Evidence-based Software Engineering Revisited: Evaluation of a Practice-driven Application

Sebastián Pizard, Diego Vallespir, Barbara Kitchenham, *Member, IEEE*

***Abstract*—*Context:* Systematic reviews (SRs) are the main method for supporting evidence-based software engineering (EBSE). However, although SRs have been widely adopted by SE researchers, so far, there have been no studies assessing the value of the broader EBSE framework. *Objective:* To evaluate an EBSE application in an industry environment. *Method:* Using the participant observation method, we conducted an EBSE-based project to address an industry problem. This includes collaborating with practitioners to diagnose the problem, collect evidence through a rapid review (RR), and transfer the results to the company. *Results:* The practitioners utilized some recommendations that addressed their problems. The biggest barriers we encountered were the difficulty in finding relevant evidence, the complexity of applying evidence, and the lack of guidelines or examples of EBSE use. Factors that supported our project were: close collaboration with the company and its commitment to process improvement, appropriate dissemination of the results, using an RR, and participation of external researchers. *Conclusions:* Currently, the use of EBSE is challenging and requires both professional and research skills. While it is a valuable instrument for researchers to enhance collaboration with industry, it may not be as suitable for general use by practitioners, as initially hoped.**

***Index Terms*—Evidence-based software engineering, Evidence-based practice, Rapid review, Participant observation, Industry-academia collaboration, Knowledge transfer.**

I. Introduction

EVIDENCE-BASED software engineering (EBSE) aims to improve decision-making related to software development and maintenance by integrating the best current evidence of research with practical experience and human values [1]. Systematic reviews (SRs[1]) are a key component of EBSE as they provide a rigorous and transparent approach to searching and synthesizing the existing research evidence on a particular topic or research question. Since EBSE introduction in 2004, SRs have been widely adopted by SE researchers, allowing them to synthesize research on many software engineering (SE) topics. To illustrate, Kamei et al. identified 446 SRs published only in the top SE journals and conferences pre-2019 [2].

Although SRs support EBSE, they are not synonymous [3]. EBSE goes beyond the search and synthesis of evidence to also include (1) converting practical problems into questions that can be answered with evidence, (2) applying evidence obtained from SRs by considering the context, preferences, and expertise of stakeholders, and (3) evaluating the use of evidence and the performance of the process.

Several authors acknowledged the lack of EBSE adoption by industry [4]–[8], with some of them highlighting that it may be because SRs do not address practice-relevant problems [4], [9] or lack of useful recommendations for practitioners [8]. In addition, very few SRs involve non-academic stakeholders, without whom, additional EBSE activities are unlikely to be needed. For example, among the 169 SRs published in 2011, 2014, and 2018 identified by Kamei et al. [2], we found only three SRs that were produced jointly with industry [10]–[12] (refer to Section III). In addition, some practice-driven SRs studied practical problems, presented evidence to practitioners, and evaluated the benefits of its use (e.g., [13]–[15]). However, none of them evaluated the overall application of EBSE and studied what challenges arise when using the approach. As researchers and educators, this lack of practical examples of EBSE use and evaluation raises questions about the purpose of EBSE, and whether we should simply remain focused on academic use of SRs.

In this study, we present an evaluation of an EBSE application in an industrial environment conducted to investigate the question:

- ***RQ:*** *What issues, barriers, and facilitators arise when using EBSE in an industry setting?*

Our study focused on investigating challenges or issues that may arise when implementing EBSE in an industry setting, and any factors that facilitated its application. Specifically, we conducted an EBSE-based project to address an industry problem in a software company. This included collaborating with the practitioners to diagnose the problems, collecting evidence through a Rapid Review[2] (RR), and transferring knowledge to the company. We studied in depth the process of applying EBSE using the participant observation method and by collecting and analyzing a large qualitative data set comprising meeting audio recordings, correspondence, and personal notes. Finally, we compared our results with the early concerns identified by the researchers who proposed EBSE twenty years ago.

The remainder of this paper is structured as follows. In Section II we outline the original EBSE proposal and early

Sebastián Pizard and Diego Vallespir are with the School of Engineering, Universidad de la República, Julio Herrera y Reissig 565, Montevideo, Uruguay E-mail: spizard@fing.edu.uy; dvallesp@fing.edu.uy.

Diego Vallespir is also with Pyxis Research, Pyxis, Blvr. España 2323, Montevideo, Uruguay E-mail: diego.vallespir@pyxis.tech

Barbara Kitchenham is with the School of Computing and Mathematics, Keele University, ST5 5BG Keele, Staffordshire, U.K. E-mail: b.a.kitchenham @keele.ac.uk.

[1] In this study, we use the term SR to refer to any form of systematic review, this includes, e.g., mappings studies (MS) and rapid reviews (RR).

[2] *"A rapid review is a form of knowledge synthesis that accelerates the process of conducting a traditional systematic review through streamlining or omitting various methods to produce evidence for stakeholders in a resource-efficient manner"* [16].

concerns. Section III includes a brief analysis of the extent to which SRs are being used to support EBSE applications. Section IV presents the design and context of our research. In Section V, we present details of our EBSE application and the results (i.e., issues, barriers, and facilitators). A discussion of the results and their significance and the limitations of the study is included in Section VI. Finally, we present our final remarks in Section VII.

II. EBSE Proposal and Early Concerns

In 2004, Kitchenham et al. proposed five steps that are needed to practice EBSE [1]. SRs, the core tool of the evidence-based approach, usually support steps 1-4 steps of the process.

1) Convert a relevant problem into an answerable question.
2) Find the best evidence[3] with which to answer the question.
3) Critically appraising the evidence for its validity (closeness to the truth), impact (size of the effect), and applicability (how useful it is likely to be).
4) Integrate the critical appraisal with SE expertise and stakeholders' values and circumstances.
5) Evaluate the effectiveness and efficiency of the previous steps and seek ways to improve them.

The proposal was based on the steps used successfully in Evidence-based Medicine (EBM). Despite the apparent similarity between the high-level process steps of EBM and EBSE, the authors reflected, this would not be a guarantee that the scientific, technological, and organizational mechanisms that support EBM will apply to EBSE. Thus, the authors subsequently examined in the rest of their paper the problems or challenges that researchers would face when applying EBSE. In 2005, when EBSE was introduced to practitioners by Dybå et al [17], the authors also reflected on the challenges that practitioners might face when using EBSE. Below we present a summary of the concerns identified in both studies[4].

A. Step 1: Ask an answerable question

Specificity of questions & Small body of evidence. The challenge in this step is to translate the practical problem into a question that is specific enough for its answer to contribute to the solution of the problem and broad enough to obtain answers from the available scientific literature. More specific questions are clearer and are a means to achieve more relevant results for the problem to be addressed in its specific context. In SE, less stringent questioning may be necessary [17] due to the smaller and more diverse body of empirical research compared to other disciplines such as healthcare. In healthcare, studies often feature controlled experiments with clear treatments and control groups, which are less common in SE. This diversity and lack of structure mean that SE practitioners may struggle to find answers to highly specific questions.

B. Step 2: Find the best evidence

Difficulty finding evidence. The body of evidence is fragmented with little attempt either to summarize topics or to integrate evidence. Both articles highlighted the lack of infrastructure to support the search for evidence. At the time, EBM already had specialist digital indexing systems, such as Medine, and the Cochrane Collaboration (www.cochrane.org) which published and updated SRs from all major areas of healthcare online. In SE, there were some online databases (e.g. IEEExplore and ACM-DL) that cataloged scientific articles, but none dedicated to the search for evidence.

Lack or inadequacy of evidence. Kitchenham et al. [1] criticized SE empirical research as immature. SE research-based studies usually did not acknowledge the problem of individual skill differences when comparing SE techniques. They also suggested that the lack of standards for empirical studies and the lack of replications were reasons that existing evidence was often unreliable. Dybå et al. also pointed out the need to report research results in a manner more accessible to practitioners [17]. Finally, both studies called for encouraging the gathering of evidence from studies of industry projects (e.g., using field experiments).

C. Step 3: Critically appraise the evidence

Difficulty assessing study quality. In the context of lack of standards for empirical research, the quality of published papers is likely to be poor. However, the same context means that critical appraisal is really difficult for practitioners (and often also for researchers) [17].

Need for contextual information. Dybå et al. highlighted the need to have detailed information on the context in which empirical studies are carried out, in order to be able to better evaluate them and decide how to integrate the evidence they present. (Contextual information is also required for EBSE Step 4.)

D. Step 4: Apply the evidence

Process Improvement Commitment. Both studies suggested that EBSE would work well in an organization with a strong commitment to process improvement (something strongly promoted at that time) [1].

Complexity of applying evidence & Collaboration between practitioners and researchers. Although there are certain decisions that can be made by individual practitioners, the decision-making process in SE usually considers organizational aspects, the experience, and skills of developers, customer requirements, and project constraints, among others. For this reason, the authors reflected that the process of applying evidence would be demanding, especially for practitioners, for which Dybå et al. recommend approaching experts or collaborating directly with researchers [17].

[3]There is a certain ambiguity in using "best evidence" in this context. Because if we already found the best evidence in Step 2, there is no point in appraising it critically in Step 3. A sensible meaning of "best" in Step 2 would be evidence that best matches the EBSE question defined in Step 1.

[4]This summary is the result of analyzing the original texts using thematic analysis [18] to identify the main issues, challenges, and facilitators presented by the authors. The analysis and the results were validated by Kitchenham, one of the authors of both papers.

E. Step 5: Evaluate performance

This step seeks both to reflect on the use of EBSE as well as to confirm that the changes introduced had the expected results [17].

Difficulty isolating effects. According to Kitchenham et al. [1], it is generally difficult to evaluate the performance of a particular SE technique since our concern is not usually the specific task to which the technique is applied but the final outcome of the project of which the task is a part. The authors argued that it is difficult to isolate the impact of a technique because (1) the techniques used interact with many others during the software development process and (2) the immediate outputs of a technique will not necessarily have a strong relationship with the outputs observed at the end of the project.

Early evaluations & Postmortem analysis. From the software process improvement viewpoint, Dybå et al. also stated that given the need to adapt and learn from the rapid changes in the software development process, we should not conduct evaluations only at the end of the project [17]. So they proposed that in addition to holding postmortem meetings, after-action reviews should also be held (i.e., brief meetings to evaluate a change while it is being carried out).

F. General aspects

A general barrier to adopting EBSE is the attitude of stakeholders to SE evidence, **Research evidence ignored.** Kitchenham et al. suggested that research results were depreciated by practitioners and other stakeholders because researchers address issues not relevant to industry and they present their results in a manner inappropriate and unsuitable for decision-makers[5].

III. Existing practice-driven use of EBSE

Despite the several potential problems suggested by its proponents, we are not aware of studies that directly assess practice-driven applications of EBSE.

Ten years ago, in the only related study that we know of, Santos and da Silva surveyed the authors of 44 SRs published between 2004 and 2010 [9] to investigate the motivations for their SRs. They did not ask if the secondary studies were commissioned by non-academic stakeholders or carried out in collaboration with industry, so we cannot confirm how many SRs correspond to the practice-driven use of EBSE. However, their results suggest a lack of connection between SRs and practice. Most researchers wanted to learn more about a topic or to support their own research. A quarter of the authors indicated that their SRs sought practical solutions that could be used in industry, although it was unclear whether they had confirmed the value of their results with practitioners.

It is fair to say that the SE community has tried to promote the use of SRs results. The Voice of Evidence (VoE) column appeared for a decade in the IEEE Software magazine (2007-2017) intending to extract practical lessons from SE articles (most of them SRs) to share with practitioners. According to its editors, the main challenge was to effectively translate the articles (most of them SRs) into takeaways that connect with practitioners' concerns [19]. Another currently active initiative is the EDOS Center newsletter[6] which outlines research results for Norwegian agencies. There are no impact assessments of either of the two initiatives (only a citation analysis of the VoE column [19]).

In summary, very little is known about the impact of SE SRs on practice, and the last study on SR author motivations was conducted more than eleven years ago. Motivated by this we decided to do a brief investigation to assess the extent to which SR are being used to support EBSE.

Before our investigation, we knew of only five studies whose authors indicate practice-driven uses of EBSE, all of which were supported by some form of SR [13]–[15], [20], [21]. To understand the extent of practice-driven use of EBSE better, we analyzed a sample of SRs in SE to identify whether they report any industry involvement (as a proxy for all practice-driven use of EBSE). We took as a basis the tertiary review reported by Kamei et al. (the most recent tertiary review sharing its SRs list) [2]. In this study, the authors analyzed 446 SRs published up to 2018 in venues with a minimum h5-index (20 for conferences and 25 for journals). We selected and analyzed the motivation and use of the results of all 169 listed SRs published in 2011, 2014, and 2018 (38% of the total). These SRs provide information about more recent SRs than those investigated by Santos and da Silva and allow us to investigate whether there are any clear changes in the motivation for conducting SRs over the period.

We classified the studies using the following categories:

- **Search for knowledge.** These SRs seek to gain knowledge of a particular field, although some indicate it explicitly and others do not (but neither do they report any other use or motivation). An example is an SR carried out to understand the state of the art of SE in startups [22].
- **SRs are complemented by other evidence.** SRs in this category are also conducted to gain knowledge, but the results are complemented or validated by conducting another study (e.g., surveys). In one such study, the authors conducted an SR to learn more about the terminology used in global SE (GSE) [23]. Subsequently, the results were complemented by a study with experts to create a GSE taxonomy.
- **Investigations of the EBSE process.** In these studies, the SR process is investigated and improved. In one of them, the authors investigated the repeatability of the SRs [24]. For that, they trained novice researchers to conduct an SR. Subsequently, their results were compared with those obtained in a same-purpose SR previously published.
- **Practice-driven SRs.** In these studies, the authors engage with companies or other non-academic stakeholders who defined the SRs' purposes or intended to use their results. In one such study, researchers were asked, during a joint industry-academia project, to build a tool to measure

[5]Kitchenham et al. originally discussed this issue in the context of Step 5.

[6]Effective Digitalization of the Public Sector (EDOS) is part of Simula Metropolitan Center for Digital Engineering. Its newsletter is available at https://enedos.substack.com/

TABLE I
SUMMARY OF THE CHARACTERISTICS OF THE PRACTICE-DRIVEN SRS FOUND.

Study	[10]	[20]	[13]	[11]	[12]	[21]	[14]	[15]
Type*	SR	SR	RR	MS	MS	SR	RR	RR
Publication year	2011	2013	2018	2018	2018	2022	2022	2023
Nature of Stakeholders Engagement								
Close collaboration reported			x				x	x
Part of broader collaborative projects	x			x	x	x	x	
Use of Results								
Results presented to company	x	x	x				x	x
Results used by company			x					x
Used in subsequent collaboration				x	x	x		
Feedback from Stakeholders								
Stakeholders feedback given	x	x	x				x	x
Positive use perception of evidence	x	x	x	x	x	x	x	x
Considered role of SR in decision making			x					
Not directly relevant to users	x						x	
Feedback used to improve SR	x						x	x

* SR refers to systematic reviews, MS to mapping studies, and RR to rapid reviews. As defined by the authors.

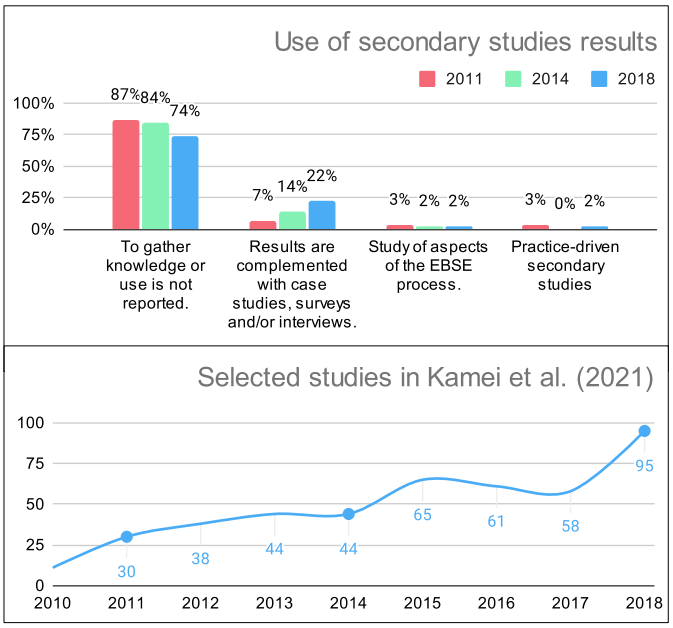

Fig. 1. Reported use of SRs selected from Kamei et al. [2].

socio-technical congruence [12]. They began their work by conducting an SR to learn more about the topic.

The results of this survey are shown in Figure 1. They confirm the lack of studies on the practice-driven use of EBSE. Only 3 of the 169 studies published in 2011, 2014 & 2018 reported SRs with practice-driven use (or motivation), all of which involved active collaborations with the industry.

In summary, we are aware of eight studies of practice-driven use of EBSE (five we knew beforehand and three that we obtained in our survey). The characteristics of those eight studies are shown in Table I.

It is also worth mentioning that: in [20] the stakeholders' requirements were surveyed through a case study, [13] includes the report of practitioners' perception of using RRs as a decision-making support method, [15] reports three RRs, all related to the development of a software selection model. Two investigated the selection of software tools (i.e., CASE tools and tools for continuous deployment/Devops, respectively), and one investigated methods of assessing the quality of tools.

Table I reveals four points of interest.

- All the studies reported collaborations with practitioners from individual companies (which helps SRs to be well-focused on stakeholders' requirements). None were commissioned by other stakeholders such as government agencies, professional bodies, or industry associations.
- Five of the eight studies reported that the companies made use of the SR results, although, in three of those cases, the results were used as part of a wider collaboration.
- RRs seem to be positioned as a method of special interest for the practice-driven use of EBSE.
- In all cases the feedback from the practitioners was positive. However, in two cases the results of the SR were not directly applicable.

This analysis has several limitations. It considers only published SRs (we have not deliberately searched for industry white papers). The SRs were classified only by the first author, and it shares the limitations of the study by Kamei et al. concerning the assembly of the list of SRs [2].

Our brief research confirms that there are few published practice-driven SRs. Although seven of the eight studies we found reported making use of the results of the SR, only one reported the practitioners' perceptions of using RRs as a decision-making support method. If we want EBSE to help bridge the gap between industry and academia, we should foster its use to improve practice. Still, we cannot recommend the adoption of EBSE if we have little evidence of its efficacy. Therefore, this indicates that our current study addresses a research gap by evaluating a practice-driven application of EBSE, identifying the naturally emerging issues, barriers, and facilitators.

IV. Study Design

Our trial of EBSE focused on assisting a small software company to address difficulties with its Knowledge Management (KM) processes.

The project team consisted of two students and one researcher (the first author) and the project was part of the students' capstone project. Students were trained on EBSE and worked to retrieve and synthesize evidence to help the company tackle its KM problems. The students carried out the project with close supervision, suggestions, and validations from the first author.

Our research method was participant observation. Participant observation is the process in which the observer remains in a social situation for the purpose of scientific research [25]. It has the *"unique strength of describing complex aspects of cognition, social interaction, and culture over time"* [26], which we believe is essential to study a complex issue such as the practice-driven use of EBSE.

Although we expected both to provide information the company could use to improve its SE process, and to observe the introduction and results of the process changes, our goal was not to introduce process changes, but to study the EBSE process (including the actions of the EBSE project team and the company staff). Our research goal meant we needed to maintain a neutral stance concerning the project's outcome. Crucial in this regard was acknowledging that participant observation studies demand researchers to be aware of, and control for, the risk of experimenter bias during when planning and conducting their study.

In our study, the first author immersed himself in an EBSE application setting to conduct the research. His role had two purposes: he led the EBSE-based project team, aiming to support a software development company while simultaneously observing the activities, people, artifacts, and interactions during the project. This approach enabled him to experience the practice-oriented use of EBSE firsthand, gaining profound insights into its complexities and nuances.

To minimize experimenter bias, the following steps were taken: (1) Data related to different activities were collected throughout the project from multiple sources. (2) Analysis of the data was validated at different stages by different researchers. Additionally, it should be noted that one of the co-authors, who held a positive bias towards EBSE, contributed solely to the post-trial discussions and presentation of results.

We used a rigorous pre-planned analysis process and systematically tracked all data to improve reliability (e.g., to avoid misinterpretation of the data) [27]. We used the O'Brien et al. checklist for reporting qualitative research to enhance reporting clarity and completeness [28]. We also considered the eight criteria for the quality of qualitative research proposed by Tracy [29], i.e., worthy topic, rich rigor, sincerity, credibility, resonance, significant contribution, ethics, and meaningful coherence. Our study addressed these criteria by detailing the context and procedures, combining diverse sources and reflections, being transparent and self-reflective, and aiming to make a valuable and coherent contribution to the field while adhering to ethical guidelines.

A. Research Program

The EBSE-based project that we conducted in collaboration with the software company served as the foundation for two interconnected yet distinct empirical studies. The first study, i.e., [30], entailed an external replication of the initial study proposing the use of RR in SE [13], *[Note for reviewers: The paper reporting this study is currently under review, and we include it as supplementary material. It includes detailed information on the RR conducted and feedback from practitioners.]* The second study is the one reported in this paper.

TABLE II
Comparison of steps of EBSE & Action Research

EBSE Steps	Action Research
1. Converting the need for information into an answerable question.	1. Diagnosis
2. Finding the best evidence with which to answer that question.	2. Planning
3. Appraising evidence validity, impact, and applicability.	3. Intervention
4. Integrating the appraised evidence with expertise and stakeholders' values and circumstances.	
5. Evaluating effectiveness and efficiency in executing previous steps and seeking ways to improve.	4. Evaluation
	5. Reflection

In the context of first study, we employed Action Research, replicating the same research method as in the original study. The action taken was the conduct of an RR to provide evidence to practitioners aimed at addressing their problems. However, from the company's perspective, their involvement was in an EBSE project (i.e., an application of all EBSE steps). We can define EBSE as a form of action research, where the goal is to adopt evidence-based actions to address a problem (refer to Table II for comparison of the steps of both). Given this comparison, it is reasonable to consider EBSE as a specific type of Action Research. This perspective enables us to utilize the Participant Observation method to evaluate the project as an example of EBSE.

It is noteworthy to mention that it was necessary to reconcile the application of both methods. According to its definitions, Participant Observation seeks to understand the phenomenon being studied, while Action Research focuses on creating positive change. In our case, we aimed to assist the company while also conducting a fair evaluation of the EBSE application. To reconcile these objectives, we strived to maintain a neutral perspective and focused our project on applying EBSE steps as they are defined. This involved explaining in advance to practitioners that our objective, in addition to trying to assist them, was to evaluate EBSE, and that we would base our suggestions for process improvements on scientific evidence collected through an RR. We also clarified to participants several times that both their positive and negative feedback was valuable. Finally, we analyzed and recorded any need for deviation from these objectives, e.g., including clarifications to the evidence obtained.

B. Research Context

Here we describe certain aspects of the research context.

Project team: The team consisted of the first author and two undergraduate students about to finish their computer science degrees. The first author has ten years of industry experience as a technical lead and software quality manager and twelve years as a member of the university. This study is part of his doctoral research that focuses on investigating EBSE adoption. Participating in the project was part of the students' capstone project. Also, both of them had full-time jobs related to software development. In particular, one of them was also part of the company's development team, and was so during the first half of the project. Both of them were trained in the planning and conduct of SRs. The training was led by the first author and based on an EBSE and SRs course he teaches [31], [32]. Finally, although Vallespir was not directly part of the team, Pizard consulted him regarding the team's decisions throughout the process.

Selection of Research Topic & Company: The EBSE team was the one that initiated the contact with the company and proposed the research project. Previously, we discussed which companies and organizations had a strong relationship with us and subsequently, we identified a research topic that was relevant to them. We chose our research topic to be about common problems present in the industry but not critical issues. In this way, practitioners could work together with our EBSE team without revealing sensitive information and with no expectation of dealing with emergencies or serious project problems. In addition, we sought a problem that did not have an *a priori* well-known solution in the software industry, but a solution to which would have a positive impact on the organization with which we worked. Finally, we verified that there were studies on the topic, carried out in non-academic contexts, that reported observations, lessons learned, or recommendations.

Given the characteristics of EBSE and the inexperience of the project members, we believe that the pre-selection of a topic area mitigated many risks related to knowledge of the domain, the stakeholders' expectations, and the existence of evidence on the topic.

In particular, we defined knowledge management (KM) as the research topic. It is a topic for which there are non-critical but important problems in the industry, especially in software development using agile methodologies (see, e.g., [33], [34]).

The company: The company was a UK company specializing in digital out-of-home (DOOH) advertising[7]. The company's IT department, located in Uruguay, was responsible for developing and maintaining a platform to manage advertising campaigns.

The requesters: The company's technical product leader and the project manager accompanied all the stages of the EBSE project, from the diagnosis of the problems to the dissemination of evidence. They answered questions, carried out intermediate validations, and received the collected evidence. For the purposes of this study, we have considered them as EBSE requesters. Their educational level was Intermediate[8] (one with *upper secondary education* and the other one with *post-secondary non-tertiary education*). As sources of information for supporting practice they usually talked to colleagues, read technology forums or blog articles, and watched technology videos (e.g., from the Microsoft Youtube channel). Neither of them consulted scientific literature.

C. Research Activities

The purpose of the project was to provide support to the company to improve its KM practices. We agreed with the company that our work would involve using EBSE, so we went through the five steps of the process. Figure 2 shows the main activities carried out (which will later be detailed in Section V).

The goal of our study to was to investigate the use of the EBSE framework. As a means of assessing the value of EBSE framework, we intended to identify potential barriers and facilitators to EBSE use that reflected the viewpoints of both the EBSE team members and the company staff members. We wanted the factors we identified both to be verifiable and to arise naturally as part of the EBSE project. To achieve our goals, we collected and recorded data throughout the EBSE project including final feedback. Data was obtained from a variety of sources: Audio recordings of all team meetings (including the final retrospective meeting) and meetings with the company, all emails and messaging app communications, a researcher's personal diary kept by the first author), and all EBSE project documents (including students' capstone project report, reports for requesters, and other artifacts).

We used a qualitative data analysis process with the following stages. First, the first author performed the initial data analysis using a method that was strongly based on thematic analysis with a realistic approach [18]. The second and third authors reviewed the results and suggested issues to review, reflect on, or expand on. Subsequently, the first author did a second analysis interpreting comparatively the results with the early concerns identified by the researchers who proposed EBSE twenty years ago. Finally, this analysis was revised and expanded by all the authors.

D. Ethical Issues

Although our university did not require our study to be approved by an ethics committee, we took care to consider possible negative impacts of the study on the participants. As recommended [35], we ensured that participation in the project was voluntary and the research process was transparent to all participants. Company members and students were informed of the characteristics of the research prior to giving their consent to participate.

There were two other major considerations. Firstly, ensuring that the students' education experience was not adversely

[7]Digital out-of-home advertising (DOOH) is advertising designed to reach consumers when they are not at home and that is also dynamically and digitally displayed. This includes digital transit, digital billboards, and digital place-based displays.

[8]According to UNESCO's ISCED 2011 classification. https://ilostat.ilo.org/resources/concepts-and-definitions/classification-education/

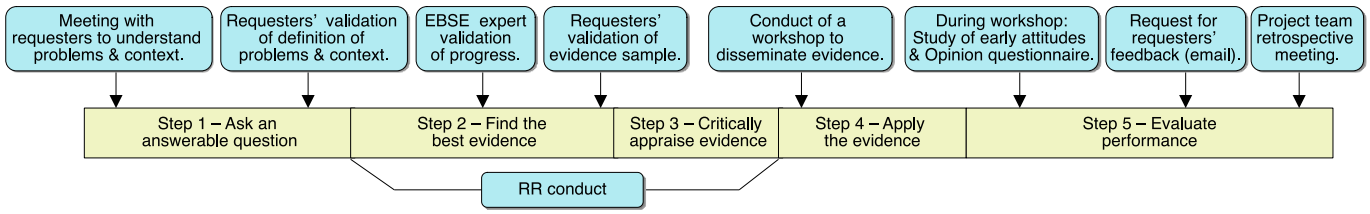

Fig. 2. Project stages.

impacted by the study: (1) Students should not be required to undertake tasks beyond their capabilities. This concern was addressed by appropriate training and supervision. (2) Students should not feel obligated to express support for the EBSE framework. The students were assured that the outcome of the study in terms of whether or not it was favorable to the use of EBSE would not impact their capstone project marks. Secondly, ensuring that the interests of the company are not adversely affected by the study: (1) The company would receive the best scientific information to help them address their process issues. This was assured by the personal experience and supervisory role of the first author. (2) Commercially or personally sensitive information would be kept confidential or anonymous as appropriate. Specifically, only the company and roles are identified, and specific comments are not attributed to specific individuals.

V. EBSE Application and Results

Below we present for each EBSE step the central characteristics of our application and the response to RQ, that is, the main issues (indicated with [·]), barriers (indicated with [–]), and facilitators (indicated with [+]) that arose during the process. As a complement to this section, the paper reporting the RR conduct includes information regarding the RR conduct and feedback from practitioners on the results of the RR, including representative quotes [30].

A. Step 1. Ask an answerable question

In this step, the team met with members of the company to understand their context and the problems they had in the KM area. We defined the following research question: *What are some empirically validated recommendations for knowledge management for software development companies?* As a validation activity, we prepared a summary of the context of the company and its KM problems which we shared with the requesters (refer to Table III).

[·] Considering requesters' needs. The requesters had no knowledge of SE research or EBSE, so we started the kickoff meeting by introducing both topics. We talked and agreed on the expectations about the results of the project. In addition, we found that the company had a continuous improvement process that involved identifying small improvements that were introduced in future sprints. Specifically, they told us that the results of our project were going to be incorporated into that improvement process. Thus, we understood that they preferred a set of small self-standing recommendations rather than a single major process change. This restriction had consequences in the subsequent stages (i.e., in the selection of primary studies, and in how we elaborated recommendations based on the evidence).

[·/-] Specificity of questions & Small body of evidence. Given the difficulty in finding evidence when testing specific questions in preliminary searches, we decided to use a rather broad question. Additional aspects (i.e., limiting the context of primary studies to small or midsize companies using agile methods, or accepting specific recommendations and not proposed frameworks or models) were considered in subsequent steps of EBSE, e.g., using them as criteria for prioritizing process changes.

[+] Strong link with company. The fact that one of the students was part of the company facilitated their willingness to work with us and also helped us better understand their problems.

TABLE III
Summary of Context and Problems related to KM.

The company is a spin-off of a UK advertising agency specializing in digital out-of-home (DOOH) and is responsible for managing the entire life cycle of a platform consisting of four products. The staff are geographically distributed, with the CEO and four account executives located outside of Uruguay, and the software development team based in Montevideo.
The team follows Scrum methodology with two-week sprints, daily stand-up meetings, and regular demos and retrospectives with stakeholders. They use various online tools for KM, including GitHub for code storage, Lucidcharts for architecture diagrams and planning, Visual Studio Online (VSO) wiki for test cases, VSO board for backlog management, Trello boards for tracking tasks, and Google Docs for architecture records and spreadsheets.
However, they face several challenges with KM. They struggle with finding the right documents, as there are duplicates or similar documents with different media and dates. Keeping the documentation up-to-date and eliminating unnecessary or outdated documents is also a challenge. The lack of standard definitions for document types creates confusion, with each person generally deciding what type of documentation to create.
The company is concerned about knowledge centralization in specific roles, particularly in the QA and DevOps manager positions, where there is only one team member for each role. This poses a risk in terms of knowledge sharing and continuity if these individuals are unavailable or leave the company.

B. Step 2. Find the best evidence

The question defined above was answered using a rapid review (RR). Its main characteristics are summarized in Table IV, following the proposal of Cartaxo et al. [6]. Before starting, we checked that there were no SRs on KM in agile methodologies that were able to answer the research question.

TABLE IV
MOST RELEVANT CHARACTERISTICS OF THE RAPID REVIEW CONDUCTED AS PART OF THE EBSE APPLICATION.

Problem	To find practical and applicable recommendations on KM for the company. The context of the studies should be similar to that of the company and the recommendations must have been empirically validated.
Research Question	What are some empirically validated recommendations for KM for software development companies?
Protocol	It was written during the first weeks of the RR. We did initial searches to validate the existence of evidence.
Stakeholders' roles	The requesters met with us during the diagnosis and validated the document that summarized its context and the problems to be addressed. They met with us to validate the evidence we found in the early stages of RR. Together with other software development team members, they participated in the results dissemination workshop.
Time Frame	The RR was done in three and a half months and the total time spent by team members was ∼150hs, including team meetings and knowledge dissemination transfer activities with practitioners.
Search Strategy	Keyword-based search in Scopus.
Selection Procedure	We considered available studies in English, with practical recommendations, addressing issues of KM in software development companies, and not presenting models or theoretical frameworks. Each reviewer assessed half of the candidate primary studies. To validate an adequate level of agreement, they reviewed the first 30 candidate primary studies together, and the kappa statistic was calculated (the value obtained was 0.618, which indicates a good level of agreement). Subsequently, the reviewers used two rounds: (1) checking titles and abstracts and (2) reading the full text. From the 425 studies returned by Scopus, we identified 21 primary studies [36]–[56].
Evidence Appraisal	We assessed the evidence's relevance to practice by studying the context and methods used for its generation.
Extraction Procedure	Each reviewer extracted data from half of the papers. Data extracted: Context of the study (agile, year of study, company), research method (experimental validation), validation result, and recommendations.
Synthesis Procedure	We used content analysis with an inductive approach (adapted from [57]). The stages carried out were: Translation into Spanish of the original texts, labeling of the different types of recommendations using open coding, grouping, and categorization of the fragments according to their codes (e.g., grouping similar or complementary codes), and, finally, creation of descriptions. As a tool to facilitate the coding, we used the freeware tool Saturate[9]. Each reviewer synthesized half of the papers. In two meetings with the first author, the results were reviewed and adjusted to obtain 21 recommendations for practice.
Report/Diffusion	We prepared an evidence briefing with the results and conducted a workshop with the development team.

[-] Difficulty finding evidence & Lack or inadequacy of evidence. In preliminary searches, we found that very few studies had adequate evidence to address our problem. The studies were difficult to find (so we had to test several questions and adjust their specificity, as discussed in the previous step). Furthermore, they did not include recommendations for practice suitable for our needs. To mitigate the latter issue, and knowing that very few SE studies usually include recommendations for practice [8], we decided to also consider those that included lessons learned, reflections, or certain observed behaviors. We analyzed those observations, as well as the context, to develop practice-oriented recommendations.

[+] Using a rapid review. We opted to conduct a RR because we had a low-resource setting. Students had a limited time period to complete their capstone project and could only allocate a limited effort each week. We also believed that the results of a RR would be adequate to address the issues identified. The RR method allowed us to work in an agile way while maintaining scientific rigor.

[+] Validations with requesters. We also conducted two validations with requesters to assess the evidence. They approved the sample evidence we presented to them. However, they wanted to obtain recommendations from contexts similar to their own. They also commented that some recommendations of the sample seemed useful to them but they did not know how to put them into practice. This activity was useful to validate the adequacy of the evidence found early in the process.

C. Step 3. Critically appraise the evidence

We evaluated the evidence's relevance to practice (i.e., applicability). This included understanding the context in which the evidence was generated and the research methods used. For example, for organizations that participated in a primary study, we identified in which country they carried out their activity and their size. In this way, we sought to provide more information to requesters so that they could better evaluate the evidence and select the one that best fits their context.

[–] Need for contextual information. Given the company context was that of a small company using agile methods, we did not need much detail about the research context of the primary studies. Even so, not all studies presented the same kind of information and some of them did not include enough context details.

D. Step 4. Applying the evidence

Since the RR was conducted by the EBSE project team, we needed a method of informing the company of the results of our RR. In addition, unlike evidence-based medicine, we were reporting a number of very different recommendations, not reporting evidence that compared two well-defined process options, so we also needed to help the company staff select the most suitable recommendations.

Firstly, following Cartaxo et al.'s reporting suggestions [4], we developed an evidence briefing to summarize the results of our RR. We, then, conducted a workshop to discuss our recommendations and to identify the specific process changes that the company could adopt. In the workshop, we presented the recommendations we derived from the evidence. Then, we assisted the company staff to prioritize the recommendations in terms of *ease of implementation* and *potential benefits*, in a hands-on exercise. We also discussed which of the recommendations the company could implement and how the process changes could be introduced. The project manager (one of the requesters) and three other members of the software development team participated in the meeting.

[–] Complexity of applying evidence. The recommendations obtained in the RR were general principles not specific

process changes. So, during the intermediate validations, the requesters expressed concerns regarding the implementation of these recommendations. Thus, although our EBSE application sought to have a strong emphasis on scientific evidence, it was necessary to include examples, suggestions, or clarifications that emerged from our professional practice and knowledge of the company. We did not add or remove any recommendations from the original set.

[+] Process improvement commitment & Collaboration between practitioners and researchers. Two factors facilitated the reception of the evidence. First, as the development team was used to improving their process, it was relatively easy for them to evaluate the evidence and consider how the recommendations could be applied to their processes. Second, the close collaboration between requesters and researchers throughout the EBSE project helped to make the results interesting to the development team, and the discussion served to better examine the recommendations and consider strategies to implement them.

[+] Appropriate dissemination. Based on the recommendation of Cartaxo et al. to incorporate discussions of the results as dissemination activities [13], we not only prepared an evidence briefing with the RR findings, but also held a workshop in which the attendees began to discuss ways of implementing the recommendations, the required effort and the potential benefits.

E. Step 5. Evaluate performance

In this step, we sought to address two issues: a) assessing whether the process change had a successful outcome and b) analyzing whether the team could have done a better job applying EBSE. We investigated both issues as follows.

We collected requesters' perceptions of the EBSE project results and the challenges faced during its conduct in three instances. First, we recorded (with prior approval) the workshop in audio to analyze the attendees' initial attitudes. At the end of the workshop, we circulated a questionnaire for participants to express their opinions about the project and its results. Finally, eight months later, we requested feedback on the process changes from the requesters via email. In addition, the EBSE team held a retrospective meeting to discuss and reflect on the EBSE application and its results.

Evaluation of results. None of the workshop participants suggested the recommendations were inappropriate. It is noteworthy that all attendees demonstrated a clear understanding of the recommendations and their potential implementation and actively participated in the prioritization process. Following the workshop, the company underwent significant changes, including an acquisition, role restructuring, shifts in organizational hierarchy, and eventual closure. However, even amidst these transformations, the requesters, in response to our email requesting feedback, confirmed that they had utilized some of the recommendations before the closure occurred. In the workshop and the subsequent feedback requests, the only negative comment regarding the EBSE process was related to the time it took to complete the RR (more detail in two paragraphs below).

Evaluation of the process. The main challenges we faced were the lack of guidelines and examples for conducting RRs for industry and the difficulty in finding appropriate evidence. At that time, there was only one previous study by Cartaxo et al. that applied RR in supporting SE practices [13]. Due to the EBSE team's lack of experience in conducting this type of secondary study, there was a substantial risk of not being able to provide evidence-based recommendations to address KM issues effectively. To mitigate this, we worked cautiously at each stage, conducting preliminary literature searches and verifying the adequacy of retrieved studies for the RR. Intermediate validations with the requesters were also carried out to ensure the relevance of the evidence found. The second challenge arose from the limited availability of studies with appropriate evidence. Given the scarcity of SE studies that offer practice recommendations, we decided to include primary studies with lessons learned or empirically validated observations.

[·] Considering requesters' needs (Step 1). One of the requesters would have preferred a shorter timeline. The limited weekly availability of the students led to the research taking three months to complete, which would have otherwise only required about two weeks of full-time work. In addition, the requesters had not emphasized EBSE project timescales, so we underestimated the need for prompt results. This highlights two key points: practitioners value quicker processes for results, and all requesters' needs are significant, not just those regarding the specific problem at hand.

[+] Early evaluations & Postmortem analysis. Both the retrospective meeting and the written feedback from requesters were useful to better understand the performance of the EBSE process and the company's perceptions of the results.

F. General aspects

Three factors affected all aspects of the EBSE project.

[–] Inexperience & Lack of guidelines and examples of using EBSE in the software industry. During our study, Cartaxo et al's research was the sole instance of employing a RR to support SE practice [13]. Also, we could not find detailed guidelines on how to apply each step of EBSE. Given our limited experience in conducting RRs and applying EBSE, we identified a substantial risk of being unable to develop scientifically-backed recommendations that effectively assist requesters in addressing their KM issues. To mitigate this, we carefully validated each EBSE activity. However, this caution resulted in extending the EBSE project timescales.

[–] Complexity when working with an industry partner. Collaborations between industry and academia usually face challenges, and our study was no exception. The most notable challenge was that we were unable to finish conducting EBSE final step, because of the changes that the company underwent. In addition, one of the requesters, the technical leader, was unable to attend the workshop (subsequently, the results were sent to him), and on a few occasions, the requesters' responses to our inquiries had a noticeable delay.

[+] Participation of external researcher. During primary study selection, we validated the progress in the use of EBSE

with an external researcher with some EBSE experience. This activity was especially useful to validate the rigor of our EBSE application and to have an objective view of the risks that were still present in the project.

VI. Discussion

This section revisits our main findings on the barriers and facilitators that we identified when applying EBSE. We also compare our results with the early concerns of EBSE proponents. Subsequently, we reflect on the implications our results have on the fit-to-purpose of EBSE. Finally, we present our assessment of the limitations of our work.

A. Barriers to EBSE application

The barriers that arose in our study could be grouped into: Difficulty in obtaining relevant evidence, Complexity of applying evidence, and Lack of guidelines and examples of using EBSE in industry.

Difficulty in obtaining relevant evidence. One of the challenges we faced was finding evidence able to address the company's problems. There were no evidence-based answers to the company's specific KM problems, and evidence related to more general issues was hard to find. This meant we had to use a rather broad, high-level research question. Lack of relevant evidence has been reported in previous studies [14], [58]. The major disadvantage of this approach is that the answers may be less specific than required.

The limitations of SE evidence affect research reliability: (1) Lack of empirical studies forces reliance on less reliable sources, (2) Absence of clear recommendations may require restructuring study results, and (3) Lack of standardized approaches may complicate aggregating results. Point 1 weakens the strength of evidence, while points 2 and 3 risk biases because other researchers, with less detailed knowledge of the primary study details, may misunderstand or misrepresent the conclusions that can be validly drawn from the evidence.

Researchers have proposed a variety of methods to address the limitations of SE evidence, i.e. using evidence considering its nature and quality (e.g., for descriptive and correlational uses and not to support claims of causality) [59], researching issues relevant to the industry [8], [60], collaborating closely with practitioners [60], including recommendations for practice [58], conducting more interdisciplinary research [61], including grey literature [62]–[65], and appropriately considering research context [66], [67].

Complexity of applying evidence. Another challenge was the complexity of applying the evidence we found. This was mainly because the evidence was in the form of high-level recommendations not directly implementable process guidelines. The nature of the evidence was a particular problem because our requesters did not themselves have any SE research experience. Thus, we needed to summarize the collected research in a way that it could be understood by practitioners.

This issue has not been raised by any of the other collaborative EBSE-related studies that we are aware of (i.e., [11]–[13], [15], [21]), because, in all those cases, the requesters either had SE research experience or specialist knowledge of the topic of interest.

Lack of guidelines and examples of using EBSE in industry. In contrast to SRs, there are no detailed guidelines for the five EBSE steps, and there are only a very limited number of examples of using secondary studies in an industry context.

First, the EBSE team had no previous experience working with EBSE and RRs in industry. Our professional experience and our strong link with the company helped us to establish a good working relationship with our requesters. However, our EBSE activities might have been both more focused and more efficient if there had been some practical SE guidelines to support our EBSE process.

As it was, since the EBSE team chose to use an RR (rather than a full SR), we followed, as far as possible, the process reported by Cartaxo et al. [13]. In particular, we prepared an evidence briefing and discussed the RR results with the requesters. However, in general, we acted in response to the project circumstances, rather than following prepared plans. For example,

- In order to define a research question we had to do several preliminary searches.
- We discussed our initial evidence with the requests, adjusting our search, selection criteria, and aggregation process to address their concerns.
- We developed a method to assist the requesters to rank the RR recommendations which took place when we presented the RR results.

Steps 3 and 5 of EBSE posed particular problems. Step 3 requires a critical assessment of the available evidence, but following Cartaxo et al.'s guidelines for RRs, we omitted any formal assessment of the identified primary studies. In practice, since we needed both to include weak forms of evidence from the primary studies (such as lessons learned), and to derive recommendations ourselves, the most important issues for the requesters were information about the type of companies generating the evidence and of the methods used to do so. It is also important to recognize that failure to assess evidence in terms of its validity, impact, and usefulness is a major deviation from the original specification of evidence-based practice defined in medical practice

Step 5 calls for reflection on the EBSE activity, but we found no examples of studies discussing the use of an SR as part of an industry-academia collaboration that reported any evaluation of the EBSE process itself. In practice, we reflected on how we performed each step, in terms of the issues that arose, how we addressed those issues, any ways in which we might have improved our performance, and the requesters' views of the EBSE project. Also, due to the closure of the company, we were unable to confirm that the recommendations were adopted and delivered the expected benefits.

Taken together, these issues seem to imply that the analogy between evidence-based health care and evidence-based software engineering is somewhat problematic. Perhaps in the case of healthcare, the steps of evidence-based practice are more straightforward, so a more detailed process is irrelevant. In SE it appears that we need more guidance to cope with broad

questions, reporting SE evidence to practitioners, developing practical context-appropriate recommendations, and evaluating EBSE activities.

Practical problems when working with an industry partner. As other researchers have pointed out, academic interactions with industry partners are difficult (see, e.g., [68]). In our case, the company placed constraints on the collaboration (i.e., defining a broad problem with several different aspects, and requesting evidence that would map to small incremental changes) that interfered with academic goals and was further complicated by company circumstances (e.g., the company changes and closing).

B. Facilitators for EBSE application

Several factors supported our EBSE project.

Close collaboration with industry partner & Considering their needs. The close collaboration with the requesters was a decisive factor in achieving our goals. This included activities in which their participation was explicit, e.g., meeting to understand their problems, validating our understanding with them later, also validating the evidence early, and jointly participating in the workshop to disseminate the results. But it also required an explicit effort on our part to consider their needs, e.g., in considering how they would apply the evidence. Even so, we did not detect that one of the requesters wanted results in less time.

Appropriate dissemination. In accordance with Cartaxo et al.'s recommendation to incorporate result discussions as part of dissemination activities [13], we not only created an evidence briefing report summarizing the findings of the RR, but we also conducted a workshop where participants engaged in discussions about the evidence and its application. This combination proved effective in disseminating our results. It seems that a one-page evidence briefing report alone is insufficient to ensure the practical use of the results. Based on our experience, it may be helpful to include the following in future reports: (1) practical implementation guidance for each recommendation, (2) dependencies between recommendations, and (3) indications of the strength of evidence supporting each recommendation.

Process improvement commitment. The company's commitment to process improvement was of great help in its willingness to reflect on its processes and to receive recommendations in a positive way. It is important to point out that, EBSE is not an alternative to process improvement but a framework that can be used to support process improvement. In particular, as suggested by Dybå et al. [17] and confirmed by our study), EBSE seems a useful method to find candidate solutions to support process improvement.

Conduct of an RR & Participation of external researcher. Two aspects that helped us achieve adequate methodological rigor were using an RR and the validation of experts. Using an RR allowed us to obtain interesting results for the company within the available low-resource setting (i.e., limited effort and use by non-experts) and without losing scientific rigor. Validation by experts helped two stages of the study, during the early stages of the RR, expert validation gave us confidence in our RR process, and during the analysis and reporting of results, the participation of an EBSE expert helped substantially improved the quality of this study.

C. Revisting EBSE after 20 years

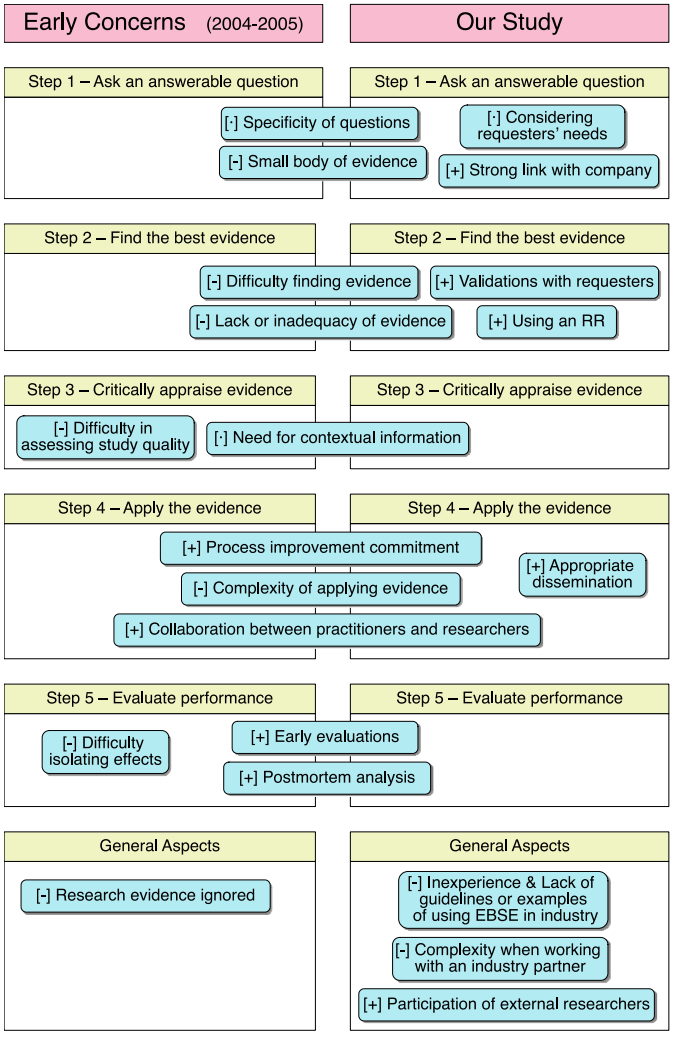

Fig. 3. Comparison of the early concerns identified by EBSE proponents and the results of our study. Notation: [·] indicates main issues, [–] represents barriers, and [+] denotes facilitators.

More than half of the barriers we identified in our use of EBSE had already been identified 20 years ago (see Figure 3). While certain aspects may have improved, there remain persistent barriers that continue to impact EBSE applications, despite their long-standing recognition. However, the major challenge, which has been identified in other studies as well, remains the lack or inadequacy of evidence. This represents a significant challenge to face when adopting EBSE. In addition, collaboration between academia and industry remains a challenge for the entire community. Thus, it is clear the EBSE has not achieved its goal of bridging the gap between industry and academia.

Our study has not exhibited all of the concerns raised by proponents of EBSE, but it is important to note that we present a specific case, that has some aspects that were not

anticipated by the EBSE proponents (i.e., the use of an RR and the lack of critical appraisal of evidence[10]). In addition, we were unable to fully evaluate the effect of implementing the recommendations, so we cannot make conclusions about the difficulty of isolating the effects of specific process changes. However, it is noteworthy that although the practitioners were not used to employing scientific research results, this did not cause a barrier to our collaboration.

Certainly, good progress has been made on some of the challenges identified 20 years ago. SRs are highly appreciated by the academic community (see, e.g., the tertiary study conducted by Kamei et al. [2]). There are standards or guidelines for several types of empirical methods (e.g., [27], [69]), including SRs. Much work has been done to improve evidence aggregation methods and their reports (see, e.g., [4], [13], [70]–[73]). We do not yet have a central repository of published SRs. However, for some time now it has been possible to pre-register reports in several venues [74]. Finally, there are also initiatives to generate more appropriate evidence for non-academic stakeholders (e.g. the ESEM's Industry, Government, and Community track[11]).

Both current research and our results indicate that EBSE has yet to fulfill its objective of effectively supporting practitioners in utilizing research evidence. It seems that proponents of EBSE underestimated the disparities between healthcare and computing, leading to challenges in its implementation:

- Healthcare options are generally more easily understood and adopted by practitioners without the need for extensive explanation, unlike SE methods. For instance, doctors prescribing a new drug, or nurses adjusting patient care practices such as coma patient turning frequency or hand-washing protocols.
- Healthcare options undergo rigorous evaluation before deployment. This is mainly due to national and international regulation, and ensures that a body of empirical studies are available as a basis for evidence-informed recommendations. However, in the computing industry, companies often deploy new methods without waiting for independent validation, or any proper understanding of their risks. There are many examples where the adoption of new computing methods to address specific problems has introduced new problems. For example, object-oriented design was supposed to ensure more reliable systems because keeping data and code together would make testing more effective. However, the emphasis on developing self-standing, independent objects led to extensive use of code clones which caused more maintenance problems and, in turn, necessitated the development of aspect-oriented design [75].

Furthermore, for certain practitioners, distinguishing between empirical evidence and their personal experience or the claims of thought leaders may not be inherently clear. For example, the practitioners who collaborated in our study learned about the concept of empirical evidence from our collaboration. Decision-making processes frequently rely on guidance from these thought leaders or consulting firms, even when disseminating questionable claims (as evidenced by the recent McKinsey case on developer productivity, refer to [76], [77]). Thus, the sole publication of a paper does not guarantee effective communication of SE evidence to practitioners. This certainly could have been another influential factor in EBSE adoption.

D. Is EBSE framework fit-for-purpose?

EBSE can be challenging for practitioners without research experience, but our study confirms that it can be effectively utilized in industry-academia collaboration. EBSE application raises several challenges and requires different skills that include academic knowledge, and professional experience. Successfully using EBSE requires the ability to convert academic recommendations into actionable process changes and the proficiency to design and utilize effective engagement mechanisms for the dissemination of knowledge.

Evidence is still somewhat limited, and our study highlights that practitioners value evidence that aligns with their practical concerns. This implies that academics should consider the needs of the industry when pursuing their research agendas. Likewise, practitioners should actively engage with academics to identify topics they consider essential and to facilitate evaluations of industry practices and trials of new technologies.

Step 1. The lack of evidence remains a significant concern, indicating that SE should adopt relatively high-level research questions. In particular, keyword searches specifying practitioner context can restrict access to potentially useful information.

Steps 2, 3 & 4. We consider these steps need to be better understood in the SE context. First, it is unclear what is meant by "finding best evidence" in Step 2, given that Step 3 of EBSE promotes the assessment of evidence in terms of validity, impact, and applicability. We suggest interpreting this as evidence that best fits the research question. Second, although EBSE separates Steps 3 and 4, in our study, we addressed Step 3 and the initial planning for Step 4 together when prioritizing recommendations. This integration of the two steps provided a means to link the selection of evidence-based process change recommendations with the process improvement capability of the company. In other words, evidence assessment can help define the required process improvement activities. Third, if we accept the need for broad questions, we should expect Step 4 to deal with prioritizing available process change options and defining how to implement and monitor the selected process changes. Our study confirms that existing procedures for managing process change mean practitioners may be prepared for change but may place restrictions on the type of change they can manage. Fourth, in terms of Step 3, better standards for empirical methods would improve the likelihood of obtaining good-quality evidence, but it would require research expertise to assess specific evidence effectively.

Steps 4 & 5. The process changes planned in Step 4 should include procedures for monitoring the process change. Our

[10]We evaluated evidence in terms of its relevance to practice (i.e., applicability).

[11]https://conf.researchr.org/track/esem-2024/esem-2024-industry-government-community

experience suggests that monitoring the process change, and therefore evaluating EBSE, can also be a complex task, and about which we have the following reflections.

Comparing EBSE. When evaluating the EBSE process, it is important to consider what to compare it against. Two possible alternatives are: (1) comparing with expert opinion via consultancy or social media posts, (2) comparing with software process improvement models such as the Capability Maturity Model. In addition, in a specific evaluation study, it is useful to determine if the organization has defined procedures to manage process change. In our study, the EBSE team students had no experience in EBSE, process improvement methods, or evaluation. As the company was taken over, this was not a problem. However, any future EBSE evaluation should consider the final EBSE step in more detail. Researchers doing this should have some experience in undertaking field evaluations of process change. Also, such an evaluation might require monitoring the process over several weeks or months to decide if the recommended changes were fully adopted and addressed the diagnosed problems.

Evaluating the use of EBSE or the value of the collected evidence. Is it possible to evaluate EBSE by applying it to a SE problem, or are we evaluating the value of the collected evidence? From practitioners' viewpoint, it does not seem easy to isolate the usefulness of EBSE from the benefit of the collected evidence. In this study, we tried to make practitioners aware of EBSE and its methods throughout the EBSE project. We also had a final meeting in which we asked questions not only about the KM recommendations, but also about the evidence-based approach, in order to encourage their reflection on the two topics separately. However, we need more EBSE evaluation studies to be able to identify cross-case results or observations that are needed for a more objective meta-evaluation of EBSE.

E. Limitations

An important limitation of our study was the fact that the company closed down, so we were unable to confirm whether or not the process changes that the company tried to introduce were successful. We can only confirm that the company identified recommendations that they planned to adopt at our dissemination workshop and the requesters reported that some changes had been adopted before the company closed. We must note that in any case, our approach to monitoring the changes would still have been indirect, since at that point, none of the members of the EBSE team worked for the company, and we had not planned any direct monitoring of the change process.

We also identified two issues related to conflicts of interest. Firstly, the affiliation of one of the students with the company and the knowledge among requesters and company members that the research was part of their capstone project raised concerns about potential positive bias in their feedback. To mitigate this, we consistently emphasized to participants the equal importance of positive and negative results. We gathered their perceptions of the research results through multiple methods and on different occasions. Secondly, the first author sought to complete his research on the adoption of EBSE (that is part of his PhD), which could be thought of as more valuable with positive results. To minimize the risk of bias in favor of positive results, the first author diligently recorded decisions and actions in a detailed journal. They also reported and sought input from the second and third authors at various stages of the study, including planning, conduct, analysis, and reporting. These measures were implemented to ensure transparency and reduce potential bias.

Although students had prior training and collaboration with Pizard, this marked their initial involvement in conducting a secondary study. To mitigate possible deviations from the RR methodology, we strictly adhered to Cartaxo et al. recommendations [13], including protocol development beforehand, while seeking guidance from Vallespir. Moreover, we validated the RR protocol and the conduct of initial stages with Fernando Acerenza, a researcher knowledgeable in EBSE.

The selection of the topic to work on with the company introduced some limitations. The primary limitation is our focus on non-critical issues to avoid disclosing sensitive information and not expecting to address serious or urgent problems. This approach may restrict the breadth and potential impact of the research findings. Additionally, it may seem that the topic addressed (i.e., KM) and the company's problems were more aligned with management than with SE. However, there are diverse and recognized studies on KM in SE indicating the importance of this topic in our field (e.g., [78]–[80]). In our case, out of the 20 selected studies, 13 were from IT and SE venues, also showing this topic's relevance to SE. Even so, dealing with a topic close to management and not a classic one within SE can be considered a limitation.

VII. Final Remarks

Currently, EBSE seems more appropriate as an approach for researchers to enhance collaboration with practitioners, rather than a freely accessible mechanism for all stakeholders as initially expected two decades ago. However, research on practice-driven use of EBSE is not so well explored, presenting certain challenges. Our study's findings contribute to addressing this gap, and although it is a single application in a specific context, we venture to present the recommendations outlined in Table V.

We are moderately optimistic that EBSE will be more widely adopted in the future. Successful experiences of applying EBSE will be the greatest attraction to motivate its use where success is closely linked to overcoming the identified barriers, especially the lack or inadequacy of evidence. However, if we want to confirm if the EBSE framework is valuable, or whether it needs to be revised, reports on practice-driven applications of EBSE are essential.

VIII. Acknowledgments

We thank Forrest Shull for telling us about the Voice of Evidence column from IEEE Software magazine and Magne Jørgensen for mentioning the EDOS newsletter. We thank all the members of the company with whom we worked and the middle and senior management who allowed the

TABLE V
RECOMMENDATIONS FOR CONDUCTING PRACTICE-DRIVEN APPLICATIONS OF EBSE.

1. Consider all EBSE steps and not just conduct an SR. Taking a broader view of evidence utilization enables consideration of various aspects, such as leveraging previously published SRs on the topic or more accurately evaluating the application of the approach. Do not dismiss the assessment of studies (Step 3) before determining whether such information is necessary.
2. Initiate early and sustained engagement with practitioners to grasp their needs and expectations about evidence, fostering collaborative decision-making throughout the process. Intermediate validations, such as those of evidence discovered, appear to be crucial for achieving successful final results.
3. Consider practitioners' values (e.g., commitment to process improvement) and constraints on required evidence (e.g., openness to recommendations for large or small process changes, or timeframes required to obtain results).
4. Anticipate the limitations of evidence by employing broad questions, and be prepared to restructure recommendations found in primary studies. For instance, converting these recommendations into actionable process changes can enhance their practical utility.
5. Facilitate the reception of evidence by considering practitioners' expectations regarding knowledge transfer and incorporate face-to-face activities where they can discuss recommendations and implementation strategies.
6. Effectively evaluate evidence utilization by monitoring process changes and preemptively considering alternatives for comparison, such as expert opinions or performance metrics before implementing changes.
7. Strive to adhere to the EBSE process and uphold scientific rigor, including conducting validations with expert researchers, to mitigate potential researcher bias introduced by current evidence limitations.
8. Anticipate challenges that may arise in academia-industry collaborations. Specifically, researchers should maintain flexibility and consider recommendations from the literature (e.g.,[81]).

conduct of our study. We also thank Fernando Acerenza for his validation of our progress during the RR conduct. We thank Joaquín Lezama and Rodrigo García for their participation in the EBSE-based project team, whose contribution is better reported in the paper about the RR conduct [30]. Finally, we thank the ICSE 2024 anonymous reviewers whose comments and suggestions improved this report.

REFERENCES

[1] B. Kitchenham, T. Dybå, and M. Jørgensen, "Evidence-based software engineering," in *International Conference on Software Engineering (ICSE)*, 2004, pp. 273–281.

[2] F. Kamei, I. Wiese, C. Lima, I. Polato, V. Nepomuceno, W. Ferreira, M. Ribeiro, C. Pena, B. Cartaxo, G. Pinto, and S. Soares, "Grey Literature in Software Engineering: A critical review," *Information and Software Technology*, 2021.

[3] S. Beecham, D. Bowes, and K.-J. Stol, "Introduction to the ease 2016 special section: Evidence-based software engineering: Past, present, and future," *Information and Software Technology*, vol. 89, pp. 14–18, 2017.

[4] B. Cartaxo, G. Pinto, E. Vieira, and S. Soares, "Evidence briefings: Towards a medium to transfer knowledge from systematic reviews to practitioners," in *International Symposium on Empirical Software Engineering and Measurement*. New York, NY, USA: Association for Computing Machinery, 2016, pp. 57:1–57:10.

[5] B. Kitchenham, D. Budgen, and P. Brereton, *Evidence-based software engineering and systematic reviews*, ser. Chapman & Hall/CRC Innovations in Software Engineering and Software Development Series. London: CRC Press, 2015.

[6] B. Cartaxo, G. Pinto, and S. Soares, *Rapid Reviews in Software Engineering*. Cham: Springer International Publishing, 2020, pp. 357–384.

[7] E. Hassler, J. Carver, N. Kraft, and D. Hale, "Outcomes of a community workshop to identify and rank barriers to the systematic literature review process," in *International Conference on Evaluation and Assessment in Software Engineering*. New York, NY, USA: Association for Computing Machinery, 2014, pp. 31:1–31:10.

[8] F. Da Silva, A. Santos, S. Soares, A. Frana, C. Monteiro, and F. Maciel, "Six years of systematic literature reviews in software engineering: An updated tertiary study," *Information and Software Technology*, vol. 53, no. 9, pp. 899–913, 2011.

[9] R. Santos and F. Da Silva, "Motivation to perform systematic reviews and their impact on software engineering practice," in *International Symposium on Empirical Software Engineering and Measurement*, 2013, pp. 292–295.

[10] T. B. Callo Arias, P. van der Spek, and P. Avgeriou, "A practice-driven systematic review of dependency analysis solutions," *Empirical Software Engineering*, 2011.

[11] M. Kalenda, P. Hyna, and B. Rossi, "Scaling agile in large organizations: Practices, challenges, and success factors," *Journal of Software: Evolution and Process*, 2018.

[12] J. M. Sierra, A. Vizcaíno, M. Genero, and M. Piattini, "A systematic mapping study about socio-technical congruence," *Information and Software Technology*, 2018.

[13] B. Cartaxo, G. Pinto, and S. Soares, "The role of rapid reviews in supporting decision-making in software engineering practice," in *International Conference on Evaluation and Assessment in Software Engineering*, 2018.

[14] Q. Song, M. Borg, E. Engström, H. Ardö, and S. Rico, "Exploring ml testing in practice: Lessons learned from an interactive rapid review with axis communications," in *International Conference on AI Engineering*, 2022.

[15] E. Bjarnason, P. Åberg, and N. bin Ali, "Software selection in large-scale software engineering: A model and criteria based on interactive rapid reviews," *Empirical Software Engineering*, 2023.

[16] C. Hamel, A. Michaud, M. Thuku, B. Skidmore, A. Stevens, B. Nussbaumer-Streit, and C. Garritty, "Defining rapid reviews: a systematic scoping review and thematic analysis of definitions and defining characteristics of rapid reviews," *Journal of Clinical Epidemiology*, vol. 129, pp. 74–85, 2021.

[17] T. Dyba, B. Kitchenham, and M. Jorgensen, "Evidence-based software engineering for practitioners," *IEEE Software*, vol. 22, no. 1, pp. 58–65, 2005.

[18] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, 2006.

[19] F. Shull, T. Dyba, H. Sharp, and R. Prikladnicki, "Voice of Evidence: A Look Back," *IEEE Software*, vol. 34, no. 4, pp. 23–25, 2017.

[20] A. Kasoju, K. Petersen, and M. Mäntylä, "Analyzing an automotive testing process with evidence-based software engineering," *Information and Software Technology*, 2013.

[21] T. Lewowski and L. Madeyski, *Code Smells Detection Using Artificial Intelligence Techniques: A Business-Driven Systematic Review*. Springer International Publishing, 2022.

[22] N. Paternoster, C. Giardino, M. Unterkalmsteiner, T. Gorschek, and P. Abrahamsson, "Software development in startup companies: A systematic mapping study," *Information and Software Technology*, vol. 56, no. 10, pp. 1200–1218, 2014.

[23] D. Šmite, C. Wohlin, Z. Galviņa, and R. Prikladnicki, "An empirically based terminology and taxonomy for global software engineering," *Empirical Software Engineering*, vol. 19, no. 1, pp. 105–153, 2014.

[24] B. Kitchenham, P. Brereton, Z. Li, D. Budgen, and A. Burn, "Repeatability of systematic literature reviews," in *15th Annual Conference on Evaluation & Assessment in Software Engineering (EASE 2011)*, 2011, pp. 46–55.

[25] M. S. Schwartz and C. G. Schwartz, "Problems in Participant Observation," *American Journal of Sociology*, vol. 60, no. 4, pp. 343–353, 1955.

[26] A. J. Ko, "A three-year participant observation of software startup software evolution," in *International Conference on Software Engineering: Software Engineering in Practice Track (ICSE-SEIP)*, 2017, pp. 3–12.

[27] P. Runeson, M. Host, A. Rainer, and B. Regnell, *Case Study Research in Software Engineering: Guidelines and Examples*, 1st ed. Wiley Publishing, 2012.

[28] B. C. O'Brien, I. B. Harris, T. J. Beckman, D. A. Reed, and D. A. Cook, "Standards for reporting qualitative research: a synthesis of recommendations," *Academic Medicine*, vol. 89, no. 9, pp. 1245–1251, 2014.

[29] S. J. Tracy, "Qualitative quality: Eight "big-tent" criteria for excellent qualitative research," *Qualitative Inquiry*, vol. 16, no. 10, pp. 837–851, 2010.

[30] S. Pizard, J. Acerenza, Lezama, R. García, D. Vallespir, and B. Kitchenham, "Using RR to support SE practice: A citation analysis and a replication study," *Submitted to Empirical Software Engineering*, 2024.

[31] S. Pizard, F. Acerenza, X. Otegui, S. Moreno, D. Vallespir, and B. Kitchenham, "Training students in evidence-based software engineering and systematic reviews: a systematic review and empirical study," *Empirical Software Engineering*, vol. 26, no. 3, p. 50, 2021.

[32] S. Pizard, D. Vallespir, and B. Kitchenham, "A longitudinal case study on the effects of an evidence-based software engineering training," in *International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET)*, 2022, pp. 1–13.

[33] S. Saito, Y. Iimura, A. K. Massey, and A. I. Antón, "Discovering undocumented knowledge through visualization of agile software development activities," *Requirements Engineering*, vol. 23, no. 3, pp. 381–399, Sep. 2018.

[34] T. S. Mendes, M. A. de F. Farias, M. Mendonça, H. F. Soares, M. Kalinowski, and R. O. Spínola, "Impacts of agile requirements documentation debt on software projects: A retrospective study," in *Proceedings of the 31st Annual ACM Symposium on Applied Computing*, ser. SAC '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 1290–1295. [Online]. Available: https://doi.org/10.1145/2851613.2851761

[35] E. T. Stringer, *Action research third edition*, 3rd ed. Sage Publications, Inc., 2007.

[36] A. Birk and T. Dingsøyr, "Trends in learning software organizations: Current needs and future solutions," in *Third Biennial Conference, WM 2005 - Professional Knowledge Management*, K.-D. Althoff, A. Dengel, R. Bergmann, M. Nick, and T. Roth-Berghofer, Eds., 2005, pp. 70–75.

[37] D. Viana, J. Rabelo, T. Conte, A. Vieira, E. Barroso, and M. Dib, "A qualitative study about the life cycle of lessons learned," in *International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE)*, 2013, pp. 73–76.

[38] K. Yglesias, "Ibm's reuse programs: knowledge management and software reuse," in *International Conference on Software Reuse*, 1998, pp. 156–165.

[39] G. Matturro and A. Silva, "Rebec: A method for capturing experience during software development projects," in *International Conference on Knowledge Engineering and Knowledge Management - Knowledge Engineering and Management by the Masses*, P. Cimiano and H. S. Pinto, Eds., 2010, pp. 524–533.

[40] D. Viana, T. Conte, S. Marczak, R. Ferreira, and C. de Souza, "Knowledge creation and loss within a software organization: An exploratory case study," in *Hawaii International Conference on System Sciences*, 2015, pp. 3980–3989.

[41] S. Vasanthapriyan, J. Xiang, J. Tian, and S. Xiong, "Knowledge synthesis in software industries: a survey in sri lanka," *Knowledge Management Research & Practice*, vol. 15, no. 3, pp. 413–430, 2017.

[42] M. L. I. Gervigny and S. D. Nagowah, "Knowledge sharing for agile distributed teams: A case study of mauritius," in *International Conference on Infocom Technologies and Unmanned Systems - Trends and Future Directions (ICTUS)*, 2017, pp. 413–419.

[43] J. Soini, T. Makinen, and V. Tenhunen, "Managing and processing knowledge sharing between software organizations: A case study," in *Portland International Conference on Management of Engineering & Technology (PICMET)*, 2007, pp. 1108–1113.

[44] V. Santos, A. Goldman, H. Filho, D. Martins, and M. Cortés, "The influence of organizational factors on inter-team knowledge sharing effectiveness in agile environments," in *Hawaii International Conference on System Sciences*, 2014, pp. 4729–4738.

[45] D. Smite and T. Dingsøyr, "Fostering cross-site coordination through awareness: An investigation of state-of-the-practice through a focus group study," in *Euromicro Conference on Software Engineering and Advanced Applications*, 2012, pp. 337–344.

[46] H. Khan, A. Ahmad, and M. Alnuem, "Knowledge management: A solution to requirements understanding in global software engineering," *Research Journal of Applied Sciences, Engineering, and Technology*, vol. 4, pp. 2087–2099, 07 2012.

[47] A. Kammani, S. Aljahdali, and H. Date, "Km capability for software development: A case study of the indian software firms," *International Journal of Business Information Systems*, vol. 12, no. 1, p. 44–67, dec 2013.

[48] M. Pulkkinen., O. Mazhelis., P. Marttiin., and J. Meriluoto., "Support for knowledge and innovations in software development: Community within company: Inner source environment," in *International Conference on Web Information Systems and Technologies*, 2007, pp. 141–150.

[49] J. Soini, "Managing information and distributing knowledge in a knowledge-intensive business environment," in *Portland International Conference on Management of Engineering & Technology (PICMET)*, 2008, pp. 889–894.

[50] A. Heredia, R. Colomo-Palacios, and P. Soto-Acosta, "Tool-supported continuous business process innovation: a case study in globally distributed software teams," *European Journal of International Management*, vol. 11, no. 4, pp. 388–406, 2017.

[51] D. Šmite, N. B. Moe, A. Šāblis, and C. Wohlin, "Software teams and their knowledge networks in large-scale software development," *Information and Software Technology*, vol. 86, pp. 71–86, 2017.

[52] A. Heredia, J. Garcia-Guzman, A. Amescua-Seco, and A. Serrano, "Study of factors influencing the adoption of agile processes when using wikis," *International Journal of Software Engineering and Knowledge Engineering*, vol. 24, no. 06, pp. 859–885, 2014.

[53] M. Humayun, C. Gang, and I. Masood, "An empirical study on investigating the role of kms in promoting trust within gsd teams," in *International Conference on Evaluation and Assessment in Software Engineering (EASE)*, 2013, p. 207–211.

[54] J. L. Jurado, A. Fernandez, and C. A. Collazos, "Applying gamification in the context of knowledge management," in *International Conference on Knowledge Technologies and Data-Driven Business*, 2015.

[55] M. Milovanović, M. Minović, V. Štavljanin, M. Savković, and D. Starčević, "Wiki as a corporate learning tool: case study for software development company," *Behaviour & Information Technology*, vol. 31, no. 8, pp. 767–777, 2012.

[56] H. Chen, M. Baptista Nunes, G. Ragsdell, and X. An, "Extrinsic and intrinsic motivation for experience grounded tacit knowledge sharing in Chinese software organisations," *Journal of Knowledge Management*, vol. 22, no. 2, pp. 478–498, 2018.

[57] J. DeFranco and P. Laplante, "A content analysis process for qualitative software engineering research," *Innovations in Systems and Software Engineering*, vol. 13, no. 2-3, pp. 129–141, 2017.

[58] D. Budgen, P. Brereton, N. Williams, and S. Drummond, "What support do systematic reviews provide for evidence-informed teaching about software engineering practice?" *E-Informatica Software Engineering Journal*, vol. 14, no. 1, pp. 7–60, 2020.

[59] S. Pizard, F. Acerenza, D. Vallespir, and B. Kitchenham, "Assessing attitudes towards evidence-based software engineering in a government agency," *Information and Software Technology*, p. 107101, 2022.

[60] V. Garousi, M. Borg, and M. Oivo, "Practical relevance of software engineering research: synthesizing the community's voice," *Empirical Software Engineering*, vol. 25, no. 3, pp. 1687–1754, 2020. [Online]. Available: https://doi.org/10.1007/s10664-020-09803-0

[61] D. Méndez Fernández and J. H. Passoth, "Empirical software engineering: From discipline to interdiscipline," *Journal of Systems and Software*, vol. 148, pp. 170–179, 2019.

[62] B. A. Kitchenham, L. Madeyski, and D. Budgen, "How should software engineering secondary studies include grey material," *IEEE Transactions on Software Engineering*, pp. 1–1, 2022.

[63] V. Garousi, M. Felderer, and M. V. Mäntylä, "Guidelines for including grey literature and conducting multivocal literature reviews in software engineering," *Information and Software Technology*, vol. 106, pp. 101–121, 2019.

[64] A. Rainer and A. Williams, "Using blog-like documents to investigate software practice: Benefits, challenges, and research directions," *Journal of Software: Evolution and Process*, vol. 31, no. 11, p. e2197, 2019, e2197 smr.2197.

[65] A. Williams and J. Buchan, "Using the case survey methodology for finding high-quality grey literature in software engineering," in *The International Conference on Evaluation and Assessment in Software Engineering 2022*, ser. EASE 2022. New York, NY, USA: Association for Computing Machinery, 2022, p. 1–9.

[66] L. Briand, D. Bianculli, S. Nejati, F. Pastore, and M. Sabetzadeh, "The case for context-driven software engineering research: Generalizability is overrated," *IEEE Software*, vol. 34, no. 5, pp. 72–75, 2017.

[67] C. Wohlin and A. Rainer, "Challenges and recommendations to publishing and using credible evidence in software engineering," *Information and Software Technology*, p. 106555, feb 2021.

[68] V. Garousi, D. C. Shepherd, and K. Herkiloglu, "Successful Engagement of Practitioners and Software Engineering Researchers: Evidence

From 26 International Industry-Academia Collaborative Projects," *IEEE Software*, vol. 7459, no. c, pp. 1–13, 2019.

[69] J. Kontio, J. Bragge, and L. Lehtola, "The focus group method as an empirical tool in software engineering," *Guide to Advanced Empirical Software Engineering*, no. 2004, pp. 93–116, 2008.

[70] V. Garousi and M. Felderer, "Experience-based guidelines for effective and efficient data extraction in systematic reviews in software engineering," in *Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering*, ser. EASE '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 170–179. [Online]. Available: https://doi.org/10.1145/3084226.3084238

[71] B. Kitchenham and P. Brereton, "A systematic review of systematic review process research in software engineering," *Information and Software Technology*, vol. 55, no. 12, pp. 2049–2075, 2013.

[72] P. Avgeriou, M. Verbeek, S. Bibi, A. Chatzigeorgiou, and A. Ampatzoglou, "Identifying, categorizing and mitigating threats to validity in software engineering secondary studies," *Information and Software Technology*, vol. 106, no. February 2018, pp. 201–230, 2018.

[73] B. A. Kitchenham, L. Madeyski, and D. Budgen, "Segress: Software engineering guidelines for reporting secondary studies," *IEEE Transactions on Software Engineering*, pp. 1–1, 2022.

[74] N. A. Ernst and M. T. Baldassarre, "Registered reports in software engineering," *Empirical Software Engineering*, vol. 28, no. 2, p. 55, 2023. [Online]. Available: https://doi.org/10.1007/s10664-022-10277-5

[75] N. Saini, S. Singh, and Suman, "Code clones: Detection and management," *Procedia Computer Science*, vol. 132, pp. 718–727, 2018, international Conference on Computational Intelligence and Data Science.

[76] C. Gnanasambandam, M. Harrysson, A. Hussin, J. Keovichit, and S. Srivastava, "Yes, you can measure software developer productivity," *McKinsey's Digital and Technology, Media & Telecommunications Practices*, no. August, 2023. [Online]. Available: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/yes-you-can-measure-software-developer-productivity#/

[77] D. North, "Mckinsey developer productivity review," Oct 2023. [Online]. Available: https://dannorth.net/mckinsey-review/

[78] F. O. Bjørnson and T. Dingsøyr, "Knowledge management in software engineering: A systematic review of studied concepts, findings and research methods used," *Information and Software Technology*, vol. 50, no. 11, pp. 1055–1068, 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950584908000487

[79] I. Rus and M. Lindvall, "Guest editors' introduction: Knowledge management in software engineering," *IEEE Software*, vol. 19, no. 03, pp. 26–38, may 2002.

[80] R. Ouriques, K. Wnuk, T. Gorschek, and R. B. Svensson, "The role of knowledge-based resources in agile software development contexts," *Journal of Systems and Software*, vol. 197, p. 111572, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0164121222002485

[81] V. Garousi, K. Petersen, and B. Ozkan, "Challenges and best practices in industry-academia collaborations in software engineering: A systematic literature review," *Information and Software Technology*, vol. 79, no. September 2014, pp. 106–127, 2016.