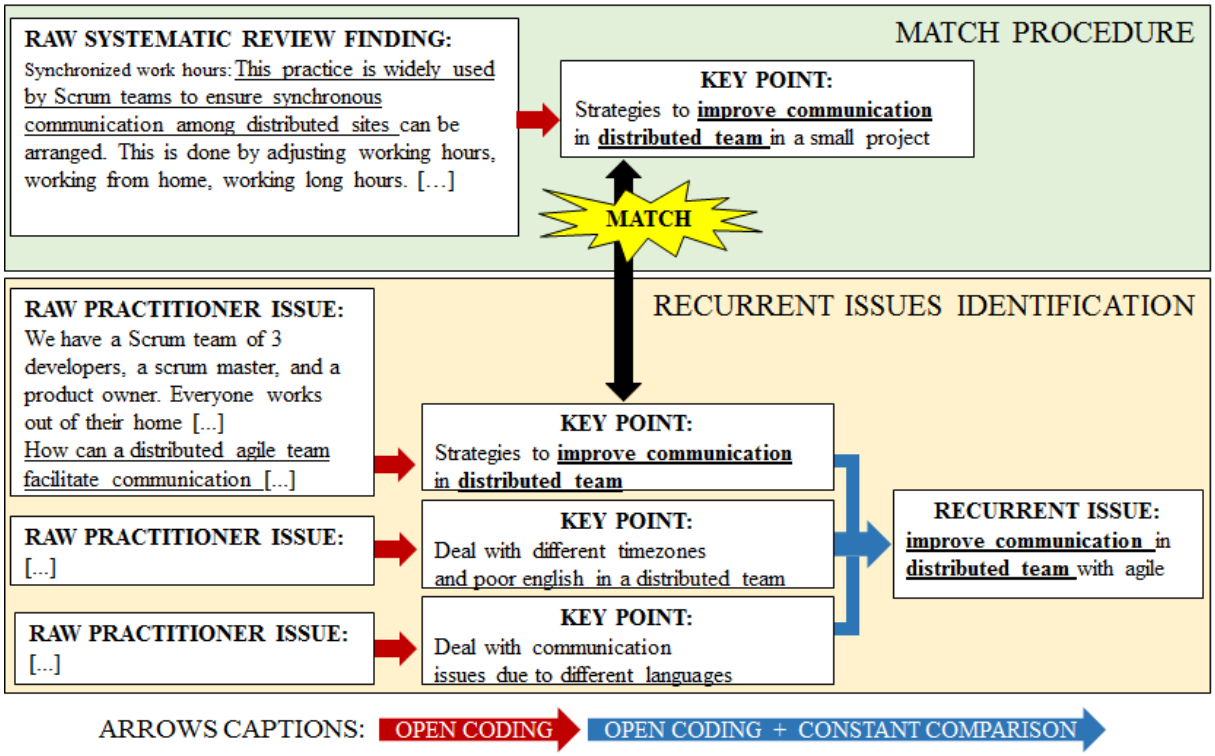


Figure 5 – Coverage analysis procedure.

Match Procedure: In the match procedure, we analyze issue per issue comparing them to the findings of the SS they are related to. Steps comprising this procedure are:

1. We extracted findings of each SS and applied open coding techniques (STOL; RALPH; FITZGERALD, 2016) to define their *Key Points*;

2. We also analyzed practitioners issues related to each SS and applied open coding techniques (STOL; RALPH; FITZGERALD, 2016) to define their *Key Points*;

3. We compared the *Key Point* of each issue against the SS findings *Key Points*, establishing whether an issue is covered or not.

The *Key Points* are codes that summarize, in few words, practitioners issues from Stack Exchange or findings of a SSs. To illustrate the match procedure, Table 7 shows some examples of issues considered as covered by findings of SS. The issue's title starts with symbol "", and its body starts with symbol "".

Table 7 – Examples of issues covered by findings of the selected secondary studies.

PRACTITIONER ISSUE	SS FINDING COVERING THE ISSUE
When does pair programming work? When to avoid it? Rather than slavishly pair program all the time, we use pair programming selectively on our team. [We] think it works best in [some] circumstances [...] When to use pair program and why? When to avoid pair programming? Why?	*"If you do not know the seniority or skill levels of your programmers, but do have a feeling for task complexity, then employ pair programming either when task complexity is low and time is of the essence, or when task complexity is high and correctness is important."* SS22 (HANNAY et al., 2009)
What makes Agile software development so appealing? Agile software development is becoming a pretty fun buzzword these days [...] what are the biggest reasons for choosing to do Agile development [...]	*"Most studies reported that agile development practices are easy to adopt and work well. Benefits were reported in the following areas: customer collaboration, work processes for handling defects, learning in pair programming, thinking ahead for management, focusing on current work for engineers, and estimation [...] benefits in projects that use agile methods because changes are incorporated more easily and business value is demonstrated more efficiently [...]"* SS11 (JORGENSEN, 2005)
How to facilitate communication and peer reviews on a distributed scrum team? We have a Scrum team of 3 developers, a scrum master, and a product owner. Everyone works out of their home [...] How can a distributed agile team facilitate communication [...]	*"Our SLR has found that Scrum teams use various practices or strategies to reduce these challenging factors to support the use of Scrum practices in globally distributed projects. This review has identified and categorized these practices as follows: Synchronous communication [...] Team Collaboration [...] Communication bandwidth [...] Tool Support [...] Team management [...] Office space [...]"* SS24 (HOSSAIN; BABAR; PAIK, 2009)

Another important step is that after the match procedure we excluded duplicated issues. Duplication occurs when more than one SS is related to the same issue. We decided to exclude duplicated issues after the match procedure to guarantee that if any SS related to that issue presents a finding helping to solve it, the issue would be considered as covered. From the 468 related issues, there were 24 duplicated. We ended up with 424 unique practitioners issues.

Recurrent Issues Identification Procedure: At the end of the *Match Procedure* we could calculate the overall coverage rate. However, due to the high number of issues, it would be hard to draw meaningful insights beyond the quantity of covered issues and the overall coverage rate. As a consequence, we conducted the *Recurrent Issues Identification Procedure* aiming to aggregate issues that report the same problem, thus we can provide a manageable list of recurrent issues practitioners face in practice. Steps comprising this procedure are:

1. We identified recurrent issues applying open coding and constant comparison techniques (STOL; RALPH; FITZGERALD, 2016);

2. We classified each recurrent issue according to SWEBOK (BOURQUE; FAIRLEY, 2014) SE areas.

A **recurrent issue** groups two or more issues about the same problem. The aggregation of issues as recurrent issues enables us to identify common problems in practice and also manage the coverage analysis. An example of a recurrent issue we identified is shown in Figure 5.

Additionally, more issues were excluded from the coverage analysis during the this procedure. Specifically, 54 issues asking for information about simple concepts like *"In pair programming, what is each role named, and why?"*. We decide to excluded them since these kind of issues are better covered by basic literature, rather than research evidence provided by SSs. From the 424 unique issues, **we ended up with 370 issues that were in fact considered to coverage analysis**.

2.2.7 Limitations

As any empirical study, this one has its particular limitations. Here we acknowledge the ones we identified.

- The selected SSs came from a tertiary study published in 2011. The most up-to-date tertiary study we could find when we conducted ES1. Thus, issues coverage might be pessimistic since newer SSs may be able to cover more issues.

- The scope is limited to how SSs cover practitioners issues since they can provide more consolidated and mature evidence to practice than primary isolated studies (LAVIS et al., 2003). Though, not covered issues can occur due to either absence of SSs or absence of primary studies. The former can be mitigated conducting a SS providing evidence that helps to solve the issue. The latter demands an effort of the research community as whole since if there are no primary studies on a specific issue, there is no chance to exist SSs covering that issue. To determine why there are few or no SSs covering specific practitioners issues is out of the scope of this study, as well as to investigate the coverage of practitioners issues by primary studies. Thus, it is possible there are primary studies that could provide evidence to cover some practitioners issues we identified, but this is beyond of our scope of investigation.

- Practitioners issues returned by Stack Exchange searches are sensible to SSs search strings. Poor search strings might lead to poor results. On the other side, identification of SSs with poor search strings is a finding itself. Researchers might test

their search strings in Stack Exchange to understand how they are connected with practice.

- The number of issues found applying the search strings can also be negatively influenced by a possible mismatch between the term used in academic research, and the ones used in SE practice. A taxonomy could make this kind of research more reliable, but it is out of scope of thesis.

- The method we proposed cannot be fully automated. The only phase that was automated was the selection of related questions (Section 2.2.4). A search engine based on Apache Lucene was built to automate this process. The remaining phases were conducted manually. To mitigate classification bias, we conducted it in pairs, with conflict resolution meetings.

- One might argue that instead of using five different Stack Exchange websites, we should favor Stack Overflow, the most popular Stack Exchange websites. However, Stack Overflow is focused on programming and, therefore, it demands questions to have a specific, concrete technical answer (*e.g.,* the best Stack Overflow questions present a code snippet (NASEHI et al., 2012)). Such questions rarely fit in topics investigated in the SE research literature.

- Our results cannot be extended to other Q&A websites (e.g. Yahoo Answers, Quora, Experts Exchange), neither all Stack Exchange websites. To mitigate the risk of not taking into account relevant Stack Exchange websites, we classified them according to all SE areas listed by SWEBOK (BOURQUE; FAIRLEY, 2014) and selected only those related to at least one SE area (Table 2).

- SSs may be useful for practitioners even if they do not provide guidelines. For instance, practitioners might get acquainted with an emerging topic or new results. However, we argue that the lack of guidelines represent a serious limitation on these studies, since their absence might leave the reader with no clear answer (*e.g.,* should I use pair programming in my context?). That is why we excluded SSs that do not present guidelines to practitioners.

- Finally, we selected Stack Exchange issues based on their score. This approach might favor old questions since it takes time for a question to have a high score. Although there are several other ways to select issues in Stack Exchange websites, score is a common property used in SE studies for filtering out relevant issues, avoiding low-quality ones (TREUDE; PROLO; FILHO, 2015).

2.3 RESEARCH RESULTS

This section presents the results of this study. Section 2.3.1 shows the overall practitioners issues coverage. Section 2.3.2 exposes each SE area coverage, considering practitioners issues were classified as belonging to at least one of the 15 SWEBOK SE areas. Section 2.3.3 exhibits the coverage of all recurrent issues identified, organized according to each SE area.

2.3.1 Overall Coverage

Figure 6 depicts overall practitioners issues coverage. Among the 370 practitioners issues, only 58 (15.6%) were considered as covered by the selected SSs. This reveals a good opportunity for conducting SSs in SE to fill the gap of not covered issues.

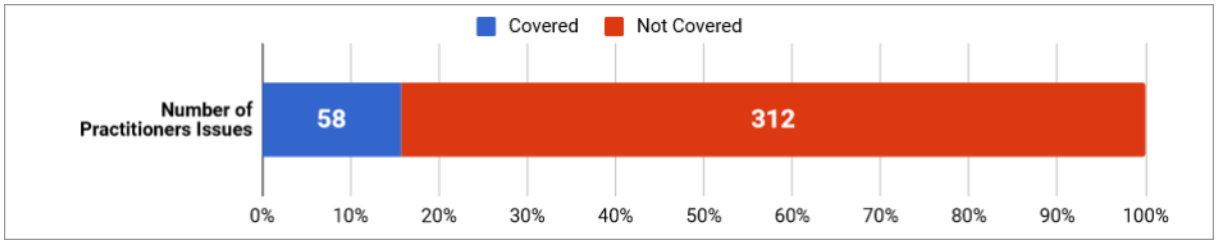


Figure 6 – Practitioners issues overall coverage.

We identified two reasons that explain why the selected SSs do not cover practitioners issues:

- **Issues reporting scenarios SSs do not fully cover**: The majority of the uncovered issues fit under this situation, when SS is related to the same SE area of a practitioner issue, but there are no findings helping to solve it. To illustrate, a user asked *"if pair programming works in case of pairs with different programming skill levels."* This issue is related to SS22 (HANNAY et al., 2009) that investigates pair programming effectiveness. However, the issue was considered not covered because the SS investigated only pairs with the same skills level. This could be avoided if the SS had analyzed pairs with different skills levels. Such analysis would, of course, be limited to the existence of primary studies.

- **Issues using approaches available/popular after the SS was conducted**: For instance, a user asked *"Is BDD actually writable by non-programmers? "*. Although related to agile methods – a topic addressed by some selected SSs – BDD was introduced only in 2012, slightly after the selected SSs publication. So, there is no chance the selected SSs would cover that kind of issue.

Nine practitioners issues particularly took our attention. They reveal practitioners are demanding scientific evidence, which might indicate "bridges" need to be built to transfer knowledge from empirical studies to SE practice (BEECHAM et al., 2014; CARTAXO, 2016; CARTAXO et al., 2016; JEDLITSCHKA; JURISTO; ROMBACH, 2014). The following two issues are given as examples:

> **Are there any studies of cross-functional teams vs. domain-based teams (*e.g.*, project-based vs. software/mechanics/etc)?**
> I work in an organization which creates many integrated systems products - i.e. it is complete products with mechanical/system/electronics/software being designed and manufactured. At the moment most teams are organized around projects in a cross functional way. The advantage of an organization like this is that people who are working closely together for a common goal are close. The disadvantages come from the isolation of engineers from their peers. Typically a project is assigned only one software engineer. This means that the projects have a high truck factor, minimal knowledge sharing and best practices, and technical development is limited. So my question is: **are there any studies comparing the cost/benefits of these two approaches?** (bold emphasis added)

> **Should there be more scientific study of the effectiveness of various hyped-up ideas in software development?**
> Everyone seems to implicitly assume that the free market of ideas will eventually converge on the "right" solutions in software development. We don't assume that in medicine - we recognise that scientific experiments are needed there - so why should we assume it in software development? I am not arguing for regulation of programmers. It is far too early to even talk about that. Before healthcare could be effectively regulated, there was a need for scientific experiments to establish which treatments worked and which didn't. Software engineering doesn't even have this scientific evidence base to back up touted methodologies such as Scrum or Agile, or programming paradigms such as functional programming or MDA. [...] **The question is, why is this scientific evidence base (for all intents and purposes) nonexistent?** (bold emphasis added)

2.3.2 Software Engineering Areas Coverage

Table 8 shows how practitioners issues are covered by the selected SSs separated by each of the 15 SE areas according to SWEBOK (BOURQUE; FAIRLEY, 2014). The "TOTAL" column stands for the number of issues classified under each SE area, and the "COVERAGE" column for the number/percentage of issues classified under each SE area and considered as covered by at least one SS. None of the SE areas presented a coverage rate

above 50%.

Table 8 – Issues coverage per SWEBOK (BOURQUE; FAIRLEY, 2014) software engineering areas.

SOFTWARE ENGINEERING AREA	**COVERED**		**TOTAL**
	#	%	#
Software Engineering Models and Methods	17	13.3%	127
Software Engineering Management	3	4.6%	64
Software Engineering Professional Practice	12	24%	46
Software Requirements	6	15.7%	38
Software Testing	8	26.6%	30
Software Engineering Process	4	20%	20
Software Construction	2	12.5%	16
Software Engineering Economics	4	36.3%	11
Software Maintenance	2	22.2%	9
Software Design	0	0%	9
Software Configuration Management	0	0%	0
Software Quality	0	0%	0
Computing Foundations	0	0%	0
Mathematical Foundations	0	0%	0
Engineering Foundations	0	0%	0

The SE area in which we found more issues is **Software Engineering Models and Methods**, with 127 issues in total. Three SSs present findings in that area, they are: SS11 (DYBÅ; DINGSØYR, 2008) , SS22 (HANNAY et al., 2009), and SS24 (HOSSAIN; BABAR; PAIK, 2009) as shown in Table 1. It is also the SE area with the highest number of covered issues, 17 in total (13.3%). All the issues from that SE area are related to agile methods/practices. The three most recurrent issues are: Applicability of agile in specific project context; Mixing agile with traditional methods/practices; and Benefits of agile methods/practices in general. More details about recurrent issues classified under Software Engineering Models and Methods area are shown in Section 2.3.3.

The second SE area in which we found more issues is **Software Engineering Management**, with 64 issues in total, less than half of what was found in Software Engineering Models and Methods. It is also the SE area with the highest number of SSs: five in total. They are: SS1 (KITCHENHAM; MENDES; TRAVASSOS, 2006), SS6 (JIMÉNEZ; PIATTINI; VIZCAÍNO, 2009), SS12 (JORGENSEN, 2005), SS14 (JøRGENSEN, 2007), and SS19 (LOPEZ; NICOLAS; TOVAL, 2009). However, only three practitioners issues were covered, resulting in the lowest coverage rate among all SE areas, 4.6%. The three most recurrent issues of this SE area are: Strategies to cost and effort estimation in specific contexts; Tools to support project management with specific features; and Strategies and metrics to measure team

productivity. More details about recurrent issues classified under Software Engineering Management area are shown in Section 2.3.3.

The third SE area in which we found more issues is **Software Engineering Professional Practice**, with 46 issues in total. Just one selected SS is related to this area, SS16 (BJøRNSON; DINGSøYR, 2008). Eleven practitioners issues were considered as covered, which corresponds to a 24% coverage rate. The three most recurrent issues of this SE area are: Strategies to deal with knowledge management in a team; Difficulties dealing with customer in specific contexts; Improving communication in a distributed team. More details about the recurrent issues classified under Software Engineering Professional Practice area are shown in Section 2.3.3.

2.3.3 Recurrent Issues Coverage

In this section, we present the coverage of each recurrent issue organized per each SE area. The five most recurrent issues — the ones that aggregate highest number of practitioners issues — are:

1. **Applicability of agile in specific project context** (19 issues);
2. **Introducing and adapting a software development process in specific context** (14 issues);
3. **Strategies to cost and effort estimation in specific contexts** (13 issues);
4. **Mixing agile with traditional methods/practices** (10 issues);
5. **Tools to support software testing with specific features** (10 issues).

We now discuss each recurrent issue per SE area. For each SE area we present a table summarizing the coverage rate. The "TOTAL" column stands for the number of issues classified under each recurrent issue, and the "COVERAGE" column for the number/percentage of issues classified under each recurrent issue and considered as covered by at least one SS. For each SE area, we quote at least one issue, and discuss why that issue was considered as covered or not. If the issue was considered as not covered, we provide discussions on how SSs could evolve to cover the issue. The *Miscellaneous* category aggregates issues that do not fit in any of the identified recurrent issues, although belonging to a SE area.

Software Requirements (6 recurrent issues): Table 9 shows recurrent issues related to software requirements area. None of the recurrent issues presented a coverage higher than 50%.

Table 9 – Coverage of software requirements recurrent issues.

RECURRENT ISSUE	SWEBOK SECTION	COVERED		TOTAL
		#	%	#
Tools to manage requirements with specific features	Software Requirements Tools. Chapter 1, Section 8	3	37%	8
Approaches to manage constant requirements change	Change Management. Chapter 1, Section 7.2	0	0%	6
Select requirements elicitation techniques in specific contexts	Elicitation Techniques. Chapter 1, Section 3.2	1	20%	5
Strategies to prioritize requirements in specific contexts	Requirements Classification. Chapter 1, Section 4.1	0	0%	3
Defining requirements attributes in specific contexts	Requirements Attributes. Chapter 1, Section 7.3	0	0%	3
User stories to specify non-functional requirements	Elicitation Techniques. Chapter 1, Section 3.2	0	0%	2
Miscellaneous	—	2	18%	11

The most recurrent issue is the need for **Tools to manage requirements with specific features**, with eight issues grouped under this classification. SS18 (NICOLáS; TOVAL, 2009) and SS21 (IVARSSON; GORSCHEK, 2009) provided evidence about tools to manage software requirements that could cover three out of eight practitioners issues which corresponds to a coverage rate of 37%. To illustrate, one example of issue follows:

> **What FOSS solutions are available to manage software requirements?**
> In the company where I work, we are starting to plan to be compliant to the software development life cycle. We already have, wiki, vcs system, bug tracking system, and a continuous integration system. The next step we want to have is to start to manage, in a structured way, software requirements. [...] **We are trying to search and we hope we can find and use a FOSS software to manage all this things.** We have about 30 people, and don't have a budget for commercial software [...] **Required features:** Software requirements divided in a structured configurable way; versioning of the requirements (history, diff, etc, like source code); Interdependency of requirements (child of, parent of, related to); Rule Based Access Control for data handling; (bold emphasis added)

To mitigate this issue, we believe SSs aimed at identifying tools, comparing their features, or assessing their effectiveness for software requirements engineering practice might play an important role. One example is the study conducted by Marshall et al. (MARSHALL et al., 2014) that identified, analyzed, and compared tools based on their features. However, the study analyzed tools to support SSs in SE, not software requirements.

The second most recurrent issue is the need for **Approaches to manage constant requirements change**, with six issues classified. Unfortunately, none SSs were able to provide useful information helping to solve the issues. To illustrate, following there is one of those issues:

> **How do you deal with the costs of too-rapid change?**
> Like most modern developers I value Agile principals like customer collaboration and responding to change, but what happens when a product-owner (or whoever determines requirements and priorities) changes requirements and priorities too often? Like several times a day? [...] **Is there some [...] in-depth study, metaphor, or quote that can help me reduce the amount of wasted effort or at least explain the costs of this chaotic behavior?** (bold emphasis added)

It is interesting to note that this second most recurrent issue may challenge one of the building blocks of agile software development, which is *"Responding to change over following a plan"* (MANIFESTO...,). We noticed that ideal balance between embracing changes and proper planning seems not to be fully accomplished in practice. SSs aimed at identifying and analyzing strategies to manage requirements change, comparing their pros and cons, might improve practitioners confidence facing this kind of scenario.

The next three most recurrent issues present a peculiar characteristic. They report particular problems but focus on the necessity of information that fits in their specific contexts, they are: Recommendations on how to **Select requirements elicitation techniques under specific contexts**; **Strategies to prioritize requirements in specific contexts**; and recommendations on **Defining requirements attributes in specific contexts**. Those three recurrent issues reinforce many claims about the importance of highly contextualized evidence to provide useful information to practitioners (DYBÅ; SJØBERG; CRUZES, 2012). A SS identified mechanisms to characterize context of primary studies in software engineering (CARTAXO et al., 2015a). However, as far as we know there are no guidelines to support context characterization of evidence for SSs.

Software Design (2 recurrent issues): We identified two recurrent issues related to software design. However, none was considered as covered by the selected SSs, as can be seen in Table 10.

Table 10 – Coverage of software design recurrent issues.

RECURRENT ISSUE	SWEBOK SECTION	COVERED		TOTAL
		#	%	#
Benefits of a service layer, when compared to libraries	Other Methods. Chapter 2, Section 7.6	0	0%	4
Strategies to structure user interface design	User Interface Design. Chapter 2, Section 4	0	0%	2
Miscellaneous	—	0	0%	3

The most recurrent issue of software design area is related to the **Benefits of a service layer, when compared to libraries**. None of the selected SS provide evidence that could cover this kind of issue. To illustrate, following there is one of the four issues under this classification:

> **How essential is it to make a service layer?**
> I started building an app in 3 layers (DAL, BL, UI), it mainly handles CRM, some sales reports and inventory. A colleague told me that I must move to service layer pattern, that developers came to service pattern from their experience and it is the better approach to design most applications. He said it would be much easier to maintain the application in the future that way. Personally, I get the feeling that **it's just making things more complex and I couldn't see much of a benefit from it that would justify that** (bold emphasis added)

The plethora of different service layers, as well as the rich set of software libraries found in any high-level programming language makes such investigation challenging. A SS aimed at studying this particular scenario might present evidence for practitioners facing this kind of situation. However, the existence of such SS is limited by the presence of primary studies related to this topic (*e.g.,* (PAPAZOGLOU, 2003)).

Software Construction (2 recurrent issues): We found two recurrent issues related to the software construction area, as shown in Table 11.

Table 11 – Coverage of software construction recurrent issues.

RECURRENT ISSUE	SWEBOK SECTION	COVERED		TOTAL
		#	%	#
Object modeling techniques	Construction Design. Chapter 3, Section 3.1	0	0%	7
Code duplication avoidance	Construction for Reuse. Chapter 3, Section 3.5	2	40%	5
Miscellaneous	-	0	0%	4

The only recurrent issue covered refers to **code duplication avoidance** with a 40% coverage rate. To illustrate, following there is one issue under that classification:

> **When is 'cloning', rather than reusing, a module acceptable design solution?**
> For this question, I'll give an example module to facilitate the discussion, Let's say the module is a calculation engine, It currently servers its purpose for its current audience. The requirement is to clone the same engine but with some tweaking for an entirely new audience. Given that, these are Considerations/Factors that will affect the design solution: [...] However, **I am still conflicted, since: It will inherently be a copy-paste solution; Duplicate code;**[...] Is the compromise acceptable in this situation, given the user expectations highlighted above? And follow up question, is there something I can add to the solution that will address the conflicting issues [...] (bold emphasis added)

Code clones have a long history in SE research, with traditional studies dating from the 1990s (*e.g.,* (BAXTER et al., 1998)). SS15 (HORDIJK; PONISIO; WIERINGA, 2009) is

an example of a SS that investigates code clones, providing evidence and guidance for practitioners. In particular, it builds a model to demonstrate under which circumstances code duplication harms system quality. It also provides strategies to mitigate each of these situations. Since this recurrent issue comprises a wide spectrum of code duplication, this might explain this recurrent issue coverage.

Software Testing Coverage (1 recurrent issue): Among the 30 practitioners issues regarding software testing, we identified only one recurrent issue, which is **Tools to support software testing with specific features**. Ten out of the 30 issues were classified in this recurrent issue, although only 3 of them are covered by the selected SSs, as can be observed on Table 12. SS5 (HAUGSET; HANSSEN, 2008) offers evidence about FiteNesse (FITNESSE...,) and other tools related to automated acceptance testing, covering all the three issues. Following there is one issue classified under this recurrent issue.

Table 12 – Coverage of software testing recurrent issues.

RECURRENT ISSUE	SWEBOK SECTION	COVERED		TOTAL
		#	%	#
Tools to support software testing with specific features	Software Testing Tools. Chapter 4, Section 6	3	30%	10
Miscellaneous	-	5	25%	20

Fitnesse vs Robot

We are choosing what system to start using in our company. it should be used for both backend (REST API, some DB checks) and UI testing; it should use a simple language so even non-programmers/tester can understand the test cases (Product Owners should be able to see whether all acceptance criteria are covered);it should support integration with Jenkins; it should support versioning of test cases so that for a particular product version we also can check out relevant test cases; right now we use TestRail (test case management SW) [...]

Software Maintenance Coverage (2 recurrent issues): We identified two recurrent issues under the area of software maintenance. Together they group nine practitioners issues as shown in Table 13.

Table 13 – Coverage of software maintenance recurrent issues.

RECURRENT ISSUE	SWEBOK SECTION	COVERED		TOTAL
		#	%	#
Developers demanding refactoring the entire legacy system	Reengineering. Chapter 5, Section 4.2	2	40%	5
Strategies to perform refactoring	Reengineering. Chapter 5, Section 4.2	0	0%	4

The most recurrent issue is the one **Developers are demanding refactoring the entire legacy system**, with five issues under this classification. This shows maintenance

activities are still problematic due to either bad software design/construction, or to a culture among software developers that prefer to spend effort reinventing the wheel instead of understanding and evolving a legacy system (HUNOLD et al., 2009). Following there is one issue that illustrates this situation:

> **Reengineering the project from scratch**
> I am currently working on a project that has been in development for the last few years used throughout the organization but the way the project has been coded the maintainability of it is completely shot. Reading the code presents with pages and pages of Anti-Patterns and trying to identify the path of a business workflow takes on occasion days. At this point I would probably classify the software in its current state as "Working by accident" rather then as intended. So I am looking for some wisdom as to the following: **At what point would you consider simply dumping the project into an abandonware pile and starting from scratch?** (bold emphasis added)

Evidence that supports decision-making during software design and construction could help software developers to design systems with higher maintainability. For cases where developers want to reinvent the wheel, proper training can help to reduce the impetus to re-implement a system from scratch (HUNOLD et al., 2009).

Software Engineering Management Coverage (6 recurrent issues): Software engineering management is the second area with more issues related to it, with a total of 64 issues. We could identify six recurrent issues among them. However, four are not covered at all, and two present a coverage rate below 50%, as shown in Table 14.

Table 14 – Coverage of software engineering management recurrent issues.

RECURRENT ISSUE	SWEBOK SECTION	COVERED		TOTAL
		#	%	#
Strategies to cost and effort estimation in specific contexts	Effort, Schedule, and Cost Estimation. Chapter 7, Section 2.3	1	7%	13
Tools to support project management with specific features	Software Engineering Management Tools. Chapter 7, Section 7	0	0%	6
Strategies and metrics to measure team productivity	Reviewing and Evaluating Performance. Chapter 7, Section 4.2	1	25%	4
Strategies to manage distributed teams	Software Engineering Management. Chapter 7	0	0%	4
Tasks that do not fit in one sprint	Effort, Schedule, and Cost Estimation. Chapter 7, Section 2.3	0	0%	4
Strategies to negotiate project scope	Determination and Negotiation of Requirements. Chapter 7, Section 1.1	0	0%	3
Miscellaneous	-	1	3.3%	30

The most recurrent issue is practitioners asking for **Strategies to cost and effort estimation in specific contexts**. Only one out of 13 issues is covered. SS12 (JORGENSEN, 2005) and SS14 (JøRGENSEN, 2007) are related to cost and effort estimation. However, SS12 did not return any issue from the search, and SS14 is focused on the comparison

of expert judgment versus formal models to estimate effort in general. On the other side, issues are context specific. An example of one issue classified under this recurrent issue is:

> **How does a team (new to product and domain) estimate user stories of a ten year old product?**
> I am the scrum master for one of the products in a software development company. Our team, including me, operates from India. However my product owner is in USA. We are working on the feature development for this product that exists for ten years now. **Our team in India started six months ago, with no product nor domain knowledge on it.** (bold emphasis added)

A SS identifying and comparing estimation techniques could support practitioners facing problems like the one we showed in the example. Additionally, it is important to note that evidence of such techniques effectiveness needs to be contextualized, so practitioners can check if they fit in their working environment. In the example we have shown, only techniques to deal with situations of low knowledge about software domain matter.

The second most recurrent issue is when practitioners ask for recommendations of **Tools to support project management with specific features**. None selected SSs could cover the six issues under this classification. In previous sections, we already discussed approaches providing evidence when recommendations about tools are demanded.

The third most recurrent issue is when practitioners ask for **Strategies and metrics to measure team productivity**. SS13 (TRENDOWICZ; MüNCH, 2009) could cover one of the four issues under this classification. To illustrate, one issue grouped under this recurrent issue is:

> **How to measure team productivity?**
> The upper management at our company has laid out a goal for our software team to be 15% more productive over the next year. Measuring productivity in a software development environment is very subjective, but we are still required to come up with a set of metrics. **What sorts of data can we capture that would measure our team's productivity?** (bold emphasis added)

A mapping study identifying software team productivity metrics could provide an interesting overview to practitioners facing that kind of problem. One good example in a different topic is the Saraiva's et al. mapping study (SARAIVA et al., 2012) that identified metrics to measure how software maintainability is affected by aspect-oriented programming. Traditional systematic literature reviews and meta-analyses can also provide rich evidence about the effectiveness and applicability of each metric.

Software Engineering Process Coverage (1 recurrent issue): We identified just one recurrent issue among the 20 issues related to software engineering process. The issues coverage is presented in Table 15.

Table 15 – Coverage of software engineering process recurrent issues.

RECURRENT ISSUE	SWEBOK SECTION	COVERED		TOTAL
		#	%	#
Introducing and adapting a software development process in specific context	Software Process Adaptation. Chapter 8, Section 2.3	3	21%	14
Miscellaneous	-	1	16%	6

The only recurrent issue identified is practitioners asking for ways to support them **Introducing and adapting a software development process in specific context**, with 14 issues under this classification. Just three issues are covered. SS11 (DYBÅ; DINGSØYR, 2008) SS16 (BJØRNSON; DINGSØYR, 2008), and SS24 (HOSSAIN; BABAR; PAIK, 2009) are the ones that presented evidence that could help to solve the three issues considered as covered. Following there is one example of issue:

> **Introducing Scrum in a distributed team**
> We would like to start using scrum [...] Until know we used a "home-grown" methodology, but we would like to switch to something more defined and mature. Scrum would be a great choice in my opinion and also the management supports us to go agile. **Where should we start this transition? Is there some guide or best practices for this transition?** (bold emphasis added)

SSs identifying processes as well as best practices during introduction or adaptation of a process/method could provide useful evidence for issues likes the one we mentioned, specially when they are highly contextualized.

Software Engineering Models and Methods Coverage (15 recurrent issues): This area is the one with more issues, 127 in total. We could identify 15 recurrent issues, all of them are related to agile software development. Table 16 shows those issues coverage.

The most recurrent issue is about **Applicability of agile in specific project context**. Only one out of the 19 issues under that classification was covered by SS11 (DYBÅ; DINGSØYR, 2008). Some issues are highly tied to their project context, and that is why contextualized evidence is important to assess whether its applicable in real environments. To illustrate, following there is one of the issues under that classification, which asks for evidence about the applicability of agile methods in a context of firmware/embedded project:

Table 16 – Coverage of software engineering models and methods recurrent issues.

RECURRENT ISSUE	SWEBOK SECTION	COVERED		TOTAL
		#	%	#
Applicability of agile in specific project context	Agile Methods. Chapter 9, Section 4.4	1	5%	19
Mixing agile with traditional methods/practices	Agile Methods. Chapter 9, Section 4.4	0	0%	10
Benefits of agile methods/practices in general	Agile Methods. Chapter 9, Section 4.4	3	33%	9
Pair programming to transfer knowledge	Agile Methods. Chapter 9, Section 4.4	0	0%	9
Pair programming with distributed pairs	Agile Methods. Chapter 9, Section 4.4	1	12%	8
Impact of low detail level or absence of documentation in agile	Agile Methods. Chapter 9, Section 4.4	0	0%	6
Benefits of agile methods/practices from a specific perspective	Agile Methods. Chapter 9, Section 4.4	2	40%	5
Low customer collaboration in agile	Agile Methods. Chapter 9, Section 4.4	1	25%	4
Pair programming hindering concentration	Agile Methods. Chapter 9, Section 4.4	0	0%	4
Benefits of agile methods/practices in specific contexts	Agile Methods. Chapter 9, Section 4.4	0	0%	4
Mixing multiple agile methods/practices	Agile Methods. Chapter 9, Section 4.4	0	0%	4
Tools for agile methods/practices	Agile Methods. Chapter 9, Section 4.4	0	0%	4
Negative impact of agile in software design	Agile Methods. Chapter 9, Section 4.4	3	100%	3
Ad-hoc software development as agile	Agile Methods. Chapter 9, Section 4.4	0	0%	3
Pair programming as replacement to code reviews	Agile Methods. Chapter 9, Section 4.4	0	0%	2
Miscellaneous	-	5	15%	33

> **How to adopt agile methodology for developing firmware/embedded-systems-software?**
> [...] how to apply agile methods in large complex embedded system software (100+ engineers). Firmware development has some unique characteristics that make it difficult to do agile (ie. Hardware is not available until late in the dev cycle; Once product is released, can't easily update firmware; etc...) The norm in this kind of development is thick documentation and grueling peer reviews. You can't get a simple code fix like renaming a variable without 2-3 signatures. (I exaggerate a little but this is typical. Additionally, a lot of people do take shortcuts and the Project Managers even approve them especially in the face of hard market deadlines.) **I would like to hear any tips or guidelines on how to adopt agile methodology for firmware development projects.** (bold emphasis added)

Studies aggregating and synthesizing evidence from cases studies, ethnographies, and action researches could provide interesting information for practitioners who want to decide which agile methods/practices fit in their contexts.

The second most recurrent issue is practitioners asking for strategies to deal with **Mixing agile with traditional methods/practices**. None of the ten issues under this classification were considered as covered by the selected SSs. To illustrate this situation, one of the issues under this classification follows:

> **How to synchronize an agile software team with a waterfall hardware team?**
> Our team is composed of both software and hardware engineers. The software team uses Scrum project management while the hardware team uses waterfall. The priority of our software requirements change quite frequently, so staying Agile makes sense for us. The priority of our hardware requirements are rather static and slow-moving, so again sticking with waterfall makes sense for us. **The tricky part is the integration of hardware and software. Are there any methodologies for deterministically synchronizing these two contrasting project management styles?** (bold emphasis added)

In some cases there is an impression that once agile is adopted, every stakeholder and process operate through agile philosophy. However as we can see, many practitioners face situations where their team is agile, but not their company as a whole. Or even when the team is agile, but not their customer. Such situations provoke many disarrangements during software development life-cycle. In a survey in Microsoft, Begel et al. (BEGEL; ZIMMERMANN, 2014) identified that agile is not adopted simultaneously by all teams. They reported one situation engineers are more worried about how agile teams coordinate dependencies and deliverables with non-agile teams. However, more evidence is demanded in this topic, specially identifying strategies to deal with such kind of scenarios.

Three recurrent issues were identified around agile methods/practices benefits. The most recurrent is when practitioners ask for information about **Benefits of agile methods/practices *in general***, which can be illustrated by the following issue:

> **What makes Agile software development so appealing?**
> Agile software development is becoming a pretty fun buzzword these days. [...] Whether it is crystal, agile methods, dsdm, rup, xp, scrum, fdd, tdd, you name it. [...] **what are the biggest reasons for choosing to do Agile development** (bold emphasis added)

For those cases, SS11 (DYBÅ; DINGSØYR, 2008) provides useful evidence since it focus on agile benefits in general, chiefly aggregating qualitative evidence.

Another recurrent issue is when practitioners ask for information about **Benefits of agile methods/practices from a specific perspective**. Following there is one issue to exemplify:

> **What are the monetary benefits of going agile?**
> Why go agile? This is the first question that comes to my mind when I think of going agile. What are the **possible financial benefits** one can achieve from going agile? (bold emphasis added)

This situation corroborates with the idea that empirical evidence should comprise not

only data about the effectiveness of an intervention, but also useful information for target audience, in this case, cost-effectiveness (ALI, 2016). We also identified other issues that report the agile benefits information from other specific perspectives, beyond monetary, such as, benefits from the perspective of developers, managers, testers, customers, and others.

The third and last recurrent issue on agile benefits is when practitioners demand for information about **Benefits of agile methods/practices in specific contexts**, which again shows the importance of contextualized evidence.

Four recurrent issues about pair programming were identified. The first most recurrent issue in that matter is the one that practitioners ask about the applicability of **Pair programming to transfer knowledge** from more skilled developers to less skilled ones. None of the nine issues classified under this recurrent issue are covered by the selected SSs. SLR22 (HANNAY et al., 2009) evaluated the impact of pair programming in many dimensions, but only considered pairs with same experience level, failing to perceive practical importance of evaluating pair programming with pairs with different levels of experience and domain knowledge. To illustrate, one issue under this classification is:

> **Pair programming when driver and observer have different skill level and experience**
> I just wonder [if] the strategy still work in the case. For example if they have a very different programming skill level. if one never experience in the problem domain while another have. Is it still OK if they have low programming skill level?

The second most recurrent issue about pair programming is practitioners demanding support to have **Pair programming with distributed pairs**. Just one out of eight issues is covered. SLR22 (HANNAY et al., 2009) was the SS that provided evidence to cover the issue. To illustrate, an issue classified under this recurrent issue is:

> **Any suggestions for pair programming with external resource?**
> [...] I was considering hiring a developer [...] to assist me. Ideally, we would be a collaborative team [...] Has anyone attempted this? [...]

SSs identifying and comparing tools to support distributed pair programming could provide interesting information to practitioners.

Another recurrent issue we identified about pair programming is practitioners reporting problems with **Pair programming hindering concentration**. None of the four issues under this classification are covered. To illustrate, an issue classified under this recurrent issue is:

How can my team reconcile flow and pair-programming?
[...] Flow is a mental state attained by creativity workers (engineers, writers, programmers, etc.) which is often described as a state of immersion in which time seems to pass unknowingly and creative work flows from the mind [...] Pair programming, advocates a two person team which functions as an single organic, programming entity to accomplish a single goal. [...] are these ideas reconcilable?

A primary study has reported that concentration decreases in long pair programming sessions due to exhaustiveness of pair dynamics (VANHANEN; LASSENIUS, 2007). However, more evidence is needed about this topic to draw more accurate conclusions.

Another recurrent issue that deserves mention is practitioners affirming to experience **Ad-hoc software development as agile**. None of the three issues under this classification could be covered by the selected SSs. To illustrate, one issue classified under this recurrent issue is:

Is the agile approach too much of a convenient excuse for cowboys
I believe that an agile approach is best for projects where the requirements are fuzzy and a lot of interaction is required to help shape the end user's ideas. However... **In my professional work, I keep ending up at companies where an "agile" approach is used as an excuse as to why no effort was put into an up front design; when the requirements are well understood**. (bold emphasis added)

Some empirical studies have observed situations like that (HILKKA; TUURE; ROSSI, 2005; CAO; RAMESH, 2007). However, SSs are demanded to investigate that kind of scenario, as well as to identify best practices to avoid it.

Another recurrent issue we identified shows practitioners asking about the **Impact of low detail level or absence of documentation in agile** and how to deal with that situation. None of the six issues under this classification are covered. To illustrate, there is the following issue:

Disillusioned with agile; how do you prepare for life after release 1.1?
My company is going full steam with the agile process, with multiple agile projects in work. [...] establish ideal documentation effort, the team was quickly disbanded [...] the next sprint leaves little documentation, little vision, and poor records of that design decisions were made [...]

This issue possibly challenges basic principles of agile software development, like *"working software over comprehensive documentation"* (MANIFESTO...,). This does not mean we should go back to over-documentation times and high costs associated with it. However, it is noticeable that an ideal balance between documentation and working software is, sometimes, far from being accomplished in practice. SSs identifying current strategies to establish ideal documentation, their pros and cons under specific contexts,

and an evaluation of those strategies, could provide evidence to practitioners facing that scenario.

Another recurrent issue challenging a basic principle of agile methods/practices emerges when practitioners report problems due to **Low customer collaboration**. The agile manifesto (MANIFESTO...,) says that *"Customer collaboration over contract negotiation"* is one of the most important values. However, situations in practice may hinder agile adoption due to difficulties collaborating with customers. Some studies also observed such kind of situations (HODA; NOBLE; MARSHALL, 2011), but there is a need of more evidence around this topic. Just one out of four issues under this classification are covered, this time by SS11 (DYBÅ; DINGSØYR, 2008).

Software Engineering Professional Practice Coverage (7 recurrent issues): This area is the third with more issues related to, with a total of 47. We could identify seven recurrent issues. Three of them present a coverage rate equal to or above 50%, two present a coverage rate below 50%, and two are not covered at all, as can be seen in Table 17.

Table 17 – Coverage of software engineering professional practice recurrent issues.

RECURRENT ISSUE	SWEBOK SECTION	COVERED		TOTAL
		#	%	#
Strategies to deal with knowledge management in a team	Team and Group Communication. Chapter 11, Section 3.3	4	50%	8
Difficulties dealing with customer in specific contexts	Interacting with Stakeholders. Chapter 11, Section 2.4	1	20%	5
Improving communication in a distributed team	Team and Group Communication. Chapter 11, Section 3.3	4	100%	4
Difficulties dealing with team members in specific contexts	Interacting with Stakeholders. Chapter 11, Section 2.4	1	33%	3
Team rotation	Dynamics of Working in Teams/-Groups. Chapter 11, Section 2.1	0	0%	4
Ideal workplace layout	Group Dynamics and Psychology. Chapter 11, Section 2	0	0%	3
Tools to support knowledge management with specific features	Team and Group Communication. Chapter 11, Section 3.3	1	50%	2
Miscellaneous	-	1	5%	17

The most recurrent issue is when practitioners ask for **Strategies to deal with knowledge management in a team**. Four out of eight issues under this classification are covered. SS16 (BJøRNSON; DINGSøYR, 2008) is responsible for offering knowledge that helps to solve those issues. This happens because SS16 investigates concepts, findings, and methods to manage knowledge in SE. To illustrate this recurrent issue, following there is an issue:

How to motivate team for knowledge sharing sessions
I work in a team with wide range of expertise and experience. I have been trying to introduce weekly knowledge sharing sessions. Sessions of 30-60 min length where everybody gets a chance to present something and talk about it [...] However, the team is not motivated towards this, either the attendance is too low or none. How to get a team work towards such an idea?

The second most recurrent issue is practitioners reporting to face **Difficulties dealing with customer in specific contexts**. Only one out of five issues under this classification is covered. To illustrate that situation there is the following issue:

Does anyone have experience with a difficult customer?
We have a reoccurring conflict with one of our larger and strategically important customers [...] I'm looking for hints on how to control the development-process (and our own economy) and still provide the customer with a product that gives the company most value-for-money [...]

To deal with stakeholders, a particular set of skills is necessary. These skills sometimes are even more important than traditional technical skills in SE world (DAMIAN; BORICI, 2012). Yaman et al. (YAMAN et al., 2016) conducted an interesting SS about benefits, challenges, methods, and tools to support customer involvement in a continuous deployment environment. A similar approach could be adopted and an SS with recommendations and best practices about customer relationship could provide useful information for practitioners facing that kind of problem.

Improving communication in a distributed team also seems to concern practitioners since it is the third most recurrent issue under Software Engineering Professional Practice area. This recurrent issue is particularly well-supported since all four issues under this classification are covered by SS24 (HOSSAIN; BABAR; PAIK, 2009).

Another recurrent issue is practitioners reporting **Difficulties dealing with team members in specific contexts**. As with issues related to difficulties dealing with customers, it is important to explore the human aspects of that kind of environment, and strategies to deal with it since team conflicts might hinder team performance (DUBINSKY et al., 2011; KLÜNDER et al., 2016).

Another recurrent issue is when practitioners ask for information about **Team rotation**. None of the four issues under this classification are covered. A primary study was published recently building a team rotation theory grounded in a qualitative case study (SANTOS et al., 2016). However, more evidence is needed to support the theory, as well as to aggregate enough empirical evidence to conduct a SS on this topic.

Finally, some practitioners ask for information about **Ideal workplace layout** for a software development team. None of the three issues classified under this recurrent issue are covered by the selected SSs. There are some primary studies under this topic, for ex-

ample, Sykes' study (SYKES, 2011) proposing workplace layout strategies to reduce level of interruption in a software development team. However, more evidence is demanded in this topic, specially addressing strategies to define workplace layout aiming at different types of goals beyond reducing interruptions.

Software Engineering Economics Practice Coverage (3 recurrent issues): All three recurrent issues we identified under this SE area are related to outsourcing/offshore software development. Table 18 shows how the selected SSs covered those issues.

Table 18 – Coverage of software engineering economics recurrent issues.

RECURRENT ISSUE	SWEBOK SECTION	COVERED		TOTAL
		#	%	#
Strategies to introduce outsourcing/offshoring in specific contexts	Offshoring and Outsourcing. Chapter 12, Section 5.4	1	20%	5
Characteristcs of a good outsourcer/offshorer	Offshoring and Outsourcing. Chapter 12, Section 5.4	1	33%	3
Improving communication with the outsourcer/offshorer	Offshoring and Outsourcing. Chapter 12, Section 5.4	1	50%	2
Miscellaneous	-	1	100%	1

The most recurrent issue is when practitioners ask for **Strategies to introduce outsourcing/offshoring in specific contexts**. One of the five issues under this classification is covered. For instance, the following issue:

> **How to outsource the UI of a dynamic Web application?**
> [...] The project is nearly completed, but I want to enhance the look and feel. This will include better graphics and some extra behaviour. I want to outsource this task [...] I don't have experience with outsourcing and don't know how to incorporate an outsider to the project [...]

SS7 (KHAN; NIAZI; AHMAD, 2009a) and SS8 (KHAN; NIAZI; AHMAD, 2009b) provide useful evidence and could cover one issue. However, more evidence is demanded for specific contexts. For instance, the one in the example demands information about the possibility to outsource user interface web development. Strategies and recommendations may be different if the intention is to outsource software testing, or other parts/components of a software system.

The second most recurrent issue is when practitioners ask about the **Characteristics of a good outsourcer/offshorer**. Again SS7 (KHAN; NIAZI; AHMAD, 2009a) and SS8 (KHAN; NIAZI; AHMAD, 2009b) provided useful evidence and could cover one issue. However, the same problem concerning context limited the level of coverage of that kind of issues. To illustrate this recurrent issue, following there is an issue:

> **What to look for in an outsourced partner**
> [...] We might need some more development help and are looking at an Indian company which comes recommended by someone we know (although they are not very technical). I'll be having an informal chat with them, and thought I'd see if people here had some wisdom regarding **what to look for** and good questions to ask. (bold emphasis added)

The third recurrent issue is the one where practitioners ask for guidance for **Improving communication with the outsourcer/offshorer**. SS7 (KHAN; NIAZI; AHMAD, 2009a) and SS8 (KHAN; NIAZI; AHMAD, 2009b) provided useful evidence and could cover one issue. However, the same problem concerning context limited the level of coverage of that kind of issues. To illustrate, there is the following issue classified under this recurrent issue:

> **Communicating requirements to offshore teams**
> Just to give a context, there is an offshore team in India for a client in San Francisco. The offshore team is about 9 developers and 4 Q&A, with one project manager. I am doing onsite coordination for this team from the client location [...] They obviously fail to deliver sprint after sprint. What would you do to get these things right? How much of adequate is adequate clarity in requirement? [...]

2.4 DISCUSSION

There are healthy Stack Exchange websites beyond Stack Overflow, focusing on non-technical issues and ready to be explored: This enhances the possibilities to discover which are the issues practitioners are facing and to explore them through research aiming to provide empirical evidence connected to software engineering practice demands. Those Stack Exchange websites approach a wide variety of SE areas, even though not all of them yet. For researchers who want to explore practitioners issues in SE through Stack Exchange, we recommend the following websites: *Programmers* (PROG) (STACKEXCHANGE..., a), *Project Management* (PM) (STACKEXCHANGE..., b), and *Software Quality Assurance & Testing* (SQA) (STACKEXCHANGE..., d). The *Reverse Engineering* (RE) (STACKEXCHANGE..., c) website needs further investigation to understand whether it is a good source or not. On the other hand, we do not recommend *Software Recommendations* (SREC) (STACKEXCHANGE..., e) website, since it is focused on tools in general, not tools to support SE practice.

Many problems were observed with SSs search strings: Some SSs present malformed search strings, or even do not report them at all. This hinders replicability, one important characteristic of systematic methods proposed by EBSE community. Part of the low quantity of practitioners issues is related to poorly defined SSs' search strings.

Seven out of 24 SSs search strings did not return any issue. This occurred due to search strings with few key terms, key terms with no synonyms, or key terms with composed synonyms only. Additionally, one main reason for the excess of false positives is also poorly defined search strings. For instance, search strings using rather common terms. Despite SSs' search strings were not originally defined to search for practitioners issues in Q&A platforms like Stack Exchange, such poorly defined search strings can lead to problems even using them with their original purpose, which is to find primary studies in search engines. On Da Silva's et al. tertiary study (SILVA et al., 2011), which is based on two other tertiary studies (KITCHENHAM et al., 2009; KITCHENHAM et al., 2010), each of the 120 SSs was evaluated on their quality. The evaluation was based on a questionnaire with four questions, and one of those questions was "Are the review's inclusion and exclusion criteria described and appropriate?". We believe a similar question should be added to their quality questionnaire: "If the review uses automatic search, are the review's search string described and appropriate?" This would increase quality of SSs, specially their potential to be fully replicated.

Only 1.75% of most relevant practitioners issues are related to the selected SSs: This shows SSs are still far from touching the whole spectrum of topics discussed by practitioners. It is important to reinforce that related issues are not the same of covered issues. One related issue needs to belong to the same SE area of the SS, whereas a covered issue needs one SS with at least one finding that helps to solve the issue. So the selection of SSs was very strict. And we considered any issue.

Only 15.6% of the 370 practitioners issues related to the selected SSs are covered: This suggests the selected SSs are also facing a difficult time to provide evidence covering specific issues faced by practitioners. Section 2.4.1 presents some guidelines to support researchers wanting to conduct SSs more connected with practitioners issues.

Practitioners issues related to nine out of the 15 SWEBOK SE areas are covered to some extent by the selected SSs: No SE area presented a coverage rate above 50%. The two SE areas with higher coverage rates are Software Engineering Economics and Software Testing with 36.3% and 26.6% respectively. No issue related to Software Design is covered by the selected SSs. Additionally, the selected SSs' search strings could not find any issue related to the following SE areas: Software Configuration Management, Software Quality, Computing Foundations, Mathematical Foundations, and Engineering Foundations.

We identified 45 recurrent issues distributed in many SE areas: The three most often recurring issues are: Applicability of agile in specific project context (19 issues); Introducing and adapting a software development process in specific context (14 issues); and Strategies to cost and effort estimation in specific contexts (13 issues).

There are practitioners explicitly asking for scientific empirical evidence in Stack Exchange websites: This shows there is interest in empirical evidence from the

practitioners side, as also observed in (CARTAXO et al., 2016).

Many practitioners issues ask for recommendations of tools with specific features: SSs identifying tools, comparing their features, and aggregating evidence about their effectiveness could help to cover that gap. One example of such study is the one conducted by Marshall et al. (MARSHALL et al., 2014) that identified analyzed and compared tools to support SSs in SE based on their features. Another approach with direct implications for tool builders is to identify features demanded for SE tools based on issues posted in Stack Exchange websites. An example of this approach can be found in Pinto and Kamei's study (PINTO; KAMEI, 2013), which investigated practitioners issues in Stack Exchange websites to identify the most demanded features for refactoring tools.

Practitioners demand contextualized information: This can be observed by looking at recurrent issues like: Benefits of agile methods/practices from a specific perspective; Applicability of agile in specific project context; and Effort estimation in agile in specific project context. This supports many claims about importance of rich and contextualized evidence (CARTAXO et al., 2015a; DYBÅ; SJØBERG; CRUZES, 2012; ALI, 2016).

Practitioners demand target oriented information: For instance, we identified the following recurrent issue: Benefits of agile methods/practices from a specific perspective. This situation corroborates the idea that empirical evidence should comprise not only data about the effectiveness of an intervention but also useful information for the target audience (ALI, 2016). For instance, cost-effectiveness. We also identified other issues that report the need of information from other specific perspectives beyond monetary, such as, from the perspective of developers, managers, testers, customers, and others.

We identified 15 recurrent issues related to agile software development comprising 127 practitioners issues. This is almost one-third of the issues we have found in this study, which means agile is still an important topic in practice. Two recurrent issues reveal practitioners are facing problems that challenge some of the basic agile principles, which can be observed when practitioners report Low customer collaboration or when they acknowledge not desirable Impacts of low detail level or absence of documentation in agile. Another revealing recurrent issue is practitioners affirming to have experienced Ad-hoc software development as agile, corroborating some evidences and claims that in some situations agile is used as an excuse for absence of software process (HILKKA; TUURE; ROSSI, 2005; CAO; RAMESH, 2007). Additionally, we identified practitioners are facing problems Mixing agile with traditional methods/practices, and demanding evidence about the applicability of Pair programming to transfer knowledge, among many other recurrent issues.

2.4.1 Guidelines for Conducting SSs Considering Practitioners Issues

After applying the proposed coverage analysis, we observed that some SSs have room for improvement and if researchers adopt a few guidelines, those SSs might address a wide range of practitioners issues. We also would like to reinforce that not all SSs need to approach practical issues. We recognize there are plenty of SSs exploring abstract and methodological aspects that do not necessarily interest practitioners and are still important to the development of EBSE. For this kind of SSs, there are no advantages of pursuing our guidelines. Following are our recommendations:

Table 19 – Mitigating actions to avoid problems that prevent coverage of practitioners issues.

PROBLEM	PROBLEM CAUSE	MITIGATING ACTION
No practitioners issues returned by the search	Search string with too much key terms	Reduce the amount of key terms, since they are connected by the AND operator, that is restrictive
	Search string with no synonyms	Use synonyms for each key term to increase the possibility to find related issues
	Search string with key term with composed synonyms only	Mix composed synonyms with not composed ones, since the former is more unlikely to happen in the exact order on the issues
Excess of False positives	SSs using rather common terms in their search strings	Avoid general key terms on search strings unless they are connected by an AND operator with specific key terms.
Not Covered Issues	Issues reporting scenarios SSs do not fully cover	If possible, include the scenario on the SS's scope, since it can enriches the study
	Issues using approaches available/popular after the SS was conducted	This should not occur when one assess Stack Exchange questions during early stages of SS planning. But if it occurs, this suggests maybe the area of research is outdated

- **Test SS's search string on Stack Exchange at early stages of study planning:** We believe one who wants to conduct a SS with useful findings to practitioners can adopt a similar research strategy we presented here: assess on early stages – protocol definition – how its research relates to the practitioners issues on Stack Exchange. The insights found at Stack Exchange can be included in SS's scope of investigation during its planning phase.

- **Test SS's search string on Stack Exchange when planning to update an already existent SS:** Querying Stack Exchange may reveal candidate updates of the original research questions and opportunities to cover issues demanded in practice.

- **Mitigate problems that prevent SS to cover practitioners issues well:** During this research we identified some key problems that hinder practitioners issues coverage as well as their causes. They are either related to poorly defined

search strings or to a research scope that does not consider practitioners issues. In Table 19 we defined actions as suggestions to mitigate those problems during SS planning phase. The table should be read as following: To avoid <problem> caused by <problem cause>, I should <mitigation action>. For example, ***To avoid*** *excess of false positives* ***caused by*** *SSs using rather common terms in their search strings,* ***I should*** *avoid general key terms on the search string unless they are connected by an AND operator with specific key terms.*

- **Select high-quality primary studies:** The proposed coverage analysis is based on SSs. However, the quality of these SSs is subject to the quality of the primary studies selected. If the primary studies are poor quality, then it will be difficult to place great confidence in outcomes. Therefore, researchers should place additional care when deriving inclusion/exclusion criteria.

2.5 CHAPTER SUMMARY

We introduced a coverage technique to compare findings of SSs and issues practitioners report in five Stack Exchange websites. From 26,000+ issues reported on these websites, just 1.75% are related to the same SE areas of the 24 selected SSs. From that set of issues related the same SE areas as the selected SSs, only 15.6% are in fact covered by those SSs – when findings of a SS offer knowledge that can solve or at least attenuate the a practitioner issue. This low percentage corroborates with the claims related to the lack of connection of SSs with SE practice. The results also suggest that, although SSs might help practitioners, there is room for studies more connected to issues faced in practice, especially on agile software development.

3 THE KNOWLEDGE TRANSFER MODEL

The results of Empirical Study (ES1), together with previous studies (HASSLER et al., 2014; SANTOS; SILVA, 2013; SILVA et al., 2011), provide evidence on the existence of a lack of connection between Secondary Studies (SSs) and Software Engineering (SE) practice. Only 1.75% of the most relevant practitioners issues reported in the Stack Exchange websites were considered as related to the same topics of selected SSs. From these issues, just 14.1% were considered as covered by the SSs.

Evidence Based Medicine (EBM) has also faced a similar problem in its early days, and it is still facing it to some extent today (BEST et al., 1997; TRICCO et al., 2017; TRICCO et al., 2015). Unlike Evidence Based Software Engineering (EBSE), which is just recently concerned about that problem, EBM identified it longer ago, and proposed various alternatives to deal with it. One of the most successful initiatives to mitigate that problem in EBM is what are called as Rapid Reviews (RRs) (TRICCO et al., 2015; TRICCO et al., 2017) and Evidence Briefings (EBs) (CARTAXO et al., 2016; YOUNG et al., 2014; CHAMBERS; WILSON, 2012; KHANGURA et al., 2012). The former is intended to deliver evidence to practitioners in a timely manner, while the latter aims to report evidence in a more appealing way (more details in sections 3.5 and 3.4).

In this context, the contributions of this chapter is to propose a model to transfer knowledge from SSs to SE practice based on RRs and EBs, as well as some scenarios of adoption (Section 3.6) and challenges (Section 3.7) likely to be faced when adopting it. Fundamental background information on knowledge transfer is presented in Section 3.1. The operational definition of *knowledge* we use throughout this entire thesis is presented in Section 3.2. The model we are proposing was published in the Information and Software Technology (IST) journal (CARTAXO; PINTO; SOARES, 2018b).

3.1 BACKGROUND

There are many terms to designate the attempt to move knowledge into action. Graham et al. (GRAHAM et al., 2006) identified many of these terms, such as: Knowledge Transfer (BOZEMAN, 2000; DIEBOLD; VETRò, 2014; GORSCHEK et al., 2006), Knowledge Translation (BUDGEN; KITCHENHAM; BRERETON, 2013; BADAMPUDI; WOHLIN, 2016; STRAUS; TETROE; GRAHAM, 2009), Knowledge Diffusion (ROGERS, 2003), and others. According to Graham et al.:

> *"**Knowledge transfer** is about transferring good ideas, research results and skills between universities, other research organizations, business and the wider*

> *community to enable innovative new products and services to be developed."* (GRAHAM et al., 2006) (underline and bold emphasis added)

This thesis is more aligned with the concept of Knowledge Transfer since we investigate and propose strategies to produce and deliver *research results* (i.e. SSs' evidence) to practitioners aiming to support their decision-making in practice. Knowledge transfer can also be understood as a type of technology transfer where the object being transfered is knowledge (BOZEMAN, 2000).

Bozeman conducted a vast literature review and proposed a reference model for knowledge transfer which became influential in various scientific fields — been cited by more than 1,500 studies according to Google Scholar (BOZEMAN, 2000). This thesis proposes an instantiation of Bozeman's model applied to SE, as presented in Section 3.3. Bozeman's model is composed by five elements:

- **Transfer Agent**: Institution or organization seeking to transfer knowledge. Government agency, university, or a private firm;
- **Transfer Medium**: Vehicle, formal or informal, which the knowledge is transferred through. License, copyright, person-to-person, or a formal literature;
- **Transfer Object**: Content and form of what is transferred; the transfer entity. Scientific knowledge, technological device, process, or know-how;
- **Transfer Recipient**: Organization or institution receiving the transfer object. Firm, agency, organization, consumer, or an informal group;
- **Demand Environment**: The characteristics of the environment the Recipient Agent is immersed in.

Complementary to knowledge transfer models, one should also explore the way knowledge flows. Two main strategies have been discussed over the years, they are: **Knowledge-push**, and **Demand-pull** (GANDER, 1986; PIPER; NAGHSHPOUR, 1996; SPIVEY et al., 1997; SCHERER, 1982; BOZEMAN, 2000; PETERS et al., 2012; BROCATO, 2010). The former assumes a supply-side-driven and mainly linear process from research to development (BROCATO, 2010). As the name suggests, it pushes knowledge the transfer agents (researchers) believe is important to the transfer recipients (practitioners). In the opposite direction, the latter strategy focuses on identifying transfer recipients' demand for knowledge (BROCATO, 2010). The idea is to investigate and produce knowledge according to transfer recipients' demand. Some authors also affirm the dichotomy between knowledge-push and demand-pull is somewhat artificial since the combination of strategies is a possible and viable approach (PETERS et al., 2012; LAVIS et al., 2003; BROCATO, 2010).

3.2 OPERATIONAL DEFINITION OF KNOWLEDGE

Discussions on the meaning of the term "Knowledge" can scale to profound philosophical/epistemological disputes that are undoubtedly out of this thesis scope. As a consequence, we decided to provide an operational definition for it. The definition is, not surprisingly, rooted on the researchers' philosophical stance, as discussed in Section 1.3. Therefore, throughout this thesis, the term "Knowledge" is always used with the following meaning, except when stated otherwise:

> **Knowledge** is an evidence presented in a scientific report, produced through a systematic method, in compliance with the ethical principles of its scientific community, and potentially useful to support decision-making towards the solution, or at least attenuation, of a practical problem in a specific practical context.

To make the operational definition more understandable, following are two examples of what we consider as valid knowledge in the scope of this thesis:

> **Practical Problem:** Low customer collaboration.
>
> **Specific Practical Context:** *"30 Agile practitioners from 16 different software development organizations [...], half of whom were from New Zealand and half from India. [...] All the teams were using Agile methods, primarily combinations of Scrum and eXtreme Programming (XP). The level of Agile experience varied across the different teams."* (HODA et al., 2010).
>
> **Scientific Report:** Agile Undercover: When Customers Don't Collaborate (HODA et al., 2010).
>
> **Systematic Method:** Grounded Theory.
>
> **KNOWLEDGE:** Just Demos Strategy - *"Despite their reluctance or inability to attend other meetings, almost all customers were interested enough to attend demonstrations (demos) as it gave them an opportunity to see new functionalities of their software. Demos were often the only regular collaboration that these Agile teams received from the customer representatives and they used this opportunity to discuss features and receive feedback."* (HODA et al., 2010).
>
> **Potential Use:** A software engineering practitioner could apply the *Just Demos* strategy aiming to improve customer collaboration.

Practical Problem: Implement Software Process Improvement (SPI) in a software development company.

Specific Practical Context: *"A random sample of 154 software and quality managers in the Norwegian IT industry with corporate membership in the Association of the Norwegian Software Industry or the Norwegian IT Technology Forum were contacted by telephone to request participation in the study prior to mailing the questionnaires. A total of 120 software and quality managers representing whole organizations or independent business units within 55 companies completed and returned the questionnaire."* (DYBA, 2005)

Scientific Report: An Empirical Investigation of the Key Factors for Success in Software Process Improvement (DYBA, 2005).

Systematic Method: Quantitative Survey.

KNOWLEDGE: Involved Leadership Strategy - *" Out of 15 correlations between the independent variables, two have a correlation coefficient larger than 0.5. The highest correlation (0.63) is between involved leadership [and SPI Success] [...] Involved leadership — the extent to which leaders at all levels in the organization are genuinely committed to and actively participate in SPI— had a strong and highly significant correlation with overall SPI success, which supports the hypothesis that SPI success is positively associated with involved leadership."* (DYBA, 2005)

Potential Use: A software development company could stimulate *Involved Leadership* aiming to achieve success implementing SPI, hoping that the correlation observed on the study reveals itself a causality relationship.

The first example reports evidence produced through grounded theory, a widely used qualitative method, while the second reports evidence produced through a quantitative survey. The examples would instigate fierce discussions if analyzed under antagonistic epistemological views (e.g. positivism vs. constructivism) aiming to define whether the evidence could be considered as valid knowledge. The quarrel might occurs because each side would probably expose many limitations inherent to the antagonistic study's method of choice. If maintained the good practices of fair debates, the exposed limitations would probably be pertinent. However, both sides would not be willing to cede their view, because for them, the limitations are not acceptable and hurt the premises governing their philosophical stances. Such kind of impasse would not occur when evidence is analyzed under the enunciated operational definition, which is settled on the premise that knowledge is always approximate and incomplete, but it is still valid when produced through a

systematic method and is considered as useful in practice.

3.3 THE KNOWLEDGE TRANSFER MODEL

The knowledge transfer model we are proposing is an instantiation of the Bozeman's model (BOZEMAN, 2000). Figure 7 illustrates each element of the model we re proposing. In our instantiation, software development companies (**Demand Environment**) are the sources used to identify the issues practitioners face. Once an issue is detected, it motivates researchers (**Transfer Agents**) to conduct a Rapid Review along with practitioners to identify evidence (**Transfer Object**) that could help practitioners (**Transfer Recipient**) to address the issue. The evidence is presented to practitioners through Evidence Briefings (**Transfer Medium**).

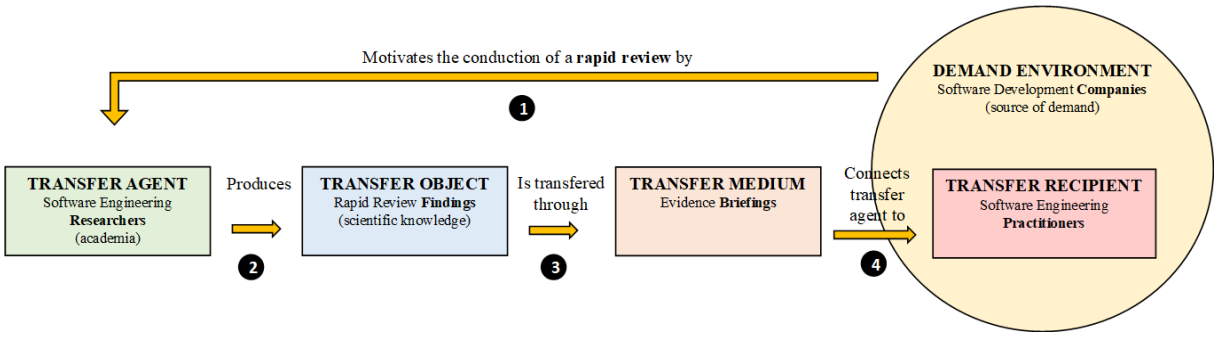


Figure 7 – Proposed knowledge transfer model.

Since Demand Environment, Transfer Agent, and Transfer Recipient are straightforward to understand, we focus on two elements that deserve more attention: the Transfer Object and the Transfer Medium. We operationalize the latter in terms of Evidence Briefings (CARTAXO et al., 2016) and the former in terms of Rapid Reviews (CARTAXO; PINTO; SOARES, 2018a).

Regarding the way knowledge flows, our model is mainly aligned with the *demand-pull* strategy because the knowledge is produced, through a RR, based on what the knowledge recipients (practitioners) demand. Nevertheless, the model can also be considered mixed strategy since the knowledge produced is always based on primary studies already conducted by other researchers, and under this perspective the knowledge is being pushed.

3.4 THE TRANSFER MEDIUM: EVIDENCE BRIEFINGS

Researchers in medicine argue that Full Secondary Studies (FSSs) often neglect practitioners needs (TRICCO et al., 2015) hindering the consumption of that kind of content. This led to the creation of alternative mediums to transfer knowledge that better fit prac-

titioners needs (CHAMBERS; WILSON, 2012; KHANGURA et al., 2012; YOUNG et al., 2014). Following the initial findings from medicine, we recently introduced the concept of Evidence Briefings (CARTAXO et al., 2016) in SE. Evidence Briefing is an one-page document that presents only the main findings of a research initiative assuming the role of transfer medium in our knowledge transfer model. We evaluated the EBs with the main stakeholders, researchers and practitioners. More details about EBs and the results of evaluating them can be found in Chapter 4.

3.5 THE TRANSFER OBJECT: RAPID REVIEWS' EVIDENCE

Rapid Reviews are lightweight SSs focused on delivering evidence in a timely manner to practitioners. To conduct RRs, researchers deliberately omit or simplify steps of FSSs. For instance, they can limit literature search, use just one person to screen studies, skip quality appraisal of primary studies, etc (TRICCO et al., 2015).

It is important to highlight that RRs are not conceived to replace FSSs. Instead, both can coexist. While FSSs are important to provide in-depth, highly-reliable evidence, RRs are important to quickly transfer knowledge to practice. As expected, RRs have obvious limitations. To cite a few: paper coverage is limited to few search sources; there might be selection bias if only one person screens papers, etc.

In spite of the limitations inherent to the less rigorous method, the interest in RRs is growing in health-care fields (TRICCO et al., 2015; MOHER et al., 2015; GARRITTY et al., 2016; TRICCO et al., 2017). Following the promising results in medicine, we conducted RRs in two software development companies, as a way to evaluate the model we proposed. The results we obtained were mainly positive. More details about RRs and the results of applying them in practical contexts can be seen in Chapter 5.

3.6 OPPORTUNITIES FOR APPLYING THE MODEL

The proposed knowledge transfer model can be applied in many circumstances. Following there are some scenarios we envision.

3.6.1 Scenarios Considering the Level of Participants Involvement

The level of involvement of the transfer agents (researchers) and the transfer recipients (practitioners) can vary depending on the practical constraints imposed by the demand environment. For instance, the selection of primary studies for a RR can be executed by researchers and practitioners, or by researchers only, if effort is a constraint to the software company. Similar arrangements are also possible to any other RR task, like protocol definition, data extraction, data synthesis, etc. What must be preserved is the

active participation of both sides, specially the practitioners, regardless they participate just providing their needs and validating decisions during the RR, or executing RRs tasks together with researchers.

3.6.2 Scenarios Considering Who Commissions the Initiative

- **Commissioned by practitioners:** A RR can be commissioned by practitioners of a software development project aiming to support decision-making towards the solution of an issue. For instance, a RR synthesizing the challenges and strategies to establish agile practices in a distributed team, for a project team pondering whether to go distributed is a good idea.
- **Commissioned by researchers:** A RR can be commissioned by researchers who want to study a specific SE topic, but with strong relation with practitioners needs. For example, researchers wanting to investigate the problems related to acceptance testing can establish partnerships with companies facing that kind of problem and conduct the RR together.
- **Commissioned by policy makers:** Regulatory agencies or companies who want to define policies can benefit from RRs. To illustrate, a company aiming to define its software improvement process to apply it institutionally can contact researchers to collaborate conducting RRs.
- **Educational purposes:** Evidence Briefings reporting RRs can be used by educators who want to approach a specific SE topic with their students. For instance, a professor wanting to show her/his students a comparison between automated software testing tools and their pros and cons can present the findings of RRs previously conducted by other researchers, or even conduct a collective RR with the students in partnership with a software company.

Regardless of who commissions an initiative to conduct RRs and produce EBs, the presence of a researcher on the entire process is crucial to guarantee that the basic scientific principles are respected. Non-specialist must not coordinate such kind of initiative, neither supervising RRs, nor creating EBs.

3.7 CHALLENGES TO APPLY THE MODEL

Despite the clear benefits of transferring empirical knowledge to practice, there are challenges that anyone who wants to adopt the model we proposed will probably face. Based on our reflections after conducting RRs and producing EBs, as well as on research literature on knowledge transfer (PRECHELT; ZIERIS; SCHMEISKY, 2015), we listed the following challenges and mitigations, that are not exhaustive.

- **Effort:** Some companies may oppose the conduction of RRs arguing the impossibility to allocate effort of their employees on that kind of initiative. A mitigation action could make researchers responsible for the execution of the RR. Researchers would involve practitioners just to present their decisions about the RR protocol an ask questions for clarification. For instance, researchers can propose RR's research questions, search strings, and exclusion/inclusion criteria and fine tune them with practitioners. If RRs would have a good reception in SE practice in future, we believe practitioners will probably recognize the benefits and would be more willing to put effort on such kind of initiative. Thus, researchers would have help to execute RRs.

- **Confidentiality:** Sometimes both practitioners and researchers agree with mutual benefits of collaboration but companies pose restrictions both on information access – preventing practitioners from providing information to researchers – and on information disclosure – preventing researchers from publishing research results to the scientific community. It can occurs, for example, when what is been investigated is one of the market advantages of the company that must not be exposed. Or when it reveals errors committed, which may damage the image of the company. A mitigation action could explore non-sensitive data, and anonymize any reference to the company and its employees when publishing the results.

- **Schedule Constraints:** Some companies may oppose the conduction of RRs to support decision-making fearing delays on the projects schedules. Assuming a situation where a researcher proposes to conduct a RR to support decision-making aiming to solve a practical issue, and that issue is in the critical path of the next release. Project managers may not approve such kind of intervention arguing this might provoke delays on the release if the development team waits for the RRs results to make their decisions. A mitigation action could be conducting RRs on issues not on the critical path, at least while practitioners are not fully convinced about the benefits of RRs. Another option is conducting RRs on issues that companies want to solve, but are not priority, which usually stay forgotten forever. For example, when a development team recognizes the benefits of continuous integration, would like to introduce it in the project, but none are assigned to implement it, because the project has been working without it since the beginning, and functional requirements would be delayed if one is allocated to do it.

- **Researchers resistance:** It is true that there are many researchers eager to be closer to SE practice. However, some are resistant to this idea arguing that this kind of partnership may hurt scientific principles or produce biased results (KELLY; MOHER; CLIFFORD, 2016b).

- **Practitioners resistance:** Not only researchers pose resistance to establish a close relationship with SE practitioners. The opposite also occurs, since some practitioners prefer expert opinion over empirical evidence, or even believe that academia does not produce relevant work to practice (RAINER et al., 2003).

In addition to the mitigation actions already mentioned, one of the most important strategies to overcome those challenges is showing the benefits of RRs. This can be done presenting to practitioners the good results that health related fields have obtained recently, as well as the ones we present in this thesis. It is important to highlight that RRs provide scientific evidence from studies that collected experiences and data from other companies. This shows, not only that researchers benefit by the conduction of RRs, but also the company itself, learning from other companies' experience.

Conducting RRs with companies that already have a relationship with the research community is probably a good first step, since they might be more prone to collaborate with researchers. As soon as the first RRs are conducted on such kind of companies, the experience can be presented to companies with weaker or no ties with the research community, as a way to show the good results and feasibility of RRs.

3.8 RELATED WORK

Knowledge Transfer is not a recent research topic. However, it experiences continuous discussions and investigations in a wide variety of research fields due to its inherently evolving nature. With SE it is not different since many studies has been conducted exploring knowledge transfer under different perspectives. In this section, we present some work related to knowledge transfer, and how it is different from the research we present in this thesis. Finally, it is important to highlight that no single model could claim to solve all knowledge transfer problems. Thus, our model is not intended to be a "one size fits all" solution, and the other models we discuss are not better or worse than ours, they are just applicable in different contexts, or sometimes are even complementary.

Many models have been proposed to transfer knowledge in SE. They usually are based on the idea of applied research, which allocates research effort aiming to directly solve a practical problem. Pfleeger's work is an example of such kind of model (PFLEEGER, 1999). It identifies practitioners issues and proposes a direct solution to them, while the model we propose in this thesis intends to identify practitioners issues and synthesize empirical evidence to support their decision-making towards the issue solution. The former model demands more effort and commitment from researchers and practitioners, which might hinder cooperation between research and practice. The model we proposed, on the other side, demands less effort and commitment, as well as stimulate practitioners to consume empirical evidence, which can foster a culture of informed decision-making in SE practice. Later, Gorschek's et al. also presented a model similar to Pfleeger's one (GORSCHEK et al.,

2006). The most relevant difference is that, Gorschek's et al. model is more collaborative. It acknowledges demand-pull as an important characteristic of knowledge transfer. An even more recent model was proposed by Mikkonen et al. (MIKKONEN et al., 2018). The crucial difference from the other is that, Mikkonen's et al. model is based on the idea of continuous technology transfer through large consortia between academia and practice, and stimulates even more participation of practitioners. RRs could easily fit in this kind of consortia. Practitioners and researches would continuously conduct RRs as issues arise in the consortia projects.

Many other studies explore the knowledge transfer territory, but instead of proposing models, they report lessons learned during academia-industry collaborations. For instance, Duarte's work, that establishes a set of academia-industry collaboration patterns based on his previous learned lessons (DUARTE, 2015). Another example is a work of Gorschek, which presents a list of lessons learned during more than ten years of collaboration with industry (GORSCHEK, 2015).

Another category of studies on knowledge transfer is the ones focused on transferring knowledge through education, usually with no formal model, for instance, the study of Laird and Yang, which presents results obtained applying the concept of Open University, where the line separating academia and industry is somewhat blurred (LAIRD; YANG, 2015). Another example, is the Skevoulis' work, which reports experiences creating and conducting a software engineering graduate program customized to a company needs (SKEVOULIS, 2011).

3.9 CHAPTER SUMMARY

In this chapter we propose a model to transfer knowledge from SSs to SE practice based on RRs e EBs. This has became crucial since evidence are pilling on SSs having low connection with SE practice, which is the main goal of EBSE. We also discuss scenarios where the model is probably applicable, the main challenges that will be probably faced when applying the model, as well as some mitigation actions. RRs are a form of SS that deliberately omit or simplify some FSSs steps in order to deliver evidence in a more timely manner, and more importantly, connected to practitioners needs. EBs, on the other side, are one-page documents that summarize the main findings of a research aiming to make evidence more appealing to practitioners. Based on the results we obtained conducting empirical studies to evaluate EBs and RRs, we believe the model we proposed has a promising future ahead.

4 THE TRANSFER MEDIUM: EVIDENCE BRIEFINGS

In this chapter, we *propose*, *evaluate*, and *discuss* the Evidence Briefings (EBs), which are the Transfer Medium composing the knowledge transfer model we presented in Chapter 3. The evaluation was conducted guided by the following research question:

> **GRQ2:** What are the perceptions of researchers and practitioners on Evidence Briefings as transfer mediums?

To achieve theses goals, we **propose** a medium to transfer knowledge to practice. We call it Evidence Briefing. It is a one-page document reporting the main findings of an empirical research. EBs are intended to support practitioners in making informed decisions. The format was defined based on the best practices observed in alternative mediums proposed in Evidence Based Medicine (EBM) as well as on information design principles, to make EBs more appealing to practitioners. For more details about EBs format see Section 4.1.

To **evaluate** the use of EBs as mediums to transfer knowledge, and answer GRQ2, we analyzed the perception of the main stakeholders, practitioners and researchers. This is important because, according to Rogers, the perception of individuals about an initiative is one of the main predictors of its adoption (ROGERS, 2003). In addition, an empirical study conducted by Grigoleit et al. identified that low perception of practitioners about mediums to transfer knowledge is the main barrier to adopt those mediums (GRIGOLEIT et al., 2015). Since the main stakeholders involved with EBs are researchers and practitioners, we evaluated their perceptions through three empirical studies.

In the *first empirical study* – which in fact is the second empirical study of this thesis (ES2) – we created 12 EBs based on empirical research published by diverse authors. We sent these EBs to 32 practitioners facing issues related to the same topics of the EBs we created. To identify these practitioners, we mined Stack Exchange websites. In particular, we searched for questions reporting problems related to the EBs we created. Finally, we analyzed their perceptions on the content and format of the EBs.

In the *second empirical study* – which in fact is the third empirical study of this thesis (ES3) – we reached out seven authors (i.e. researchers) of the studies we created EBs based on. We investigated their perceptions on the EBs' content and format. This was important to asses whether the EBs we created were befitting their original studies. Additionally, since the first empirical study evaluated the practitioners position, this second one evaluated the researchers side.

These two studies revealed that both practitioners and researchers are generally positive on EBs. We pondered on the lessons learned, and decided to conduct a third and more ambitious research to strengthen the evidence we already had with the two first

studies. The idea was to introduce the EBs in SE conferences, so we could have a full scale scenario where researchers create their own EBs.

In the *third empirical study* – which in fact is the fourth empirical study of this thesis (ES4) – we presented EBs to the steering committee of four SE research conferences. They agreed an encouraged the initiative giving us permission to invite the authors (i.e. researchers) of the accepted papers to create EBs of their research. Authors of 44 papers created EBs. They also reported their perceptions about the creation process. This study is particularly important because researchers are the most active part of such initiative. They are the ones that put more effort creating EBs, while practitioners just consume them. Thus, the success of the EBs is strongly related not only if practitioners like them, but also whether researchers have positive perceptions on creating EBs themselves, and are willing to do so.

The results of ES2 and ES3 were published in the 10th International Conference on Empirical Software Engineering and Measurement (ESEM 2016) (CARTAXO et al., 2016). The extended version, with the results of ES4, was submitted and is under review in the Empirical Software Engineering (EMSE) journal.

4.1 EVIDENCE BRIEFINGS STRUCTURE

In this section, we provide details about how we defined the EBs format. To strengthen the method of this research and offer professional high quality EBs, we invited a design researcher to define, together with us, the EBs template.

The primary objective is to develop documents that are comprehensible, accurately retrievable, natural, and pleasant. Therefore, we took advantage of well-established principles of Information Design (TONDREAU, 2011) and Gestalt Theory (LUPTON; PHILLIPS, 2015). Another important phase on designing EBs' template is to verify the Best Practices (AMBROSE; HARRIS, 2009) that could be inherited from others projects. The graphic design is inspired by a mix of characteristics from various mediums proposed in medicine. For instance:

- We limited our Evidence Briefings to one page, similar to Young et al. with their contextual summaries (YOUNG et al., 2014);
- We summarized the main findings in one section, likewise the briefings proposed by Chambers and Wilson (CHAMBERS; WILSON, 2012);
- We also used an informative box separated from the main text to highlight the audience and nature of the briefings' content, like the evidence summaries of Khangura and et al. (KHANGURA et al., 2012).

Figures 8– 10 show the characteristics of the EB template we are proposing. Figure 8 shows numbers within squares denoting the structure of the EB. They are: (1) The title of the briefing; (2) a short paragraph to present the goal of the briefing; (3) the main section that present the findings of an empirical study; (4) informative box that outlines the intended audience and explains the nature of the briefings' content; (5) the reference to the original paper; and (6) the logos of the research group, university, or company responsible for creating the EB.

Figure 9 shows numbers within circles denoting where Tondreau's principles of Information Design were applied (TONDREAU, 2011). They are: (1) Hierarchy of Information principle which states that important information should be large, bold and used with bullets to be distinguished; (2) Space Between Elements principle states that space communicates volume, sets off the message, and give appropriate room for reading; (3) Typography principle states that fonts should be friendly and widely recognized, that is why we used Calibri, a friendly reading sans-serif type that is widely known (it is recognized as the main Windows font); (4) the Color principle states that color is a way to make modules stand out, as occurs with the colorful box that help to organize elements; and (5) Rhythm and Flow principle that are followed when we present the information in one page document showing a sense of security and variation in size and positions of images and typography.

Figure 10 shows numbers within diamonds denoting where Gestalt Principles were applied (LUPTON; PHILLIPS, 2015). They are: (1) Similarity principle which states that elements that are similar are more likely to be organized together; (2) Proximity principle which states that closer elements are more likely to be perceived as a group; (3) Continuation principle which states that elements will be grouped as a whole if they are co-linear; and (4) Unity principle which states that elements that have a visual connection should belong to a uniform group.

The EBs' template can be found at our research group website.[1] We encourage researchers who want to transfer knowledge from empirical studies to practitioners to use the template we developed. The template is open-source under CC-BY license.

[1] `http://cin.ufpe.br/eseg/briefings`

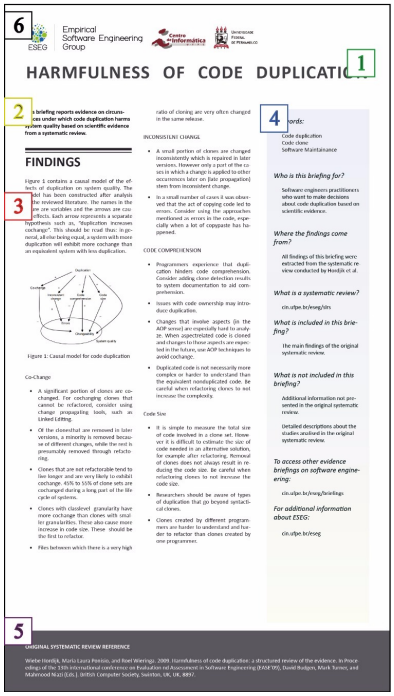


Figure 8 – Overview of each part that compose the briefings structure.

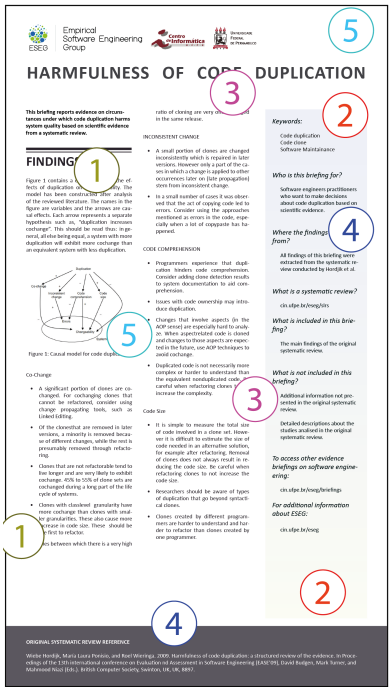


Figure 9 – Overview of the Information Design elements (TONDREAU, 2011)

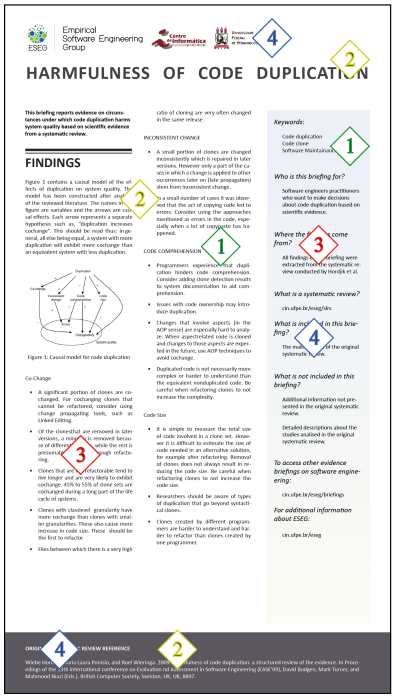


Figure 10 – Overview of the Gestalt Theory elements (LUPTON; PHILLIPS, 2015)

4.2 BACKGROUND AND RELATED WORK

According to Bozeman, the Transfer Medium is the vehicle through which knowledge is transferred from the transfer agents (researchers) to the transfer recipients (practitioners) (BOZEMAN, 2000). Transfer mediums have been proposed in a variety of scientific areas, specially aiming to make knowledge more appealing to practitioners. Medicine and other health-care fields presented many studies around this topic. For instance, the Contextual Summaries by Young et al.(YOUNG et al., 2014); the Briefings proposed by Chambers and Wilson (CHAMBERS; WILSON, 2012); or even the Evidence Summaries by Khangura et al. (KHANGURA et al., 2012). This research is strongly influenced by transfer mediums proposed in health-care related fields.

In the context of SE, however, there are few investigations targeting transfer mediums. An example is the study of Jedlitschka et al. proposing a model to report experiments according to practitioners' needs (JEDLITSCHKA; JURISTO; ROMBACH, 2014). Nevertheless, the model defines what kind of information should be reported, not the format of the medium. Additionally, the model is only applicable to experiments, not covering other kinds of empirical studies (e.g. case studies, surveys, systematic reviews, and others). In a succinct discussion on transfer mediums, Beecham et al. suggested that researchers should

write shorter evidence-based reports (BEECHAM et al., 2014). The one-page EBs we are proposing here are in accordance with this recommendation. Furthermore, Grigoleit et al. grouped the types of mediums in two categories: (1) Artifacts, including publications and documents, or (2) Human Intensive Mediums, including conferences and workshops (GRIGOLEIT et al., 2015). The EBs we are proposing can be classified as an artifact medium.

Some initiatives are related to the EBs, like the Structured Abstracts (BUDGEN; BURN; KITCHENHAM, 2011; BUDGEN et al., 2008; BUDGEN et al., 2007) and Visual Abstracts (STOREY et al., 2017). Structured abstracts emerged in SE, as a way to make scientific papers' abstracts more complete and clear, in particular the ones reporting experiments(BUDGEN et al., 2007), all of this targeting researchers who want to improve SSs process (BUDGEN et al., 2008). EBs, on the other side, target practitioners as primary audience, and due to be the larger space (one page, against few words of structured abstract) are ideal to more in-depth discussions about results of, not only experiments, but also more qualitative studies. Regarding visual abstracts, they address the same challenges of the structured abstracts, but in a visual form (STOREY et al., 2017). In fact, the proponents of visual abstracts suggested that they are suitable to be included in the EBs as they say *"we suggest that the visual abstract is more suited to posters or presentations, or as an addition to blog posts [...] and* ***evidence briefings****" (bold emphasis added)* – citing our first paper about Evidence Briefings (CARTAXO et al., 2016). Thus, unlike EBs, both structured and visual abstracts are not complete self-contained transfer mediums, and as so, they are more complementary to EBs than replacements/competitors. Due to those inherent differences, we decided to not compare EBs with structured or visual abstracts, avoiding comparing oranges with apples.

4.3 RESEARCH METHOD

Here we state the research questions (Section 4.3.1); the stages of this research (Section 4.3.2); and the method we followed in each empirical study (Sections 4.3.3, 4.3.4 and 4.3.5).

4.3.1 Research Questions

The GR2, is split in three more specific questions, they are:

RQ1: What are the practitioners perceptions on Evidence Briefings format and content?

RQ2: What are the researchers perceptions on Evidence Briefings format and content?

RQ3: What are the researchers' perceptions on the process of creating Evidence Briefings?

RQ1 and **RQ2** intend to investigate the perceptions of practitioners and researchers about EBs as a medium to transfer knowledge to SE practice. Answering these questions is crucial to understand whether the EBs have a chance to be accepted as a knowledge transfer medium, since according to Rogers (ROGERS, 2003), the perception of individuals about an initiative is one of the main predictors of its adoption.

With **RQ3** – differently from **RQ1**, which is just concerned with researchers perceptions on EBs format and content – we intend to discover the perceptions of researchers creating the EBs themselves, and with this, uncover the challenges associated with this process, as well as strategies to mitigate them.

4.3.2 Research Stages

In **Stage I**, we created the Evidence Briefings template, as well as conducted empirical studies to answer **RQ1** and **RQ2**, one per question, ES2 and ES3, respectively. Stage I was conduced in the following four steps:

1. **Secondary Studies Selection:** This step was in fact conducted during the practitioners issues coverage study (ES1). Thus, our initial set of SSs is composed by the 24 SSs, which were considered as related to at least one practitioner issue reported on Stack Exchange websites. We randomly selected half of the 24 SSs to create EBs. Table 20 presents details about each of the 12 selected SS and their respective SWEBOK areas.

Table 20 – Selected papers to create Evidence Briefings for STAGE I (ES2 and ES3).

TITLE	SWEBOK AREA
Automated Acceptance Testing: a Literature Review and an Industrial Case Study	Software Testing
Critical Barriers for Offshore Software Development Outsourcing Vendors: A Systematic Literature Review	Software Engineering Economics
Critical Success Factors for Offshore Software Development Outsourcing Vendors: A Systematic Literature Review	Software Engineering Economics
Definitions and approaches to model quality in model-based software development – A review of literature	Software Quality
Empirical studies of agile software development: A systematic review	Software Engineering Models and Methods
Factors Influencing Software Development Productivity – State-of-the-Art and Industrial Experiences	Software Engineering Economics
Forecasting of software development work effort: Evidence on expert judgement and formal models	Software Engineering Management
Harmfulness of Code Duplication - A Structured Review of the Evidence	Software Construction
Knowledge management in software engineering: A systematic review of studied concepts, findings and research methods used	Software Engineering Professional Practice
On the generation of requirements specifications from software engineering models: A systematic literature review	Software Requirements
The effectiveness of pair programming: A meta-analysis	Software Engineering Models and Methods
Using Scrum in Global Software Development: A Systematic Literature Review	Software Engineering Models and Methods

2. **Secondary Studies Data Extraction:** For each of the selected SSs, we extracted the following data: paper title, research goal(s), research findings, and paper references. We simplified the **paper titles** in order to make the EBs more appealing to practitioners, putting effort to not compromise the meaning. In many cases, we just removed terms that explicitly mention the research method. For instance, we created an EB entitled "*The effectiveness of pair programming*" from a paper entitled "*The effectiveness of pair programming: A meta-analysis*" (HANNAY et al., 2009). The **research goals** were re-written without changing semantics, fitting the following template: "*This briefing reports evidence on <GOAL> based on scientific evidence from a systematic review.*" For example, the goal of the EB created from paper (KHAN; NIAZI; AHMAD, 2009b) is: "*This briefing reports evidence on critical success factors that have a positive impact on software outsourcing clients in the selection process of offshore software development outsourcing vendors based on scientific evidence from a systematic review.*" We kept, as much as possible, the **research findings** as the authors of the SSs have written. Modifications were made only to assure the EBs fluency of reading. The **papers references** were extracted as usual, from their publishers' websites.

3. **Evidence Briefings Creation:** To create EBs, we used the extracted data to fill the EBs template (see section 4.1).

4. **Evidence Briefings Evaluation:** We evaluated the EBs according to their *content* and *format* surveying two distinct groups: (1) users of Stack Exchange websites that asked questions related to the selected SSs, chiefly practitioners; and (2) the authors of those reviews, chiefly researchers. Section 4.3.3 presents methodological details about the first survey, whereas Section 4.3.4 discusses the second one.

In **Stage II**, we conducted an empirical study to answer **RQ3**. The focus of the second stage was on the perceptions of researchers on creating EBs for their own papers. Additionally, after enriching discussions during ESEM 2016, at Stage I, we were encouraged by many peers to use EBs to also report primary studies' results, not only SSs, as in Stage I. Thus, we introduced the idea to the steering committee of four SE conferences:

- **EASE:** 21st Evaluation and Assessment in Software Engineering Conference;[2]
- **SBES:** 31st Brazilian Symposium on Software Engineering;[3]
- **SBCARS:** 11th Brazilian Symposium on Components, Architectures and Software Reuse;[4]

[2] http://ease2017.bth.se/
[3] http://www.lia.ufc.br/ cbsoft2017/programacao-sbes/
[4] http://www.lia.ufc.br/ cbsoft2017/en/programacao-sbcars/

- **SAST:** 2nd Brazilian Symposium on Systematic and Automated Software Testing.[5]

We selected these conferences because: (1) they happen on yearly basis, so there is a constant flow of new studies being conducted regularly; (2) they are well-established in their domain (although SAST is the newest one, it occurred for nine times as a workshop before becoming a symposium; the SBCARS symposium occurred six times as a workshop before becoming a symposium); (3) they have a selective review process (in 2017, EASE, SBES, SBCARS, and SAST had, respectively, 32%, 34%,31%, and 46.6% of acceptance rate); and (4) they are well-aligned with the purposes of EBs: research papers published on these conferences often have evidence that practitioners are interested in. We ran this initiative on the 2017 edition of each conference. With the permission of the steering committee, we invited the authors of the accepted papers to create EBs of their research. Our goals were to (1) invite researchers create their own EBs and (2) evaluate their perception on creating them — differently from Stage I, which we created the EBs based on already published papers, and asked the authors about their perceptions.

4.3.3 Empirical Study 2 (ES2) Method: Practitioners Perceptions on Evidence Briefings

The **goal** of this cross-sectional survey is to acquire evidence on how practitioners perceive the content and format of EBs.

As **instrument**, we created a questionnaire mixing open and closed questions, divided in four sections. The first with demographic questions that helps to understand the characteristics of our sample; the second with questions to discover what mediums practitioners use to acquire knowledge; the third with questions to discover the perceptions on EBs content; and the fourth to examine the perceptions about the EBs format. The questionnaire is available in Appendix A.

The target **population** of this survey is the 473 Stack Exchange users who posted questions related to at least one of the 12 SSs. We found these questions by applying SSs' search strings on five Stack Exchange websites related to SE – details about the Stack Exchange websites are available in Table 2. The initial search on the Stack Exchange returned 1,738 questions. The title and body of each question were analyzed in pairs to exclude false-positives. This process resulted on the 473 questions that compose the population of this survey. Questions were considered as false-positives when they are out of the SSs' topics – according to the SWEBOK topics, as shown on Table 20. Also, we had meetings for conflict resolution in order to avoid classification bias. The Kappa value was 0.72, which means *Substantial Agreement* according to Kappa's reference table (VIERA; GARRETT, 2005). The classification of questions, after the conflict resolution meetings, is

[5] http://www.lia.ufc.br/ cbsoft2017/programacao-sast/

available online.[6] Our intention was to invite all users that created the 473 questions to participate in the survey. However, we contacted only users that provided contact information on their Stack Exchange public profiles, following the privacy policies established by Stack Exchange.[7] Therefore, we could contact 146 (30.8%) users out of the 473 initial ones. To improve response rate, we applied majority of the principles listed by Smith et al. (SMITH et al., 2013). For instance, the Reciprocity principle, when we offered a raffle of 100 USD gift on Amazon for the respondents. The Social Benefit principle, when we highlighted the importance of understanding and reducing the gap between research and practice, and also when we advertised a donation of 1 USD for the Brazilian Red Cross for each participant. The Brevity principle, applied by asking closed and direct questions as much as possible.

Our **sample** is composed by 32 Stack Exchange users who responded the questionnaire. This corresponds to 21.9% of the 146 invitations. It is considerably above the 5% that Singer et al. affirmed to consistently observe as response rate to SE surveys (SINGER; SIM; LETHBRIDGE, 2008). Among the Stack Exchange users that answered our survey, 56% are software developers, 21% are software architects, 9% are project managers, 3% are software testers, and 9% hold other positions. In terms of experience, 18% are on their current position from 8 to 12 years, 21% from 5 to 8 years, and 40% from 2 to 5 years. Most of them work for the software industry (87.5%), 9% work for open source initiatives, 9% for the government, and 6% are academics. They are highly educated — 10% have a Ph.D. degree, 48% have a Master degree, and 35% have a Bachelor degree.

Regarding the **limitations** of ES2, not all Stack Exchange users were contacted. It occurred because Stack Exchange does not provide a way do directly contact their users. Therefore, we sent the survey only to the 146 users that we could manually find contact information in their Stack Exchange public profile. Moreover, we did not compare EBs to traditional research paper format. This would reveal if practitioners prefer to consume EBs over research papers, or not. We decided to avoid this comparison because it would demand practitioners to read research papers with more than ten pages, sometimes 20. We believe this would reduce drastically the response rate.

4.3.4 Empirical Study 3 (ES3) Method: Researchers Perceptions on Evidence Briefings

The **goal** of this cross-sectional survey is to acquire evidence on how the authors of the selected SSs perceive the content and format of the EBs created based on their SSs.

As **instrument**, we created a questionnaire mixing open and closed questions that we divided in three sections. The first with questions to understand to what extent the authors of the selected SSs are interested in sharing research results with practitioners;

[6] Stack Exchange Questions Classification is available in `http://bit.ly/1M3cZvY`
[7] Stack Exchange Privacy Policy `http://stackexchange.com/legal/privacy-policy`

the second with questions to discover the perceptions about the EBs contents; and the third focused on the perceptions about the EBs format. The questionnaire is available on Appendix B.

The target **population** of this survey is the 22 authors of the 12 selected secondary studies.

Our **sample** is composed by 7 authors that responded the questionnaire, which corresponds to 31% of the 22 invitations. Again it is considerably higher than the 5% that Singer et al. consistently observed as response rate to SE surveys (SINGER; SIM; LETHBRIDGE, 2008)

Regarding the **limitations** of ES3, not all authors were contacted. It occurred because not all authors still maintain the email address used in the papers. We tried to find their personal websites or public profiles and also asked others co-authors whether any of their colleagues are using another email address. Additionally, we are not the authors of the research we created EBs based on. Therefore, chances are we have misunderstood or not included some findings of the original papers. On the other side, on ES3 we asked the authors of the original papers "How does the briefing that we sent to you cover the main findings of your paper?". About 72% of the respondents said that is was "Good" or "Very Good". The remaining 28% said that it is "Acceptable". Finally, we did not evaluate the "bad design" case, which would be mediums with poorly designed interface. Our EBs were created by a design researcher specialist, using established design techniques. Evaluating a poor graphical design comparing with the one we proposed would expose the significance of the presentation of EBs vs the Content of EBs.

4.3.5 Empirical Study 4 (ES4) Method: Researchers Perceptions on Creating Evidence Briefings

The **goal** of this cross-sectional survey is to acquire evidence on the perception of researchers about the process of creating an EB.

As **instrument** for this survey we created a questionnaire mixing open and closed questions that we divided in three sections. The first with questions to understand to what extent researchers are interested in sharing results with practitioners. The second with questions to capture the perceptions of researchers creating EBs for their papers. The third, with questions to focused on the perceptions of researchers about EBs initiative as a whole. The complete list of questions of this questionnaire is in Appendix C.

The target **population** of this survey is the authors of the 115 papers published on the four conferences (EASE, SBES, SBCARS, and SAST) that supported the EBs initiative. Similar to our first two surveys, we applied some of the principles introduced by Smith et al. (SMITH et al., 2013). For instance, the Reciprocity principle (e.g., the authors who responded the questionnaire participated in a raffle of 75,00 USD), or the

Brevity principle, applied by asking closed and direct questions as much as possible.

Our **sample** is composed by the authors of 44 papers that created EBs for their papers. This corresponds to 38% of the 115 invitations. Just one author per paper responded the questionnaire. Half (22) of the authors that created EBs are PhD students, while 13 (29%) are professors. The remaining ones (9) reported to be undergrad students, MSc students, working in industry, or unemployed. Table 21 presents the number of papers published and EBs created per conference. PP stands for Published Papers, and EBC stands for Evidence Briefings Created. It is interesting to note the researchers adherence with the initiative, in particular because the initiative demanded a couple of hours and dedication of the authors to create the EBs, as we shall see in Section 4.4.3. On average, the authors spent 2.5 hours to create an EB. References to the EBs created, as well as to the 44 original papers can be seen in Appendix D.

Table 21 – Evidence Briefings created per conference.

CONFERENCE	**# PP**	**# EBC**	**%**
EASE	50	6	12%
SBES	42	26	61%
SBCARS	12	6	50%
SAST	11	6	54%
TOTAL	115	44	38%

Regarding the **limitations** of ES4, the author of this thesis is also authors of three papers published on two of the four conferences. To avoid bias, he did not participated on the creation of EBs or answered the questionnaire related to their three papers. We asked other co-authors to create the EBs and answer the questionnaire in order to avoid bias. Additionally, authors of more than half of the papers published on three – SBES, SBCARS, and SAST – out of four conferences created EBs of their papers. On the other side, just 12% of EASE's papers had an EB created. Nonetheless, we believe 12% is not a low rate in SE (SINGER; SIM; LETHBRIDGE, 2008), considering the high involvement demanded from the researchers to create Evidence Briefings, there is a possible explanation to the difference between EASE and the other three conference rates. We sent the invitation via e-mail (June 13th, 2017) to the EASE authors, only two days before the beginning of the conference (June 15th, 2017). So, the authors did not had much time to create EBs and were probably traveling or organizing their activities for the conference. The invitations to the other three conferences were sent about one month before the conferences.

4.4 RESEARCH RESULTS

In this section, we present the results of each of the three empirical studies (ES2, ES3 and ES4) aiming to evaluate EBs.

4.4.1 Empirical Study 2 (ES2) Results: Practitioners Perceptions on Evidence Briefings

In this section we present the results of the survey conducted with 32 Stack Exchange users (i.e. practitioners). Throughout this section, the questions of this survey are referred following this pattern: **S1Q<N>**, where **S1Q** stands for **S**urvey **1** **Q**uestion, and **N** refers to the question number in the questionnaire. For instance, the first question of this questionnaire is represented as **S1Q1**. Further on you will note we did not present the results of S1Q1 to S1Q4 in this section. These are demographic questions, thus their results were already shown in Section 4.3.3.

Mediums to acquire knowledge: We asked respondents how often do they refer to Stack Exchange websites (S1Q5) and SE research papers (S1Q6) to support their decision-making. Figure 11 presents the results.

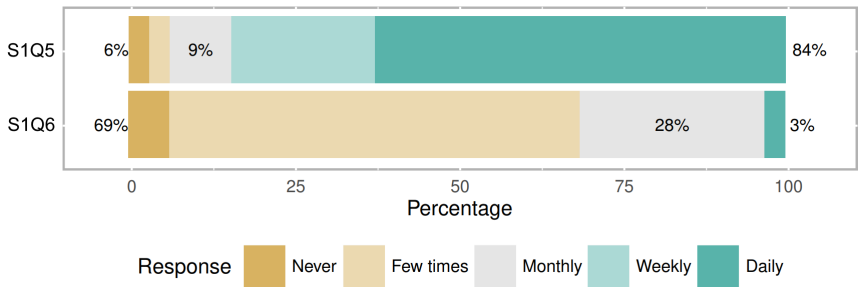


Figure 11 – How often software engineers refer to Stack Exchange websites (S1Q5) and research papers (S1Q6) (%).

As observed, 62% of the respondents use Stack Exchange websites on daily basis and another 22% use it weekly (S1Q5). On the other hand, we found that 62% of the respondents said they have read only a few SE research papers, but it is not common, and 6% of the respondents had never read a single SE research paper (S1Q6). We also found that only 28% of those who have read SE research papers have read a SS paper (S1Q7). Among them, 44% did it for research purposes, 33% for decision making on practice, and 11% for teaching (S1Q8). It means that only 9% (3 out of 32) of all Stack Exchange users that participated in this survey have read SS papers to support their decision-making in practice. Thus, we believe SE researchers are in need of better ways to disseminate their research findings to practitioners. This fact motivates us to create and evaluate our EBs in terms of content and format. Additionally, 78% of the respondents believe that a platform such as

Stack Exchange to discuss EBs of SE research is "Important" or "Very important" (S1Q9).

Briefings Contents: We asked six questions to evaluate the briefings contents. We start by asking "To what degree do you think the information available in the briefing we sent to you can answer your question on Stack Exchange?" (S1Q10). Among the responses, 10% said that the EB has "Totally" answered, and another 20% said that it has "Partially" answered their Stack Exchange questions. Another 32% said the EB touches a related topic, but does not help to answer their questions. The remaining 38% said that the EB is not related to the their Stack Exchange questions and, therefore, it does not help to answer them. In the following question (S1Q11), we asked respondents who said the EBs does NOT "Totally" answered their Stack Exchange question to explain why they think so. We categorized their responses into five main reasons:

- **The question is too specific:** Three Stack Exchange users reported that their questions are highly specific, whereas the SS is rather general. For instance, one user said *"My question is if FDD is compliant to most famous Agile methodologies, not about advantages/disadvantages of Agile methods over not-Agile methods."*

- **The question expects more than one answer:** In fact, this happens with just one question. In this question the Stack Exchange user asked *"What are the alternatives to gather requirements from large specification files with Scrum? Should the PM take care of it with a specification team? Or the development team should be more suitable for this kind of task? The specification analysis should be time boxed into the sprint or it comes before Scrum can be applied?"* Although the EB is capable of answering the technical side of the questions, it does not answer the human side. In the survey, the user recognized it, saying *"The question is also about the human side of the problem. Even with the right approach, we need to take into account the large amount of repetitive work that needs to be done for large requirements and how it can be divided in teams."*

- **The question touches a slightly different issue:** Sometimes the question is about a slightly different topic, which prevents the EB from totally answering the question. We found six occurrences of this pattern. For instance, one respondent said that *"The briefing provides interesting information about Fit, but my question was more about deployment issues rather than testing of itself."*

- **The briefing lacks details:** Six Stack Exchange users reported the EBs lack important details. For instance: "*My question was how one might introduce agile methods in a startup. The briefing confirms the gut feeling that it is easier to introduce agile methods in small, non-complex companies, but it doesn't go into detail as to which steps to take and how to organize it.*"

- **Not related at all:** Four respondents suggested the EB has nothing to do with the question they asked, which lead us to expected situations when Stack Exchange users do not share the same perception of us about the topic of the SSs and, in consequence, the EBs' topics.

Due to the high rate of users affirming their questions are not answered and also considering not related to the EBs, we took a moment to investigate why it occurs. First, we found four of those questions can be also classified as related to the subject. For instance, the following question *"Why is Feature Driven Development considered an Agile methodology?"* asked on PROG website, and the EB related to this question is about agile methods (DYBÅ; DINGSØYR, 2008), which we believe is related to the question. In the survey, the respondent mentioned that *"My question is if FDD is compliant to most famous Agile methodologies, not about advantages/disadvantages of Agile methods over not-Agile methods."* Thus, as we mentioned the perception of which questions are considered as related to the briefings may vary from person to person.

In the next question, we asked "Regardless the briefing answers or not your question, how important do you think is the research presented on the briefing?" (S1Q12). We found that 62% of the respondents said the research presented in the EBs are "Very important" or "Important" (25% and 6% believe the research is "Moderately important" and "Slightly Important" respectively). The remaining 6% believe the research are "Unwise." Here we can observe a paradox between the high rate (62%) of practitioners considering the research important and the low rate (28%) of the same researches that could help practitioners answer their questions (S1Q11). This finding suggests that if we as researchers want to produce evidence more useful to practice, it is important to focus on research that answers practitioners questions.

The respondents that have answered *"Unwise"* were asked to describe their reasons in S1Q13. Two of them did so. The first one said *"I truly believe the research is a good path, and is closer to Moderately Important, or even Important. But it is misleading and dogmatic."* According to that Stack Exchange user, code duplication is not always harmful. For instance, he/she mentioned that "*code duplication should be defined more acutely, as it is a necessity in many cases. Not a necessity for poor reasons such as being in a hurry, but a necessity for good reasons such as following SOLID Principles."* The second said *"Agile is not a one size fits all methodology. To make it work you need to see what works for you and your team. [...] Making bold high level statistical statements about Agile software development will only hurt it where as it can shine in truly Agile organizations."* For this particular case, we believe no research finding would easily change her/his mind. This lead us to areas of SE that are based on strong beliefs, and as verified by Devambu et al., strong evidence is required to change strong beliefs (DEVANBU; ZIMMERMANN; BIRD, 2016).

Also, 62% of the respondents prefer the answers of Stack Exchange websites over the findings presented in the EBs (S1Q14). One of the reasons that might explain this is the gamification mechanism employed on Stack Exchange websites, which motivates practitioners to contribute (DUBOIS; TAMBURRELLI, 2013), in contrast to the static-based format of EBs. Yet, it is encouraging that respondents kept open to EBs since 78% affirmed that a platform like Stack Exchange to discuss EBs is "Important" or "Very Important" (S1Q6). To increase the influence of EBSE in practice, we believe that researchers can take advantage of these well-known techniques, and use in favor of SE.

Briefings Format: In this final set of questions, we asked respondents what they think about the structure of the EBs. Figure 12 shows the overall impression.

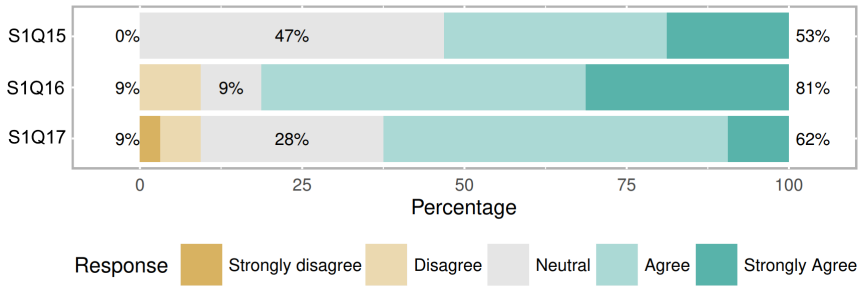


Figure 12 – Practitioners perceptions on the briefings format (%).

First, we observed that 53% of the respondents "Agree" or "Strongly Agree" that it was easy to find the information in the briefing (S1Q15), and 82% of the respondents "Agree" or "Strongly agree" that the briefing interface is clear and understandable (S1Q16). For these two questions, we believe that the great impression about the EB interface is due to the well-known design principles we used for creating them, namely those from Gestalt and Design Information theories. Second, we found that 62% of the respondents "Agree" or "Strongly agree" the EBs look reliable (S1Q17). In addition to the research findings available in the briefing, we highlight the original paper in the footnote. Moreover, in order to increase reliability, we put institutional logos from our research group, department, and university on the briefings. Institutional websites are also provided.

4.4.2 Empirical Study 3 (ES3) Results: Researchers Perceptions on Evidence Briefings

In this section, we present the results of the survey conducted with authors (i.e. researchers) of seven of the selected SSs, aiming to investigate their opinions about the EBs we created. Throughout this section, the questions of this survey are referred following this pattern: **S2Q<N>**, where **S2Q** stands for **S**urvey **2** **Q**uestion, and **N** refers to the

question number in the questionnaire. For instance, the first question of this questionnaire is represented as **S2Q1**.

We start asking "How important for you is to share research results to practitioners?" (S2Q1). We found that 100% of the respondents said it is "Very important". However, 29% (2) of the them affirmed to have shared a few times only, but it is not common. The remaining ones share research findings on weekly (42%) or monthly (28%) basis (S2Q2). For those who affirmed to have shared research evidence with practitioners, we asked how they do so (S2Q3). It was an open question, so we applied open coding and constant comparison techniques to group the results in topics (STOL; RALPH; FITZGERALD, 2016). The sum of topics counting is greater than the number of respondents because some researchers mentioned more than one way to share knowledge. We identified five ways researchers share knowledge with practitioners (the frequencies are in parenthesis), they are:

- **Scientific Writing (4):** This was the most cited way the respondents of this survey mentioned to use to share knowledge with practitioners;
- **Seminars and Meetings (2):** One example, is a respondent affirming to share knowledge through *" [...] Seminars [...] Informal meetings/discussion;"*
- **Teaching and Advising (2):** To illustrate, a researcher mentioned s/he shares knowledge through *" Teaching [...] Advisory work;"*.
- **Applied Research (2):** Here, an example is the researcher affirming that *"Practitioners do PhD research with me, applying their research in their daily practice;"*
- **Social Media (1):** The one that mentioned this way to share knowledge said that *"Mainly twitter, but also research gate [...]."*

Next we asked "How does the briefing cover the main findings of your paper?" (S2Q4). We found 72% (5) of the respondents describe as "Good" or "Very good". The remaining 28% (2) said that it is "Acceptable". This shows an interesting impression of the authors of the selected SSs, and suggests that even though we are not the authors of the research papers, we were capable of creating, on the worst scenario, acceptable EBs.

The next and final group of questions is the same to the ones related to format of EBs in the first survey (ES2). Figure 13 shows the overview of the answers.

As we can see from this figure, similarly to what we found in the survey with practitioners (ES2), there is a consensus around the EB interface. We observed that 71% (5) of them "Agree" or "Strongly agree" that it is easy to find information in the briefings (S2Q6). Another 71% (5) "Agree" or "Strongly agree" the EB interface is clear and understandable (S2Q7). Finally, 56% (4) "Agree" or "Strongly agree" that EBs look reliable (S2Q8).

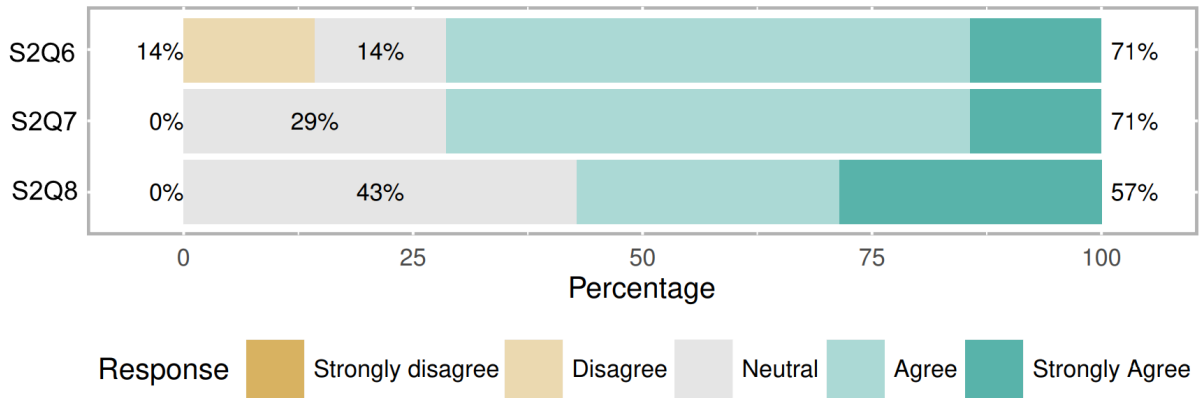


Figure 13 – Researchers perceptions on the briefings format (%).

4.4.3 Empirical Study 4 (ES4) Results: Researchers Perceptions on Creating Evidence Briefings

In this section we present the results of the survey conducted with authors of 44 papers published on four SE conferences: EASE, SBES, SBCARS, and SAST.

Throughout this section, the questions asked in this study are referred following this pattern: **S3Q<N>**, where **S3Q** stands for **S**urvey **3** **Q**uestion, and **N** refers to the number of the question in the questionnaire. For instance, the first question of this questionnaire will be referred as S3Q1. Since some respondents answered open questions in Portuguese, some of our quotes are free translations to English. Further on you will note we did not present the results of S3Q1 in this section. It is a demographic question, thus its result is shown in Section 4.3.5.

Knowledge Sharing Habits: More than 90% (41) of the respondents considered "Important" or "Very Important" to share research evidence with practitioners. None of them considered it as "Not Important" (S3Q2). These results are in line with ES3, where 100% of the seven researchers answered "Very Important" to the same question.

Although almost all researchers consider it important, to share knowledge with practitioners, 59% (26) of the them declared to have never shared research results with practitioners, or have shared only few times (S3Q3). The remaining 41% declared to share results with practitioners on yearly, monthly, weekly, or daily basis. This shows that proper mediums like EBs and incentives from conferences to create them can be one, among others, initiative to change this situation. For those who confirmed to share research evidence with practitioners, we asked how they do so (S3Q4). It was an open question, so we applied open coding and constant comparison techniques to group the results in topics (STOL; RALPH; FITZGERALD, 2016). The sum of topics counting is greater than the number of respondents because some researchers mentioned more than one way to share knowledge.

We identified seven ways researchers share knowledge with practitioners (the frequencies are in parentheses), they are:

- **Seminars and Meetings (15):** This is the most cited way to share knowledge with practitioners. To illustrate, a researcher mentioned that *"Occasionally they [practitioners] came to me [...] we have talked about results informally"*, and another researcher answered that knowledge is shared through *"Meetings, [...] and at practitioners conferences;"*

- **Scientific Writing (6):** One example, is when a researcher mentioned to share knowledge *"Only through traditional mediums like papers and posters"* and another researcher simply answered *"Researcher papers;"*

- **Practitioner Oriented Writing (2):** The EBs we are proposing can be considered a practitioner-oriented writing medium. Only two researchers mentioned to share knowledge this way. For instance, a researcher affirmed to share evidence *"[...] creating blog posts, creating website [...];"*

- **Social Media (2):** To illustrate, a researcher answered s/he uses *"[...] linkedin, [...] twitter [...];"*

- **Applied Research (1):** A researcher mentioned that the way s/he is used to sharing knowledge with practitioners is by *"Developing research within an industrial partner;"*

- **Consulting (1):** Another researcher stated that s/he shares knowledge through consulting activities;

- **Teaching and Advising (1):** One researcher affirmed to share knowledge through lectures or advising research.

Perceptions on Evidence Briefings Creation: About 86% (38) of the respondents reported to have taken up to three hours to create their EBs (S3Q5), although about 88% (39) of the respondents found it "Not Difficult" or "Somewhat Difficult" to create the EBs (S3Q6). The remaining 12% found it "Difficult" or "Very Difficult". For those that reported difficulties regarding the process of creating their EBs, we invited them to elaborate and discuss what were their difficulties (S3Q7). Since it was an open question, we applied open coding and constant comparison techniques to group the results in topics (STOL; RALPH; FITZGERALD, 2016). We identified four types of problems that the respondents faced when creating their EBs (the frequencies are in parenthesis), they are:

- **Difficulties to Summarize the Findings (5):** The most common difficulty was to summarize the findings of the original papers to fit the EBs. For instance, a

respondent declared that *"[to] synthesize the most relevant information from a 8 to 10 pages paper, it is a hard task"*.

- **Difficulties with the Template (5):** The box on the bottom of the EBs template disappeared when opening with different file readers. Two respondents suggested to create a LaTeX template; we only provided an OpenDocument (odt) and a Microsoft Word (docx) format.
- **Difficulties with English (2):** This difficulty is not inherent to the EBs creation process, since to write a traditional research paper they would face the same problem. Additionally, we believe that, as researchers, practitioners may also face difficult reading EBs written in a non-native language. Thus, EBs should be written in various languages in order to impact more practitioners.
- **Difficulties Due to Research Type (2):** Two researchers affirmed their research does not fit in EBs format. It occurred because we invited the authors of all papers published on the four conferences. However, this two were position papers with little or no empirical evidence. EBs are focused on presenting just research with empirical evidence. Thus, it would naturally occur.

Despite the fact that some authors faced some difficult, the majority of them (63%) considered the EB they created as "Good" or "Very Good" (S3Q8). This percentage grows to 95% considering respondents that considered the EB as "Acceptable." Two researchers considered their EBs as "Poor." We asked why they thought so (S3Q9). One respondent was not sure if s/he included all the essential findings on her/his EB, as can be seen when s/he answered *"I do not know if it contains all essential information in Findings."* While the other believed her/his research does not fit into the idea of EB since it was published on the "New Ideas" track of the conference, which usually does not provide strong empirical evidence.

Yet, 90% of the respondents affirmed that the one-page EB format was enough to summarize the results of their original papers (S3Q10). This reveals that with some effort and training, researchers can present the main results of their research in an accessible way to practitioners.

Perceptions on Evidence Briefings Initiative: Although EBs were initially thought of as a medium to transfer knowledge to practitioners, they also might be useful knowledge sources for researchers too. In particular, 75% (33) of the researchers affirmed to have interest in reading EBs before the original papers, if conferences start to publish EBs along with research papers at their official websites (S3Q11). In the next question (S3Q12), we asked respondents that did not support this initiative (six respondents) to justify their

decision. We identified two situations where researchers favor the original papers instead of EBs, they are:

- **High Interest on the Research Topic (4):** Four researchers affirmed preferring to read the original paper instead of EBs when they are highly interested in the research topic. For example, a researcher mentioned *"If paper's title seems of my obviously interest, I would go directly to the paper. I would take a look on the briefings only for the papers that I am not sure or I am curious about."*

- **Abstract Provides Enough Information (2):** Two respondents affirmed preferring to read only the abstract of the papers when it provides enough information. However, they said they would also read the EBs if they like to read a bit more. For instance, a researcher mentioned *"If abstract and summary are not sufficient to give me an idea about the paper then I would read the briefing."*

As we can see, although these respondents do not strongly support the use of EBs for academic purposes, they still would like to refer to an EB in particular occasions.

When asked how likely it would be for the researchers to create EBs, if they are commonplace in SE conferences, yet not mandatory (S3Q13), 88% (39) answered "Likely" or "Very Likely". This shows that, despite the effort demanded to create EBs, researchers are willing to do it.

When asked whether practitioners would benefit from the EBs (S3Q14), about 70% of the respondents mentioned that "Likely" or "Very Likely". In comparison, the results of our survey with practitioners (ES2) are in sharp agreement with this researchers' perception (Section 4.3.4).

In S3Q15 we asked researchers which medium they believe is more suitable to transfer knowledge to practitioners. Evidence Briefings and traditional research paper format drew, with 17 (38%) researchers voting for each of them. We leave an open option on this question, so the respondents could insert any other medium they believe is the best to transfer knowledge to practice, and 25% of them chose a different option. Some of the options are: "Magazines like nature" and "Face-to-face conversations".

When asked about the benefits of the EBs initiative (S3Q16), we identified eight, they are:

- **Quick and Simple Source of Information (26):** This is the most cited benefit. To illustrate, according to one respondent *"The [EB] is a quick and friendly form of sharing the main results of a research"*. Another respondent said that *"[EBs] are a quick and simple way to understand a research, that sometimes is long and complex;"*

- **Stimulates Reading the Original Paper (3):** To illustrate, one respondent mentioned EBs promote *"a faster reading of the achievements of the paper, which may*

generate a greater interest in reading the full paper." Another researcher affirmed *"The briefings allow both researchers and practitioners to have a summarized view of a research paper. In this perspective, the advantages are indeed the raison d'être of briefings. If one wants to know more details, he/she can read the complete version of the paper;"*

- **Stimulates Practitioners to Consume Empirical Evidence (2):** Two respondents believe this is the main benefit of EBs. For instance, one of them mentioned that *"the main advantage is to summarize for an industry professional the results related to a research [...] These briefings may stimulate a practitioner to directly adopt or not a certain research approach;"*

- **Increase Research Visibility (2):** One example, is a respondent affirming that EBs can *"make my research more visible."*

Other benefits mentioned only once are: (1) **Low effort to create,** as one researcher said that EBs allow one to *"share results faster and more effectively with little additional work";* (2) **Serve as Executive Summary,** as a respondent mentioned that *"for the industry, [EBs] can be used as an executive summary to propose changes;"* (3) **Richer than abstracts**, since a respondent mentioned an EB "*allows reaching the target audience (richer than abstract)*;" and finally (4) **A better way to share some findings**, a researcher declared *"in my case, by elaborating the briefing, I had another perspective. I perceive better ways to indicate a finding."*

We also asked the researchers what are the drawbacks of the EBs (S3Q17). We could identify three, they are:

- **Lack of Method Information (8):** The most cited drawback was the lack of information about the research method. For example, a researcher complained that *"[...] the methodology is not explained;"*

- **Additional Effort (5):** To illustrate, a respondent said *"[we still] need to write an additional document [beyond the paper itself];"*

- **Size restriction (4):** One example is a researcher saying that *"for some studies it is complicated to summarize [the findings] in one page."*

In addition to the three mentioned drawbacks, eight researchers said they do not see any relevant drawback. Thus, the number of researchers seeing no relevant drawbacks on the EBs initiative drew with the most cited drawback (i.e. "Lack of Method Information").

4.5 DISCUSSION

Researchers want to transfer knowledge to practice but not all of them do so: In ES3, we found that 100% (7) of the respondents believe transferring knowledge to practitioners is "Very important" but only 42% do it at least on weekly basis. We found similar results in ES4, where 90% (42) of the researchers consider "Important" or "Very Important" to share knowledge with practitioners, but 59% of them never did it, or did it only few times. In this context, evidence is pilling suggesting that researchers are having a hard time to transpose research findings to practitioners (SANTOS; SILVA, 2013; SILVA et al., 2011; DEVANBU; ZIMMERMANN; BIRD, 2016; OSTERWEIL et al., 2008; CARTAXO et al., 2017). Additionally, based not only on the results of this research, we could also see that researchers consider this situation an important issue for the SE community (BEECHAM et al., 2014; JEDLITSCHKA; JURISTO; ROMBACH, 2014; MEYER, 2018c; STOREY et al., 2017). Thus, we believe this research and others on the same direction are crucial to reduce the gap between research and practice in SE.

Software engineering practice still has beliefs with no evidence basis: A practitioner in ES2 considered one of the research results reported in an EB as unwise, and justified her/his perception saying *"Agile is not a one size fits all methodology. To make it work you need to see what works for you and your team. [...] Making bold high level statistical statements about Agile software development will only hurt it where as it can shine in truly Agile organizations."* For these cases, we believe no research finding would easily change her/his mind. This lead us to SE topics that are soaked with strong beliefs, and as verified by Devambu et al. (DEVANBU; ZIMMERMANN; BIRD, 2016), strong evidence is required to change strong beliefs.

Some researchers consider scientific writing is an effective way to transfer knowledge to practice, but few practitioners read research papers: Six respondents approached on ES4 mentioned that, by writing and publishing research papers, they are sharing knowledge with practitioners. On the same survey, traditional research paper was considered the best way to transfer knowledge to practice, by 38% of the researchers. The belief on traditional research paper format from researchers' side is understandable. It is the most spread and formal source of information in academia. However, results from ES2 shows that practitioners barely read research papers as source of information: 69% percent declared to have never read a researcher paper or have read only few times. Therefore, the development of mediums like EBs are important to transfer knowledge to practitioners. On the other hand, we are aware that this is a long term cultural change. Multiple and complimentary initiatives are demanded to cross the bridge that makes research in SE so distant from practice.

The effort to create Evidence Briefings is non-trivial but researchers still see benefits of creating them: In ES4, the majority of researchers (88%) created their EBs

in up to three hours. However, the researchers considered EBs as a quick and simple source of information that stimulates practitioners and researchers to consume more empirical evidence, demanding a low effort to create. When asked how likely researchers would create EBs if they are common place in SE conferences, yet not mandatory, 88% (39) answered "Likely" or "Very Likely." This shows that, despite the effort demanded to create EBs, researchers are willing to do it. Still, eight out of 44 researchers reported none or only few difficulties to create EBs. Five researchers considered the extra effort one of the main drawbacks of EBs. Additionally, 88% (39) of researchers believe practitioners would "Likely" or "Very Likely" be benefited with EBs. One important point is that, the success of EBs initiative is strongly related to how motivated researchers are to create them. Since they believe practitioners would be benefited, it may raises their motivation.

Researchers should learn how to summarize research findings if they want to transfer knowledge to practitioners: In ES4, when we asked researchers how difficult was to created the EBs, only 11% considered it as "Difficult" or "Very Difficult". However, the majority that have difficulties was related to summarize the results. For instance, a researcher mentioned *"I had some difficulties to summarize the findings."* Additionally, one of the researchers spontaneously sent as a feedback e-mail affirming that *"[...] it was hard to present the results in a simple and attractive way [...] this made me think whether my paper was written in a attractive and straightforward way [...] I learned a lot with this exercise [...]"*. This may reveal it is important to invest more time in practitioners oriented writing since usually researchers write to their peers which is a different audience.

Evidence Briefings are not applicable to all kinds of research: Some researchers faced problems to create EBs to their positional or non-evidence based papers. Thus, conferences interested in adopting EBs should advise authors that only evidence based papers should be considered.

Both researchers and practitioners positively evaluated the Evidence Briefings: Based on the three empirical studies (ES2, ES3, and ES4), there are evidences about the promising future of EBs as one of the viable mediums to transfer knowledge to SE practice. The majority of the practitioners on ES3 considered EBs are "Easy to find information" (71%), have "clear and understandable interface" (71%), and also "Look reliable" (56%). Additionally, the majority of the researchers on ES4 (63%) considered the EBs they created are "Good" or "Very Good." Moreover, 90% of the researchers on ES4 considered that the one-page format is enough to present their results in straight-forward way. This is particularly relevant since one of the biggest challenges to make EBs feasible is to have a concise format to be attractive to practitioners.

Researchers are also interested in consuming Evidence Briefings, although they are not the main target audience: The majority of the researchers on ES4 (75%) affirmed they would read Evidence Briefings before reading the original papers, if they would be available at conferences' websites.

Researchers believe the lack of research method information is the main drawback of Evidence Briefings: The majority of the drawbacks mentioned in ES4 were related to missing information. However, the information researchers claimed to be missing is related to research method, which EBs deliberately omit. This occurs because the idea with EBs focuses on practitioners, which usually are not interested in methodological details. On the other hand, we envision opportunities to create EB templates focused on academic audience, that could be more concerned about methodological issues.

The software engineering research community welcomed the Evidence Briefings. Besides filling the questionnaire, some respondents of ES3 and ES4 (researchers) also spontaneously sent us positive feedbacks about EBs. For example, a researcher highlighted the importance of such kind of initiative saying *"congratulations for the initiative, I liked it a lot, and I find this type of research dissemination is extremely important."* Other envisioned an opportunity to use EBs to disseminate knowledge in her/his research group when said *"I liked a lot the Evidence Briefings, I have never heard about it. I'm thinking in apply it with my undergraduate students, so we can have a clear and objective view of the main findings of our research. Very nice."*

4.5.1 Implications

Researchers: Can observe software engineers believe in the importance of software engineering research, although are not used to refer to it. Thus, creating EBs might increase the visibility and impact of a research.

Practitioners: Can take advantage of EBs in different ways. For instance, fostering SE research results between their peers, and also discussing high-level research results that might motivate them to find additional information in the original papers.

Educators: Can be useful for undergraduate and graduate educators who teach computer science courses. In particular, educators can better motivate students to read and discuss research findings highlighted in EBs.

Conference and Publicity Chairs. In addition to regular research papers, EBSE conferences and publicity chairs could invite authors of accepted papers to submit EBs of their research. These briefings, therefore, can be promoted through the official website, as well as through social networks, potentially increasing the visibility of the conference.

4.6 CREATING AN EVIDENCE BRIEFING

Based on the responses received in the three empirical studies (ES2, ES3, and ES4) and the authors' own experience creating EBs, in this section we present a set of guidelines to support researchers who want to create their EBs.

1. The **original research** paper should report an empirical study, either primary or secondary. Positional papers do not fit well in the Evidence Briefings' template since they might not be mature enough to have evidence based on empirical methods.

2. **Logos** of researcher's institutions should be placed at the very top of the Evidence Briefing document. This publicizes the institutions producing Evidence Briefings, and might make practitioners search for more in the institutions websites.

3. The **title** of the Evidence Briefing does not need to be the same of the original paper. On the contrary, we encourage researchers to make the title as concise as possible when creating the Evidence Briefing. One or two lines title is ideal. Titles with more than two lines should be avoided.

4. To fill the Evidence Briefing's **abstract**, we suggest researchers to adopt the structure as in the template file, which is: *This briefing reports scientific evidence <RESEARCH GOAL>*. The abstract should span few lines. Following is an example of Evidence Briefing's abstract: *"This briefing reports scientific evidence on the challenges involved in using Scrum for global software development (GSD) projects, and strategies available to deal with them."*

5. Researchers should not put information about the **research method**. The idea of the Evidence Briefing is to quickly communicate the main findings of an empirical research, chiefly to practitioners. If they have interest they can search for the original paper since there is a reference to it on the briefing.

6. The **findings** section is the most important one. It should list the main findings of the original research paper. Findings of a research are usually on paper's result, discussion, or sections alike. When writing the findings, we recommend to use one finding per paragraph. Bullets to highlight important points as well as charts, figures, and tables are welcomed since they make the findings even easier to understand. Findings should be short sentences, straight to the point.

7. The **box at the right side** of the Evidence Briefing should be filled with information about the Evidence Briefing's target audience, clarifications about what information is included and what is not on the Evidence Briefing. The template has a complete set of suggestions to structure information at the right box.

8. The **reference** to the original paper should be placed at the bottom of the Evidence Briefing. It should be concise, so we suggest to follow the structure provided on the template.

For more information and examples of Evidence Briefings refer to our research group website.[8]

[8] `http://cin.ufpe.br/eseg/briefings`

4.7 CHAPTER SUMMARY

In this chapter, we proposed Evidence Briefings (EBs) – inspired on Evidence Based Medicine (EBM) – as a medium to transfer knowledge from research to Software Engineering (SE) practice. An extensive evaluation – three empirical studies – was also conducted with practitioners (32), researchers (7), and also introducing EBs in four SE conferences (44 researches created briefings). Our results show that EBs were positively evaluated by researchers and practitioners. Most of the researchers and practitioners believe that it is easy to find information on EBs. Yet, most of them believe EBs are clear, understandable, and reliable. On the other side, some researchers consider scientific writing is the most effective way to transfer knowledge to practice, but few practitioners read research papers (69% of practitioners reported to have never read research papers or read only few times). Additionally, we observed that the majority of researchers (90%) consider "Important" or "Very Important" to share research evidence with practitioners, but few of them do it frequently. Our results also revealed that it is non-trivial to create EBs (88% of researchers took up to three hours) but researchers sill see benefits creating them, since 88% (39) researchers answered they "Likely" or "Very Likely" create EBs if they are common place in SE conferences, yet not mandatory. This shows that, despite the effort demanded to create EBs, researchers are willing to do it.

In this context, we believe EBs can play an important role on transferring knowledge from research to SE practice. Indeed, we believe no transfer medium is definitive and one-for-all solution. Instead, multiple transfer mediums should be used aiming to transfer different kinds of knowledge, targeting different audiences. EBs are one of these possible mediums.

5 THE TRANSFER OBJECT: RAPID REVIEWS' EVIDENCE

In this chapter, we *propose*, *evaluate* and *discuss* the use of Rapid Reviews (RRs) as the Transfer Objects of the knowledge transfer model we presented in Chapter 3. The evaluation was conducted guided by the following research question:

GRQ3: What are the perceptions of practitioners on the knowledge produced with Rapid Reviews as transfer objects?

To answer this question, we present the first introduction of RRs in a Software Engineering (SE) setting. We conducted two RRs in two different companies situated in Recife, Brazil. To evaluate the practitioners perceptions about the RRs, we conducted empirical studies, based on action research principles (SANTOS et al., 2011), in both software companies. Several interviews were conducted during the course of those studies. In Company 1, we observed the main problem they were facing was dealing with low costumer collaboration in one of their projects, while in Company 2 the problem was related to low team motivation. We proposed as action the conduction of RRs on those topics, changing the way they make decisions, which usually is based on informal sources or expert opinion. The 17 and 35 primary studies included in the RRs to Company 1 and Company 2, respectively, were synthesized and the findings presented in (1) workshops and (2) printed as Evidence Briefings (EBs) (CARTAXO et al., 2016). During the workshops, we conducted another interview, which was aimed at understanding the practitioners perception about the RRs.

Practitioners of both companies perceived many benefits, such as the novelty of the approach, the applicability of the RRs to their problem, and the reliability of the content. As a shortcoming, they found some findings were not clear in the EB form – although they became clearer after discussing with researchers during the workshop. After two months the RR was introduced in Company 1, we conducted another interview to assess whether practitioners applied the knowledge obtained with the RR. Fortunately, practitioners applied some of the knowledge in their daily work habits, although some findings that were initially thought as useful were not adopted. Part of the results of this chapter were published at the 22th International Conference on Evaluation and Measurement in Software Engineering (EASE 2018) (CARTAXO; PINTO; SOARES, 2018b).

5.1 BACKGROUND AND RELATED WORK

RRs are lightweight Secondary Studies (SSs) focused on delivering evidence to practitioners in a timely manner (TRICCO et al., 2015; HARTLING et al., 2015). Some steps of Full

Secondary Studies (FSSs) are deliberately omitted or simplified in RRs to achieve their proposed goal.

Although RRs are not well-known in SE, there is a growing interest in health-related fields. Tricco et al. (TRICCO et al., 2015) mapped 100 RRs published between 1997 and 2013, and observed that 56% were published on the last five years of investigation. Additionally, major medicine venues, such as the prestigious Systematic Reviews journal, are officially recognizing RRs as one Evidence Based Practice (EBP) method (MOHER et al., 2015). Moreover, Cochrane — a global renowned group of researchers and practitioners specialized in evidence diffusion in health-care — announced in 2016 a group to play a leading role guiding the production of RRs (GARRITTY et al., 2016; COCHRANE...,). Due to the increasing importance of RRs, the Canadian Agency for Drugs and Technologies in Health (CADTH), promoted the Rapid Review Summit in 2015, which was focused on the evolving role and practices of RRs to support informed health care policy and clinical decision-making (POLISENA et al., 2015). Even the World Health Organization (WHO) has recently published a guide presenting the importance of RRs (TRICCO et al., 2017)

The emerging character of RRs can be explained in terms of its benefits. For instance, a study observed RRs saved approximately $ 3 millions when implemented in a hospital (MCGREGOR; BROPHY, 2005). Moreover, a survey exploring the use of 15 RRs revealed that 67% were used as reference material and 53% were used to directly apply their results to support decision-making in practice (HAILEY, 2009). Additionally, Lawani et al. reported RRs enabled the development of clinical tools more rapidly than with FSSs (LAWANI et al., 2017). Taylor-Phillips et al. reported to have read 2,176 fewer title/abstracts and 129 fewer full texts when performing a RR (TAYLOR-PHILLIPS et al., 2017). Other studies have also demonstrated positive impact of RRs in practice (HAILEY et al., 2000; BATTEN, 2012; ZECHMEISTER; SCHUMACHER, 2012; TRICCO et al., 2015).

One may argue SE research is not as close to practice as medicine research is, thus RRs would not thrive in SE as thrived in medicine. However, this kind of initiative is exactly to make SE research closer to practice. The results presented in this research are the first evidence of that. Additionally, and more important, if medicine research – which is well known by its rigor and criticality — is embracing RRs as source of evidence, we believe SE — which is way less critical than medicine, and does not provoke life losses, at least in most cases – could also be benefited by RRs.

5.1.1 Characteristics

Since RRs are a recent phenomenon in Evidence Based Medicine (EBM), it is known there are many variations on their methods. This can be observed on the results of Featherstone's et al. study, which analyzed the method of many published RRs (FEATHERSTONE et al., 2015). Another comprehensive study about RRs reinforces the method variability by interviewing 40 RR's producers and applying Delphi approach with 113 RRs stakeholders

(TRICCO et al., 2016). These two studies identified high heterogeneity among RRs, from varying time frames, to ambiguous definitions of what is a RR.

Despite RRs high variability, the majority of them share common core characteristics (KELLY; MOHER; CLIFFORD, 2016a; TRICCO et al., 2015):

- They reduce costs of heavyweight methods (Section 5.1.2);
- They deliver evidence in a timely manner (Section 5.1.2);
- They are performed in close collaboration with practitioners (Section 5.1.3);
- They aim to provide evidence to issues practitioners are facing in practice (Section 5.1.3);
- They report results through appealing mediums (Section 5.1.4).

The study of Hartling et al. (HARTLING et al., 2015) defines a taxonomy to classify RRs, they are: Inventories, listing what evidence is available; Rapid Responses, present best available evidence with no formal synthesis; pure Rapid Reviews, synthesizing findings from the primary studies; and Automated Approaches generating meta-analyses in response to user-defined queries. As we shall see in Appendix E and F, in this thesis we employed the pure rapid review approach. Table 22 compares the main characteristics of RRs and FSSs. The RRs characteristics are based on many medicine studies and guidelines (TRICCO et al., 2017; KHANGURA et al., 2012; ABOU-SETTA et al., 2016; TAYLOR-PHILLIPS et al., 2017), while the FSSs characteristics are based on some SE studies and guidelines (KITCHENHAM; CHARTERS, 2007; CRUZES; DYBå, 2011; SANTOS; SILVA, 2013).

5.1.2 Timely Evidence and Reduced Costs

Many strategies have been applied to RRs in health-related fields to reduce costs and time, such as: (1) limiting search strategy by date of publication and/or search source; (2) using just one person to screen studies; (3) not conducting quality appraisal of primary studies; (4) or presenting results with no formal synthesis (TRICCO et al., 2016; TRICCO et al., 2015).

5.1.3 Collaboration with Practitioners

The argument to have more lightweight secondary studies like RRs holds only if time and costs are hard constraints. This kind of scenario is typically observed in the practice of many fields. Therefore, RRs are only conceived to be conducted bounded to a practical problem, inside a practical context. In consequence, practitioners should be willing to devote part of their busy schedule to participate on RRs.

Table 22 – Comparison of rapid reviews with full secondary studies characteristics.

CHARACTERISTIC	RAPID REVIEWS	FULL SECONDARY STUDIES
Problem	Must be a practical problem emerged from a practical context.	Can both emerge from research and practical contexts (KITCHENHAM; CHARTERS, 2007). However the lack of focus on practical problems has been the rule in software engineering (SANTOS; SILVA, 2013).
Research Questions	Must lead to answers that helps solving or at least attenuating the practitioners problem. Exploratory questions aiming to identify which are the strategies and their effectiveness to deal with the practitioners problem are one of the gold standards.	Admit questions aiming to support practitioners decision-making, but also ones that are primarily of interest to researchers (KITCHENHAM; CHARTERS, 2007).
Protocol	Ideally have a document do formalize the protocol, but there are some that do not have one.	Must have a document formalizing the protocol.
Execution	Must be conducted in close collaboration with practitioners, sometimes even having practitioners responsible for executing some of the steps.	Despite practitioners participation is possible, researchers usually conduct the entire process.
Time Frame	Days or Weeks	Months or Years
Search Strategy	- May use few or just one single source searching for primary studies (e.g. Scopus). - May limit search by publication year, language, and study design.	- Multiple sources to search for primary studies are recommended. - May also limit search by publication year, language, and study design, although more comprehensive search is recommended.
Selection Procedure	- Can be conducted by a single person. - The inclusions/exclusion criteria can be more restrictive aiming to find primary studies conducted in contexts similar to the one that motivated the conducting of the rapid review. (e.g. small/medium/large companies, companies situated in a specific country, etc) (TRICCO et al., 2017)	- Must be conduced in pairs to avoid selection bias. - Usually is less restrictive regarding specificities of primary studies context, specially when it is a mapping study, broader in scope.
Quality Appraisal	Conducted by a single person, or not conducted at all (TRICCO et al., 2017).	Conducted in pairs to avoid severe threats to validity due to low primary studies quality.
Extraction Procedure	Usually conducted by a single person to reduce time and effort.	Conducted in pairs to avoid extraction bias.
Synthesis Procedure	Narrative summaries are the most common way to synthesize evidence (TRICCO et al., 2015).	More systematic methods should be applied (e.g. meta-analysis, meta-ethnography, thematic analysis, etc), although it is not always the case (CRUZES; DYBå, 2011).
Results Report	Alternative mediums that better fit practitioners needs (e.g. Evidence Briefings).	Traditional research paper format.

Examples of that intrinsic characteristic can be observed in RRs on health-related fields. They are conducted with a practical problem as premise. For instance, the work of Best et al. described the experience of conducting 63 RRs for decision-making through the Development and Evaluation Committee in the UK (BEST et al., 1997). Or Bambra's et al. study, which described the experience of conducting nine RRs for the Secretary of State for Health (BAMBRA et al., 2010). Or even the research of Jahangirian et al., that discusses the experience of conducting five RRs for the Research into Global Healthcare Tools consortium (JAHANGIRIAN et al., 2011). Practitioners collaboration is crucial to the thrive of RRs.

5.1.4 Appealing Mediums

One important aspect of RRs is the way they are reported. Many authors argue that alternative mediums should be used, instead of the traditional research paper format (BEECHAM et al., 2014; GRIGOLEIT et al., 2015; CARTAXO et al., 2016). To substantiate this claim, Tricco (TRICCO et al., 2015) observed that, although RRs present several variations on their methods and terminologies, 78% present results as a narrative summary reported in mediums that better fit practitioners' needs. Examples of alternative mediums include: the Contextual Summaries of Young et al. (YOUNG et al., 2014), that limits the report to a one-page document; the Briefings presented by Chambers and Wilson (CHAMBERS; WILSON, 2012), that summarize the main findings of a SS in one section; or even the Evidence Summaries by Khangura et al. (KHANGURA et al., 2012), which use an informative box separated from the main text to highlight the audience and nature of the report.

In SE context, there are few studies proposing alternative mediums. The EBs we proposed (Chapter 4) are employed to report the results of the RRs we conducted in this research.

5.1.5 Rapid Reviews and Full Secondary Studies

Some studies were conducted to evaluate the impact of RRs methodological adaptations by comparing them with FSSs. A recent scoping study found nine studies comparing the results of RRs and FSSs. The conclusion is that both are generally similar (ABOU-SETTA et al., 2016). To illustrate, Corabian et al. compared six RRs with their FSSs peer reviewed publications (CORABIAN; HARSTALL, 2002). The conclusions differed only in one case. Another example is the study of Best et al., where two of the RRs they conducted were in agreement with FSSs published later on the same topic (BEST et al., 1997). Still, Taylor-Phillips et al. conducted a RR and a FSS about the same topic in order to compare them (TAYLOR-PHILLIPS et al., 2017). The comparison shows that RRs can provide similar results compared to FSSs. In that case, both methods identified the same papers.

On the opposite direction, other researchers affirm that conflicting results can be found (VELDE et al., 2011). For instance, the work of Van de Velde et al. compared results from their RR to a FSSs that was conducted by another group, on the same topic, and conflicting results were observed (VELDE et al., 2011).

Although there is more evidence reporting the similarity of results obtained with RRs and FSSs, there is also evidence on the opposite side. Thus, further investigations are needed to draw more conclusive results. In consequence, instead of considering RRs as replacements for FSSs, we believe RRs should be understood as a complementary EBP product. While FSSs are important to provide in-depth insights, RRs are important to easily and quickly transfer established knowledge to practice.

5.1.6 On Action Research

Action research is a flexible research method well-suitable to support EBP (DICK, 2011; ELSEY; LATHLEAN, 2006; BRIDGES; MEYER, 2007) and knowledge transfer initiatives (GAROUSI et al., 2016). These characteristics are clearly aligned with the goal of this thesis. Thus, we conducted empirical studies based action research principles to evaluate the perception of practitioners on RRs, and also to learn and improve how to apply RRs in SE practice. However, it is important to highlight that the method we followed cannot be classified as a canonical action research since it breaks some of the traditional characteristics of it, as we detail in section 5.2.3.

Action research has its roots on philosophical stances aligned to the critical theory. According to Easterbrook, from that perspective, researchers judges scientific knowledge by its ability to free people from restrictive systems of thought. So, critical theorists tend to choose which research to conduct based on whom it helps, and prefer participatory approaches like action research (EASTERBROOK et al., 2008).

In that sense, the studies presented in this chapter are closely related to the philosophical premises of action research, and it can be better understood through the Ivory Tower metaphor. This figure of speech is commonly used to highlight the distance that academia maintain from the mundane everyday life, deliberately or not (SHAPIN, 2012). That distance is usually discussed in terms of its negative impacts on academia and practice. Some argue that that distance can make academic research irrelevant due to its disconnection from the practical problems. On the opposite side, that distance also manifests the exclusiveness of the academic world, and sometimes its resistance to make their discoveries accessible to the society (LURIA; LURIA, 1970; ETZKOWITZ et al., 2000; SHAPIN, 2012; GITTELL, 2018).

In this context, the studies we present in this chapter are clearly not exempt from socio-political-ideological considerations. We intend to pose a perspective by trying deliberately to confront the status quo in academia, proposing a way to make research closer and more relevant to practice. At the same time, by adapting the way we conduct and present results of literature reviews in SE – in close collaboration with practitioners – we intentionally are trying to make the knowledge produced in academia more accessible and useful to practice. In consequence, we are acting politically in a way that empowers practitioners, specially because knowledge is one of the building blocks to empowering people in a way they can consciously exert their citizenship. So, this kind of initiative shakes the structures of the Ivory Tower, making it reachable to the ordinary lives, and acting as a solution in the middle of those worlds. Making relevant knowledge accessible to a group of people that once had no access to it is certainly one of the most classical ways to unlock empowerment. Shapin discusses how it is important to track the consequential changes in how we think about the nature of knowledge; the conditions for the production, maintenance and transmission of knowledge; the proper agency of knowledge making; and

the relations between knowledge and virtue, both individual and political (SHAPIN, 2012). This view is in accordance with other researchers that already envisioned how action research is well-suitable to support EBP (DICK, 2011; ELSEY; LATHLEAN, 2006; BRIDGES; MEYER, 2007) and knowledge transfer initiatives (GAROUSI et al., 2016), like the one we are proposing in this thesis.

An interesting illustrative example of solution in the middle is the one of professor Jody Gittell, which narrates her detour from what she considered an Ivory Tower career in Harvard Business School to the one where she embraces the relationship with practice, and how it was possible by adopting an action research oriented approach (GITTELL, 2018). One of the important lessons she mentioned to have learned is to share ownership of ideas and theories, which is, according to her, a ground breaking paradigm change for the academic world, which places great value on establishing ownership of ideas. The studies we present in this chapter are also aligned to the shared ownership concept, and this can be clearly seem by looking to some of the practitioners testimony who participated with us co-creating a way to make scientific evidence closer and more useful to practice. For instance, a practitioner envisioned a use to RRs and EBs, different from what we – the researchers – were thinking in the first place, and promptly posed it in a way that influenced us to also embrace that kind of use. They suggested that RRs and EBs could be conducted in a regular basis inside a software company as a way to foster internal discussions around their topics of interest, or even as technical clippings, not only as way to provide evidence to a problem they are actually facing in a specific project, which was our idea in the first place.

Apart from the more philosophical considerations, the following paragraphs present discussions to make clear the differences between action research like methods to other scientific methods. This is particularly important, because some confusions have been reported in the literature about the correct use of methods terminology, as you can on th following paragraphs.

Action research aims at studying a complex social process by introducing changes into that process and observing the effects of those changes (BASKERVILLE, 1997). This interventionist perspective is closely related to the basic principles of experiments, where an intervention is deliberately made in a system to observe the outputs (EASTERBROOK et al., 2008). However, experiments are reductionist by nature and demand controlled environments aiming to claim causal relationships between the interventions and the outputs (EASTERBROOK et al., 2008). This characteristic goes against one of the fundamentals of action research, which presumes that complex social systems cannot be reduced for meaningful study, and therefore should be studied within its real-life context (BASKERVILLE, 1997).

This characteristic, on the other side, approximates action research to case studies and other ethnographically oriented methods at a point that causes confusions (BASKERVILLE,

1997; SHARP et al., 2016). According to Yin, a case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context, specially when the boundaries between phenomenon and context are not clearly evident (YIN, 2008). In this context, one of the notable differences between case studies and action research is that, case studies are mainly concerned to augment scientific knowledge, not necessarily with a direct impact in practice. Action research, conversely, aims to contribute both to the practical concerns of people in an immediate problematic situation, and to the goals of scientific research community by joint collaboration (BASKERVILLE, 1997). Researchers in an action research disturb the environment under study aiming to solve a practical problem and directly change the state of practice (BASKERVILLE, 1997). On case studies, on the other side, researchers usually observe the environment without disturbing it, aiming to build, confirm, or refute scientific theories, not committed to directly change the state of practice (EASTERBROOK et al., 2008).

This leads to the last misunderstanding, which is confounding action research with consultancy. Consultancy is a process of temporarily introducing external expertise in order to obtain objective analysis, specialist knowledge, or the benefits of cross-organizational experience without permanently acquiring additional organizational members (STEELE, 1975). Consultants and researchers conducting an action research are both concerned about solving a problem practitioners are facing. However, the resemblance ends here. In action research, it is emphasized the need to collaborate in the discovery of beneficial changes. Practitioners must assume a posture of collaboration with researchers, and recognize that experimental changes must be introduced into the organization in the quest for a successful solution. The organization has to explicitly accept the uniqueness of their problem, and the necessity of experimental solutions (BASKERVILLE, 1997). In consultancy, oppositely, the organization usually expects consultants to bring their knowledge and experience with similar problem into the organization, and suggest the solutions, with high probability of success (BASKERVILLE, 1997). Consultants want to propose solutions to practical problems and directly change the state of practice, but they are not concerned to contribute by expanding scientific knowledge. Action research on the other side, aims to provide insights for both worlds.

In this thesis, we follow the action research guidelines (SANTOS et al., 2011) – with some discrepancies, as shown in Section 5.2.3 – and intervene in software projects conducting RRs aiming to change the way practitioners make decisions towards the solution of their problems. We also enrich scientific community by providing insights on how to better conduct RRs in practical settings as a way to transfer knowledge from scientific research to SE practice. Researchers both observe and participate in the phenomena under study. And finally, as already discussed, the motivations and philosophical views behind the studies of this chapter are highly aligned with the ones that guides action research like initiatives.

5.2 RESEARCH METHOD

This section presents the methodological details of the action research conducted in Company 1 and 2. Since the focus of this research is on the perceptions of practitioners on RRs (the results of the action research), not the RRs themselves, the protocol of RRs with Company 1 and 2 are not presented in this section. See Appendix E and F for the RRs protocols.

Figure 14 shows the steps we followed conducting both the two action researches. The numbers denote the order of each step. Looking to major activities (large gray areas), we started visiting companies and interviewing their employees searching for problems they were facing. Then, we defined the RR protocol with the practitioners. Next, we intervene on the practitioners usual way to make decisions, and conducted, with their close participation, the RR as planned in the protocol, as well as reported the RR results in an EB and discussed it in a workshop. Just after the workshop, we conducted an interview to evaluate how the practitioners perceived the introduction of the RR in their practical context. Two months after the introduction interview, we conducted an adoption interview aiming to evaluate if the RR evidence was in fact applied by the practitioners. The adoption interview was just conducted in Company 1 due to problems we have faced with Company 2, as will be explained latter in this section. Finally, we reflected about all we have learned during the entire process and structured the knowledge that will be discussed throughout this chapter.

5.2.1 Research Context: Company 1

The company: We conducted this study with an applied-research institute in Recife, Brazil. The mission of the institute is to "increase Brazilian industry competitiveness" providing services such as software development, applied research, and consultancy. The institute was founded in 2013, and had 16 innovation projects under development, and 21 employees.

The project: We first contacted the projects coordinator, who is responsible for coordinating all project managers. After presenting the goal of this research, a project manager joined us and discussed problems regarding low customer collaboration he was facing in one of his projects. Besides the project manager, this project has three software developers and one designer. The project started in August 2016 aiming to develop a system that monitors reusable packages during the entire production chain — from suppliers to factories — of the automotive industry. In this context, there are packages more expensive than the object they carry. Eventually, one of those packages get lost, which provokes high monetary losses. Hereafter, we refer this project as "Project 1."

The participants: The projects coordinator has about 15 years of experience in software development and holds a master degree in computer science. The project manager has

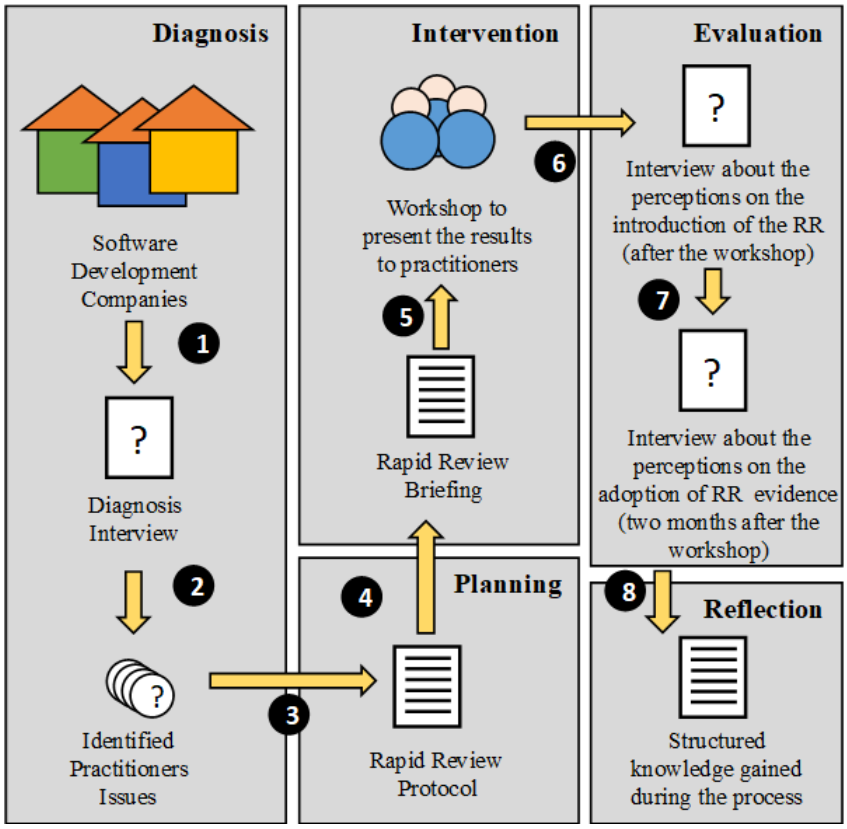


Figure 14 – Research steps (ES5 and ES6).

about 2.5 years managing software development projects and also holds a master degree in computer science. Hereafter, we refer to them as "The Participants of Company 1". The researcher who conducted this study together with the practitioners is a PhD candidate. Two professors supervised the entire study giving suggestions and feedbacks.

5.2.2 Research Context: Company 2

The company: We conducted this study with a software development company in Recife, Brazil. The company develops educational software products as well as musical software products. The company was founded in 2001 and has won several national and international awards.

The project: We first contacted the IT director, who is responsible for all the technological aspects of the company. After presenting the goal of this research, a project manager joined us and discussed problems regarding low team motivation he was facing in one of the projects. Besides the project manager, this project has four software developers and one software tester. The project aims to develop a platform to create and publish e-books for basic schools in the entire Brazilian territory. Hereafter, we refer this project as "Project 2."

The participants: The IT director has about 12 years of experience in software devel-

opment and holds a master degree in computer science. The project manager has about eight years in software development and holds an undergraduate degree in computer science. Hereafter, we refer to them as "The Participants of Company 2". The researcher who conducted this study together with the practitioners is a PhD candidate. Two professors supervised the entire study giving suggestions and feedbacks.

5.2.3 Research Steps

We followed the guidelines to conduct action research (SANTOS et al., 2011) in both Company 1 and 2, which encompass five main steps: diagnosis, planning, intervention, evaluation, and reflection. However, it is important to highlight that we did not follow the canonical action research method. Specially, because there were neither cycles for refining the prosed solution – as prescribes traditional action research guidelines – nor the solution itself was build entirely from scratch with practitioners. Albeit, they actively participated adapting the prosed solution (RR) to their reality in way they believed is more proper.

Diagnosis: This step intends to explore and define the problem the participants are facing in their environment (SANTOS et al., 2011). To identify the participants problem, we conducted semi-structured face-to-face interviews. The interviews were divided in four parts:

1. We explained the purpose of this study;
2. We asked questions regarding the background of Projects 1 and 2 (e.g. an overview of the project and the characteristics of the development team);
3. We asked questions to the participants of Company 1 and 2 regarding the challenges faced according to each SE area, as defined by SWEBOK;
4. We ended the interview asking whether the participants of Company 1 and 2 had additional comments that were not covered by the previous questions.

The complete interview script is in Appendix G. The interview was recorded and lasted ∼1h15min in Company 1, and ∼1h01min in Company 2. The audio was transcribed. The development team of Company 1 reported problems due low customer collaboration, while the development team of Company 2 reported problems due to low team motivation.

Planning: During this step, we decided to conduct RRs to provide empirical evidence to the participants. Therefore, we built the RRs' protocols in close collaboration with the participants of project 1 and 2 (see Appendix E and F for details). We also agreed with a deadline to finish the RRs and present the findings to the participants in a workshop. The execution of both RRs were planned to last one week with one researcher dedicated

full-time. Each aspect of the RRs protocols were discussed with the participants (e.g. research questions to guide the RRs, primary studies inclusion/exclusion criteria, etc). Online channels such as Skype and e-mail were frequently used during this step.

Intervention: This step performs the planning phase. We also executed the RRs protocols in close collaboration with the participants through online channels (e.g., during the selection procedure we showed to practitioners the title and abstract of primary studies we were in doubt whether to include or not). After conclude the RR, we had a face-to-face workshop with the participants to present the entire RRs process[1]. We discussed the evidence found, as well as how they can take advantage of that evidence on their job. The RR of Company 1 lasted six days, while the RR of Company 2 lasted eight days, considering they started on the diagnosis interview, and ended on the workshop to present the results.

Evaluation: To evaluate the participants perceptions on the RRs, we conducted and transcribed two semi-structured interviews. The first aims to evaluate the process of introducing RRs in practice. It happened during the workshop to report and discuss the results of the RRs. The complete interview script is in Appendix H. The interviews were recorded and lasted ∼45 minutes for the Company 1, and ∼32 minutes for the Company 2. The audios were transcribed. The results are presented in Section 5.3.1. The second interview was conducted about two months after we introduced the RR in Company 1. It enabled us to evaluate whether the knowledge introduced with the RR was in fact adopted by the participants. The second interview was also recorded and transcribed. It lasted ∼21 minutes. The results are presented in Section 5.3.2. This follow-up interview could not be conducted in Company 2 because both participants quit their jobs.

Reflection: According to Santos and Travassos, the reflection step involves the dissemination of acquired knowledge among participants and other organizations (SANTOS et al., 2011). Our reflections and reasoning about the experience of conducting these firsts RRs in SE can be found in section 5.4.

5.2.4 Limitations

Stringer (STRINGER, 2013) suggests to establish valid action research, it is important to assess the following attributes of the study:

Credibility: The participants commitment can only be reached if they trust in the integrity of the research process. During this study we maintained close connection with

[1] The workshop presentation to Company 1 can be found at <http://bit.ly/2s1exh5> and to Company 2 at <http://bit.ly/2FvqYfy>

practitioners, discussing every aspect of the study with them, which stimulated their engagement.

Transferability: Unlike experimental studies, action research results cannot be fully generalized to contexts out of the research. It does not mean this study cannot be applied in different settings. We detailed each method step and provided access to all materials we have used. This could encourage researchers and practitioners alike to conduct/replicate similar investigations.

Confirmability: Researchers must be able to confirm the procedures described actually took place. We provided the scripts of all face-to-face interviews, as well as spreadsheets with all the material we have used to conduct the RRs. This exposes our study to full scrutiny by other researchers judgment.

An important limitation on the results of the action researches we conducted with Company 1 and 2 is that we evaluated the perceptions of the participants, but not the effects of using RRs directly. This means if they affirm that a RR finding solved or attenuated their problem, it is their perceptions, but not necessarily the truth. The focus of this thesis is on the perceptions of practitioners, not in direct observations, because RRs are new in scientific community. Measuring the perceptions of the stakeholders are important to discover the probability of adoption of such new approach. According to Rogers (ROGERS, 2003), the perception of individuals about an initiative is one of the main predictors of its adoption. In addition, the interviews in both companies were conducted with all participants together, which might be a threat to validity. We could not interview each participant separated due to the limited time they have conceded.

5.3 RESEARCH RESULTS

This section presents the results of the empirical study we conducted in Company 1 and 2. It presents the participants perceptions on RRs supporting decision-making in practice. Since the results of the RRs are not the main goal of this research, we decided to omit them in this thesis and make it more concise. Nevertheless, the EBs we created to report the results of each RR – as were presented to the participants – are available online.[2,3]

5.3.1 Perceptions on Introducing Rapid Reviews

Here we present the perceptions of the participants of Company 1 and 2 regarding the introduction of the RRs in their practical environment. Their perceptions were obtained

[2] Evidence Briefing reporting the results of RR with Company 1: <https://bit.ly/2s0EP7O>
[3] Evidence Briefing reporting the results of RR with Company 2: <https://bit.ly/2Ghar0e>

through interviews after the workshop we presented and discussed the RRs results. The interviews were conducted in Portuguese, thus all quotes are free translations. The perceptions were grouped under positive and negative points. Table 23 presents the positive points. When presenting each positive point, "C1" stands for a point raised by participants of Company 1, "C2" stands for a point raised by participants of Company 2. If the two codes are present, the point was raised by the participants of both companies.

Table 23 – Positive points of RRs according to the participants of Company 1 and 2.

POSITIVE POINTS	COMPANY 1	COMPANY 2
Applicable to software engineering practice	X	X
Novel approach to support decision-making	X	X
Offer reliable content	X	X
Increase team confidence	X	X
Reduce time and cost for decision-making	X	X
Fast and easy way to find information	X	X
Interest in consuming Evidence Briefings regularly	X	X
Fostered the learning of new concepts	X	-
Problem-oriented	X	-
Improve problem comprehension	X	-
Avoid reading multiple sources	X	-
Flexible knowledge transfer medium	X	-
Non-applicable evidence can support other problems	X	-
Practitioners are willing to embrace rapid reviews, even considering the inherent limitations	-	X
The participation of practitioners on the rapid review is important	-	X

Applicable to software engineering practice (C1,C2): The participants considered RRs are applicable to SE practice. To illustrate, in Company 1, the project manager mentioned:

> *"[RRs] are very applicable, specially to institutes that conduct applied research in collaboration with industry [...] but I [the project manager] find it [RR] useful even for software factories."*

While, the projects coordinator believes they can apply:

> *"Just Demos, Change Priority, and Story Owner [strategies to improve customer collaboration identified with the RR] to our project."*

In Company 2, the IT director also believes RRs are applicable to SE practice, he mentions:

> *"I really believe this [RR] is very good to a company. It would be interesting if all companies do something like this. It should be part of the culture of any company to do this. I think it is very useful!"*

Novel approach to support decision-making (C1,C2): The participants mentioned they have never heard about such approach to support decision-making. For example, in Company 1, the projects coordinator told us they:

> *"usually we conduct research [in informal sources] to support decision-making, but never with the systematic you [the interviewer] have shown us."*

While the project manager mentioned he:

> *"never heard about anything similar [to RRs]."*

On Company 2, the IT director affirmed:

> *"When we have a problem, we go to the internet search for blogs and forums. But something [RR] created from scientific papers are better because it was certainly created more carefully."*

Offer reliable content (C1,C2): The participants considered the evidences provided by the RRs more reliable than the information they use to consume. For instance, in Company 1, the projects coordinator mentioned:

> *"it [RR] is an optimized way to find rich content, because usually scientific papers are rigorously analyzed to be published, which increases the confidence, [...] and avoid you to find false answers [on informal sources] [...] so, it is really useful [...] it transmits confidence since there are links to the primary studies."*

In Company 2, the IT director affirms:

> *"something [RR] created from scientific papers are better because it was certainly created more carefully. You can see several people already thought about this [motivation in software teams], they tried to solve this problem [...] You were able to cover more than 7000 papers. Thus, we are more confident there not much more things [strategies to deal with team motivation] out there."*

Increase team confidence (C1,C2): The participants mentioned RRs evidence gave them confidence about decisions they have taken or thought in the past. For instance, in Company 1, the projects manager said:

> *"this Change Priority [a strategy to improve customer collaboration identified with the RR] is something I have wondered before [but was not confident]. So, you realize that it worked [the strategy] and other people have used in practice, you see it would be the best solution."*

In Company 2, the IT director stated:

> *"we are more confident there are not much more things [strategies to deal with team motivation] out there. [...] There are things [findings of the RR] confirming our intuition. We implicitly believe it could be like that [the way to deal with low team motivation], and having evidence from scientific papers is more interesting. It confirms some of our thoughts, gives us courage to do it! [apply the strategies to deal with low team motivation]."*

Reduce time and cost for decision-making (C1,C2): The participants mentioned the RRs avoided the cost to define a solution from scratch and also reduced the time to decision-making. For example, on Company 1, the project manager told us:

> *"it is better to search for a solution someone else already proposed and validated in a paper [...] than reinvent the wheel".*

The projects coordinator mentioned:

> *"the main benefit is the lighter protocol [of RRs] [...] you can do it in one week. This is a good timing for us [...] why not applying a solution used by other people? You gain time. It [RR] helps you to optimize it [the decision-making process]."*

Fast and easy way to find information (C1,C2): The participants mentioned the time gained reading the EB when compared to a traditional research papers is high. For example, in Company 1, the project manager affirmed he:

> *"read it [the EB] in 15 minutes, everything condensed in just one page [...] It would take hours and hours of reading if we would read it direct from the papers [the primary studies]. Impressive the time we saved!"*

While the projects coordinator mentioned:

> *"the information is very clear, in topics, and the topics that need further information present that information."*

In Company 2, the project manager said:

> *"it is easy to find information on it [EB]. I like the format, the fact that it is summarized is nice because I can refer to the links [the primary studies] to read more. If I am not interest in everything, I can refer specifically to it [the primary study he is interested in]."*

While the IT director mentioned:

> *"the content is very succinct, but at the same time it offers a lot of information. For sure it is not easy to summarize that amount of papers in just one page. [...] The fact it is a one page document is interesting because you take a look and see it is just this you have to read. Maybe if it has too much pages I would just briefly screen the text. This way [with just one page] I could read it more attentively, aware of what I was reading. I really enjoyed the format. It has bullets, figures, tables..."*

Interest in consuming Evidence Briefings regularly (C1,C2): When we asked the participants if they would be interested in consuming EBs like the one we presented to them but about other SE topics, they answered positively. To illustrate, in Company 1, the projects coordinator suggested to do it *"like a newsletter."* While in Company 2, the IT director affirmed:

> *"If there are things like that [EBs] about each theme [SE topics], and a place where we can search for, it would be awesome!*

Recommend Rapid Reviews to other practitioners (C1,C2): When asked whether they recommend their peers to adopt RRs to better understand and mitigate a real problem they would be facing, the participants of both companies answered positively.

Fostered the learning of new concepts (C1): The participants of Company 1 affirm to have learned new concepts with the findings of the RR. For example, the projects coordinator told us he has never heard about:

> *"Story Owner, or Just Demos [strategies to improve customer collaboration identified in the RR]."*

Problem-oriented (C1): The participants of Company 1 reported that the RR is problem oriented, since it supports decision-making for a specific problem they are facing. For instance, the project manager mentioned:

> *"the information is condensed based on the problem we are trying to solve."*

While the projects coordinator affirmed:

> *"it [RR] is interesting, specially because the problem is well defined, so you can adopt the systematic process you have described in a very pragmatic way."*

Improve problem comprehension (C1): The participants of Company 1 noted the RR helped them to better comprehend and structure the problem their were facing. For instance, the project manager told us:

> *"it [the RR] helps to organize the ideas, what you want to search for."*

While the projects coordinator mentioned:

> *"it [the RR] makes you more confident to stratify the ideas that will help you to solve your problem, instead of conduct a blind search."*

Avoid reading multiple sources (C1): The participants of Company 1 considered RRs are more suitable as sources of knowledge than primary studies. The former aggregate evidence from various studies preventing them to take time reading multiple primary studies. This could be observed when the projects coordinator mentioned:

> *"it [the RR] is more straight-forward [compared to primary studies] and captures the important information of many studies more objectively."*

Flexible knowledge transfer medium (C1): The participants of Company 1 suggested other possible contexts to use EBs. The projects coordinator told us he:

> *"wants to buy the idea and use them [EBs] to build technical clippings."*

We then asked *"like internal and external communication? for instance if you want to diffuse good programming practices to enhance technical level of your teams, you can use the briefings to guide a discussion, is this it?"*. The project manager confirmed this intuition, and still, he stated:

> *"it [the RR] is really good to transfer knowledge."*

Non-applicable evidence can support other problems (C1): The participants of Company 1 also suggested that even evidence not direct applicable to their problem can support the decision-making of other problems they are facing in different projects. The projects coordinator mentioned:

> *"some information are not applicable [to Project 1] but almost all evidence could be applied [in different projects]."*

Practitioners are willing to embrace rapid reviews, even considering the inherent limitations (C2): In Company 2, the IT director affirmed:

> *"Sometimes we search for solutions in just one source [...] Then we do it exactly it was recommended by that source but it may not work for us. When we do it like this [the RR], we can have more possibilities [the strategies identified with the RR], even considering it was conducted faster [the rapid review compared to FSSs], and maybe many things [papers] could be lost just because of the title [the first round of selection procedure, which we analyzed only the titles of the papers], because someone put a bad title. That is ok, who cares?"*

The participation of practitioners on the rapid review is important (C2): The IT director affirmed:

> *"[..] exactly because we felt participating in the entire cycle, so we could understand why things are made like that. To understand the method, the strategies, I believe it is important."*

Table 24 shows the negative points. Similar to the positive points, when presenting a negative point, "C1" stands for a point raised by participants of Company 1, "C2" stands for a point raised by participants of Company 2, and if the two codes are present, the point was raised by the participants of both companies.

Table 24 – Negative points of RRs according to the participants of Company 1 and 2.

NEGATIVE POINTS	COMPANY 1	COMPANY 2
Discussing the rapid reviews' findings is needed	X	-
Primary studies context far from their findings	X	-
Evidence briefing printed in black-and-white	X	-
Absence of graphical information	X	-
Absence of researchers recommendation	-	X

Discussing the rapid reviews' findings is needed (C1): The participants of Company 1 mentioned that some findings of the RR only became clearer after the discussion during the workshop. For instance, the project manager declared:

> *"things I found difficult to assimilate [only reading the EB], I assimilated during the discussion [the workshop]"*

Primary studies context far from their findings (C1): The participants of Company 1 suggested that primary studies context should come near their findings, instead of grouped together with the context of all other primary studies on the first paragraph of the RR report, as made in the EB we created for Company 1. For instance, the projects coordinator mentioned:

> *"the first paragraph is very bureaucratic. I think maybe there are better ways to present the domain [context of the primary studies] which findings came from."*

When we asked what do they think if we have a paragraph briefly describing the context of a study, and immediately after, the findings related to that study, the projects coordinator answered:

> *"Yes, that is it, the context should come near to the findings."*

Evidence briefing printed in black-and-white (C1): The participants of Company 1 affirmed that some information was not highlighted as it should be, because the EB was printed in black-and-white. The projects coordinator thought:

> *"there were no links to contact you [the research group that created the EB] but then I saw the link on the right-side box, it is not much visible."*

Then the interviewer (researcher) asked if *"you [the projects coordinator] believe it is because the Evidence Briefing was printed in black-and-white and the right-side box that is highlight in gradient blue lost its highlight?"* The projects coordinator agreed.

Absence of graphical information (C1): The participants of Company 1 suggested that charts and figures would make the RR report even more appealing. For instance, the projects coordinator mentioned:

> *"if you could represent that evidence in a graphic way it would be even better."*

Absence of researcher recommendation (C2): The participants of Company 2 demanded a conclusion section in the EB, briefly pointing and discussing the strategy we (the researchers) recommend to deal with their problem. To illustrate, the IT director said:

> *"since it [the RR] was focused on our problem, maybe if there was something saying which one [strategy identified with the RR] you recommend [...] this is what is missing [...] maybe it is missing a conclusion, the researcher's comments."*

5.3.2 Perceptions on Adopting Rapid Reviews Evidence

To evaluate whether the knowledge introduced with the RR was applied, we conducted an interview with the participants of Company 1, two months after the introduction of the RR. The interview was in Portuguese, thus all quotes are free translations. We could not

have a follow-up interview with the participants of Company 2 because both quit their jobs.

What Worked: The participants mentioned the RR was positive when it comes to mitigating their main problem. They stated that they **used the Evidence Briefing as reference material** many times throughout the weeks after the introduction of the RR. To illustrate, the project manager said he *"analyzed it [the EB] after our discussion [the one during the workshop to introduce the RR] and we started to test some strategies."*

They also **used the Evidence Briefing to discuss with team members** how to properly deal with their problem. More importantly, the participants applied the knowledge provided by the RR. Among the strategies employed, they highlighted that Story Owner, Change Priority, and Risk Assessment Up Front were introduced in their work habits. In particular, the participants said they introduced the Story Owner as an additional role for one member of the development team. This member was then in charge to deal with the costumer, which alleviated the burden on the projects coordinator.

When asked whether RRs could be applied in other projects inside Company 1, the participants shared a positive enthusiasm. For instance, the project manager said *"the customer background varies a lot, but the problems with collaboration and communication are recurrent."*

Additionally, the participants mentioned the RR **motivated them to seek additional knowledge** that would help to cover corner cases. Despite the low customer collaboration in Project 1 — the participants mentioned that *"emails requesting clarification about requirements take one or two weeks for customer to reply"* — the customer used to demand prompt attention when the communication started from them. For instance, before the RR introduction, the participants often had endless/non-focused meetings with customers. Nonetheless, after the RR, they blocked an one hour meeting, and the meetings became more straight to the point. In summary, the participants mentioned the RR was useful to decrease the tension between company and costumer. They also mentioned that *"in the end, the customer was satisfied with the system we developed [...] he is thinking in contract us again in the future for the next phases of the project."*

What Did Not Work: Some evidence was not useful to the participants. For instance, the strategy to leverage **e-collaboration tools** was not necessary since the participants already use these kind of tools in Company 1. According to them *"[...] the problem lies on the quality of the communication [...] delays to answer the development team, and meetings that did not have focus and time limit."*

Similarly, although **Just Demos** was initially considered as an interesting strategy, turns out the participants were not able to implement it. The customers did not accept meetings only to present a release and make demos. They demanded for more intermediary meetings.

5.4 DISCUSSION

The participants of Company 1 and 2 perceived many positive points regarding the use of RRs. They found that RRs **offer reliable content**, which is crucial from the scientific point of view. Even though we acknowledge RRs may be not as reliable as FSSs, from the eyes of the participants, they are more reliable than expert opinion or informal sources, which they often use to support their decision-making. The participants also mentioned RRs are **problem-oriented**. This means that when the topic of a research is aligned to a practical problem, chances are the results are more useful and aligned to practitioners needs. The participants also affirmed that RRs have **fostered the learning of new concepts**, particularly relevant to deal with the problem they were facing. This feedback might encourage other researchers to conduct further RRs. The participants also mentioned RRs **reduce the time and cost** related to decision-making activities, when compared to the informal way they execute those kinds of activities. This finding corroborates with EBM literature, which states RRs deliver evidence in a timely manner. Still, the participants mentioned that RRs helped them to **improve their comprehension about the problem** they were facing. Discussing the problem with the researcher, as well as formalizing the RR protocol through explicit research questions, have sharpened the problem on practitioners perceptions. They also considered RRs can be **applicable in a practical settings**, and also they **recommend rapid reviews to other practitioners** as a source of information. In addition, **the participants are willing to embrace rapid reviews, even considering the inherent limitations**. It reveals practitioners are willing to take the risks related to the less rigorous methods of RRs. This is not a surprise since – in this research and others – we observed that practitioners are used to rely on knowledge from sources even less rigorous than RRs, such as blogs and forums (CARTAXO et al., 2016; CARTAXO; PINTO; SOARES, 2018a).

Despite the positive impressions, participants also highlighted the need **to discuss the rapid reviews findings** with the researchers who conducted the RR. This close connection helped to clarify unclear items while reading the EB. This finding endorses the claims of Grigoleit et al. which argue it is important to use artifact mediums (e.g. EBs) together with human-intensive mediums (e.g. face-to-face discussions) in order to effectively transfer knowledge (GRIGOLEIT et al., 2015).

Moreover, the participants considered positive the use of EBs to report RRs findings. Yet, the participants pointed it was **easy to find information** in the EBs, and also that, compared to traditional research papers, it takes considerably **less time to read**. They reported to need about 15 minutes to read the EB. This finding is particularly relevant to SE practice since practitioners are often under time pressure and are not used to consume SE research (CARTAXO et al., 2016). Despite the positive points, the participants also suggested some improvements to the EBs. For instance, the participants of Company 1 suggested that **information about the context of each finding should**

come near the finding itself. This reveals the importance of contextual information about the primary studies — like the number of cases investigated, the characteristics of participants, the countries which the primary studies were conducted, etc. Based o the lesson learned with the RR with Company 1, we fixed this situation and created the EB to Company 2 making the context of the primary studies nearer to the findings. No complains or comments arose with the participants of Company 2 about this point. It is important to highlight that, much have been discussed about context characterization in primary studies (DYBÅ; SJØBERG; CRUZES, 2012; CARTAXO et al., 2015a), but few on how to present the context of evidence in SSs. This is particularly challenging since findings of SSs come from diverse primary studies. Additionally, the participants of Company 1 affirmed that **Graphic information is needed** in EB. However, the challenge here is associated with the effort required to derive such graphical information, e.g., summarizing/synthesizing qualitative data from diverse primary studies in figures/charts. In the RR conducted with Company 2, we put tables and figures on the EB. However, they were extracted from individual primary studies, not figures or tables that synthesizes primary studies. This may point to future challenges on RRs.

Apart from the mostly positive perceptions, we also observed that the participants of Company 1, in fact, **applied the knowledge acquired from the rapid review** in practice. The usage of some findings of the RR led the practitioners to experience improvements on their problem. The participants also took advantage of the RR and **used the Evidence Briefing as reference material**. Similarly, they **used the evidence briefing to discuss evidence with the team**, which might suggest the benefits of RRs are not restricted to the first impressions in the workshop. Finally, the RR **motivated the team to seek additional knowledge** to solve their problem. Since the findings of the RR were condensed in a single one-page document, when practitioners did not find the evidence they needed (e.g., to deal with corner cases not covered by the RR), they had to figure out for themselves. This only occurred because the RR helped them to learn some of the fundamentals. As a shortcoming, not all the strategies identified in the RR were useful to their project, although practitioners mentioned they could apply some of the findings in other projects.

5.4.1 Asking the Right Questions

One of the most important characteristics of RRs is they should offer useful knowledge to support decision-making for a practical problem. Therefore, in order to provide useful answers, one has to ask the right questions. Based on the reflections made after and during the two RRs we conducted, we argue that RRs should ask at least two kinds of research questions.

Exploratory Questions: A certain type of question we believe is crucial to a RR is the

exploratory ones, in special aiming to identify strategies available in scientific literature that claims to solve or reduce practitioners problem. We consider this kind of question as the cornerstone of RRs, because the most important thing to practitioners under time constraints is to discovery strategies to solve their problems (YOURDON, 1995). RQ4 of the RR with Company 1, and RQ2 of the RR in Company 2 are examples of such kind research question (see Appendix E, Section **??**, and Appendix F, Section **??**). Other arbitrary examples are: "What are the strategies to improve software development team motivation?"; or "What are the strategies to introduce agile practices in a ad-hoc development team?"

Motivational Questions: Questions discussing why is important to solve a particular problem. In the presence of skeptical stakeholders, these questions provide evidence that could convince them. RQ1 and RQ2 (Section **??**) are examples of such questions. In the RR with Company 1, we asked questions to identify the benefits of solving practitioners problem and the shortcomings of not solving it. It was important for the participants of Company 1 to convince their customers to change their attitudes towards a better customer collaboration. One can also define more specific questions depending on the stakeholder interest. For instance, *"What are the benefits of unit testing on software quality?"* would be useful, for example, to convince a Software Quality Analyst (SQA) about the importance of adopting unit tests. More specific questions are aligned to Ali's observation (ALI, 2016). He identified that empirical SE research usually search for general benefits of proposed interventions. However, to support informed decision-making, we should comprise not just effectiveness, but also the evidence of cost-effectiveness. His claims consider only the cost dimension, but other dimensions can also be considered, such as quality, customer satisfaction, and security.

5.4.2 Implications

Research: Those interested in applied research can identify companies and projects facing problems related to their research topics and conduct RRs together with practitioners. This might help researchers to identify what kind of issues practitioners are facing and better guide their research. Moreover, researchers can promote RR tracks in SE conferences. We believe EBSE community as a whole can be benefited from regular RRs, ultimately, approximating SE practice and research, establishing fruitful collaborations, and enacting a common agenda.

Practice: Practitioners can use RRs to get a source of information based on evidence, beyond the informal ones they used to consume. If time and personnel are hard constrains, software companies can, for instance, commission researchers to conduct RRs to guide the definition of institutional company policies, as suggested by Kitchenham (KITCHENHAM;

CHARTERS, 2007); or even to guide technical clippings for discussions within the company, as suggested by the participants of Company 1.

5.5 CHAPTER SUMMARY

In this chapter we report the first experiences applying Rapid Reviews (RRs) in Software Engineering (SE) settings. RRs aim to support decision-making in SE practice, by omitting or simplifying the Full Secondary Studies (FSSs) process. We conducted and evaluated RRs in two software companies, both in Recife, Brazil, using action research like methods to evaluate the practitioners perceptions. Among the findings, we found that the majority of practitioners impressions were positive. They affirmed to have learned new concepts about the problem they were facing, as well as declared to trust in the findings provided with the RRs, even aware of the inherent limitations. According to them, the RR helped to reduce time and costs related to decision-making. The practitioners also suggested some improvements to the EBs, which are the documents used to present the findings of the RRs. In conclusion, we believe RRs can play an important role on promoting knowledge transfer from scientific empirical evidence to SE practice.

6 CONCLUDING REMARKS

In this chapter we conclude this thesis showing the main contributions, as well as why we believe this thesis is important to Software Engineering (SE) community.We also present directions we believe one can follow aiming to conduct further investigations on the topic of this thesis.

6.1 CONCLUSIONS

A new era of SE has emerged, and it is changing the way we think about empirical research in SE. In a recent series of posts at Communications of ACM blog, Bertrand Meyer (MEYER, 2018c; MEYER, 2018a; MEYER, 2018b) precisely framed this era throughout a vision where empirical evidence and practice orientation are pivotal elements:

> *"As long as empirical software engineering was a young, fledgling discipline, it made good sense to start with problems that naturally landed themselves to empirical investigation. But now that the field has matured, it may be time to reverse the perspective and start from the consumer's perspective: for practitioners of software engineering, what problems, not yet satisfactorily answered by software engineering theory, could benefit, in the search for answers, from empirical studies?"* (MEYER, 2018a)

Meyer's voice certainly is not alone. Many other researchers are starting to recognize practice orientation as the next long way ahead (BEECHAM et al., 2014; DUARTE, 2015; LAIRD; YANG, 2015; SANTOS; SILVA, 2013). Unfortunately, considering the results we obtained analyzing how Secondary Studies (SSs) cover practitioners issues posted in Stack Exchange websites, researchers are still scratching the surface of the problem. From 26,000+ issues reported on these websites, just 1.75% are related to the same SE areas of the 24 selected SSs. From that set of issues related the same SE areas as the selected SSs, only 15.6% are in fact covered by those SSs – when findings of a SS offer knowledge that can solve or at least attenuate the a practitioner issue.

Rapid Reviews (RRs) and alternative mediums like Evidence Briefings (EBs) are gaining attention from mature scientific fields like medicine and other health-care related fields. Due to the great efforts, studies have been continuously showing positive experiences regarding the use of RRs and EBs in medical settings. In this thesis we provide the first evidence of applying RRs and EBs in SE; Initial results suggest SE can also benefit from these approaches.

Regarding EBs, results revealed that most of the practitioners and researchers surveyed believe it is easy to find information on EBs (53% and 71% respectively); that EBs are

clear and understandable (82% and 71% respectively); and that EBs look reliable (62% and 56% respectively). We also invited authors of papers in four SE conferences (EASE, SBES, SBCARS, and SAST) to create EBs of their papers. Authors of 44 papers adhered and created EBs. Our results revealed that although it is time consuming to create EBs (88% of researchers took up to three hours), researchers still see benefits creating them, since 88% of the researchers affirmed they "Likely" or "Very Likely" are willing to create EBs for their papers.

In terms of RRs, practitioners from two software companies located in Recife, Brazil, reported that the RRs were important for them to learn new concepts, reduce time and cost of decision-making, improve their understanding about the problem they were facing, among other benefits.

Even looking for all these good results, one to be fair has to highlight that it is not always a bed of roses. RRs and EBs have their limitations, and this must be considered carefully. They are certainly not silver bullets nor are getting in to substitute Full Secondary Studies (FSSs) and the traditional research paper format.

In this scenario, and based on the mostly positive evidence we have obtained with six empirical studies, we believe the model we proposed to transfer knowledge to SE practice is a relevant contribution towards a SE closer to practice. However, as any SE approach, this one could not be seen as a silver bullet. Limitations are known and stakeholders interested in applying our model should consider them carefully.

6.2 FUTURE DIRECTIONS

This thesis is just the starting point of RRs and EBs in SE; there is much work ahead. Following, we present a list of directions naturally derived from this thesis:

- Conduct more RRs, create more EBs. Specially researchers from other research groups, which certainly are less biased than us;
- Compare the results of RRs with FSSs in SE. As soon as more RRs will be conducted, this comparison will become viable and significant.
- Investigate the perception of researchers about RRs in SE. In this thesis we focused just on practitioners perceptions because they are the most important stakeholders in a RR scenario, the consumers of evidence. However, to RRs thrive in real world, researchers have to be willing to conduct them. Actually, they have to recognize that RRs are a good choice of knowledge transfer in spite of their limitations;
- Directly measure the impacts of applying RRs in practical context. This would constitute important complementary evidence to the one we obtained by investigating the perceptions of practitioners on RRs. The former objectively measure how RRs

improved, or not, a current practical issue. On the other side, the latter reveals the subjective perceptions of practitioners on RRs, which are usually good predictors of adoption success of an initiative;

- Create more EBs templates focused on different audiences. The one we proposed targets practitioners, thus some information like methodological details are deliberately missing. However, in one of our empirical studies, researches who created EBs for their papers published in SE conferences affirmed they would like to have EBs in other SE conferences so they can also be consumers, not only practitioners. The idea is the EB to serve as a medium between the shorter abstract and the longer paper. For this to happen, the EBs template must be adapted to that different audience, more preoccupied with scientific method, for instance;

- Investigate how to graphically synthesize qualitative and quantitative evidence. Specially qualitative ones since quantitative has certainly more established methods. Practitioners from our action research to analyze RRs highlighted how important it is to have a simple and concise image summarizing the main results relevant to them;

- To have a preciser view of how well certain SE topic is covered by SSs, we believe more coverage analysis like the one we conducted with Stack Exchange issues should be conducted, but focusing in specific topics. Is the pair programming issues practitioners face in practice well covered by SSs? How about distributed teams? By answering these questions one can fine tune its FSSs to be more aligned with practice. The same can also be done with primary studies.

REFERENCES

ABOU-SETTA, A. M. et al. Methods for developing evidence reviews in short periods of time: A scoping review. *PloS one*, 2016.

ALI, N. b. Is effectiveness sufficient to choose an intervention?: Considering resource use in empirical software engineering. In: *Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*. New York, NY, USA: ACM, 2016. (ESEM '16), p. 54:1–54:6. ISBN 978-1-4503-4427-2. Available at: <http://doi.acm.org/10.1145/2961111.2962631>.

ALI, N. B.; PETERSEN, K. Evaluating strategies for study selection in systematic literature studies. In: *Proceedings of the 8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*. New York, NY, USA: ACM, 2014. (ESEM '14), p. 45:1–45:4. ISBN 978-1-4503-2774-9. Available at: <http://doi.acm.org/10.1145/2652524.2652557>.

AMBROSE, G.; HARRIS, P. *Basics Design 08: Design Thinking*. [S.l.]: AVA Publishing, 2009. (Basics Design). ISBN 978-2940411177.

ANDERSON, N. B. Evidence-based practice in psychology. *American Psychologist*, v. 61, n. 4, p. 271–285, 2006.

BADAMPUDI, D.; WOHLIN, C. Bayesian synthesis for knowledge translation in software engineering: Method and illustration. In: *2016 42th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*. [S.l.: s.n.], 2016. p. 148–156. ISSN 2376-9505.

BAMBRA, C. et al. Reducing health inequalities in priority public health conditions: using rapid review to develop proposals for evidence-based policy. *Journal of Public Health*, Oxford University Press, 2010.

BANO, M.; ZOWGHI, D.; IKRAM, N. Systematic reviews in requirements engineering: A tertiary study. In: *2014 IEEE 4th International Workshop on Empirical Requirements Engineering (EmpiRE)*. [S.l.: s.n.], 2014. p. 9–16. ISSN 2329-6348.

BASKERVILLE, R. L. Distinguishing action research from participative case studies. *Journal of Systems and Information Technology*, 1997.

BATTEN, J. *Comment on Editorial Literature Reviews as a Research Strategy*. 2012.

BAXTER, I. D.; YAHIN, A.; MOURA, L.; SANT'ANNA, M.; BIER, L. Clone detection using abstract syntax trees. In: *Proceedings of the International Conference on Software Maintenance*. [S.l.: s.n.], 1998. (ICSM '98), p. 368–. ISBN 0-8186-8779-7.

BEECHAM, S.; O'LEARY, P.; BAKER, S.; RICHARDSON, I.; NOLL, J. Making software engineering research relevant. *Computer*, IEEE Computer Society Press, Los Alamitos, CA, USA, v. 47, n. 4, p. 80–83, Apr. 2014. ISSN 0018-9162. Available at: <http://dx.doi.org/10.1109/MC.2014.92>.

BEGEL, A.; ZIMMERMANN, T. Analyze this! 145 questions for data scientists in software engineering. In: *36th ICSE*. [S.l.: s.n.], 2014. ISBN 978-1-4503-2756-5.

BERNARD, H. R.; KILLWORTH, P.; KRONENFELD, D.; SAILER, L. The problem of informant accuracy: The validity of retrospective data. *Annual review of anthropology*, JSTOR, p. 495–517, 1984.

BEST, L. et al. Rapid and responsive health technology assessment: The development and evaluation process in the south and west region of england. *Journal of Clinical Effectiveness*, 1997.

BJøRNSON, F. O.; DINGSøYR, T. Knowledge management in software engineering: A systematic review of studied concepts, findings and research methods used. *Information and Software Technology*, v. 50, n. 11, p. 1055 – 1068, 2008. ISSN 0950-5849. Available at: <http://www.sciencedirect.com/science/article/pii/S0950584908000487>.

BOOTH, W. C.; COLOMB, G. G.; WILLIAMS, J. M. *The craft of research*. [S.l.]: University of Chicago press, 2003.

BORGES, A.; FERREIRA, W.; BARREIROS, E.; ALMEIDA, A.; FONSECA, L.; TEIXEIRA, E.; SILVA, D.; ALENCAR, A.; SOARES, S. Support mechanisms to conduct empirical studies in software engineering. In: *Proceedings of the 8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*. New York, NY, USA: ACM, 2014. (ESEM '14), p. 50:1–50:4. ISBN 978-1-4503-2774-9. Available at: <http://doi.acm.org/10.1145/2652524.2652572>.

BORGES, A.; FERREIRA, W.; BARREIROS, E.; ALMEIDA, A.; FONSECA, L.; TEIXEIRA, E.; SILVA, D.; ALENCAR, A.; SOARES, S. Support mechanisms to conduct empirical studies in software engineering: A systematic mapping study. In: *Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering*. [S.l.: s.n.], 2015. (EASE '15), p. 22:1–22:14. ISBN 978-1-4503-3350-4.

BOUCHER, S. C. What is a philosophical stance? paradigms, policies and perspectives. *Synthese*, Springer, v. 191, n. 10, p. 2315–2332, 2014.

BOURQUE, P.; FAIRLEY, R. E. (Ed.). *SWEBOK: Guide to the Software Engineering Body of Knowledge*. Version 3.0. Los Alamitos, CA: IEEE Computer Society, 2014. ISBN 978-0-7695-5166-1. Available at: <http://www.swebok.org/>.

BOZEMAN, B. Technology transfer and public policy: a review of research and theory. *Research Policy*, v. 29, n. 4–5, p. 627 – 655, 2000.

BRIDGES, J.; MEYER, J. Exploring the effectiveness of action research as a tool for organizational change in health care. *Journal of Research in Nursing*, 2007.

BROCATO, D. Push and pull marketing strategies. *Wiley International Encyclopedia of Marketing*, Wiley Online Library, 2010.

BUDGEN, D.; BURN, A. J.; KITCHENHAM, B. Reporting computing projects through structured abstracts: a quasi-experiment. *Empirical Software Engineering*, v. 16, n. 2, p. 244–277, Apr 2011. ISSN 1573-7616. Available at: <https://doi.org/10.1007/s10664-010-9139-3>.

BUDGEN, D.; KITCHENHAM, B.; BRERETON, P. The case for knowledge translation. In: *ESEM*. [S.l.: s.n.], 2013.

BUDGEN, D.; KITCHENHAM, B.; CHARTERS, S. M.; TURNER, M.; BRERETON, P.; LINKMAN, S. G. Preliminary results of a study of the completeness and clarity of structured abstracts. In: *EASE*. [S.l.: s.n.], 2007.

BUDGEN, D.; KITCHENHAM, B. A.; CHARTERS, S. M.; TURNER, M.; BRERETON, P.; LINKMAN, S. G. Presenting software engineering results using structured abstracts: a randomised experiment. *Empirical Software Engineering*, v. 13, n. 4, p. 435–468, Aug 2008. ISSN 1573-7616. Available at: <https://doi.org/10.1007/s10664-008-9075-7>.

BURRELL, G.; MORGAN, G. *Sociological paradigms and organisational analysis: Elements of the sociology of corporate life*. [S.l.]: Routledge, 2017.

CAO, L.; RAMESH, B. Agile software development: Ad hoc practices or sound principles? *IT Professional*, v. 9, n. 2, p. 41–47, March 2007. ISSN 1520-9202.

CARTAXO, B. Integrating evidence from systematic reviews with software engineering practice through evidence briefings. In: *Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering*. New York, NY, USA: ACM, 2016. (EASE '16), p. 6:1–6:4. ISBN 978-1-4503-3691-8. Available at: <http://doi.acm.org/10.1145/2915970.2915973>.

CARTAXO, B. Supporting knowledge transfer from secondary studies to software engineering practice. In: *International Doctoral Symposium of Empirical Software Engineering (IDoESE 2017)*. [S.l.: s.n.], 2017.

CARTAXO, B. Supporting researchers to transfer knowledge from systematic reviews to software engineering practice. In: *Experimental Software Engineering Latin American Workshop (ESELAW 2017)*. [S.l.: s.n.], 2017.

CARTAXO, B.; ALMEIDA, A.; BARREIROS, E.; SARAIVA, J.; FERREIRA, W.; SOARES, S. Mechanisms to characterize context of empirical studies in software engineering. In: *Experimental Software Engineering Latin American Workshop (ESELAW 2015)*. [S.l.: s.n.], 2015. p. 1–14.

CARTAXO, B.; BORBA, P.; SOARES, S.; FUGIMOTO, H. Improving performance and maintainability of object cloning with lazy clones: An empirical evaluation. In: *2015 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*. [S.l.: s.n.], 2015. p. 1–8. ISSN 1949-3770.

CARTAXO, B.; PINTO, G.; RIBEIRO, D.; KAMEI, F.; SANTOS, R. E. S.; SILVA, F. Q. B. da; SOARES, S. Using q&a websites as a method for assessing systematic reviews. In: *2017 IEEE/ACM 14th International Conference on Mining Software Repositories (MSR)*. [S.l.: s.n.], 2017. p. 238–242.

CARTAXO, B.; PINTO, G.; SOARES, S. The role of rapid reviews supporting decision-making in software engineering practice. In: *22nd International Conference on Evaluation and Assessment in Software Engineering (EASE)*. [S.l.: s.n.], 2018.

CARTAXO, B.; PINTO, G.; SOARES, S. Towards a model to transfer knowledge from software engineering research to practice. *Information and Software Technology*, 2018. ISSN 0950-5849. Available at: <http://www.sciencedirect.com/science/article/pii/S0950584918300028>.

CARTAXO, B.; PINTO, G.; VIEIRA, E.; SOARES, S. Evidence briefings: Towards a medium to transfer knowledge from systematic reviews to practitioners. In: *Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*. New York, NY, USA: ACM, 2016. (ESEM '16), p. 57:1–57:10. ISBN 978-1-4503-4427-2. Available at: <http://doi.acm.org/10.1145/2961111.2962603>.

CARVER, J. C.; DIESTE, O.; KRAFT, N. A.; LO, D.; ZIMMERMANN, T. How practitioners perceive the relevance of esem research. In: *Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*. New York, NY, USA: ACM, 2016. (ESEM '16), p. 56:1–56:10. ISBN 978-1-4503-4427-2. Available at: <http://doi.acm.org/10.1145/2961111.2962597>.

CHAMBERS, D.; WILSON, P. A framework for production of systematic review based briefings to support evidence-informed decision-making. *Systematic reviews*, 2012.

COCHRANE Rapid Reviews Methods Group (RRMG). Http://methods.cochrane.org/rapidreviews/. Acessed in: Mar. 27, 2018.

CORABIAN, P.; HARSTALL, C. Rapid assessments provide acceptable quality advice. *Annu Meet Int Soc Technol Assess Health Care*, 2002.

CRESWELL, J. W. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications, 2009. ISBN 9781412965576. Available at: <http://books.google.com.br/books?id=bttwENORfhgC>.

CRUZES, D. S.; DYBå, T. Research synthesis in software engineering: A tertiary study. *Information and Software Technology*, v. 53, n. 5, p. 440 – 455, 2011. ISSN 0950-5849. Special Section on Best Papers from {XP2010}. Available at: <http://www.sciencedirect.com/science/article/pii/S095058491100005X>.

DAMIAN, D.; BORICI, A. Teamwork, coordination and customer relationship management skills: As important as technical skills in preparing our se graduates. In: *2012 First International Workshop on Software Engineering Education Based on Real-World Experiences (EduRex)*. [S.l.: s.n.], 2012. p. 37–40.

DAVIES, P. What is evidence-based education? *British journal of educational studies*, Wiley Online Library, v. 47, n. 2, p. 108–121, 1999.

DEVANBU, P.; ZIMMERMANN, T.; BIRD, C. Belief & evidence in empirical software engineering. In: *ICSE*. [S.l.: s.n.], 2016.

DEVEREAUX, P.; SCHÜNEMANN, H. J.; RAVINDRAN, N.; BHANDARI, M.; GARG, A. X.; CHOI, P. T.-L.; GRANT, B. J.; HAINES, T.; LACCHETTI, C.; WEAVER, B. et al. Comparison of mortality between private for-profit and private not-for-profit hemodialysis centers: a systematic review and meta-analysis. *Jama*, American Medical Association, v. 288, n. 19, p. 2449–2457, 2002.

DICENSO, A.; CULLUM, N.; CILISKA, D. Implementing evidence-based nursing: some misconceptions. *Evidence Based Nursing*, BMJ Publishing Group Ltd and RCN Publishing Company Ltd, v. 1, n. 2, p. 38–39, 1998.

DICK, B. Action research literature 2008—2010: Themes and trends. *Action Research*, 2011.

DIEBOLD, P.; VETRò, A. Bridging the gap: Se technology transfer into practice: Study design and preliminary results. In: *Proceedings of the 8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*. New York, NY, USA: ACM, 2014. (ESEM '14), p. 52:1–52:4. ISBN 978-1-4503-2774-9. Available at: <http://doi.acm.org/10.1145/2652524.2652552>.

DRURY, G. Opinion piece: Social media: Should marketers engage and how can it be done effectively? *Journal of Direct, Data and Digital Marketing Practice*, v. 9, n. 3, p. 274–277, 2008. ISSN 1746-0174. Available at: <http://dx.doi.org/10.1057/palgrave.dddmp.4350096>.

DUARTE, C. H. C. Patterns of cooperative technology development and transfer for software-engineering-in-the-large. In: *2015 IEEE/ACM 2nd International Workshop on Software Engineering Research and Industrial Practice*. [S.l.: s.n.], 2015. p. 32–38.

DUBINSKY, Y.; RAVID, S.; RAFAELI, A.; BAR-NAHOR, R. Governance mechanisms in global development environments. In: *2011 IEEE Sixth International Conference on Global Software Engineering*. [S.l.: s.n.], 2011. p. 6–14. ISSN 2329-6305.

DUBOIS, D. J.; TAMBURRELLI, G. Understanding gamification mechanisms for software development. In: *Proceedings of the 2013 9th Joint Meeting on Foundations of Software Engineering*. [S.l.: s.n.], 2013. (ESEC/FSE 2013), p. 659–662. ISBN 978-1-4503-2237-9.

DYBA, T. An empirical investigation of the key factors for success in software process improvement. *IEEE Trans. Softw. Eng.*, IEEE Press, Piscataway, NJ, USA, v. 31, n. 5, p. 410–424, May 2005. ISSN 0098-5589. Available at: <http://dx.doi.org/10.1109/TSE.2005.53>.

DYBÅ, T.; DINGSØYR, T. Empirical studies of agile software development: A systematic review. *Information and Software Technology*, v. 50, n. 9–10, p. 833 – 859, 2008. ISSN 0950-5849. Available at: <http://www.sciencedirect.com/science/article/pii/S0950584908000256>.

DYBÅ, T.; SJØBERG, D. I.; CRUZES, D. S. What works for whom, where, when, and why?: On the role of context in empirical software engineering. In: *Proceedings of the ACM-IEEE International Symposium on Empirical Software Engineering and Measurement*. New York, NY, USA: ACM, 2012. (ESEM '12), p. 19–28. ISBN 978-1-4503-1056-7. Available at: <http://doi.acm.org/10.1145/2372251.2372256>.

EASTERBROOK, S.; SINGER, J.; STOREY, M.-A.; DAMIAN, D. Selecting empirical methods for software engineering research. In: *Guide to advanced empirical software engineering*. [S.l.]: Springer, 2008. p. 285–311.

ELSEY, H.; LATHLEAN, J. Using action research to stimulate organisational change within health services: experiences from two community-based studies. *Educational Action Research*, 2006.

ETZKOWITZ, H.; WEBSTER, A.; GEBHARDT, C.; TERRA, B. R. C. The future of the university and the university of the future: evolution of ivory tower to entrepreneurial paradigm. *Research policy*, Elsevier, v. 29, n. 2, p. 313–330, 2000.

FARRINGTON, D. P.; MACKENZIE, D. L.; SHERMAN, L. W.; WELSH, B. C. et al. *Evidence-based crime prevention*. [S.l.]: Routledge, 2003.

FEATHERSTONE, R. M. et al. Advancing knowledge of rapid reviews: an analysis of results, conclusions and recommendations from published review articles examining rapid reviews. *Systematic reviews*, 2015.

FEREDAY, J.; MUIR-COCHRANE, E. Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, 2006.

FERREIRA, W.; BALDASSARRE, M. T.; SOARES, S.; CARTAXO, B.; VISAGGIO, G. A comparative study of model-driven approaches for scoping and planning experiments. In: *Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering*. New York, NY, USA: ACM, 2017. (EASE'17), p. 78–87. ISBN 978-1-4503-4804-1. Available at: <http://doi.acm.org/10.1145/3084226.3084258>.

FITNESSE - The fully integrated standalone wiki and acceptance testing framework. <http://www.fitnesse.org>. Acessed in: Nov. 4, 2016.

GANDER, J. P. The economics of university-industry research linkages. *Technological Forecasting and Social Change*, Elsevier, v. 29, n. 1, p. 33–49, 1986.

GAROUSI, V.; MäNTYLä, M. V. A systematic literature review of literature reviews in software testing. *Information and Software Technology*, v. 80, p. 195 – 216, 2016. ISSN 0950-5849. Available at: <http://www.sciencedirect.com/science/article/pii/S0950584916301446>.

GAROUSI, V. et al. Challenges and best practices in industry-academia collaborations in software engineering. *IST*, 2016.

GARRITTY, C. et al. Cochrane rapid reviews methods group to play a leading role in guiding the production of informed high-quality, timely research evidence syntheses. *Systematic reviews*, 2016.

GITTELL, J. H. An unexpected detour from ivory tower to action research. In: *Cultivating Creativity in Methodology and Research*. [S.l.]: Springer, 2018. p. 71–77.

GOLDER, S.; MACY, M. Social science with social media. *ASA footnotes*, v. 40, n. 1, p. 7, 2012.

GORSCHEK, T. How to increase the likelihood of successful transfer to industry – going beyond the empirical. In: *2015 IEEE/ACM 3rd International Workshop on Conducting Empirical Studies in Industry*. [S.l.: s.n.], 2015. p. 10–11.

GORSCHEK, T.; GARRE, P.; LARSSON, S.; WOHLIN, C. A model for technology transfer in practice. *IEEE Softw.*, IEEE Computer Society Press, Los Alamitos, CA, USA, v. 23, n. 6, p. 88–95, Nov. 2006. ISSN 0740-7459. Available at: <http://dx.doi.org/10.1109/MS.2006.147>.

GOULÃO, M.; AMARAL, V.; MERNIK, M. Quality in model-driven engineering: a tertiary study. *Software Quality Journal*, v. 24, n. 3, p. 601–633, 2016. ISSN 1573-1367. Available at: <http://dx.doi.org/10.1007/s11219-016-9324-8>.

GRAHAM, I. D.; LOGAN, J.; HARRISON, M. B.; STRAUS, S. E.; TETROE, J.; CASWELL, W.; ROBINSON, N. Lost in knowledge translation: Time for a map? *Journal of Continuing Education in the Health Professions*, Wiley Subscription Services, Inc., A Wiley Company, 2006.

GRIGOLEIT, F.; VETRO, A.; FERNANDEZ, D. M.; BOHM, W.; DIEBOLD, P. In quest for proper mediums for technology transfer in software engineering. In: *ESEM*. [S.l.: s.n.], 2015.

HAILEY, D. A preliminary survey on the influence of rapid health technology assessments. *International journal of technology assessment in health care*, Cambridge Univ Press, 2009.

HAILEY, D. et al. The use and impact of rapid health technology assessments. *International journal of technology assessment in health care*, Cambridge Univ Press, 2000.

HANNAY, J. E.; DYBå, T.; ARISHOLM, E.; SJøBERG, D. I. The effectiveness of pair programming: A meta-analysis. *Information and Software Technology*, v. 51, n. 7, p. 1110 – 1122, 2009. ISSN 0950-5849. Special Section: Software Engineering for Secure SystemsSoftware Engineering for Secure Systems. Available at: <http://www.sciencedirect.com/science/article/pii/S0950584909000123>.

HARTLING, L. et al. A taxonomy of rapid reviews links report types and methods to specific decision-making contexts. *Journal of Clinical Epidemiology*, 2015.

HASSLER, E.; CARVER, J. C.; KRAFT, N. A.; HALE, D. Outcomes of a community workshop to identify and rank barriers to the systematic literature review process. In: *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*. New York, NY, USA: ACM, 2014. (EASE '14), p. 31:1–31:10. ISBN 978-1-4503-2476-2. Available at: <http://doi.acm.org/10.1145/2601248.2601274>.

HAUGSET, B.; HANSSEN, G. K. Automated acceptance testing: A literature review and an industrial case study. In: IEEE. *Agile, 2008. AGILE'08. Conference*. [S.l.], 2008. p. 27–38.

HILKKA, M.-R.; TUURE, T.; ROSSI, M. Is extreme programming just old wine in new bottles: A comparison of two cases. *Journal of Database Management*, IGI Global, v. 16, n. 4, p. 41, 2005.

HODA, R.; NOBLE, J.; MARSHALL, S. The impact of inadequate customer collaboration on self-organizing agile teams. *Information and Software Technology*, v. 53, n. 5, p. 521 – 534, 2011. ISSN 0950-5849. Special Section on Best Papers from {XP2010}. Available at: <http://www.sciencedirect.com/science/article/pii/S0950584910001941>.

HODA, R. et al. Agile undercover: When customers don't collaborate. In: _____. *XP Conference*. [S.l.: s.n.], 2010.

HODA, R.; SALLEH, N.; GRUNDY, J.; TEE, H. M. Systematic literature reviews in agile software development: A tertiary study. *Information and Software Technology*, p. –, 2017. ISSN 0950-5849. Available at: <http://www.sciencedirect.com/science/article/pii/S0950584917300538>.

HOLDEN, M. T.; LYNCH, P. Choosing the appropriate methodology: Understanding research philosophy. *The marketing review*, Westburn Publishers Ltd, v. 4, n. 4, p. 397–409, 2004.

HORDIJK, W.; PONISIO, M. L.; WIERINGA, R. Harmfulness of code duplication: A structured review of the evidence. In: *Proceedings of the 13th International Conference on Evaluation and Assessment in Software Engineering*. Swinton, UK, UK: British Computer Society, 2009. (EASE'09), p. 88–97. Available at: <http://dl.acm.org/citation.cfm?id=2227040.2227051>.

HOSSAIN, E.; BABAR, M. A.; PAIK, H. y. Using scrum in global software development: A systematic literature review. In: *2009 Fourth IEEE International Conference on Global Software Engineering*. [S.l.: s.n.], 2009. p. 175–184. ISSN 2329-6305.

HUGHES, J. A.; SHARROCK, W. W. *The philosophy of social research*. [S.l.]: Longman London, 1997.

HUNOLD, S.; KRELLNER, B.; RAUBER, T.; REICHEL, T.; RÜNGER, G. Pattern-based refactoring of legacy software systems. In: SPRINGER. *International Conference on Enterprise Information Systems*. [S.l.], 2009. p. 78–89.

INSFRAN, E.; FERNANDEZ, A. A systematic review of usability evaluation in web development. In: _____. *Web Information Systems Engineering – WISE 2008 Workshops: WISE 2008 International Workshops, Auckland, New Zealand, September 1-4, 2008. Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. p. 81–91. ISBN 978-3-540-85200-1. Available at: <http://dx.doi.org/10.1007/978-3-540-85200-1_10>.

IVARSSON, M.; GORSCHEK, T. Technology transfer decision support in requirements engineering research: a systematic review of rej. *Requirements Engineering*, 2009. ISSN 0947-3602.

JAHANGIRIAN, M. et al. A rapid review method for extremely large corpora of literature: Applications to the domains of modelling, simulation, and management. *International Journal of Information Management*, Elsevier, 2011.

JAMES, W. *Pragmatism, a New Name for Some Old Ways of Thinking, Popular Lectures*. [S.l.: s.n.], 1907.

JEDLITSCHKA, A.; JURISTO, N.; ROMBACH, D. Reporting experiments to satisfy professionals' information needs. *Empirical Softw. Engg.*, Kluwer Academic Publishers, Hingham, MA, USA, v. 19, n. 6, p. 1921–1955, Dec. 2014. ISSN 1382-3256. Available at: <http://dx.doi.org/10.1007/s10664-013-9268-6>.

JIMÉNEZ, M.; PIATTINI, M.; VIZCAÍNO, A. Challenges and improvements in distributed software development: A systematic review. *Adv. Soft. Eng.*, Hindawi Publishing Corp., New York, NY, United States, v. 2009, p. 3:1–3:16, Jan. 2009. ISSN 1687-8655. Available at: <http://dx.doi.org/10.1155/2009/710971>.

JORGENSEN, M. Evidence-based guidelines for assessment of software development cost uncertainty. *Software Engineering, IEEE Transactions on*, 2005.

JøRGENSEN, M. Forecasting of software development work effort: Evidence on expert judgement and formal models. *International Journal of Forecasting*, v. 23, n. 3, p. 449 – 462, 2007. ISSN 0169-2070. Available at: <http://www.sciencedirect.com/science/article/pii/S016920700700074X>.

KAMEI, F.; PINTO, G.; CARTAXO, B.; VASCONCELOS, A. On the benefits/limitations of agile software development: An interview study with brazilian companies. In: *Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering*. New York, NY, USA: ACM, 2017. (EASE'17), p. 154–159. ISBN 978-1-4503-4804-1. Available at: <http://doi.acm.org/10.1145/3084226.3084278>.

KAVALER, D.; POSNETT, D.; GIBLER, C.; CHEN, H.; DEVANBU, P.; FILKOV, V. Using and asking: Apis used in the android market and asked about in stackoverflow. In: *Proceedings of the 5th International Conference on Social Informatics - Volume 8238*. New York, NY, USA: Springer-Verlag New York, Inc., 2013. (SocInfo 2013), p. 405–418. ISBN 978-3-319-03259-7. Available at: <http://dx.doi.org/10.1007/978-3-319-03260-3_35>.

KELLY, S. E.; MOHER, D.; CLIFFORD, T. J. Defining rapid reviews: a modified delphi consensus approach. *International journal of technology assessment in health care*, Cambridge University Press, v. 32, n. 4, p. 265–275, 2016.

KELLY, S. E.; MOHER, D.; CLIFFORD, T. J. Expediting evidence synthesis for healthcare decision-making: exploring attitudes and perceptions towards rapid reviews using q methodology. *PeerJ*, v. 4, p. e2522, Oct. 2016. ISSN 2167-8359. Available at: <https://doi.org/10.7717/peerj.2522>.

KHAN, S.; NIAZI, M.; AHMAD, R. Critical barriers for offshore software development outsourcing vendors: A systematic literature review. In: *APSEC*. [S.l.: s.n.], 2009. ISSN 1530-1362.

KHAN, S.; NIAZI, M.; AHMAD, R. Critical success factors for offshore software development outsourcing vendors: A systematic literature review. In: *ICGSE*. [S.l.: s.n.], 2009.

KHANGURA, S. et al. Evidence summaries: the evolution of a rapid review approach. *Systematic Reviews*, 2012.

KITCHENHAM, B.; BRERETON, O. P.; BUDGEN, D.; TURNER, M.; BAILEY, J.; LINKMAN, S. Systematic literature reviews in software engineering - a systematic literature review. *Inf. Softw. Technol.*, Butterworth-Heinemann, Newton, MA, USA, v. 51, n. 1, p. 7–15, Jan. 2009. ISSN 0950-5849. Available at: <http://dx.doi.org/10.1016/j.infsof.2008.09.009>.

KITCHENHAM, B.; CHARTERS, S. *Guidelines for performing Systematic Literature Reviews in Software Engineering*. 2007.

KITCHENHAM, B.; MENDES, E.; TRAVASSOS, G. H. A systematic review of cross- vs. within- company cost estimation studies. In: *10th EASE*. [S.l.: s.n.], 2006.

KITCHENHAM, B.; PRETORIUS, R.; BUDGEN, D.; BRERETON, O. P.; TURNER, M.; NIAZI, M.; LINKMAN, S. Systematic literature reviews in software engineering – a tertiary study. *Information and Software Technology*, v. 52, n. 8, p. 792 – 805,

2010. ISSN 0950-5849. Available at: <http://www.sciencedirect.com/science/article/pii/S0950584910000467>.

KITCHENHAM, B. A.; DYBA, T.; JORGENSEN, M. Evidence-based software engineering. In: *Proceedings of the 26th International Conference on Software Engineering*. Washington, DC, USA: IEEE Computer Society, 2004. (ICSE '04), p. 273–281. ISBN 0-7695-2163-0. Available at: <http://dl.acm.org/citation.cfm?id=998675.999432>.

KLÜNDER, J.; SCHNEIDER, K.; KORTUM, F.; STRAUBE, J.; HANDKE, L.; KAUFFELD, S. Communication in teams - an expression of social conflicts. In: _____. *Human-Centered and Error-Resilient Systems Development: IFIP WG 13.2/13.5 Joint Working Conference, 6th International Conference on Human-Centered Software Engineering, HCSE 2016, and 8th International Conference on Human Error, Safety, and System Development, HESSD 2016, Stockholm, Sweden, August 29-31, 2016, Proceedings*. Cham: Springer International Publishing, 2016. p. 111–129. ISBN 978-3-319-44902-9. Available at: <http://dx.doi.org/10.1007/978-3-319-44902-9_8>.

LAIRD, L.; YANG, Y. Transferring software engineering research into industry: The stevens way. In: *2015 IEEE/ACM 2nd International Workshop on Software Engineering Research and Industrial Practice*. [S.l.: s.n.], 2015. p. 46–49.

LAVIS, J. N.; ROBERTSON, D.; WOODSIDE, J. M.; MCLEOD, C. B.; ABELSON, J. How can research organizations more effectively transfer research knowledge to decision makers? *Milbank quarterly*, Wiley Online Library, v. 81, n. 2, p. 221–248, 2003.

LAWANI, M. A. et al. Five shared decision-making tools in 5 months: use of rapid reviews to develop decision boxes for seniors living with dementia and their caregivers. *Systematic reviews*, 2017.

LOPEZ, A.; NICOLAS, J.; TOVAL, A. Risks and safeguards for the requirements engineering process in global software development. In: *4th ICGSE*. [S.l.: s.n.], 2009.

LUPTON, E.; PHILLIPS, J. C. *Graphic Design: The New Basics*. 2nd edition. ed. [S.l.]: Princeton Architectural Press, 2015.

LURIA, S.; LURIA, Z. The role of the university: Ivory tower, service station, or frontier post? *Daedalus*, JSTOR, p. 75–83, 1970.

MAMYKINA, L.; MANOIM, B.; MITTAL, M.; HRIPCSAK, G.; HARTMANN, B. Design lessons from the fastest q&a site in the west. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM, 2011. (CHI '11), p. 2857–2866. ISBN 978-1-4503-0228-9. Available at: <http://doi.acm.org/10.1145/1978942.1979366>.

MANIFESTO for Agile Software Development. <http://agilemanifesto.org>. Acessed in: Feb. 1, 2017.

MARQUES, A. B.; RODRIGUES, R.; CONTE, T. Systematic literature reviews in distributed software development: A tertiary study. In: *2012 IEEE Seventh International Conference on Global Software Engineering*. [S.l.: s.n.], 2012. p. 134–143. ISSN 2329-6305.

MARSHALL, C. et al. Tools to support systematic reviews in software engineering: A feature analysis. In: *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*. New York, NY, USA: ACM, 2014. (EASE '14), p. 13:1–13:10. ISBN 978-1-4503-2476-2. Available at: <http://doi.acm.org/10.1145/2601248.2601270>.

MCGREGOR, M.; BROPHY, J. M. End-user involvement in health technology assessment (hta) development: a way to increase impact. *International Journal of Technology Assessment in Health Care*, 2005.

MCKIBBON, K. Evidence-based practice. *Bulletin of the Medical Library Association*, Medical Library Association, v. 86, n. 3, p. 396, 1998.

MERTENS, D. M. *What comes first? The paradigm or the approach?* [S.l.]: SAGE Publications Sage CA: Los Angeles, CA, 2012.

MEYER, B. Blog, *Empirical Answers to Important Software Engineering Questions (part 1 of 2)*. 2018. <https://cacm.acm.org/blogs/blog-cacm/224351-empirical-answers-to-important-software-engineering-questions-part-1-of-2/fulltext>.

MEYER, B. Blog, *Empirical Answers to Important Software Engineering Questions (part 2 of 2)*. 2018. <https://cacm.acm.org/blogs/blog-cacm/224677-empirical-answers-to-important-software-engineering-questions-part-2-of-2/fulltext>.

MEYER, B. Blog, *The End of Software Engineering and the Last Methodologist*. 2018. <https://cacm.acm.org/blogs/blog-cacm/224352-the-end-of-software-engineering-and-the-last-methodologist/fulltext>.

MIKKONEN, T.; LASSENIUS, C.; MäNNISTö, T.; OIVO, M.; JäRVINEN, J. Continuous and collaborative technology transfer: Software engineering research with real-time industry impact. *Information and Software Technology*, v. 95, p. 34 – 45, 2018. ISSN 0950-5849. Available at: <http://www.sciencedirect.com/science/article/pii/S0950584917304007>.

MOHER, D. et al. All in the family: systematic reviews, rapid reviews, scoping reviews, realist reviews, and more. *Systematic Reviews*, 2015.

MORRISON, P.; MURPHY-HILL, E. Is programming knowledge related to age? an exploration of stack overflow. In: *2013 10th Working Conference on Mining Software Repositories (MSR)*. [S.l.: s.n.], 2013. p. 69–72. ISSN 2160-1852.

MURTHY, D. Towards a sociological understanding of social media: Theorizing twitter. *Sociology*, Sage Publications, v. 46, n. 6, p. 1059–1073, 2012.

NASEHI, S. M.; SILLITO, J.; MAURER, F.; BURNS, C. What makes a good code example?: A study of programming q amp;a in stackoverflow. In: *2012 28th IEEE International Conference on Software Maintenance (ICSM)*. [S.l.: s.n.], 2012. p. 25–34. ISSN 1063-6773.

NEWMAN, I.; BENZ, C. *Qualitative-quantitative Research Methodology: Exploring the Interactive Continuum*. Southern Illinois University Press, 1998. ISBN 9780809321506. Available at: <http://books.google.com.br/books?id=xumf1ABFz8cC>.

NICOLáS, J.; TOVAL, A. On the generation of requirements specifications from software engineering models: A systematic literature review. *Information and Software Technology*, v. 51, n. 9, p. 1291 – 1307, 2009. ISSN 0950-5849. Available at: <http://www.sciencedirect.com/science/article/pii/S0950584909000378>.

OSTERWEIL, L. J.; GHEZZI, C.; KRAMER, J.; WOLF, A. L. Determining the impact of software engineering research on practice. *Computer*, v. 41, n. 3, p. 39–49, March 2008.

PAPAZOGLOU, M. P. Service-oriented computing: concepts, characteristics and directions. In: *Proceedings of the Fourth International Conference on Web Information Systems Engineering, 2003. WISE 2003.* [S.l.: s.n.], 2003. p. 3–12.

PETERS, M.; SCHNEIDER, M.; GRIESSHABER, T.; HOFFMANN, V. H. The impact of technology-push and demand-pull policies on technical change–does the locus of policies matter? *Research Policy*, Elsevier, v. 41, n. 8, p. 1296–1308, 2012.

PFLEEGER, S. L. Understanding and improving technology transfer in software engineering. *J. Syst. Softw.*, Elsevier Science Inc., New York, NY, USA, v. 47, n. 2-3, p. 111–124, Jul. 1999. ISSN 0164-1212. Available at: <http://dx.doi.org/10.1016/S0164-1212(99)00031-X>.

PINTO, G.; CASTOR, F.; LIU, Y. D. Mining questions about software energy consumption. In: *Proceedings of the 11th Working Conference on Mining Software Repositories*. New York, NY, USA: ACM, 2014. (MSR 2014), p. 22–31. ISBN 978-1-4503-2863-0. Available at: <http://doi.acm.org/10.1145/2597073.2597110>.

PINTO, G. H.; KAMEI, F. What programmers say about refactoring tools?: An empirical investigation of stack overflow. In: *Proceedings of the 2013 ACM Workshop on Workshop on Refactoring Tools*. New York, NY, USA: ACM, 2013. (WRT '13), p. 33–36. ISBN 978-1-4503-2604-9. Available at: <http://doi.acm.org/10.1145/2541348.2541357>.

PIPER, W. S.; NAGHSHPOUR, S. Government technology transfer: the effective use of both push and pull marketing strategies. *International Journal of Technology Management*, Inderscience Publishers, v. 12, n. 1, p. 85–94, 1996.

POLISENA, J. et al. Rapid review summit: an overview and initiation of a research agenda. *Systematic reviews*, 2015.

POSNETT, D.; WARBURG, E.; DEVANBU, P.; FILKOV, V. Mining stack exchange: Expertise is evident from initial contributions. In: *2012 International Conference on Social Informatics*. [S.l.: s.n.], 2012. p. 199–204.

PRECHELT, L.; ZIERIS, F.; SCHMEISKY, H. Difficulty factors of obtaining access for empirical studies in industry. In: *Proceedings of the Third International Workshop on Conducting Empirical Studies in Industry*. Piscataway, NJ, USA: IEEE Press, 2015. (CESI '15), p. 19–25. Available at: <http://dl.acm.org/citation.cfm?id=2819303.2819312>.

RADDON, A. Early stage research training: Epistemology & ontology in social science research. *Generic Skills Training for Research Students*, 2010.

RAINER, A. et al. Persuading developers to 'buy into' software process improvement: Local opinion and empirical evidence. In: *Proceedings of the 2003 International Symposium on Empirical Software Engineering*. [S.l.: s.n.], 2003. (ISESE '03).

RALPH, P.; CHIASSON, M.; KELLEY, H. Social theory for software engineering research. In: *Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering*. New York, NY, USA: ACM, 2016. (EASE '16), p. 44:1–44:11. ISBN 978-1-4503-3691-8. Available at: <http://doi.acm.org/10.1145/2915970.2915998>.

REBOU, M.; PINTO, G.; EBERT, F.; TORRES, W.; SEREBRENIK, A.; CASTOR, F. et al. An empirical study on the usage of the swift programming language. In: IEEE. *2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER)*. [S.l.], 2016. v. 1, p. 634–638.

RIAZ, M.; MENDES, E.; TEMPERO, E. A systematic review of software maintainability prediction and metrics. In: *Proceedings of the 2009 3rd International Symposium on Empirical Software Engineering and Measurement*. Washington, DC, USA: IEEE Computer Society, 2009. (ESEM '09), p. 367–377. ISBN 978-1-4244-4842-5. Available at: <http://dx.doi.org/10.1109/ESEM.2009.5314233>.

ROBILLARD, P. N. The role of knowledge in software development. *Commun. ACM*, ACM, New York, NY, USA, v. 42, n. 1, p. 87–92, Jan. 1999. ISSN 0001-0782. Available at: <http://doi.acm.org/10.1145/291469.291476>.

ROBINSON, H. et al. Collaboration, communication and co-ordination in agile software development practice. In: _____. *Collaborative Software Engineering*. [S.l.: s.n.], 2010.

ROGERS, E. *Diffusion of Innovations, 5th Edition*. Free Press, 2003. ISBN 9780743258234. Available at: <https://books.google.com.br/books?id=9U1K5LjUOwEC>.

SANTOS, A. C. C. d.; JUNIOR, I. H. d. F.; MOURA, H. P. d.; MARCZAK, S. A systematic tertiary study of communication in distributed software development projects. In: *2012 IEEE Seventh International Conference on Global Software Engineering*. [S.l.: s.n.], 2012. p. 182–182. ISSN 2329-6305.

SANTOS, P. S. M. D. et al. Action research can swing the balance in experimental software engineering. *Advances in computers*, 2011.

SANTOS, R. E. S.; SILVA, F. Q. B. d. Motivation to perform systematic reviews and their impact on software engineering practice. In: *2013 ACM / IEEE International Symposium on Empirical Software Engineering and Measurement*. [S.l.: s.n.], 2013. p. 292–295. ISSN 1949-3770.

SANTOS, R. E. S.; SILVA, F. Q. B. da; aES, C. V. C. de M.; MONTEIRO, C. V. F. Building a theory of job rotation in software engineering from an instrumental case study. In: *Proceedings of the 38th International Conference on Software Engineering*. New York, NY, USA: ACM, 2016. (ICSE '16), p. 971–981. ISBN 978-1-4503-3900-1. Available at: <http://doi.acm.org/10.1145/2884781.2884837>.

SARAIVA, J.; BARREIROS, E.; ALMEIDA, A.; LIMA, F.; ALENCAR, A.; LIMA, G.; SOARES, S.; CASTOR, F. Aspect-oriented software maintenance metrics: A systematic mapping study. In: *16th International Conference on Evaluation Assessment in Software Engineering (EASE 2012)*. [S.l.: s.n.], 2012. p. 253–262.

SAUNDERS, M.; LEWIS, P.; THORNHILL, A. Understanding research philosophies and approaches. *Research methods for business students*, Pearson London, v. 4, p. 106–135, 2009.

SCHERER, F. M. Demand-pull and technological invention: Schmookler revisted. *The Journal of Industrial Economics*, Wiley, v. 30, n. 3, p. 225–237, 1982. ISSN 00221821, 14676451. Available at: <http://www.jstor.org/stable/2098216>.

SHANNON-BAKER, P. Making paradigms meaningful in mixed methods research. *Journal of Mixed Methods Research*, SAGE Publications Sage CA: Los Angeles, CA, v. 10, n. 4, p. 319–334, 2016.

SHAPIN, S. The ivory tower: the history of a figure of speech and its cultural uses. *The British Journal for the History of Science*, Cambridge University Press, v. 45, n. 1, p. 1–27, 2012.

SHARP, H. et al. The role of ethnographic studies in empirical software engineering. *IEEE TSE*, 2016.

SILVA, F. Q. B. da; SANTOS, A. L. M.; SOARES, S. C. B.; FRANÇA, A. C. C.; MONTEIRO, C. V. F. A critical appraisal of systematic reviews in software engineering from the perspective of the research questions asked in the reviews. In: *Proceedings of the 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement*. New York, NY, USA: ACM, 2010. (ESEM '10), p. 33:1–33:4. ISBN 978-1-4503-0039-1. Available at: <http://doi.acm.org/10.1145/1852786.1852830>.

SILVA, F. Q. da; SANTOS, A. L.; SOARES, S.; FRANçA, A. C. C.; MONTEIRO, C. V.; MACIEL, F. F. Six years of systematic literature reviews in software engineering: An updated tertiary study. *Information and Software Technology*, v. 53, n. 9, p. 899 – 913, 2011. ISSN 0950-5849. Studying work practices in Global Software Engineering. Available at: <http://www.sciencedirect.com/science/article/pii/S0950584911001017>.

SINGER, J.; SIM, S. E.; LETHBRIDGE, T. C. Software engineering data collection for field studies. In: *Guide to Advanced Empirical Software Engineering*. [S.l.]: Springer, 2008.

SKEVOULIS, S. Engineering a successful partnership between academia and the financial industry: A software engineering program for it professionals. In: *2011 24th IEEE-CS Conference on Software Engineering Education and Training (CSEE T)*. [S.l.: s.n.], 2011. p. 343–350. ISSN 1093-0175.

SMITH, E.; LOFTIN, R.; MURPHY-HILL, E.; BIRD, C.; ZIMMERMANN, T. Improving developer participation rates in surveys. In: *CHASE*. [S.l.: s.n.], 2013.

SPENCER, D. *Card sorting: Designing usable categories*. [S.l.]: Rosenfeld Media, 2009.

SPIVEY, W. A.; MUNSON, J. M.; NELSON, M. A.; DIETRICH, G. B. Coordinating the technology transfer and transition of information technology: A phenomenological perspective. *IEEE Transactions on engineering management*, IEEE, v. 44, n. 4, p. 359–366, 1997.

STACK Overflow Search – Now 81% Less Crappy. 2011. <https://stackoverflow.blog/2011/01/stack-overflow-search-now-81-less-crappy/>. Acessed in: Nov. 4, 2016.

STACKEXCHANGE Programmers Community. <http://programmers.stackexchange.com>. Acessed in: Nov. 4, 2016.

STACKEXCHANGE Project Management Community. <http://pm.stackexchange.com>. Acessed in: Nov. 4, 2016.

STACKEXCHANGE Reverse Engineering Community. <http://reverseengineering.stackexchange.com>. Acessed in: Nov. 4, 2016.

STACKEXCHANGE Software Quality & Testing Community. <http://sqa.stackexchange.com>. Acessed in: Nov. 4, 2016.

STACKEXCHANGE Software Recommendations Community. <http://softwarerecs.stackexchange.com>. Acessed in: Nov. 4, 2016.

STACKEXCHANGE Survey 2016. <http://stackoverflow.com/research/developer-survey-2016>. Acessed in: Nov. 4, 2016.

STACKEXCHANGE Tour. <http://stackexchange.com/tour>. Acessed in: Nov. 4, 2016.

STEELE, F. *Consulting for organizational change*. [S.l.]: Univ of Massachusetts Press, 1975.

STOL, K.-J.; RALPH, P.; FITZGERALD, B. Grounded theory in software engineering research: A critical review and guidelines. In: *Proceedings of the 38th International Conference on Software Engineering*. New York, NY, USA: ACM, 2016. (ICSE '16), p. 120–131. ISBN 978-1-4503-3900-1. Available at: <http://doi.acm.org/10.1145/2884781.2884833>.

STOREY, M.-A.; ENGSTRÖM, E.; HÖST, M.; RUNESON, P.; BJARNASON, E. Using a visual abstract as a lens for communicating and promoting design science research in software engineering. In: IEEE–INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS INC. *11th International Symposium on Empirical Software Engineering and Measurement (ESEM), 2017*. [S.l.], 2017.

STRAUS, S. E.; TETROE, J.; GRAHAM, I. Defining knowledge translation. *Canadian Medical Association Journal*, Can Med Assoc, v. 181, n. 3-4, p. 165–168, 2009.

STRINGER, E. T. *Action research*. [S.l.]: Sage Publications, 2013.

SYKES, E. R. Interruptions in the workplace: A case study to reduce their effects. *International Journal of Information Management*, v. 31, n. 4, p. 385 – 394, 2011. ISSN 0268-4012. Available at: <http://www.sciencedirect.com/science/article/pii/S0268401210001568>.

TAYLOR-PHILLIPS, S. et al. Comparison of a full systematic review versus rapid review approaches to assess a newborn screening test for tyrosinemia type 1. *Research Synthesis Methods*, 2017.

TONDREAU, B. *Layout Essentials: 100 Design Principles for Using Grids (Design Essentials)*. [S.l.]: Rockport Publishers, 2011. (Design Essentials). ISBN 978-1592537075.

TRENDOWICZ, A.; MüNCH, J. Chapter 6 factors influencing software development productivity—state-of-the-art and industrial experiences. In: . Elsevier, 2009, (Advances in Computers, v. 77). p. 185 – 241. Available at: <http://www.sciencedirect.com/science/article/pii/S0065245809012066>.

TREUDE, C.; BARZILAY, O.; STOREY, M.-A. How do programmers ask and answer questions on the web? (nier track). In: *Proceedings of the 33rd International Conference on Software Engineering*. New York, NY, USA: ACM, 2011. (ICSE '11), p. 804–807. ISBN 978-1-4503-0445-0. Available at: <http://doi.acm.org/10.1145/1985793.1985907>.

TREUDE, C.; PROLO, C. A.; FILHO, F. F. Challenges in analyzing software documentation in portuguese. In: *Software Engineering (SBES), 2015 29th Brazilian Symposium on*. [S.l.: s.n.], 2015. p. 179–184.

TRICCO, A. C. et al. A scoping review of rapid review methods. *BMC Medicine*, 2015.

TRICCO, A. C. et al. An international survey and modified delphi approach revealed numerous rapid review methods. *Journal of clinical epidemiology*, 2016.

TRICCO, A. C. et al. Rapid reviews to strengthen health policy and systems: a practical guide. *World Health Organization. http://apps. who. int/iris/bitstream/10665/258698/1/9789241512763-eng. pdf*, 2017.

VANHANEN, J.; LASSENIUS, C. L. Perceived effects of pair programming in an industrial context. In: *33rd EUROMICRO Conference on Software Engineering and Advanced Applications (EUROMICRO 2007)*. [S.l.: s.n.], 2007. p. 211–218. ISSN 1089-6503.

VASILESCU, B.; CAPILUPPI, A.; SEREBRENIK, A. Gender, representation and online participation: A quantitative study of stackoverflow. In: *2012 International Conference on Social Informatics*. [S.l.: s.n.], 2012. p. 332–338.

VASILESCU, B.; FILKOV, V.; SEREBRENIK, A. Stackoverflow and github: Associations between software development and crowdsourced knowledge. In: *Social Computing (SocialCom), 2013 International Conference on*. [S.l.: s.n.], 2013. p. 188–195.

VELDE, S. Van de et al. Medicinal use of potato-derived products: conclusions of a rapid versus full systematic review. *Phytotherapy Research*, 2011.

VERNER, J. M.; BRERETON, O. P.; KITCHENHAM, B. A.; TURNER, M.; NIAZI, M. Risks and risk mitigation in global software development: A tertiary study. *Inf. Softw. Technol.*, Butterworth-Heinemann, Newton, MA, USA, v. 56, n. 1, p. 54–78, Jan. 2014. ISSN 0950-5849. Available at: <http://dx.doi.org/10.1016/j.infsof.2013.06.005>.

VIERA, A. J.; GARRETT, J. M. Understanding interobserver agreement: The kappa statistic. *Journal of the Society of Teachers of Family Medicine*, 2005.

WANG, W.; GODFREY, M. W. Detecting api usage obstacles: A study of ios and android developer questions. In: *Proceedings of the 10th Working Conference on Mining Software Repositories*. Piscataway, NJ, USA: IEEE Press, 2013. (MSR '13), p. 61–64. ISBN 978-1-4673-2936-1. Available at: <http://dl.acm.org/citation.cfm?id=2487085.2487100>.

WEBB, S. A. Some considerations on the validity of evidence-based practice in social work. *British journal of social work*, BASW, v. 31, n. 1, p. 57–79, 2001.

WOHLIN, C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In: *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*. New York, NY, USA: ACM, 2014. (EASE '14), p. 38:1–38:10. ISBN 978-1-4503-2476-2. Available at: <http://doi.acm.org/10.1145/2601248.2601268>.

YAMAN, S. G.; SAUVOLA, T.; RIUNGU-KALLIOSAARI, L.; HOKKANEN, L.; KUVAJA, P.; OIVO, M.; MÄNNISTÖ, T. Customer involvement in continuous deployment: A systematic literature review. In: _____. *Requirements Engineering: Foundation for Software Quality: 22nd International Working Conference, REFSQ 2016, Gothenburg, Sweden, March 14-17, 2016, Proceedings*. Cham: Springer International Publishing, 2016. p. 249–265. ISBN 978-3-319-30282-9. Available at: <http://dx.doi.org/10.1007/978-3-319-30282-9_18>.

YEN, C.-L. A discussion of paradigms and research methods. In: . [S.l.: s.n.], 2011.

YIN, R. K. *Case Study Research: Design and Methods (Applied Social Research Methods)*. Fourth edition. [S.l.]: Sage Publications, 2008.

YOUNG, I. et al. A guide for developing plain-language and contextual summaries of systematic reviews in agri-food public health. *Foodborne pathogens and disease*, 2014.

YOURDON, E. When good enough software is best. *IEEE Software*, 1995.

ZECHMEISTER, I.; SCHUMACHER, I. The impact of health technology assessment reports on decision making in austria. *International journal of technology assessment in health care*, 2012.

ZHOU, Y.; ZHANG, H.; HUANG, X.; YANG, S.; BABAR, M. A.; TANG, H. Quality assessment of systematic reviews in software engineering: A tertiary study. In: *Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering*. New York, NY, USA: ACM, 2015. (EASE '15), p. 14:1–14:14. ISBN 978-1-4503-3350-4. Available at: <http://doi.acm.org/10.1145/2745802.2745815>.

APPENDIX A – QUESTIONNAIRE OF PRACTITIONERS PERCEPTION ON EVIDENCE BRIEFINGS - ES2

In this appendix, we present the questionnaire used as instrument to conduct the Stack Exchange users survey (ES2). The questions are grouped in four sections. Open questions are marked with an asterisk "*".

DEMOGRAPHICS

S1Q1: What is your current position?
☐ Software developer
☐ Software requirements engineer
☐ Software architect
☐ Software test engineer
☐ Software project manager
☐ Software quality assurance engineer
☐ Other _____

S1Q2: How many years of experience do you have in your current position?
☐ Less than 1 year
☐ 1 to 2 years
☐ 2 to 5 years
☐ 5 to 8 years
☐ 8 to 12 years
☐ More than 12 years

S1Q3: You work for:
☐ The software industry
☐ The academia
☐ The government
☐ Open source software
☐ Other _____

S1Q4: What is your level of educational attainment?
☐ High School
☐ Bachelor's degree

☐ Master's degree
☐ Doctorate
☐ Other _____

MEDIUMS TO ACQUIRE KNWOLEDGE

S1Q5: How often do you use StackExchange websites?
☐ Daily
☐ Weekly
☐ Montly
☐ I have used only a few times, but it is not common
☐ I never refer to StackExchange websites

S1Q6: How often do you read software engineering research papers?
☐ Daily
☐ Weekly
☐ Montly
☐ I have read only a few papers, but it is not common
☐ I never read a software engineering research papers

S1Q7: Have you ever read a systematic review paper?
☐ Yes
☐ No

S1Q8: If you answered Yes for the last question, for what reason you read a systematic review paper?
☐ Research purposes
☐ Decision-making on practice
☐ Policy-making
☐ Teaching
☐ Other _____

S1Q9: How important do you think is to have a platform such as StackExchange but to discuss briefings of software engineering research?
☐ Very important
☐ Important
☐ Moderately Important
☐ Slightly Important
☐ Not Important

PERCEPTIONS ON BRIEFINGS CONTENTS

S1Q10: To what degree do you think the information available in the briefing we sent to you can answer your question on StackExchange?
☐ Totally
☐ Partially
☐ It does not helped, but it is related to the question
☐ It does not helped and it is also not related to the question

***S1Q11**: If you have NOT answered "Totally" in the previous question, please describe why is it the case.

_____

S1Q12: Regardless the briefing answers or not your question, how important do you think is the research presented on the briefing?
☐ Very important
☐ Important
☐ Moderately Important
☐ Slightly Important
☐ Unwise

***S1Q13**: If you have answered "unwise" in the previous question, please describe why is it the case.

_____

S1Q14: How do you compare the answers from the StackExchange community to the findings presented in the briefing?
☐ I prefer the answers from the StackExchange community
☐ Neutral
☐ I prefer the findings presented in the briefing

PERCEPTIONS ON BRIEFINGS FORMAT

S1Q15: How do you think it was to find the information in the briefing?
☐ Very easy
☐ Easy
☐ Neutral
☐ Difficult

☐ Very difficult

S1Q16: Is the briefing interface clear and understandable?

☐ Strongly agree

☐ Agree

☐ Neutral

☐ Disagree

☐ Strongly disagree

S1Q17: Does the briefing look reliable?

☐ Strongly agree

☐ Agree

☐ Neutral

☐ Disagree

☐ Strongly disagree

APPENDIX B – QUESTIONNAIRE OF RESEARCHERS PERCEPTIONS ON EVIDENCE BRIEFINGS - ES3

In this appendix, we present the questionnaire used as instrument to conduct the survey with authors of the SSs (ES3). The questions are grouped in three sections. Open questions are marked with an asterisk "*".

SHARING RESEARCH WITH PRACTITIONERS

S2Q1: How important for you is to share research results to practitioners?
☐ Very important
☐ Important
☐ Moderately Important
☐ Slightly Important
☐ Not Important

S2Q2: How often do you share research results to practitioners?
☐ Daily
☐ Weekly
☐ Monthly
☐ Yearly
☐ I have shared only a few times, but it is not common
☐ I have never shared research results to practitioners

***S2Q3**: If you have ever shared research results to practitioners, how did you do that?
_____

PERCEPTIONS ON BRIEFINGS CONTENTS

S2Q4 How does the briefing that we sent to you cover the main findings of your paper?
☐ Very good
☐ Good
☐ Acceptable
☐ Poor
☐ Very poor

***S2Q5**: If you have answered "poor" or "very poor" in the previous question, please describe why is it the case and mention which paper are you talking about.

_____

PERCEPTIONS ON BRIEFINGS FORMAT

S2Q6 How do you think it was to find the information in the briefing?
☐ Very easy
☐ Easy
☐ Neutral
☐ Difficult
☐ Very difficult

S2Q7: Is the briefing interface clear and understandable?
☐ Strongly agree
☐ Agree
☐ Neutral
☐ Disagree
☐ Strongly disagree

S2Q8 Does the briefing look reliable?
☐ Strongly agree
☐ Agree
☐ Neutral
☐ Disagree
☐ Strongly disagree

APPENDIX C – QUESTIONNAIRE OF RESEARCHERS PERCEPTIONS ON EVIDENCE BRIEFINGS CREATION - ES4

This appendix presents the questionnaire used as instrument to conduct the survey with the researchers who created Evidence Briefings from their papers published on the four selected conferences (ES4). Open questions are marked with an asterisk "*".

DEMOGRAPHICS

S3Q1: What is your current position?
☐ Undergrad Student
☐ Grad Student
☐ PostDoc
☐ Professor
☐ Practitioner

KNOWLEDGE SHARING HABITS

S3Q2: How important for you is to share research results to practitioners?
☐ Very important
☐ Important
☐ Somewhat important
☐ Not important

S3Q3: How often do you share research results with practitioners?
☐ Daily
☐ Weekly
☐ Monthly
☐ Yearly
☐ It is not common
☐ I've never shared

***S3Q4:** If you have ever shared research results to practitioners, how did you do that?

_____

PERCEPTIONS ON EVIDENCE BRIEFING CREATION

S3Q5: How much time was needed to create the briefing?
☐ <1h
☐ Up to 2h
☐ Up to 3h
☐ Up to 4h
☐ 4h>

S3Q6: How difficult was to create the briefing?
☐ Very Difficult
☐ Difficult
☐ Somewhat difficult
☐ Not difficult

***S3Q7:** If you DID NOT answered "Not Difficult" in the last question, could you please elaborate over it?

_____

S3Q8: How do you evaluate your final briefing?
☐ Very Good
☐ Good
☐ Acceptable
☐ Poor
☐ Very Poor

***S3Q9:** If you answered "It is poor" in the last question, could you please elaborate over it?

_____

S3Q10: Was one single page enough to summarize your results?
☐ Yes
☐ No

***S3Q11:** If you answered NO in the last question, could you please elaborate over it?

_____

PERCEPTIONS ON EVIDENCE BRIEFINGS INITIATIVE

S3Q12: If conferences start to publish briefings at their official web-pages, would you

consider reading the briefings of accepted papers BEFORE the actual papers?
☐ Yes
☐ No
☐ Maybe

***S3Q13:** If you answered NO or MAYBE in the last question, could you please elaborate over it?

_____

S3Q14: If briefings were commonplace in software engineering conferences (yet not mandatory), how likely would you create such briefings?
☐ Very probably
☐ Probably
☐ Probably not
☐ Definitely Not

***S3Q15:** In your opinion, what are the main ADVANTAGES of briefings usage?

_____

***S3Q16:** In your opinion, what are the main DISADVANTAGES of briefings usage?

_____

***S3Q17:** Do you have any suggestions on how to improve the briefings?

_____

S3Q18: How likely do you think practitioners would benefit from Evidence Briefings?
☐ Very Likely
☐ Likely
☐ Maybe
☐ Not Likely

***S3Q19:** In your opinion, which approach is more suitable for practitioners to acquire knowledge based on scientific evidence?
☐ Traditional research papers
☐ Evidence briefings
☐ Other _____

APPENDIX D – PAPERS WITH EVIDENCE BRIEFINGS CREATED - ES4

Table 25 lists the papers which Evidence Briefings were created for ES4, as well as the conferences each paper was published.

Table 25 – Papers which Evidence Briefings were created for ES4.

TITLE	CONF.	EB
'2TScrum': A Board Game to Teach Scrum	SBES	<http://bit.ly/2F7S5xS>
A Comparative Study of Model-Driven Approaches For Scoping and Planning Experiments	EASE	<http://bit.ly/2CRJpWB>
Adding human interaction aspects in the writing of User Stories: a perspective of software developers	SBES	<http://bit.ly/2oC37RQ>
Analysing Requirements Communication Using Use Case Specification and User stories	SBES	<http://bit.ly/2GUnXCF>
An Automated Refactoring Approach to Remove Unnecessary Complexity in Source Code	SAST	<https://bit.ly/2GvdW38>
Aspects influencing feature-oriented software comprehension: observations from a focus group	SBCARS	<https://bit.ly/2Guy4CB>
A Strategy for Functional Defect Prediction in Homogeneous Datasets: A case study in the SIGAA academic system	SAST	<https://bit.ly/2GQ9gnL>
A structured survey on the usage of the issue tracking system provided by the GitHub platform	SBCARS	<http://bit.ly/2FIyqBJ>
A systematic review on the use of Definition of Done on agile software development projects	EASE	<http://bit.ly/2CQQjeM>
Better similarity coefficients to identify refactoring opportunities	SBCARS	<https://bit.ly/2GQ7jYt>
Challenges to the Development of Smart City Systems: A System-of-Systems View	SBES	<http://bit.ly/2FbAowX>
Characterizing big data software architectures: a systematic mapping study	SBCARS	<https://bit.ly/2GT6Epm>
Coding Dojo as a transforming practice in collaborative learning of programming: an experience report	SBES	<http://bit.ly/2F2p1Il>
Comparing Configuration Approaches for Dynamic Software Product Lines	SBES	<http://bit.ly/2t7l6nQ>
Evaluating an Automatic Text-based Test Case Selection using a Non-Instrumented Code Coverage Analysis	SAST	<https://bit.ly/2GOh3Tj>
Feasibility of using Source Code Changes on the Selection of Text-based Regression Test Cases	SAST	<https://bit.ly/2GOh3Tj>
Feature-Based Test Oracles to Categorize Synthetic 3D and 2D Images of Blood Vessels	SAST	<https://bit.ly/2uPjyzW>
From Statecharts into Model Checking: A Hierarchy-based Translation and Specification Patterns Properties to Generate Test Cases	SAST	<https://bit.ly/2H6woMr>
Hearing the Voice of Developers in Mobile Software Ecosystems	SBES	<http://bit.ly/2FjaG9T>
How are Conceptual Models used in Industrial Software Development? A Descriptive Survey	EASE	<http://bit.ly/2oIB3LT>
How Do Software Developers Identify Design Problems?: A Qualitative Analysis	SBES	<http://bit.ly/2oJkBLs>
How Has the Health of Software Ecosystems Been Evaluated?: A Systematic Review	SBES	<http://bit.ly/2FKbZw3>
Improving the structure of KDM instances via refactorings: An experimental study using KDM-RE	SBES	<http://bit.ly/2oKEm5d>

Incremental Strategy for Applying Mutation Operators Emphasizing Faults Difficult to be Detected by Automated Static Analyser	SBES	<http://bit.ly/2FbyMmT>
Investigating the Effectiveness of Peer Code Review in Distributed Software Development	SBES	<http://bit.ly/2oK4SeZ>
Investigating the variability impact on the recovery of software product line architectures: an exploratory study	SBCARS	<https://bit.ly/2uIkR3F>
Is There a Demand of Software Transparency?	SBES	<http://bit.ly/2F6WL3g>
Late Decomposition of Applications into Services through Model-Driven Engineering	SBES	<http://bit.ly/2t4k8c2>
On the Benefits/Limitations of Agile Software Development: An Interview Study with Brazilian Companies	EASE	<http://bit.ly/2t5JKFF>
On the Evaluation of Effort Estimation Models	EASE	<http://bit.ly/2oFyF8G>
PBL Integration into a Software Engineering Undergraduate Degree Program Curriculum: An Analysis of the Students' Perceptions	SBES	<http://bit.ly/2GXJj2d>
Problem-Based Learning to Align Theory and Practice in Software Testing Teaching	SBES	<http://bit.ly/2ozTl2u>
Programming Language Adoption as an Epidemiological Phenomenon	SBES	<http://bit.ly/2CRSLBR>
Retrospective for the Last 10 years of Teaching Software Engineering in UFC's Computer Department	SBES	<http://bit.ly/2tckgqe>
Reuse of model-based tests in mobile apps	SBES	<http://bit.ly/2FKcMNx>
Revealing design problems in stinky code: a mixed-method study	SBCARS	<https://bit.ly/2q4lvCH>
Spotify characterization as a software ecosystem	SBCARS	<https://bit.ly/2El00TC>
Software Interoperability Analysis in Practice: A Survey	EASE	<http://bit.ly/2HYrj8X>
Test case prioritization: a systematic review and mapping of the literature	SBES	<http://bit.ly/2GUqBbx>
Testing context-aware software systems: Unchain the context, set it free!	SBES	<http://bit.ly/2t3gHCk>
Testing Game: An Educational Game to Support Software Testing Education	SBES	<http://bit.ly/2oKD0r9>
Tweaking Association Rules to Optimize Software Change Recommendations	SBES	<http://bit.ly/2oLb4mX>
Understanding Technical Debt at the Code Level from the Perspective of Software Developers	SBES	<http://bit.ly/2t6CmK1>
What are Software Engineers asking about Android Testing on Stack Overflow?	SBES	<http://bit.ly/2CRdaqG>

APPENDIX E – RAPID REVIEW PROTOCOL WITH COMPANY 1 - ES5

Here we present the protocol of the Rapid Review (RR) conducted with Company 1. The entire protocol was defined in close relationship with the participants of Company 1.

THE PRACTITIONERS PROBLEMS

Previous studies revealed that about 75% of Full Secondary Studies (FSSs) authors in Software Engineering (SE) were motivated by academic purposes only (SANTOS; SILVA, 2013). Thus, few FSSs are in fact motivated by an actual problem faced in practice, or under practitioners restrictions. This RR is an exception to this trend.

The participants reported problems related to *low customer collaboration*. They would like to have access to evidence that could improve their relationship with their costumer and, consequently, mitigate this problem.

RESEARCH QUESTIONS

Together with the participants, we formulated the following research questions to guide this RR:

RQ1: What are the benefits of customer collaboration in software development practice?

RQ2: What are the problems caused by low customer collaboration in the software development practice?

RQ3: What are the challenges to establish customer collaboration in software development practice?

RQ4: What are the strategies to improve customer collaboration in software development practice?

The first two RQs are intended to provide evidence about the benefits of adequate customer collaboration. As a consequence, the participants can have better arguments to convince their customer to increase collaboration. RQ3 seeks for a better understanding of the challenges that hinder customer collaboration in order to mitigate this problem. Finally, RQ4 is intended to provide evidence about strategies to mitigate low customer collaboration.

SEARCH STRATEGY

To abbreviate the search for primary studies, and conduct the RR under the agreed time frame, we used only the Scopus[1] search engine. It searches in many of the most relevant SE digital libraries. We tested many different versions of the search string until we found a set that returned relevant papers. Before conducting the search, we present the possible search string to practitioners, and through a feedback loop with them, we refined and defined the following search string:

(customer OR "product owner" OR stakeholder)
AND
(collaborat* OR participat* OR cooperat* OR relation* OR involvement OR engagement)
AND
("software development" OR "software engineering" OR "software project")

SELECTION PROCEDURE

After discussing with the participants, the selection procedure was based on the following criteria:

1. The study must be in the context of software engineering;
2. The study must be a primary study (i.e., We do not consider secondary or tertiary studies);
3. The study must present evidence based on scientific empirical methods (e.g., interviews, surveys, case studies, etc). Propositional or anecdotal studies were not considered;
4. The study must provide answers to at least one of the rapid review's research questions.

Figure 16 depicts the selection procedure results. The search in Scopus returned 1,973 studies. In the first round, a solo researcher analyzed the studies' titles and excluded those that clearly did not meet the criteria, resulting in 84 studies. Although the wild-cards used in the search string might provoked high number of false-positives, we kept them to have a more comprehensive search. In the second round, a solo researcher analyzed the studies' abstract, reducing the number of studies to 47. In the third round, a solo researcher analyzed the entire papers content and excluded those that could not answer any of the

[1] <https://www.scopus.com>

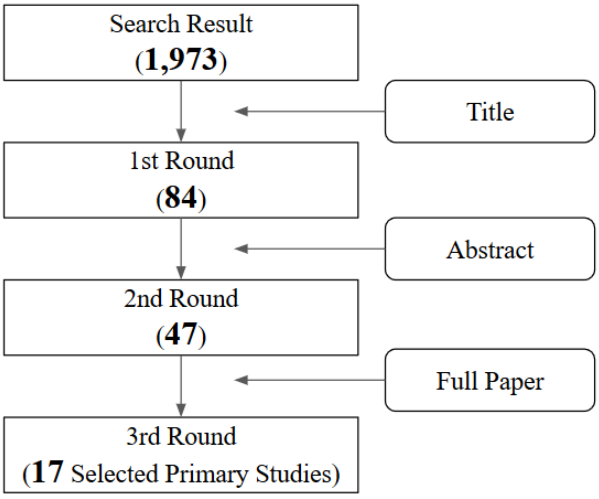


Figure 15 – Selection procedure results (ES5).

research questions. Thus, we ended up with 17 selected primary studies, as shown in Table 26. The selection procedure is available online[2].

Table 26 – Selected primary studies (ES5).

PAPER TITLE
A case study of customer communication in globally distributed software product development
A comparative case study on clients participation in a 'traditional' and in an Agile software company
A competency model for customer representatives in agile software development projects
A prototype tool to support extended team collaboration in agile project feature management
Agile undercover: When customers don't collaborate
Collaboration, communication and co-ordination in agile software development practice
Customer involvement in requirements management: lessons from mass market software development
Customer involvement in service production, delivery and quality: the challenges and opportunities
Customer-developer links in software development
Engaging stakeholders with agent-oriented requirements modelling
How Stakeholders' Commitment May Affect the Success of Requirements Elicitation
Mutual development: A case study in customer-initiated software product development
Relationship Research Between Communication Activities and Success Indexes in Small and Medium Software Projects
The connection of the stakeholder cooperation intensity and team agility in software development
The impact of agile practices on communication in software development
The relationship between customer collaboration and software project overruns
Towards Customer-Centric Software Development: A Multiple-Case Study

Selection in pairs was not considered due the participants' time and resource constraints. The selection procedure was also conducted in close collaboration with the participants. For instance, we showed to practitioners the title and abstract of primary studies that we were in doubt whether to include or not.

EXTRACTION PROCEDURE

[2] The selection procedure is available at <http://bit.ly/2su8xC8>

In this step, we extracted all relevant findings that could help to answer any of the research questions. The extraction was also conducted by a solo researcher, as was the selection procedure.

SYNTHESIS PROCEDURE

Usually, RRs perform just a narrative summary to synthesize evidence aiming to reduce time/effort. However, we decided to conduct a formal Thematic Analysis because the number of selected primary studies was not high. We followed the phases of thematic analysis, as presented by Fereday and Cochrane (FEREDAY; MUIR-COCHRANE, 2006).

1. **Familiarization with data:** At this phase, we analyzed each selected primary study. When a study introduced new concepts, we searched on the literature to better familiarize with them.

2. **Creating initial codes:** Here we gave a code for each finding of the primary studies. These codes summarize the core of each finding. For instance, the finding *"[...] participants expressed their frustration over not being able to choose the customer reps [...] It is not enough to have a customer rep for the project, it is also important for that rep to be effective in providing requirements and feedback to the team"*, from the study of Hoda et al. (HODA et al., 2010) was coded as *Ineffective customer representative*.

3. **Searching for themes:** In this step, we already had a list of initial themes (e.g., *Lack of commitment*, as a challenge to establish customer collaboration), but we begin to focus on broader patterns in the data, combining coded data with proposed themes.

4. **Reviewing themes:** Here we have a potential set of themes. We then searched for data that supports or refutes our themes. For instance, we initially themed the finding *"the role of customer is rarely taken by the ideal individual, and the individual circumstances of that person affects the nature of collaboration and communication. For example, how much authority the customer has in making decisions; how much knowledge of the domain the customer has [...]."* from the study of Robinson et al. (ROBINSON et al., 2010) as *"The customer representative is rarely ideal."*. However, we later realized that this finding would fit as *"Ineffective customer representative"*, so we merged the two themes.

5. **Producing the final report:** This process leads to the themes that composed the final report of the RR.

RAPID REVIEW REPORT

To make the RR more appealing to practitioners, we reported its result through an Evidence Briefing (CARTAXO et al., 2016). It is an one-page document that presents only the main findings of a research. The Evidence Briefing document is also available online [3].

LIMITATIONS AND THREATS TO VALIDITY

RRs usually present more threats to validity than FSSs due to its lightweight methodology. Some of the threats and limitations of this RR include:

- Only one search engine was used, which may limited the number of primary studies found;
- The entire selection procedure was conducted by a solo researcher, which may introduce selection bias;
- We did not conduct quality assessment on the selected primary studies. This might limit the reliability on the evidence we have found;
- The selection of papers by title consists in another threat;
- The RR was presented to practitioners through an Evidence Briefing. Because it is an one-page document focused only on the findings, no information about RR's method was presented;
- The participants have master degree in computer science, who are more likely to consume scientific evidence. However, as reported in a recent survey, 48% of the StackExchange users — practitioners oriented QA website — holds a master degree (10% holds a Ph.D.) (CARVER et al., 2016).

[3] The Evidence Briefing can be found at <http://bit.ly/2rHiQDd>

APPENDIX F – RAPID REVIEW PROTOCOL WITH COMPANY 2 - ES6

Here we present the protocol of the Rapid Review (RR) conducted with Company 2. The entire protocol was defined in close relationship with the participants of Company 2.

THE PRACTITIONERS PROBLEM

Previous studies revealed that about 75% of Full Secondary Studies (FSSs) authors in Software Engineering (SE) were motivated by academic purposes only (SANTOS; SILVA, 2013). Thus, few FSSs are in fact motivated by an actual problem faced in practice, or under practitioners restrictions. This RR is an exception to this trend.

The practitioners reported problems related to *low team motivation*. They would like to have access to evidence that could improve the team motivation and, consequently, mitigate this problem.

RESEARCH QUESTIONS

Together with the participants, we formulated the following research questions to guide the RR:

RQ1: What are the factors that impact software development teams motivation?

RQ2: What are the strategies to enhance software development teams motivation?

RQ1 seeks for a better understanding on the factors that impact team motivation, so the participants can identify which of them are occurring in Project 2. RQ2 is intended to provide evidence about strategies to mitigate low team motivation.

SEARCH STRATEGY

To abbreviate the search for primary studies, and conduct the RRs under the agreed time frame, we used only the Scopus[1] search engine. It searches in many of the most relevant SE digital libraries. We tested many different versions of the search string until we found a set that returned relevant papers. Before conducting the search, we present the possible search string the participants, and through a feedback loop with them, we refined and defined the following search string:

[1] <https://www.scopus.com>

(motivat* OR demotiv* OR satisf* OR inspir* OR enthusias*) AND “software engineering”

SELECTION PROCEDURE

After discussing with the participants, the selection procedure was based on the following inclusion/exclusion criteria:

1. The study MUST be related to software engineering;
2. The study MUST answer at least one of the research questions;
3. The study MUST present evidence based on scientific empirical methods (e.g., interviews, surveys, case studies, etc). Propositional or anecdotal studies must be excluded;
4. The study MUST be a primary study. Secondary or tertiary studies must be excluded;
5. The study MUST be written in english;
6. The study CANNOT be a retracted paper;
7. The study CANNOT be related to open source software development;
8. The study CANNOT be related crowd source software development;
9. The study CANNOT have students as subjects;
10. The study CANNOT be related to distributed teams;
11. The study CANNOT be related to large companies;

Figure 16 depicts the selection procedure results. The search in Scopus returned 7,022 studies. In the first round,a solo researcher analyzed the studies' titles and excluded those that clearly did not meet the criteria, resulting in 393 studies. Although the wild-cards used in the search string might provoked high number of false-positives, we kept them to have a more comprehensive search. In the second round, a solo researcher analyzed the studies' abstract, reducing the number of studies to 147. In the third round, a solo researcher analyzed the entire papers content and excluded those that could not answer any of the research questions. Thus, we ended up with 35 selected primary studies, as shown in Table 27. The selection procedure is available online.[2]

[2] The selection procedure is available at <https://bit.ly/2J2YUPZ>

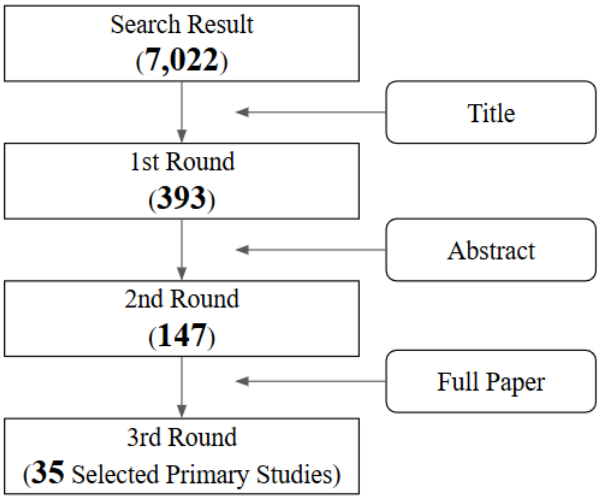


Figure 16 – Selection procedure results (ES6).

Selection in pairs was not considered due the participants' time and resource constraints. The selection procedure was also conducted in close collaboration with the participants. For instance, we showed to practitioners the title and abstract of primary studies that we were in doubt whether to include or not.

EXTRACTION PROCEDURE

In this step, we extracted all relevant findings that could help to answer any of the research questions. The extraction was also conducted by a solo researcher, as was the selection procedure. To make this step more systematic and replicable, we created a extraction spreadsheet, which is available online.[3]

SYNTESIS PROCEDURE

Usually, RRs perform just a narrative summary to synthesize evidence aiming to reduce time/effort. However, we decided to conduct a formal Thematic Analysis because the number of selected primary studies was not high. We followed the phases of thematic analysis, as presented by Fereday and Cochrane (FEREDAY; MUIR-COCHRANE, 2006).

1. **Familiarization with data:** At this phase, we analyzed each selected primary study. When a study introduced new concepts, we searched on the literature to better familiarize with them.

2. **Creating initial codes:** Here we gave a code for each finding of the primary studies. These codes summarize the core of each finding.

[3] The extraction procedure can be found at <https://bit.ly/2Ibr0Hb>

Table 27 – Selected primary studies (ES6).

PAPER TITLE
A framework for gamification in software engineering
A preliminary evaluation of a gamification framework to jump start collaboration behavior change
Adopting collaborative games into Open Kanban
An initial investigation of software practitioners' motivation
Benefits and limitations of project-to-project job rotation in software organizations: A synthesis of evidence
Career goals of software development professionals and software development students
Challenges and strategies for motivating software testing personnel
Characterizing and predicting mental fatigue during programming tasks
Designing motivation strategies for software engineering teams: An empirical study
Do XP customer-developer interactions impact motivation? Findings from an industrial case study
Effects of agile practices on social factors
Empirical insights into the perceived benefits of agile software engineering practices: A case study from SAP
Exploring motivational differences between software developers and project managers
Factors affecting team motivation: A survey of Finnish software engineers
Factors influencing group member satisfaction in the software industry
Factors that motivate software engineering teams: A four country empirical study
Feedback: How does it impact software engineers?
How does radical collocation help a team succeed?
How temporal work styles and product modularity influence software quality and job satisfaction
Individual empowerment of agile and non-agile software developers in small teams
Job satisfaction in agile development teams: Agile development as work redesign
Motivating factors in agile and traditional software development methods: A comparative study
Motivation and cohesion in agile teams
Motivation in software engineering industrial practice: A cross-case analysis of two software organisations
Motivation of software engineers: A qualitative case study of a research and development organisation
Motivational and de-motivational factors for software engineers: An empirical investigation
On the unhappiness of software developers
REfine: A gamified platform for participatory requirements engineering
Scrum hero: Gamifying the scrum framework
Socialness in the recruiting of software engineers
Software engineers' perceptions of factors in motivation
Software project team characteristics and team performance: Team motivation as a moderator
Teamwork quality and project success in software development: A survey of agile development teams
Towards understanding the underlying structure of motivational factors for software engineers to guide the definition of motivational programs

3. **Searching for themes:** In this step, we already had a list of initial themes, but we begin to focus on broader patterns in the data, combining coded data with proposed themes.

4. **Reviewing themes:** Here we have a potential set of themes. We then searched for data that supports or refutes our themes.

5. **Producing the final report:** This process leads to the themes that composed the final report of the RR.

RAPID REVIEW REPORT

To make the RR more appealing to practitioners, we reported its result through an Evidence Briefing (CARTAXO et al., 2016). It is an one-page document that presents only the main findings of a research. The Evidence Briefing document is also available online [4].

LIMITATIONS AND THREATS TO VALIDITY

RRs usually present more threats to validity than FSSs due to its lightweight methodology. Some of the threats and limitations of this RR include:

- Only one search engine was used, which may limited the number of primary studies found;
- The entire selection procedure was conducted by a solo researcher, which may introduce selection bias;
- We did not conduct quality assessment on the selected primary studies. This might limit the reliability on the evidence we have found;
- The selection of papers by title consists in another threat;
- The RR was presented to practitioners through an Evidence Briefing. Because it is an one-page document focused only on the findings, no information about RR's method was presented;
- The participants have master degree in computer science, who are more likely to consume scientific evidence. However, as reported in a recent survey, 48% of the StackExchange users — practitioners oriented QA website — holds a master degree (10% holds a Ph.D.) (CARVER et al., 2016).

[4] The Evidence Briefing can be found at <http://bit.ly/2rHiQDd>

APPENDIX G – DIAGNOSIS INTERVIEW SCRIPT - ES5 AND ES6

The script presented here refers to the semi-structured interview we have conducted with the participants of Company 1 and 2, in order to identify the problems they were facing in practice. The interview was conducted during the diagnosis step of the action researches we conducted to evaluate the perceptions of practitioners about Rapid Reviews (RRs).

GENERAL QUESTIONS

Q1: What is your current position in your company?
Q2: How many years of experience do you have in your current position?
Q3: What is your level of educational attainment?
Q4: How often do you read software engineering research papers?
Q5: Have you ever read a secondary study paper?
Q6: If you answered Yes for the last question, for what reason you read a secondary study paper?

DIAGNOSING QUESTIONS

We are studying the applicability of rapid reviews as a method to support decision-making in industrial real-world software development projects. Is it OK to record our talk?

Q7: Could you give an overview of your project?
Q8: Could you describe the characteristics of your team?
Q9: Could you talk about the main challenges your team is facing at this moment?

Q10: Does your team facing any challenges related to software requirements at this moment? <MENTION AN EXAMPLE> If so, could you talk about it?

Q11: Does your team facing any challenges related to software design/architecture at this moment? <MENTION AN EXAMPLE> If so, could you talk about it?

Q12: Does your team facing any challenges related to software coding at this moment? <MENTION AN EXAMPLE> If so, could you talk about it?

Q13: Does your team facing any challenges related to software testing at this moment?

<MENTION AN EXAMPLE> If so, could you talk about it?

Q14: Does your team facing any challenges related to software maintenance at this moment? <MENTION AN EXAMPLE> If so, could you talk about it?

Q15: Does your team facing any challenges related to software configuration management at this moment? <MENTION AN EXAMPLE> If so, could you talk about it?

Q16: Does your team facing any challenges related to software management at this moment? <MENTION AN EXAMPLE> If so, could you talk about it?

Q17: Does your team facing any challenges related to software process at this moment? <MENTION AN EXAMPLE> If so, could you talk about it?

Q18: Does your team facing any challenges related to software quality at this moment? <MENTION AN EXAMPLE> If so, could you talk about it?

Q19: Do you have any other comment or question?

APPENDIX H – INTERVIEW SCRIPT TO EVALUATE THE INTRODUCTION OF RAPID REVIEWS IN PRACTICE - ES5 AND ES6

The script presented here refers to the semi-structured interview we have conducted with participants of Company 1 and 2, in order to identify their perceptions about the Rapid Reviews we conducted. The interview was conducted during the evaluation step of the action research at the workshop with the participants of Company 1 and 2.

INTERVIEW SCRIPT

Based on the previously interview, we conducted a Rapid Review to provide scientific evidence aiming to support the decision-making towards the solution of the problem you have mentioned to be facing. Thus we would like to present the Evidence Briefing with the findings. Is it OK to record our talk?

Q1: What do you think about this idea of Rapid Reviews? Have you heard about something similar before?

Q2: What do you think about the process of creating a Rapid Review? Does that make sense to you?

Q3: What do you think about the idea of transferring research evidence through one-page documents? Do you think it should be more common? Would you like to see more Evidence Briefings like this?

Q4: What do you think about the evidence found by the Rapid Review? Did you learn something new? What? Do you concur with them? Do you believe you can apply (some of) them?

Q5:Does the Rapid Review evidence have any direct potential impact on your daily activities? Why?

Q6: To what extent do you think the evidence provided by the Rapid Review can support you on decision-making? Why?

Q7: Regardless if the Rapid Review helps to mitigate your needs, how important do you

think is the research presented? Why?

Q8: Would you encourage your peers to read Evidence Briefings like this?

Q9: How easy/hard it was to find the information in the Evidence Briefing?

Q10: Is the Evidence Briefing interface clear and understandable?

Q11: Does the Evidence Briefing looks reliable? Do you believe in the evidence available on it? Why?

Q12: Would you like to receive Evidence Briefings like that on regular basis?

Q13: Do you have any additional questions/comments?

APPENDIX I – EVIDENCE BRIEFINGS OF THE EMPIRICAL STUDIES

In the following pages there are three Evidence Briefings reporting the main findings of the six empirical studies of this thesis. The first reports the main findings of ES1, the empirical study about the coverage of practitioners issues by Secondary Studies. The second reports the main findings of ES2, ES3, and ES4, the three empirical studies about Evidence Briefings. And finally, the third reports the main findings of ES5 and ES6, the two empirical studies about rapid reviews.

HOW SYSTEMATIC REVIEWS IN SOFTWARE ENGINEERING COVER PRACTITIONERS ISSUES

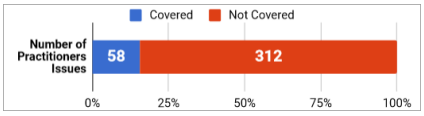
This briefing reports scientific evidence on how systematic reviews cover the issues practitioners face in software engineering practice.

FINDINGS

The results presented here were obtained following a mining software repository approach. The results of 24 systematic reviews exploring diverse topics of software engineering were compared with the issues reported by practitioners in five Stack Exchange websites (PROG, PM, SQA, RE, SREC). The issues were considered as covered when at least one finding of one of the 24 systematic reviews offered knowledge that could solve or at least attenuate that issue.

• **Only 1.38% of most relevant practitioners issues are related to the systematic reviews.** Among the 26,687 more relevant issues posted on the five Stack Exchange websites, only 370 were considered as related to the same topic of at least one of the 24 systematic reviews. This shows that systematic reviews are still far from touch the whole spectrum of topics discussed by practitioners.

• **Only 15.7% of the practitioners issues related to the systematic reviews are covered them.** Among the 370 issues related to the same topic of the 24 systematic reviews, only 58 were considered as covered by those reviews. This shows that even those few systematic reviews touching topics that interest practitioners are not exploring issues practitioners are really facing in practice.



• **None software engineering topic has a coverage rate above 50%.** Table 1 presents the total number of issues grouped in 15 software engineering topics, and their coverage rate. For instance, only 26.6% (8) of the 30 issues related to *Software Testing* are covered by the 24 systematic reviews.

SOFTWARE ENGINEERING TOPIC	COVERAGE		TOTAL
	%	#	#
Software Engineering Economics	36.3%	4	11
Software Testing	26.6%	8	30
Software Engineering Professional Practice	24.0%	12	46
Software Maintenance	22.2%	2	9
Software Engineering Process	20.0%	4	20
Software Requirements	15.7%	6	38
Software Engineering Models and Methods	13.3%	17	127
Software Construction	12.5%	2	16
Software Engineering Management	4.6%	3	64
Software Design	0.0%	0	9
Software Configuration Management	0.0%	0	0
Software Quality	0.0%	0	0
Computing Foundations	0.0%	0	0
Mathematical Foundations	0.0%	0	0
Engineering Foundations	0.0%	0	0
TOTAL	15.7%	58	370

Looking to the coverage rate on Table 1 it is possible to see which software engineering topics are more neglected by the systematic reviews. Thus, researchers planning to conduct a systematic review could take that into consideration.

• **45 recurrent issues were identified distributed in many software engineering topics.** A recurrent issue groups two or more issues about the same problem. It enables one to identify the most common problems in practice. The *Software Models and Methods* topic was the one with more recurrent issues, all of them related to agile methodologies.

• **15 recurrent issues related to agile software development were identified.** They group 127 issues in total. This is about one-third of the issues we have found in this study, which means agile is still an important topic in practice. Table 2 presents the coverage of each of the 15 recurrent issues about agile.

RECURRENT ISSUE	COVERAGE		TOTAL
	#	%	#
Negative impact of agile in software design	3	100%	3
Benefits of agile methods/practices from a specific per spective	2	40%	5
Benefits of agile methods/practices in general	3	33%	9
Low customer collaboration in agile	1	25%	4
Pair programming with distributed pairs	1	12%	8
Applicability of agile in specific project context	1	5%	19
Mixing agile with traditional methods/practices	0	0%	10
Pair programming to transfer knowledge	0	0%	9
Impact of low detail level or absence of documentation in ag ile	0	0%	6
Pair programming hindering concentration	0	0%	4
Benefits of agile methods/practices in specific contexts	0	0%	4
Mixing multiple agile methods/practices	0	0%	4
Tools for agile methods/practices	0	0%	4
Ad-hoc software development as agile	0	0%	3
Pair programming as replacement to code reviews	0	0%	2
Miscellaneous	6	15%	33
TOTAL	17	13.3%	127

• **Some recurrent issues reveal practitioners are facing problems that challenge some of the basic agile principles.** This can be observed when practitioners report *Low Customer Collaboration* or when they acknowledge not desirable *Impacts of Low Detail Level or Absence of Documentation* in agile.

• **Looking to the recurrent issues, one can see that practitioners have experiencing *Ad-hoc Software Development as Agile.*** This evidence may support some claims that in some situations agile is used as an excuse for absence of software process.

• **Still, it is possible to see that practitioners are facing problems Mixing Agile with Traditional Methods/Practices.** In some cases there is an impression that once agile is adopted, every stakeholder and process operate through agile philosophy. However as we can see, many practitioners face situations where their team is agile, but not their company as a whole. Or even when the team is agile, but not their customer. Such situations provoke disarrangement during software development life-cycle, and mixing agile with traditional methodologies is imperative.

• **There are practitioners explicitly asking for scientific empirical evidence in Stack Exchange websites.** This shows that there is interest in empirical evidence from practitioners side. So, researchers have space if they want to make research closer to practice.

• **Many practitioners ask for recommendations of tools with specific features to support various software engineering tasks.** Thirty issues were identified asking for tools. Thus, systematic reviews identifying tools, comparing their features, and aggregating evidence about their effectiveness could help to cover that gap. Another approach with direct implications for tool builders is to identify features demanded for software engineering tools based on issues posted in Stack Exchange websites.

• **Practitioners demand contextualized evidence.** This can be observed by looking at recurrent issues like: *Benefits of agile methods/practices from a specific perspective; Applicability of agile in specific project context; and Effort estimation in agile in specific project context*. This supports many claims about importance of rich and contextualized evidence.

• **Practitioners demand target oriented information.** For instance, we identified the following recurrent issue: *Benefits of agile methods/practices from a specific perspective*. This situation shows that empirical evidence should comprise not only data about the effectiveness of an intervention but also useful information for the target audience. For instance, cost-effectiveness. We also identified other issues reporting the need of information from other specific perspectives beyond monetary, such as, from the perspective of developers, managers, testers, customers, and others.

IN SUMMARY, the low coverage rate (15.7%) of practitioners issues by systematic reviews corroborates with the claims related to the lack of connection between systematic reviews and software engineering practice. The results also suggest that, although systematic reviews might help practitioners, there is room for studies more connected to issues faced in practice, especially on agile software development.

Who is this briefing for?

Software engineering researchers who want to conduct a systematic review covering the kind of issues that practitioners face in software engineering practice.

Where the findings come from?

All findings of this briefing were extracted from the mining repository study conducted by Cartaxo et al.

What is included in this briefing?

Some results and discussions about the how systematic reviews cover practitioners issues.

What is not included in this briefing?

The list of recurrent issues of other topics beyond agile software development. For more details and discussions about the other recurrent issues, see the full document.

To access other evidence briefings on software engineering:

http://cin.ufpe.br/eseg/evidence-briefings

For additional information about ESEG:

http://cin.ufpe.br/eseg

ORIGINAL RESEARCH REFERENCE

Cartaxo, Bruno et al. ***A Model To Transfer Knowledge From Research To Software Engineering Practice Based On Rapid Reviews And Evidence Briefings***. 2018.

PERCEPTIONS ON EVIDENCE BRIEFINGS IN SOFTWARE ENGINEERING

This briefing reports scientific evidence on the perceptions of practitioners and researchers on using Evidence Briefings as knowledge transfer mediums in software engineering.

FINDINGS

PERCEPTIONS OF PRACTITIONERS ON EVIDENCE BRIEFINGS

The perceptions here come from a survey (Survey 1 - S1) with 32 Software Engineering (SE) practitioners who asked questions on Stack Exchange websites related to the same topics of 12 Evidence Briefings (EBs) we have created based on one systematic review each.

We asked the practitioners how often do they refer to Stack Exchange websites (S1Q5) and SE research papers (S1Q6) to support their decision-making. Figure 1 presents the results.

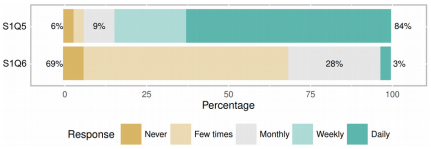


Figure 1: Frequency of mediums to acquire knowledge.

• **Practitioners do not read research papers often**: 84% of the practitioners use Stack Exchange websites on daily or weekly basis, while almost 70% have never read a research paper, or read only few times.

• **Practitioners consider software engineering research important**: 62% of the practitioners affirmed that the research presented in the EBs are important or very important.

We asked the practitioners if they agree it is easy to find information in the EBs (S1Q15), if the EBs interface is clear (S1Q16), and if the EBs look reliable (S1Q17). Figure 2 presents the results.

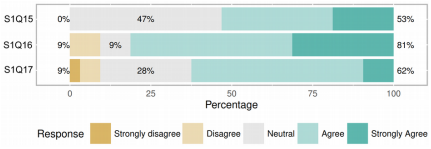


Figure 2: Perceptions of practitioners on evidence briefings format.

• **Practitioners are positive towards the evidence briefings format**: 53% of the practitioners agree or strongly agree that it is easy to find the information in the EBs (S1Q15), 81% agree or strongly agree that the EBs interface is clear (S1Q16), and 62% agree or strongly agree that the EBs look reliable.

We believe that the great impression about the EBs format is due to the well-known design principles we used for creating them, namely those from Gestalt and Information Design.

PERCEPTIONS OF RESEARCHERS ON EVIDENCE BRIEFINGS

The perceptions here come from a survey (Survey 2 - S2) with 7 authors (researchers) of the 12 systematic reviews we created Evidence Briefings (EBs) based on.

• **Researchers consider it is very important to share research knowledge with practitioners, but not all of them do so:** 100% of the researchers answered it is very important. However, only 43% share knowledge in weekly basis, and none of them in daily basis.

• **Researchers consider that the evidence briefings we created well cover their original research papers:** 72% of the researchers describe as good or very good. This shows that even though we are not the authors of the research papers, we were capable of creating, on the worst scenario, acceptable EBs.

We asked the researchers if they agree it is easy to find information in the EBs (S2Q6), if the EBs interface is clear (S2Q7), and if the EBs look reliable (S2Q8). Figure 3 presents the results.

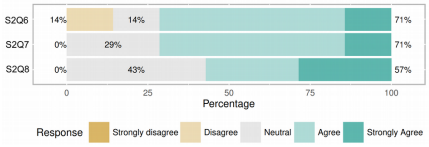


Figure 3: Perceptions of researchers on evidence briefings format.

• **Researchers, as practitioners, are positive towards the evidence briefings format:** 71% of them agree or strongly agree it is easy to find information in the EBs (S2Q6), 71% agree or strongly agree the EBs interface is clear (S2Q7), 56% agree or strongly agree the EBs look reliable (S2Q8).

PERCEPTIONS OF RESEARCHERS ON CREATING EVIDENCE BRIEFINGS

The perceptions here come from a survey (Survey 3 - S3) with 44 researchers who created, themselves, Evidence Briefings for their papers published in four software engineering conferences (EASE, SBES, SBCARS, SAST).

• **The effort to create evidence briefings is non-trivial but researchers are willing to create them**: 86% of the researchers reported to have taken up to three hours to create their EBs. When asked how likely they would create EBs, if they are common place in SE conferences, yet not mandatory, 88% answered likely or very likely.

• **Researchers consider scientific writing is an effective way to transfer knowledge to practice, but few practitioners read research papers**: 38% of the researchers consider that traditional research paper is the best way to transfer knowledge to practice, but on S1 we could se that practitiers rarely consume research papers.

IN SUMMARY, researchers and practitioners are very positive about Evidence Briefings, which might show they can play an important role as knowledge transfer medium to software engineering practice.

Who is this briefing for?

Software engineering researchers who want to assess the viability of using Evidence Briefings a knowledge transfer medium.

Where the findings come from?

All findings of this briefing were extracted from three surveys conducted by Cartaxo et al.

What is included in this briefing?

Some results and discussions about the perceptions of practitioners and researchers on using Evidence Briefings as knowledge transfer mediums.

What is not included in this briefing?

The results and discussions of some questions of the three surveys were not presented here. For more details see the original research document.

To access other evidence briefings on software engineering:

http://cin.ufpe.br/eseg/evidence-briefings

For additional information about ESEG:

http://cin.ufpe.br/eseg

ORIGINAL RESEARCH REFERENCE

Cartaxo, Bruno et al. ***A Model To Transfer Knowledge From Research To Software Engineering Practice Based On Rapid Reviews And Evidence Briefings***. 2018.

RAPID REVIEWS IN SOFTWARE ENGINEERING

This briefing reports scientific evidence on the perceptions of practitioners actively participating in a process of conducting rapid reviews to produce knowledge to support their decision-making in practice.

FINDINGS

Two Rapid Reviews (RRs) were conducted in two different companies.

Company 1 is an applied-research institute in Recife, Brazil. The project develops a system that monitors reusable packages during the entire production chain — from suppliers to factories — of the automotive industry. Two members of the project participated on the RR, the projects coordinator (the manager of all project managers in the company), and the project manager. The problem they were facing was related to low customer collaboration, so the RR aimed to search for strategies to deal with it.

Regarding the Company 2, it is a software development company also in Recife, Brazil. It develops educational software products as well as musical software. The project aims to develop a platform to create and publish e-books for basic schools in the entire Brazilian territory. Two members of the project participated on the RR, the IT director, and the project manager. The problem they were facing was related to low team motivation, so the RR aimed to search for strategies to deal with it.

PERCEPTIONS INTRODUCING RAPID REVIEWS

Here there are the perceptions of the participants of Company 1 and 2 regarding the introduction of the RRs in their practical environment. Their perceptions were obtained through interviews after the workshop we presented and discussed the RRs results.

Table 1 presents the positive points raised by the participants of Company 1 and 2 about their experience participating on a RR conducted to support their decision-making regarding a problem they were facing in practice.

POSITIVE POINTS	COMPANY	
	1	2
Applicable to software engineering practice	X	X
Novel approach to support decision-making	X	X
Offer reliable content	X	X
Increase team confidence	X	X
Reduce time and cost for decision-making	X	X
Fast and easy way to find information	X	X
Interest in consuming Evidence Briefings regularly	X	X
Fostered the learning of new concepts	X	-
Problem-oriented	X	-
Improve problem comprehension	X	-
Avoid reading multiple sources	X	-
Flexible knowledge transfer medium	X	-
Non-applicable evidence can support other problems	X	-
Practitioners are willing to embrace rapid reviews, even considering the inherent limitations	-	X
The participation of practitioners on the rapid review is important	-	X

Table 1: Positive points of RRs according to practitioners of Company 1 and 2.

• **Offer reliable content**: The participants considered the evidences provided by the RRs more reliable than the information they use to consume. This is crucial from the scientific point of view. Even though we acknowledge RRs may be not as reliable as full systematic reviews, from the eyes of the practitioners, they are more reliable than expert opinion or informal sources, which they often use to support their decision-making.

• **Fostered the learning of new concepts**: The participants of Company 1 affirm to have learned new concepts with the findings of the RR.

• **Problem-oriented**: The participants of Company 1 reported that the RR is problem oriented, since it supports decision-making for a specific problem they are facing.

• **Improve problem comprehension**: The participants of Company 1 noted the RR helped them to better comprehend the problem they were facing.

Table 2 presents the negative points raised by the participants of Company 1 and 2 about their experience participating on a RR conducted to support their decision-making regarding a problem they were facing in practice.

NEGATIVE POINTS	COMPANY	
	1	2
Discussing the rapid reviews' findings is needed	X	-
Primary studies context far from their findings	X	-
Evidence briefing printed in black-and-white	X	-
Absence of graphical information	X	-
Absence of researchers recommendation	-	X

Table 2: negative points of RRs according to practitioners of Company 1 and 2.

• **Discussing the rapid reviews' findings is needed**: The participants of Company 1 mentioned that some findings of the RR only became clearer after the discussion during the workshop. This close connection helped to clarify unclear items while reading the EB. Thus, it is important to use artifact mediums (e.g. EBs) together with human-intensive mediums (e.g. face-to-face discussions) in order to effectively transfer knowledge.

• **Absence of researcher recommendation**: The participants of Company 2 demanded a conclusion section in the EB, briefly pointing and discussing the strategy we (the researchers) recommend to deal with their problem.

PERCEPTIONS ADOPTING RAPID REVIEWS EVIDENCE

Here there are the perceptions of the participants of Company 1 regarding the adoption of the RR's evidence in their practical environment. Their perceptions were obtained through interviews two months after the introduction of the RRs.

The participants mentioned the RR was positive when it comes to mitigating their main problem. They affirmed to use the EB as a **reference material** many times throughout the weeks after the introduction of the RR. They also used the EB to **discuss with team members** how to properly deal with their problem. More importantly, the participants applied the knowledge provided by the RR.

IN SUMMARY, practitioners are very positive about RRs, which might show they can play an important role in knowledge transfer in software engineering.

Who is this briefing for?

Software engineering researchers and practitioners who want to assess the viability of using Rapid Reviews to produce knowledge to support decision-making in software engineering practice.

Where the findings come from?

All findings of this briefing were extracted from two action researh conducted by Cartaxo et al.

What is included in this briefing?

Some results and discussions about the perceptions of practitioners on using Rapid Reviews to produce knowledge to support decision-making in software engineering practice.

What is not included in this briefing?

Some results and discussions of the two action research were not presented here. For more details see the original research document. Additionally, there are no evidence about the perceptions of researchers about Rapid Reviews.

To access other evidence briefings on software engineering:

http://cin.ufpe.br/eseg/evidence-briefings

For additional information about ESEG:

http://cin.ufpe.br/eseg

ORIGINAL RESEARCH REFERENCE

Cartaxo, Bruno et al. ***A Model To Transfer Knowledge From Research To Software Engineering Practice Based On Rapid Reviews And Evidence Briefings***. 2018.