# Evidence Briefings: Towards a Medium to Transfer Knowledge from Systematic Reviews to Practitioners

Bruno Cartaxo*†, Gustavo Pinto‡, Elton Vieira†, Sérgio Soares*
*Federal University of Pernambuco - Recife, Pernambuco, Brazil
†Federal Institute of Pernambuco - Olinda, Pernambuco, Brazil
‡Federal Institute of Pará - Santarém, Pará, Brazil
{bfsc, scbs}@cin.ufpe.br, gustavo.pinto@ifpa.edu.br, elton.vieira@olinda.ifpe.edu.br

## ABSTRACT

**Context**: Integrate research evidence with practice is one of the main goals of evidence-based software engineering. However, recent studies show that the connection between systematic reviews and practitioners has not fully established.
**Goal**: This paper presents the first steps towards a medium to transfer knowledge acquired from systematic reviews to practitioners.
**Method**: We selected a set of systematic reviews identified by a tertiary study and extracted their findings to generate one-page Evidence Briefings to serve as mediums. A design specialist defined the briefings structure based on information design and gestalt principles. To evaluate the *format* and *content* of the briefings we conducted personal opinion surveys based on two groups: StackExchange users that posted questions in topics related to the reviews, and the authors of the selected reviews themselves. The former had a response rate of 21.9% (32 out 146) and the latter 31.8% (7 out of 22).
**Results**: Practitioners rarely use systematic review research papers as mediums to acquire knowledge, since just 9% have told to do so. Both researchers and practitioners positively evaluated the evidence briefings, since 71% and 82% of the StackExchange users and systematic review authors, respectively, agreed or strongly agreed that the briefings' interface is clear.
**Conclusions**: Researchers and practitioners were positive about the content and format of the evidence briefings we proposed. It is also possible to say that there is a gap between practitioners and systematic reviews due to the low percentage of practitioners that consume systematic reviews. The good reception of the evidence briefings from both sides show a possible route to reduce that gap.

## Keywords

Software Engineering; Evidence-Based Software Engineering; Systematic Reviews; Evidence Briefings



*ESEM '16 Ciudad Real, Spain*



DOI: 10.1145/1235

## 1. INTRODUCTION

Inspired by the recent and promising results of Evidence-Based Medicine (EBM), in 2004 Kitchenham and colleagues coined the Evidence-Based Software Engineering (EBSE) research field [30]. According to the authors, the goal of EBSE is to provide the best research evidence to practice by means of secondary studies such as Systematic Literature Reviews, Systematic Mapping Studies and Meta-Analyzes — hereafter generically referred as Systematic Reviews. Since then, more than 10 years have passed and EBSE has spread and grown as a mature practice with many secondary studies been conducting in a regular basis.

However, some studies suggest that the goal of EBSE is not being fully achieved [10, 21, 40]. For instance, Hassler *et al.* [21] found that the lack of connection with industry is one of the top barriers of systematic reviews. In a survey with 44 authors of 120 systematic reviews, it was found that only 6 of them had direct impact on industrial practice [40]. In a tertiary study of da Silva *et al.* [10], the authors found that only 32 out 120 of systematic reviews provide guidelines to practitioners.

Although EBM be a more mature practice, it also suffered from similar problems on its early years and nowadays still suffers in a certain degree [20]. To mitigate this problem, EBM researchers developed methods with close relationship with practitioners [20]. Those studies have been called as *Rapid Reviews*. They not only use strategies to reduce effort on some steps of the traditional systematic reviews, but also provide more appealing mediums to transfer knowledge than the traditional research paper. Those mediums are in form of brief documents with few pages focusing only on the findings that are useful to practice. The growing interest about Rapid Reviews can be observed in a recent scoping study by Tricco et al. [46], which mapped 100 rapid reviews published between 1997 and 2013, and detected that 56% of them were published on the last five years of investigation. The effectiveness of those reviews can be noted in the results of a survey conducted by Hailey [17], that examined the use of 15 rapid reviews in practice, revealing that 67% of them were used as reference material and 53% were used to directly apply their results for decisions in practice. Other studies have also demonstrated positive impact of Rapid Reviews and their mediums on practice [3, 18, 33, 51].

Considering the recent and promising results of Rapid Reviews in medicine, we believe that this method can also play an important role in software engineering. In particular, **the goal of this research is** *to present, discuss and evaluate a medium to transfer knowledge acquired from systematic re-*

*views to software engineering practitioners.* The medium is what we call "Evidence Briefing", which is an one-page document, extracted from a systematic review, that contains findings useful for practitioners.

To achieve this goal, we first selected a representative set of systematic reviews identified in the tertiary study conducted by da Silva *et al.* [10], extracted their findings, and generated one-page Evidence Briefings per systematic review. To create high-quality briefings, a design research specialist defined the structure of the briefings following well-established design principles (*e.g.*, information design and gestalt principles). To evaluate the format and content of the Evidence Briefings we conducted two personal opinion surveys: one with the authors of the selected systematic reviews, chiefly researchers, and another with users of StackExchange communities that posted questions related to the topics of those systematic reviews, chiefly practitioners.

We chose StackExchange to contact practitioners due to its well-known high quality professional Questions & Answers (Q&A) communities [32]. Not only increasingly useful for software practitioners, its active and expert community has been the target of many empirical software engineering studies [36, 32, 39, 38]. Also, StackExchange communities have very diverse participants distributed on many countries around the globe, with wide variety of ages, genders, and professional experience as can bobserved in their last annual survey [1].

Next, we highlight the main contributions of this paper:

- We *propose* Evidence Briefings, one-page documents used as mediums to transfer knowledge acquired from systematic reviews to practitioners.
- We *evaluate* Evidence Briefings, in terms of content and format, with two surveys conducted with researchers and practitioners.
- We *open-source* the template used to generate Evidence Briefings. Therefore, it can be used by any researcher who conducts systematic reviews and wants to share the findings with practitioners.

## 2. RELATED WORK

***Rapid Reviews and Briefings:*** In EBM, Rapid Reviews use strategies to reduce effort on parts of the traditional systematic review method, such as: limit the literature search, use just one person to screen studies, not conduct quality appraisal, or present results with no formal synthesis. Some comparisons between rapid and systematic reviews have been made but are not conclusive since some researches claim that the results do not differ substantially [8, 5], while others affirm that conflicting results were observed [47]. Researches discovered that rapid reviews have positive impact on practice. For instance, the study of McGregor and Brophy [33] detected that rapid reviews saved approximately $ 3 millions when implemented in a hospital. Another interesting finding comes from the Tricco's scoping study [46] which shows that, although rapid reviews present several variations on their methods and terminologies, 78% of them present results as a narrative summary presented to practitioners in mediums that better fit their needs than the traditional research paper. For instance, the Briefings presented by Chambers and Wilson [9], the Evidence Summaries by Khangura *et al.* [29], the Contextual Summaries by Young *et al.* [50] and others.

The Evidence Briefings proposed in this study are more similar to the one-page summaries of Young et al. [50], since we generated each of them based on the main findings extracted from selected systematic reviews, while rapid reviews usually are conducted from scratch with a specific practitioner need in mind, adapting the systematic review method to fit time constraints, which is not our case.

***Empirical Studies based on StackExchange:*** Many studies use StackExchange communities to explore the state of practice in software engineering. For instance, Wang and Godfrey [49] detected iOS and Android API usage obstacles based on questions of a StackExchange community; Pinto and Kamei [37] discovered flaws and desirable features of refactoring tools from questions of StackExchange communities; or Vasilescu *et al.* [48], which investigated the interplay between StackExchange activities and the development process. However, to the best of our knowledge, none of such studies tried to identify practitioners needs and link them with actual research to share findings with practitioners, like we did in this work.

***Knowledge Transfer in Software Engineering:*** Several authors have been proposing ways to disseminate research results with practitioners. Graham *et al.* [15] identified various concepts used to share knowledge, such as *knowledge translation* and *knowledge transfer*. In software engineering, Budgen *et al.* [7] identified key issues to be addressed in order to apply *knowledge translation* and Grigoleit *et al.* [16] reported experiences of *knowledge transfer* using mediums between the ones that provide the knowledge (transferors) and the ones that receive it (transferee). They mention that mediums can be artifacts, like publications and documents, or human intensive mediums, like conferences and workshops. There are still other researches intending to make empirical evidence more connected with practice, like the one of Jedlitschka and colleagues [25] which proposed and evaluated a model to report software engineering experiments according to practitioners' needs.

Our study has similarity with the concept of artifact medium as stated by Grigoleit *et al.* [16], since Evidence Briefings are documents intended to transfer knowledge acquired from systematic review to practitioners. The one-page Evidence Briefings we are proposing are also in accordance with one of various recommendations of Beecham et al. [4] to make software engineering research more relevant to practice, which is to write shorter evidence-based reports.

## 3. METHOD

In this section we first discuss how we selected the systematic reviews in 3.1. Then we extracted data from those reviews as shown in 3.2. With the extracted data we generated Evidence Briefings for each review as presented in 3.3. Finally we applied personal opinion surveys to evaluate the content and format of the Evidence Briefings we generated, as detailed in 3.4.

### 3.1 Systematic Reviews Selection

In order to select a representative set of reviews, we started with 120 systematic reviews identified in the tertiary study of Da Silva *et al.* [10]. We first reduced that amount to the only 32 reviews classified by the authors of the tertiary

[1] stackoverflow.com/research/developer-survey- 2016

study as having guidelines to practitioners. Afterwards we decreased to just 24 reviews, since some of them do not present their search string, which are crucial to search for practitioners' questions on StackExchange communities. In the end, we randomly selected half of the 24 systematic reviews to generate Evidence Briefings. Table 1 shows the 12 selected systematic reviews.

## 3.2 Systematic Reviews Data Extraction

To generate the Evidence Briefings, we extracted the following data from each 12 selected systematic reviews: paper title, research goal, research findings, and paper references.

We simplified the **paper titles** in order to make the briefings more appealing to practitioners. However, we put an effort to not compromise the meaning. In many cases we just removed terms that explicitly mention the research method (*e.g.,* a systematic review or a meta-analysis). As an example, we generated a briefing entitled "*The effectiveness of pair programming*" from the paper "*The effectiveness of pair programming: A meta-analysis*" [19].

The **research goals** were re-written, without change semantics, to fit the following template: "*This briefing reports evidence on <GOAL> based on scientific evidence from a systematic review.*" For example, the goal of the briefing generated from paper [28] 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 maintained, as much as possible, the **research findings** as the authors of the systematic reviews have written. Modifications were made only to assure the briefings' fluency of reading. Information like references or mentions to primary studies were removed. It is also important to mention that we only extracted findings emerged from secondary studies processes. Thus, we have not extracted findings from primary studies processes as is the case of the paper from Haugset and Hanssen [22] that report a systematic review but also an industrial case study (*i.e.,* a primary study).

The **papers references** were extracted as usual, from their publishers' websites.

## 3.3 Evidence Briefings Generation

To strengthen the method of this research and also offer professional high quality briefings we invited a design researcher specialist, Elton Viera, that was in charge of the Evidence Briefings generation based on the data extracted from the selected systematic reviews. An important decision was understanding the briefings as a graphic product and uses the principles of *Information Design* [44] and Gestalt theory [31]. The primary objective is to develop documents that are comprehensible, accurately retrievable, natural, and pleasant. Since an important phase of design methodology is to verify the best practices that could be inherited from others projects [2], the graphic design was initially based on some rapid reviews mediums from medicine. For instance, we limited our Evidence Briefings to one page, like Young et al. [50] with their contextual summaries; we summarized the main findings in one section, like the briefings proposed by Chambers and Wilson [9]; and we also used a informative box separated from the main text to highlight the audience and nature of the briefings' content, like the evidence summaries of Khangura and colleagues [29].

Figure 1 shows numbers within squares denoting parts that composes the structure of the briefings. (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 extracted from the original systematic review. (4) Informative box that outlines the intended audience and explains the nature of the briefings' content. (5) The reference to the original systematic review. (6) The logos of our research group and university.

Figure 2 shows numbers within circles denoting where Tondreau's principles of *Information Design* [44] were applied to generate the Evidence Briefings. (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 wide recognized, that is why we used Calibri, a friendly reading sans serif type that is wide recognized as been 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. (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 3 shows numbers within diamonds denoting where *Gestalt Principles* [31] were applied to develop an efficient graphic design, through perceptual organization of visual elements. (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. (4) *Unity* principle which states that elements that have a visual connection should belong to a uniform group.

All concepts obtained in the previous steps were analyzed and applied in a template. All the 12 evidence briefings as well as the briefing's template can be found in http://cin.ufpe.br/eseg/briefings. We encourage researchers who want to share systematic reviews' findings with practitioners to use the briefing's template we developed. The evidence template is open-sourced under CC-BY license.

## 3.4 Evidence Briefings Evaluation

We evaluated the evidence briefings according to their *content* and *format* surveying two distinct groups: the users of StackExchange that asked questions related to the selected systematic reviews, chiefly practitioners, and the authors of those reviews, chiefly academics. Subsection 3.4.1 presents methodological details about the first survey, whereas Subsection 3.4.2 discusses the second one.

### 3.4.1 Survey with StackExchange Users

The **goal** of this cross-sectional survey is to acquire evidence on how StackExchange users that asked questions related to the selected systematic reviews perceive the content and format of the briefings.

As **instrument** for this survey we created a questionnaire mixing open and closed questions that we divided in four sections. The first with demographic questions that helps us to understand the characteristics of our sample; the second

**Table 1: Selected systematic reviews topics.**

| REF. | TITLE | TOPIC |
|---|---|---|
| [22] | Automated Acceptance Testing: a Literature Review and an Industrial Case Study | Software Testing |
| [27] | Critical Barriers for Offshore Software Development Outsourcing Vendors: A Systematic Literature Review | Global Software Development |
| [28] | Critical Success Factors for Offshore Software Development Outsourcing Vendors: A Systematic Literature Review | Global Software Development |
| [34] | Definitions and approaches to model quality in model-based software development – A review of literature | Model-Based Software Development |
| [14] | Empirical studies of agile software development: A systematic review | Agile Software Development |
| [45] | Factors Influencing Software Development Productivity – State-of-the-Art and Industrial Experiences | Software Development Productivity |
| [26] | Forecasting of software development work effort: Evidence on expert judgement and formal models | Cost and Effort Estimation |
| [23] | Harmfulness of Code Duplication - A Structured Review of the Evidence | Code Duplication |
| [6] | Knowledge management in software engineering: A systematic review of studied concepts, findings and research methods used | Software Engineering Knowledge Management |
| [35] | On the generation of requirements specifications from software engineering models: A systematic literature review | Software Requirements |
| [19] | The effectiveness of pair programming: A meta-analysis | Pair Programming |
| [24] | Using Scrum in Global Software Development: A Systematic Literature Review | Global Software Development |

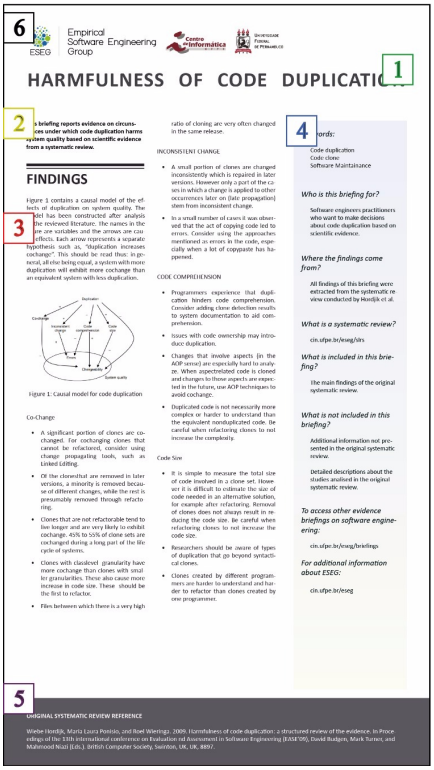


**Figure 1: Overview of each part that compose the briefings structure.**

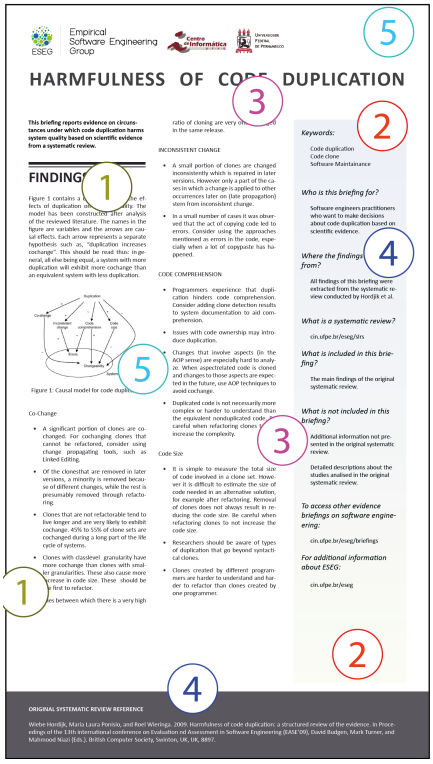


**Figure 2: Overview of the briefings and the elements of Design Information and Grids [44]**

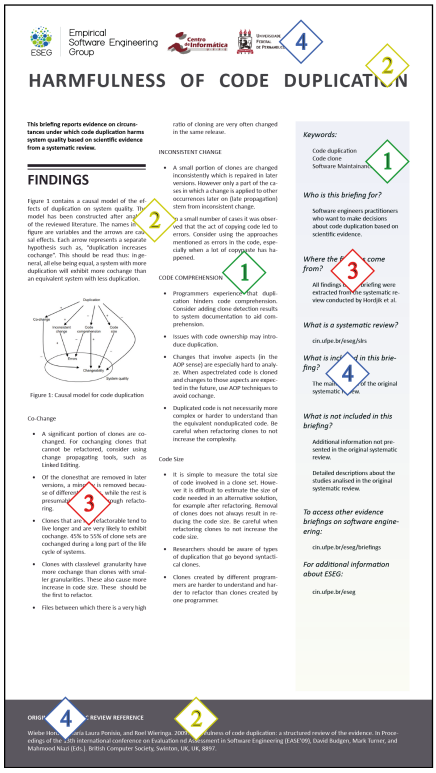


**Figure 3: Overview of the briefings and the elements of gestalt theory [31]**

with questions to discover what mediums practitioners use to acquire knowledge; the third with questions to discover the perceptions about the briefings contents; and the fourth to examine the perceptions about the briefings format. Due to space limitations the full list of questions of this questionnaire is available on http://bit.ly/1Xr0ivL. Hereafter each question of this survey will be referred following this pattern: `USQ<N>`, where `US` stands for `U`ser questionnaire, and `N` refers to the number of the question in the questionnaire. For instance, the first question of this questionnaire is represented as `USQ1`.

The target **population** of this survey is the 473 StackExchange users who asked questions related to one of the 12 selected systematic reviews. To find these related questions, we applied their search strings into StackExchange communities that discuss software engineering topics. We use their search string instead of, for instance, the title of the study, because search strings tend to have more key terms than the title, which is usually short and objective. We selected communities that are related to at least one topic of software engineering as presented in the SWEBOK [1]. Following are the five selected communities and their description as appear in their own websites:

- *Programmers* (PROG)[2], intended to conceptual questions about software development;
- *Quality Assurance & Testing* (SQA)[3], intended to software quality control experts, automation engineers and software testers;
- *Project Management* (PM)[4], intended to project managers;
- *Reverse Engineering* (RE)[5], intended to researchers and developers who explore the principles of a system through analysis of its structure, function, and operation;
- *Software Recommendations* (SR)[6], intended to people seeking specific software recommendations.

The initial search on StackExchange communities returned 1,738 questions related to the selected systematic reviews. We manually analyzed and excluded those that are not related to the reviews (false-positives), resulting in the 473 questions that compose the population of this survey. We consider questions as not related when they are out of the topic of the systematic review. We used open card sort technique to define topics for each review [43]. Table 1 present each selected systematic review together with their defined topics. In order to effectively understand if a question is related or not to the review topic, we analyzed each one of the 1,738 questions in pairs. The Kappa value was 0.72, which means *Substantial Agreement* according to Kappa's reference table. Also, we had meetings for conflict resolution in order to avoid classification bias. The final classification of all questions, after the conflict resolution meetings is in a spreadsheet on http://bit.ly/1M3cZvY. We classified the questions based on their title and body, but the spreadsheet has only the title of each question. Thus, one can use the question's ID to look the body of any question in the online websites.

Our intention was to invite all users to answer the survey. However, StackExchange communities do not provide a way to directly contact their users. Thus, we could invite just users that provide contact information in their StackExchange public profile, which were 146 (30.8%) out of the 473. The contact platforms varied from e-mail, Facebook, Google+, Twitter, LinkedIn, GitHub, and personal websites. As an ethical consideration, we contacted only those users that provided contact information on their public profiles, following the privacy policies established by StackExchange [7]

Our **sample** is composed by a total of 32 StackExchange users that responded the questionnaire. This corresponds to 21.9% of the 146 invitations. This response rate is also considerably above what Singer *et al.* [41] mentioned, as well as within the range of others software engineering surveys.

In order to improve response rate, we employed majority of the principles listed by Smith *et al.* [42]. We applied the *reciprocity* principle offering a raffle of 100 USD gift on Amazon for the respondents. We also employed the *authority and credibility* principle by presenting ourselves as researchers with PhD in Computer Science and University professors. We used the *liking* principle when we sent personalized e-mails using the participants' names and mentioning that we selected the most relevant questions on StackExchange communities and that their questions were ones of them. The *scarcity* principle was used when we defined a deadline of two weeks to complete the questionnaire. The *brevity* principle was done by asking closed and direct questions as much as possible. The *social benefit* principle was used by highlighting the importance to understand and reduce the gap between research and practice and also when we advertised that a donation of 1 USD for the Brazilian Red Cross would be performed for each participant, inspired on the interesting results reported by de Mello *et al.* [11].

[2]programmers.stackexchange.com
[3]sqa.stackexchange.com
[4]pm.stackexchange.com
[5]reverseengineering.stackexchange.com
[6]softwarerecs.stackexchange.com
[7]http://stackexchange.com/legal/privacy-policy

### 3.4.2 Survey with Systematic Review Authors

The **goal** of this cross-sectional survey is to acquire evidence on how the authors of the selected systematic reviews perceive the content and format of the briefings we generated based on their reviews.

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 the authors of the selected systematic reviews are interested in share research results with practitioners; the second with questions to discover the perceptions about the briefings contents; and the third focused on the perceptions about the briefings format. Due to space limitations the full list of questions of this questionnaire is available on http://bit.ly/1nJzI4F. Hereafter each question of this survey will be referred following this pattern: ASQ<N>. For instance, the first question of this questionnaire will be referred as ASQ1, and so forth.

The target **population** of this survey is the 22 authors of the 12 selected systematic reviews.

Our **sample** is composed by 7 authors that answered the questionnaire, which corresponds to 31% of the 22 invitations. We again believe this is a good response rate since Singer *et al.* [41] found that questionnaires in software engineering exhibit consistent low rates around 5%.

# 4. RESULTS

This section reports the results of the Survey with StackExchange Users in Subsection 4.1 and the results of the Survey with Systematic Reviews Authors in Subsection 4.2.

## 4.1 Survey with the StackExchange Users

We highlight demographic information about the subjects in Subsection 4.1.1. Discuss with what mediums they acquire knowledge in Subsection 4.1.2. And present results related to the briefings content in Subsection 4.1.3 and format in Subsection 4.1.4.

### 4.1.1 Demographics

Among the StackExchange 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 (USQ1). 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 (USQ2). Most of them work for the software industry (87.5%), 9% work for open source initiatives, 9% for the government, and 6% are academics (USQ3). They are highly educated — 10% have a Ph.D. degree, 48% have a Master degree, and 35% have a Bachelor degree (USQ4).

### 4.1.2 Mediums to acquire knowledge

We asked respondents how often do they refer to StackExchange websites and software engineering research papers to solve their problems. Figure 4 presents the results.

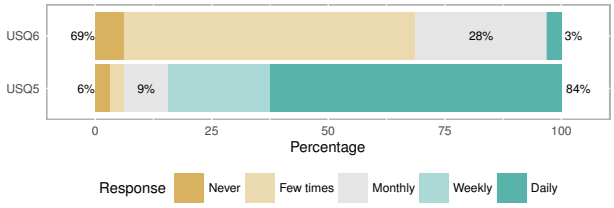


**Figure 4: How often do software engineers refer to StackExchange websites (USQ5) and research papers (USQ6) (%).**

As observed, 62% of the respondents use StackExchange websites on daily basis and another 22% use it weekly (USQ5). On the other hand, we found that 62% of the respondents said that they have read only a few software engineering research papers, but it is not common, and 6% of the respondents had never read a single software engineering research paper (USQ6). We also found that only 28% of those who have read software engineering research papers have read a systematic review paper (USQ7). Among them, 44% did it for research purposes, 33% for decision making on practice, and 11% for teaching (USQ8). That means that only 9% (3 out of 32) of all StackExchage users that participated in this survey read systematic review papers to help them on decision-making processes. Therefore, we believe that software engineering researchers are in need of better ways to disseminate their research findings to practitioners. This fact motivates us to create and evaluate our Evidence Briefings in terms of content and format.

Finally, 78% of the respondents believe that a platform such as StackExchange to discuss briefings of software engineering research is "Very important" or "Important" (USQ9).

### 4.1.3 Briefing Content

We asked six questions to evaluate the briefing's content. 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 StackExchange?" (USQ10). Among the answers, 10% said that the briefing has *totally* answered, and another 20% said that it has *partially* answered their StackExchange questions. Another 32% said that the briefing touches a related topic, but does not help to answer the question. The remaining 38% said that the briefing is not related to the question and, therefore, it does not help to answer it. In the following question (USQ11), we asked respondents that said that the briefing does not *totally* answered their StackExchange question to describe why it is the case. We categorized their response into five main reasons:

- **The question is too specific:** Three StackExchange users reported that their questions are highly specific whereas the systematic review 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 (6342, from the PM community) the StackExchange 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 briefing is capable of answering the technical side of the questions, it does not answer the human side of them. In the survey, the user recognized it, saying that "*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 briefing 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:** We found that six StackExchange users reported that the briefings are lacking 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 that the briefing has nothing to do with the question asked, which lead us to expected situations when StackExchange users do not share the same perception of us about the topic of the systematic reviews and, in consequence, the briefings' topics.

Due to the high rate of users which affirmed that their questions are not answered and also are considered not related to the briefings, we took a moment to investigate why it occurs. First, we found that four of those questions can be also judged as related to the subject. For instance, question 199021, from the PROG community, asked "Why is Feature Driven Development considered an Agile methodology?", and the briefing related to this question is about agile methods [14], 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/disavantages 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?" (USQ12). We found that 62% of the respondents said that the researches presented in the briefings are "Very important" or "Important". Also, 25%, 6% believe they are "Moderately important" and "Slightly Important" respectively. The remaining 6% believe they are "Unwise". Here we can observe a paradox between the high rate (62%) of practitioners that consider the researches important and the low rate (28%) of these same researches that could help practitioners answer their questions (USQ11). This finding suggests that if we as researchers want to produce evidence more useful to practice, it is important to focus not only on important research but also on research that answer practitioners' questions.

The respondents that have answered "Unwise" were asked to describe their reasons in question USQ13. Two of them did so. The first one said that "*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 StackExchange 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 one said that "*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 that no research finding would easily change his/her mind. This lead us to areas of software engineering that are based on strong beliefs, and as verified by Devambu and colleagues [12], strong evidence is required to change strong beliefs.

Also, 62% of the respondents prefer the answers of StackExchange communities over the findings presented in the briefings (USQ13). One of the reasons that might explain this is the gamification mechanism employed on StackExchange communities, which motivates one to contribute [13], in contrast to the static-based form of Evidence Briefings. Yet, 78% of respondents leaved the doors opened when said that a platform like StackExchange to discuss research briefings are "Important" or "Very Important" is encouraging (USQ9). 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 software engineering.

### 4.1.4 Briefing Format

In this final set of questions, we asked respondents what they think about the structure of the briefings. Figure 5 shows the overall impression.

First, we observed that 53% of the respondents "Agree" or "Strongly Agree" that it was easy to find the information in the briefing (USQ15), and 82% of the respondent "Agree" or "Strongly agree" that the briefing interface is clear and understandable (USQ16). For these two questions, we believe that the great impression about the briefing interface is due to the well-known design principles we used for creating them, namely those from Gestalt and Design Information theories. With those principles, we believe others can replicate our briefing with high success rate. Second, we found that 62% of the respondents "Agree" or "Strongly agree" that the briefings look reliable (USQ17). Besides the research findings available in the briefing, we highlight the original paper in the footnote. In addition, in order to increase reliability, we put institutional logos from our research group, department, and university on the briefings. Institutional webpages are also provided.

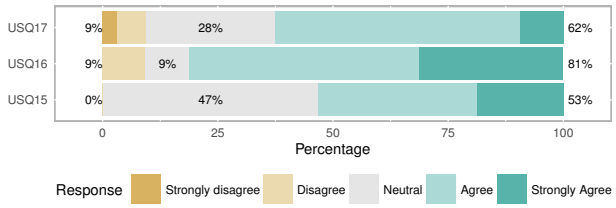


**Figure 5: The impression that practitioners had about the briefing's format (%).**

## 4.2 Survey with Systematic Reviews Authors

In this survey, we asked the authors of the selected systematic reviews their opinions about the briefings. We start asking "How important for you is to share research results to practitioners?" (ASQ1). We found what 100% of the respondents said that it is "Very important". However, 29% (2) of the them said that they have shared a few times only, but it is not common. The remaining ones share research findings on weekly (42% - 3) or monthly (28% - 2) basis (ASQ2). For those who shared research findings with practitioners, we asked how they do so (ASQ3). The respondents reported five main ways for sharing knowledge:

- Teaching;
- Seminars (*e.g.,* Network meetings, Informal meetings, Conferences);
- Writing (*e.g.,* Newspaper/magazine writing, Distributing research reports);
- Advisory work (*e.g.,* supervising practitioners, research with practical case studies);
- Social networks (*e.g.,* Twitter and ResearchGate);

Next we asked "How does the briefing cover the main findings of your paper?" (ASQ4). We found that 72% (5) of the respondents describe as "Very good" or "Good". The remaining 28% (2) said that it is "Acceptable". This shows a good impression of authors of systematic reviews, and suggests that even though we are not the authors of the research papers, we were capable of creating, at least, acceptable briefings.

The next and final group of questions is the same to the ones discussed in Section 4.1.4. Figure 6 shows the overview of the answers.

As we can see from this figure, similarly to what we found in the practitioners survey, there is a consensus around the briefing interface. We observed that 71% (5) of them "Strongly agree" or "Agree" that it is easy to find information in the briefings (ASQ6). Another 71% (5) "Strongly agree" or "Agree" that the briefing interface is clear and understandable (ASQ7).

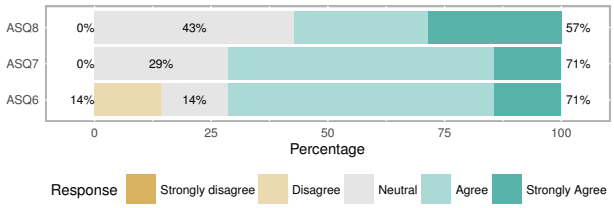


**Figure 6: The impression that researchers had about the briefing's format (%).**

Finally, 56% (4) "Strongly agree" or "Agree" that the briefings look reliable (ASQ8).

# 5. DISCUSSION

In this section we revisit the findings of this study (Section 5.1). Next, we provide additional insights on the data presented in the previous sections (Section 5.2), rise some implications of this research (Section 5.3), and also provide some limitations of this study (Section 5.4).

## 5.1 Revisiting Findings

***Industry focused, highly educated and considerably experienced.*** These are the main characteristics of the StackExchange users that responded our survey. This finding suggests that StackExchange communities can be an interesting source to investigate the needs of highly qualified and experienced practitioners.

***Practitioners rarely use research papers as mediums to acquire knowledge.*** In fact, practitioners hardly ever use systematic reviews to aid them in decision-making process. This shows that if we as EBSE researchers want to transfer knowledge to practice we need to find alternative mediums along with research papers.

***Important researches are not necessarily useful to practice.*** This finding shows that practitioners perceive the importance and relevance of basic and theoretical research, but recognize that they are hard to directly apply in their contexts. Thus, if we as EBSE researchers want to provide evidences more related with practice it is important to promote methods and mediums that fit their needs.

***Software engineering practice still has many beliefs with no evidence basis.*** Thus, if we want to promote changes in areas soaked by strong beliefs, we need to provide strong evidences, which favor multiple, rigorous and large scale empirical studies together with appealing mediums to spread those evidences.

***Both researchers and practitioners positively evaluated the evidence briefings.*** This constitute the first indicative that Evidence Briefings can be one of the mediums to transfer knowledge acquired from systematic reviews to practice. Among practitioners, most of them agree that the briefing was easy to read, clear, understandable and reliable. Similar results were found when we asked the same questions for researchers. In addition, some researchers gone beyond just filling the survey and spontaneously sent us e-mails congratulating this study. One said *"Many thanks [...] This is a good direction! I support this."*, another said *"What a nice way to put it! Thank you for having taken a look at our paper. "* and other said '*"I think your format captures the essence really well, in a way that practitioners may find useful and understandable!"*.

***The briefings we generated well covered the main findings of the original systematic reviews.*** Only one author suggested an improvement, as mentioned *"The summary of the findings is ok, but I would try to summarize so that there is easy to read "take home" messages there"*, which leaded to minor changes. This grows confidence that the information passed through the Evidence Briefings to practitioners are accurate and fair with their originals.

## 5.2 The Yin-Yang of Research and Practice

Research and practice not always speak the same language. Here we discuss some contrary forces of this problem.

***Researchers want to transfer knowledge. But not all of them do so.*** We found that although 100% of the respondents believe that it is very important to transfer research knowledge to practitioners, only 42% of the authors surveyed do so on, at least, weekly basis. Therefore, although there is an interest, there is also a lack of appropriated infrastructure to help researchers to easily and broadly disseminate research findings. We believe that evidence briefings can help to bridge the chasm of this problem.

***Practitioners want to be more aware of software engineering research. But few of them do so.*** Although 62% of the respondents believe that the researches presented in the briefings are important, only 28% of the them read software engineering research papers on regular basis (*i.e.,* monthly at least). This finding also open space to put evidence briefings in a place to play a role as a medium to transfer knowledge from research to practice.

## 5.3 Implications

***Researchers.*** Can observe that software engineers believe in the importance of software engineering research, although are not used to refer to it. Thus, creating evidence briefings of their systematic reviews might increase the visibility and impact of their research. We also believe that researchers who want to conduct systematic reviews more practical-oriented could search for questions on StackExchange related to their reviews' topic on early stages of protocol definition in order to investigate the problems practitioners are facing and fine-tune their investigations to be more practical.

***Practitioners.*** Can take advantage of evidence briefings in different ways. For instance, fostering software engineering research results between their peers, and also discussing high-level research results that might motivate them to find additional information in the original systematic review papers.

***Tool builders.*** Our findings suggest that both researchers and practitioners are interested in creating and divulging evidence briefings. In particular, about 78% of the practitioners contacted believe that a platform such as StackExchange might be important to promoting software engineering research results. Thus, tool builders can benefit from this findings in order to create and promote specialized platforms to discuss software engineering research findings.

***Educators.*** The set of findings that we include in this paper 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 evidence briefings.

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

## 5.4 Limitations

First, we are not the authors of the systematic review reviews we generated the Evidence Briefings. Therefore, chances are that we had misunderstood or not included some findings of of the original papers. To mitigate this problem, we conducted a survey with the authors of the original systematic reviews. In particular, we had a question that explicitly asked *"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".

Second, not all StackExchange users or systematic reviews authors were contacted. The former because StackExchange 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 StackExchange public profile. The latter because not all authors still maintain the email address used in the reviews. Similarly, we tried to find their personal webpages or public profiles and also asked others co-authors whether any of their colleagues are using another email address.

Third, we do not evaluate the "bad design" case, which would be mediums with poor designed interface. Our Evidence Briefings 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 Evidence Briefings vs the Content of Evidence Briefings.

# 6. CONCLUDING REMARKS

In this paper we proposed what we call "Evidence Briefings", which is a medium to transfer knowledge acquired from systematic reviews to practitioners. We evaluate Evidence Briefings with both researchers and practitioners. Our results suggest that Evidence Briefings was positively evaluated. For instance, most of the researchers and practitioners believe that it is easy to find information on Evidence Briefings. Yet, most of them believe that they clear, understandable, and reliable. We believe that Evidence Briefings can play a role on transferring knowledge from systematic reviews to practice — currently, only 40% of the researchers respondents said that they share research results on weekly basis, and 62% of the practitioners respondents said that they have read only few software engineering research papers, but it is not common. For future research, we plan to improve our evaluation methodology by observing how practitioners use Evidence Briefings.

# 7. ACKNOWLEDGEMENTS

We are thankful to all StackExchange Users and Systematic Reviews' Authors that gently participated on the survey. We also would like to thank the anonymous reviewers for their helpful comments. Sergio Soares is partially supported by CNPq, grants 304581/2013-5 and 471381/2012-8. This work was partially supported by the National Institute of Science and Technology for Software Engineering (INES), funded by CNPq grant 573964/2008-4.

# 8. REFERENCES

[1] A. Abran, P. Bourque, R. Dupuis, and J. W. Moore, editors. *Guide to the Software Engineering Body of Knowledge - SWEBOK*. 2001.

[2] G. Ambrose and P. Harris. *Basics Design 08: Design Thinking*. Basics Design. AVA Publishing, 2009.

[3] J. Batten. Comment on editorial literature reviews as a research strategy, 2012.

[4] S. Beecham, P. O'Leary, S. Baker, I. Richardson, and J. Noll. Making software engineering research relevant. *Computer*, 47(4):80–83, 2014.

[5] L. Best, A. Stevens, and D. Colin-Jones. Rapid and responsive health technology assessment: the development and evaluation process in the south and west region of england. *Journal of Clinical Effectiveness*, 1997.

[6] F. O. Bjørnson and T. Dingsøyr. Knowledge management in software engineering: A systematic review of studied concepts, findings and research methods used. *IST*, 2008.

[7] D. Budgen, B. Kitchenham, and P. Brereton. The case for knowledge translation. In *ESEM*, 2013.

[8] A. Cameron. Rapid versus full systematic reviews: an inventory of current methods and practice in Health Technology Assessment. Technical report, 2007.

[9] D. Chambers and P. Wilson. A framework for production of systematic review based briefings to support evidence-informed decision-making. *Systematic reviews*, 2012.

[10] F. Q. B. da Silva, A. L. Santos, S. Soares, A. C. C. FranÃğa, C. V. Monteiro, and F. F. Maciel. Six years of systematic literature reviews in software engineering: An updated tertiary study. *IST*, 2011.

[11] R. M. de Mello, K. T. Stolee, and G. H. Travassos. Investigating samples representativeness for an online experiment in java code search. In *ESEM*, 2015.

[12] P. Devanbu, T. Zimmermann, and C. Bird. Belief & evidence in empirical software engineering. In *ICSE*, 2016.

[13] D. J. Dubois and G. Tamburrelli. Understanding gamification mechanisms for software development. In *Proceedings of the 2013 9th Joint Meeting on Foundations of Software Engineering*, ESEC/FSE 2013, pages 659–662, 2013.

[14] T. Dybå and T. Dingsøyr. Empirical studies of agile software development: A systematic review. *IST*, 2008.

[15] I. D. Graham, J. Logan, M. B. Harrison, S. E. Straus, J. Tetroe, W. Caswell, and N. Robinson. Lost in knowledge translation: Time for a map? *Journal of Continuing Education in the Health Professions*, 2006.

[16] F. Grigoleit, A. Vetro, D. M. Fernandez, W. Bohm, and P. Diebold. In quest for proper mediums for technology transfer in software engineering. In *ESEM*, 2015.

[17] D. Hailey. A preliminary survey on the influence of rapid health technology assessments. *International journal of technology assessment in health care*, 2009.

[18] D. Hailey, P. Corabian, C. Harstall, and W. Schneider. The use and impact of rapid health technology assessments. *International journal of technology assessment in health care*, 2000.

[19] J. E. Hannay, T. Dybå, E. Arisholm, and D. I. Sjøberg. The effectiveness of pair programming: A meta-analysis. *IST*, 2009.

[20] L. Hartling, J.-M. Guise, E. Kato, J. Anderson, S. Belinson, E. Berliner, D. M. Dryden, R. Featherstone, M. D. Mitchell, M. Motu'apuaka, H. Noorani, R. Paynter, K. A. Robinson, K. Schoelles, C. A. Umscheid, and E. Whitlockg. A taxonomy of rapid reviews links report types and methods to specific decision-making contexts. *Journal of Clinical Epidemiology*, 2015.

[21] E. Hassler, J. C. Carver, N. A. Kraft, and D. Hale. Outcomes of a community workshop to identify and rank barriers to the systematic literature review process. EASE, 2014.

[22] B. Haugset and G. Hanssen. Automated acceptance testing: A literature review and an industrial case study. In *AGILE*, 2008.

[23] W. Hordijk, M. L. Ponisio, and R. Wieringa. Harmfulness of code duplication: A structured review of the evidence. In *EASE*, 2009.

[24] E. Hossain, M. Babar, and H. young Paik. Using scrum in global software development: A systematic literature review. In *ICGSE*, 2009.

[25] A. Jedlitschka, N. Juristo, and D. Rombach. Reporting experiments to satisfy professionals' information needs. *Empirical Software Engineering*, 19(6):1921–1955, 2014.

[26] M. Jørgensen. Forecasting of software development work effort: Evidence on expert judgement and formal models. *International Journal of Forecasting*, 2007.

[27] S. Khan, M. Niazi, and R. Ahmad. Critical barriers for offshore software development outsourcing vendors: A systematic literature review. In *APSEC*, 2009.

[28] S. Khan, M. Niazi, and R. Ahmad. Critical success factors for offshore software development outsourcing vendors: A systematic literature review. In *ICGSE*, 2009.

[29] S. Khangura, K. Konnyu, R. Cushman, J. Grimshaw, and D. Moher. Evidence summaries: the evolution of a rapid review approach. *Systematic Reviews*, 2012.

[30] B. A. Kitchenham, T. Dyba, and M. Jorgensen. Evidence-based software engineering. ICSE, 2004.

[31] E. Lupton and J. C. Phillips. *Graphic Design: The New Basics*. Princeton Architectural Press, 2nd edition edition, 2015.

[32] L. Mamykina, B. Manoim, M. Mittal, G. Hripcsak, and B. Hartmann. Design lessons from the fastest Q&A site in the west. In *SIGCHI*, 2011.

[33] M. McGregor and J. M. Brophy. End-user involvement in health technology assessment (hta) development: a way to increase impact. *International Journal of Technology Assessment in Health Care*, 2005.

[34] P. Mohagheghi, V. Dehlen, and T. Neple. Definitions and approaches to model quality in model-based software development – a review of literature. *IST*, 2009.

[35] J. Nicolás and A. Toval. On the generation of requirements specifications from software engineering models: A systematic literature review. *IST*, 2009.

[36] G. Pinto, F. Castor, and Y. D. Liu. Mining questions about software energy consumption. 11th MSR, 2014.

[37] G. H. Pinto and F. Kamei. What programmers say about refactoring tools?: An empirical investigation of stack overflow. WRT, 2013.

[38] D. Posnett, E. Warburg, P. Devanbu, and V. Filkov. Mining stack exchange: Expertise is evident from initial contributions. In *SocialInformatics*, 2012.

[39] M. Reboucas, G. Pinto, F. Ebert, W. Torres, A. Serebrenik, and F. Castor. An empirical study on the usage of the swift programming language. In *23rd SANER*, 2016.

[40] R. Santos and F. da Silva. Motivation to perform systematic reviews and their impact on software engineering practice. In *ESEM*, 2013.

[41] J. Singer, S. E. Sim, and T. C. Lethbridge. Software engineering data collection for field studies. In *Guide to Advanced Empirical Software Engineering*. Springer, 2008.

[42] E. Smith, R. Loftin, E. Murphy-Hill, C. Bird, and T. Zimmermann. Improving developer participation rates in surveys. In *CHASE*, 2013.

[43] D. Spencer. *Card sorting: Designing usable categories*. Rosenfeld Media, 2009.

[44] B. Tondreau. *Layout Essentials: 100 Design Principles for Using Grids (Design Essentials)*. Design Essentials. Rockport Publishers, 2011.

[45] A. Trendowicz and J. Münch. Factors influencing software development productivity-state-of-the-art and industrial experiences. Advances in Computers. 2009.

[46] A. Tricco, J. Antony, W. Zarin, L. Strifler, M. Ghassemi, J. Ivory, L. Perrier, B. Hutton, D. Moher, and S. E. Straus. A scoping review of rapid review methods. *BMC Medicine*, 2015.

[47] S. Van de Velde, E. De Buck, T. Dieltjens, and B. Aertgeerts. Medicinal use of potato-derived products: conclusions of a rapid versus full systematic review. *Phytotherapy Research*, 2011.

[48] B. Vasilescu, V. Filkov, and A. Serebrenik. Stack overflow and github: Associations between software development and crowdsourced knowledge. In *2013 IEEE International Conference on Social Computing (SocialCom)*, pages 188–195. IEEE, 2013.

[49] W. Wang and M. W. Godfrey. Detecting api usage obstacles: A study of ios and android developer questions. In *MSR*, 2013.

[50] I. Young, J.-M. Guise, E. Kato, J. Anderson, S. Belinson, E. Berliner, D. M. Dryden, R. Featherstone, M. D. Mitchell, M. Motu'apuaka, H. Noorani, R. Paynter, K. A. Robinson, K. Schoelles, C. A. Umscheid, and E. Whitlockg. A guide for developing plain-language and contextual summaries of systematic reviews in agri-food public health. *Foodborne Pathogens and Disease*, 2014.

[51] I. Zechmeister and I. Schumacher. The impact of health technology assessment reports on decision making in austria. *International journal of technology assessment in health care*, 2012.