Recruitment, Engagement and Feedback in Empirical Software Engineering Studies in Industrial Contexts

Norsaremah Salleh[a,*], Rashina Hoda[b], Moon Ting Su[c], Tanjila Kanij[d], John Grundy[e]

[a]*Department of Computer Science, International Islamic University Malaysia, Kuala Lumpur, Malaysia*
[b]*SEPTA Research, Electrical and Computer Engineering, The University of Auckland, Auckland, New Zealand*
[c]*Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia*
[d]*Department of Computer Science and Software Engineering, Swinburne University of Technology, Melbourne, Australia*
[e]*School of Information Technology, Deakin University, Geelong, Australia*

ABSTRACT

Context: Research carried out in industrial contexts studies are recognized as important to the advancement of software engineering knowledge and practice. However, several challenges present themselves in the three key phases of research carried out in industrial contexts, recruitment, engagement and feedback.

Objective: The aim of this paper is to report the challenges related to each of the three phases of research carried out in industrial contexts, and the associated solutions we have found useful from our combined body of industrial empirical software engineering research studies spanning four case studies, five grounded theory studies, seven survey studies and two quasi-experimental studies involving a total of over 400 industrial participants in the past decade.

Method: We designed an instrument to gather details of our studies carried out in industrial contexts studies and performed thematic analysis to synthesise and draw out the most prominent challenges faced.

Results: We present a set of recommendations around study design, conduct and reporting to try and mitigate some of these challenges as they apply specifically to industrial empirical research.

Conclusion: These recommendations can guide researchers, novice and experienced, working in close collaboration with industry stakeholders to make the most of their industrial software engineering research.

Keywords: *Empirical software engineering, industry, research, grounded theory, survey, case study, quasi-experiment, challenges, solutions, recommendations.*

Contents

*Corresponding author: Tel. +603 6196 5669; fax: +603 6196 5179 Email address: norsaremah@iium.edu.my (N. Salleh)
Physical address: Department of Computer Science, IIUM, P.O. Box 10, 50728 Kuala Lumpur, Malaysia.

1 INTRODUCTION

Interest in conducting empirical software engineering studies in industrial contexts, or empirical studies involving industry participants, has greatly increased over recent decades [1]. A major driver has been the need to better understand how different aspects of software engineering (SE) are practiced in "real world" settings, to learn from industrial challenges in applying these methods and tools, and to form a better understanding of industrial practices and needs. An additional benefit can be to positively influence industry practice.

Over the years, the SE community has increasingly emphasized the importance of conducting such empirical software engineering studies in industrial contexts, improving the knowledge base of real SE practices and helping software professionals make better decisions [22,23]. However carrying out such empirical software engineering studies in industrial contexts is very challenging. Limited guidelines exist to inform the process, particularly for novice researchers and also for experienced empirical software engineering researchers working in new domains or with new approaches. This has led in recent times to several papers and even an ICSE workshop series to report on key experiences and lessons learned. These are very valuable and help other researchers in designing their studies and avoiding common mistakes, and assist practitioners in improving collaboration with researchers, the quality of the studies carried out with them, and the impact of the studies. While these papers make valuable contributions, they are usually focused on one particular technique e.g. survey or experiment, limited in scope e.g. one or small number of studies, and anecdotal i.e. researchers report particular experiences and lessons of interest rather than systematically.

We analyse our experiences in study design and conduct and the key lessons learned in working with a range of empirical research methods, including case studies, Grounded Theory research method (referred to as GT studies), surveys and quasi-experiments. The authors have collectively carried out many different empirical studies with industry. These include studies of:

- agile software teams [2-5],
- agile project management and task allocation [6-8],
- customer and manager roles in agile teams [9-11];
- testing practices including work log analysis and performance appraisal [12-14],
- influence of tester personality on testing [15-16],
- personality differences between testers and non-testers [17];
- understanding current industry mobile app testing practices [18];
- studying the relationship between team climate and personality traits [19],
- adoption of agile methods [20], and
- usage of architectural documents for different tasks [21].

Using a specially designed instrument, we collected details of 18 of our studies carried out in industrial contexts and performed thematic analysis to synthesize and draw out the most prominent common challenges faced during three distinct phases of research carried out in industrial contexts as follows:

- **Recruitment** - this phase includes challenges in designing studies of interest to industry, finding suitable companies and industry participants for a study; and gaining ethics approval to collect data from the industry.
- **Engagement** - this phase includes challenges related to effective use of participants' time; approach to *designing* industrial data collection instruments and techniques; and approach to *conducting* industrial data collection.
- **Feedback to industry** - this phase includes challenges associated with sharing findings with the industry, including both participants and non-participating industry practitioners.

This paper describes these three critical areas of industrial empirical research, using examples of challenges and lessons learned from our collective body of work in the area. We identify key, common and recurring challenges that researchers are likely to face when trying to design, carry out and report back empirical SE studies involving industrial stakeholders. We identify for each area the solutions that have worked for us and also where possible those that have been reported by other researchers. We then for each make recommendations for fellow empirical SE researchers. We hope that our experiences will be useful for others working in this increasingly important yet challenging domain.

The rest of the paper is structured as follows: Section 2 discusses the key motivation and related work. Section 3 presents the research context and method. Section 4, 5, and 6 present the challenges of industrial stakeholder recruitment, engagement and feedback respectively, along with our solutions and recommendations for empirical software engineering studies in each. Section 7 concludes the paper.

2 RELATED WORK

Jain et al. [23] report some of their experiences and lessons learned from conducting *surveys* and *case studies* in industry. Among the significant challenges they faced include time factors, limited research skills and the difficulty in achieving a balance between methodological rigor and industrial needs [23]. In their experience report Fernández and Wagner [1] describe a range of practical challenges and lessons learned from over 30 *case studies* conducted in academia-industry collaborations. They structured the report into three topics: approaching a case study, conducting the case study and cross-cutting challenges. When approaching a case study, they provide some strategies on how to go about finding and convincing industry contacts, planning the study and dealing with uncertainties. In conducting a case study, it is important to properly identify the context, instruments, and continuously practice empirical and social skills [1]. Based on their experiences, they provided a list of success factors for case studies in industry that would help guide young researchers in dealing with the same.

In a previous study, Kanij et al. [24] report the experiences of handling industrial *surveys* focusing on software testing issues, including personality assessment, performance measurement, and factors influencing tester performance. A number of common challenges in designing and running such practitioner surveys were identified, including low response rates, targeted invitations and informed consent collection, questionnaire length, participant motivation and engagement, data analysis and security, and ethics approval. They also recommend a number of techniques that might be adopted to mitigate some of the risks and challenges.

Similarly, Torchiano et al. [25] provide a set of pitfalls and lessons learned from their experiences in conducting *survey* studies with software practitioners. Their lessons learnt cover different aspects of the survey process, from designing of initial objectives to designing questionnaire, emphasizing on the design issues. These include approaches to better defining survey experiment goals, identifying target populations for the survey, sampling the target population, questionnaire design, and participant recruitment.

Other related works focused on experiences in conducting *experiments* in the software industry. Vegas et al. [26] and Juristo [27] describe the difficulties experienced in running experiments at several companies. A systematic literature review of software industry experiments by Dieste et al. [28] describe 15 experimental studies run in the industry. Among the major challenges raised in these studies include time and cost issues, workload, and planning aspects. The results suggested that experimenting in industry is generally perceived to be problematic. To increase the likelihood of getting more industry participation, researchers should plan an experiment on a topic that is directly useful to the company (i.e. aligned with business goals), optimize human resources and provide a flexible schedule. Embedding an experiment in a training course can also help in getting industry participants but these professionals would be novices in the technologies being assessed [27].

These contributions to practical lessons learned from conducting empirical SE studies with industry practitioners are valuable. Despite a number of studies reporting on the diversity of experiences in conducting empirical study in industry, most of these studies focused on a specific type of empirical method, such as controlled experiment (e.g. [26], [27], [28]), survey (e.g. [25]), and case study (e.g. [1]).

In this paper, we have systematically analysed our experiences from multiple study methods employed in our studies over the past decade, which include case studies, grounded theory, surveys and quasi experiments. Table 1 highlights key differences of this work from related work. Our aim is to share our collective experiences based on the practical challenges we encountered in conducting empirical studies in the industry and to provide recommendations in three focus areas or phases of research carried out in industrial contexts: *recruitment*, maintaining the participants' involvement (*engagement*), and enabling *feedback* to industrial practice. We frame our lessons learned as a set of practical recommendations for fellow researchers, grounded in our own experiences of the challenges of conducting such studies and having to overcome those challenges. While the existing related studies focus on issues related to the design and conduct of a particular empirical research method, this study complements the existing studies by reporting challenges that are applicable to research carried out in industrial contexts targeting the areas of recruitment, engagement, and feedback to industry. Furthermore, our recommendations can benefit not only novice researchers but also experienced researchers aiming to conduct industrial empirical research.

Table 1
Comparison with Related Literature [1], [25], [26], [27]

Elements	Fernández & Wagner, 2016 **[1]**	Torchiano et al., 2017 **[25]**	Vegas et al., 2015 **[26]**	Juristo, 2016 **[27]**	**Our study**
Empirical method(s) used	Case study	Survey	Controlled experiment	Controlled experiment	Collective data from 18 industrial studies using case study, survey, grounded theory and quasi-experiment

Scope/target audience	Inexperienced researcher	Inexperienced researcher	Not reported	Not reported	Any researcher aiming to conduct empirical study in industry
Overall contribution	Discusses practical challenges and lessons learnt in conducting case studies in industry	Provides lessons learnt covering different aspects of survey processes, with emphasis on design issues	Discusses difficulties in running controlled experiments in industry	Challenges and lessons learnt in recruiting participants for experiments, designing and running experiments, and transferring results.	Presents key challenges that are particularly related to industry participants and the solutions to address them.
Investigation aspects	Three topics: i) Approaching a case study ii) Conducting a case study iii) Cross-cutting challenges	Reports experiences from four survey study phases: i) defining research objectives and target population ii) sampling, iii) designing a questionnaire iv) recruiting	Identified difficulties based on four topics: i) company involvement ii) experiment planning iii) experiment execution iv) data analysis and reporting	Three topics: i) recruiting participants ii) designing experiment iii) running the experiment iv) transferring the result	Presents challenges and recommendations based on three distinct phases of research carried out in industrial contexts: i) recruitment ii) engagement iii) feedback to industry

3 CONTEXT AND METHOD

3.1 Study Context

In this paper we draw upon our collective industrial empirical software engineering research experience. In particular, we collate and present challenges of research carried out in industrial contexts and related solutions, based on four case studies (findings described in [5, 7, 8, 12, 18]), five Grounded Theory research studies (findings described in [2-4, 6, 9, 11]), seven survey studies (findings described in [10, 13-15, 17, 19, 20]), and two quasi-experimental studies (findings described in [16, 21]).

In this study, we refer to the definition of *case study* in SE by Runeson et al. [55] as "*an empirical enquiry that draws on multiple sources of evidence to investigate one instance (or a small number of instances) of a contemporary software engineering phenomenon within its real-life context, especially when the boundary between phenomenon and context cannot be clearly specified*." (pg. 12). Since case studies are (by definition) conducted in real-world settings, they have a high degree of realism and typically low level of control. Replication is nearly impossible [55, 56, 57]. Data collection for case studies include ethnographic methods such as interviews and observations [55,56].

Grounded theory, or GT studies, refer to research studies conducted using the Grounded Theory research method [30, 52]. GT is complete research method that includes procedures such as literature review, data collection, analysis, theory formulation and reporting [4, 53, 54]. It is particularly well-suited to studying industrial practice as it enables the collection and analysis of industrial data from the practice field [30,

53]. Glasserian, or Classic GT, and Straussian are two popular versions of GT adopted in practice. A distinguishing feature of Glasserian GT, the version adopted in the included studies, is the absence of a clear research question upfront and its focus on theory generation rather than validation of existing theories [52-54]. GT can be adopted as a wrapper around other empirical methods such as case studies where GT analysis and theory formulation procedures are applied to data collected from case studies [63]. Our included studies, however, refer to interviews and observations based standalone GT studies. GT is increasingly being used to study software teams as it facilitates the investigation of social and human aspects [4, 5, 7, 8, 12, 18, 53].

We refer to *survey* as an empirical method used to collect information from or about people to describe, to compare or to explain their knowledge, attitudes and behavior [58]. Data are gathered from a sample, which is representative of the population of interest and the results are used to identify patterns, which can be generalized to the overall population [59, 60, 61].

Quasi-experiments refers to an empirical inquiry that aims to measure the causal impact of an intervention, i.e effects of "*manipulating one variable on another variable*" [59] and the subjects (participants) are not randomly assigned to treatments [62]. Table 2 summarizes these research methods, along with references to their seminal texts and popular references in SE literature.

Table 2
Research Method Description and References

Research Method	**Description**	**Seminal texts**	**Software engineering references**
Case Study	Investigation of contemporary phenomenon in real-world settings	Yin [56]	Runeson et al. [55]
Grounded Theory (GT)	GT is a complete research method, focusing on theory generation.	Glaser & Strauss [30, 52]	Hoda et al. [53] Stol et al. [54]
Survey	Gather data from sample representing population of interest	Fink [58] Robson [59]	Kitchenham & Pfleeger [60], Molléri [61]
Quasi- experiments	Measure causal-effect without random assignment of subjects to treatments.	Cook & Campbell [62]	Wohlin et al. [44]

3.2 Data Collection

We designed a data collection instrument in the form of an Excel spreadsheet on Google Site. Each of the five authors recorded details from their various industrial studies in this instrument, including information such as type of study (survey, experiment, Grounded Theory, case study), description of the aims and a summary of the findings, details of the engagement such as data collection method (e.g. face-to-face interviews), location in which the study was conducted or location of participant in case of digital contact, number of companies and individual participants involved, recruitment process, challenges with recruitment, engagement and feedback faced, and related publications.

Table 3 summarizes some of the relevant details of our included empirical studies, including the study type (i.e. the research method employed, e.g. case study, grounded theory, surveys or quasi-experiments),

the authors and year of publication, a short description of the paper topic, location of the participants of the studies, number of companies participating in the study (#C), and number of participants (#P). In the rest of the paper, we refer to these as 'included studies'.

Table 3
List of Included Empirical Studies (#C: Number of Companies; #P: Number of Participants)

Study Type	Author(s)/ year	Description	Location(s)	#C	#P
Grounded Theory	Hoda & Noble (2017) [4]	Investigating how software teams transition into agile software development	NZ, Australia, USA, India and Portugal.	18	31
	Hoda et al. (2010 & 2013) [2,3]	Investigating how agile software teams self-organize	NZ, India, Canada and USA.	23	58
	Hoda et al. (2011) [9]	Studying the level of customer involvement in agile projects and its effect on the self-organizing ability of the team.	NZ and India	16	30
	Hoda et al. (2016) [6]	Exploring agile project management challenges.	India	6	21
	Shastri et al. (2017) [11]	Exploring the role of the manager on agile teams.	NZ, India, US, Australia	18	20
Case Studies	Masood et al. (2017) [7,8]	Exploring task allocation in agile teams.	India	1	12
	Andriyani et al. (2017) [5]	Exploring how reflection happens during agile retrospectives.	NZ	1	16
	Kanij et al. (2014) [12]	An Empirical Study to Review and Revise Job Responsibilities of Software Testers.	Worldwide	6	6
	Zein et al. (2015) [18]	Exploring mobile apps testing methods and challenges faced by developers.	Palestine	5	13
Surveys	Shastri et al. (2016) [10]	Exploring whether the title 'project manager' still exists on agile teams.	31 countries	NA	97
	Kanij et al. (2015) [17]	An Empirical Investigation of Personality Traits of Software Testers	Worldwide	NA	182
	Kanij et al. (2014) [14]	Performance appraisal of software testers	Worldwide	NA	18
	Kanij et al. (2012) [13]	Performance Assessment Metrics for Software Testers	Worldwide	NA	104
	Kanij et al. (2014) [15]	A Preliminary Survey of Factors Affecting Software Testers	Worldwide	NA	104

	Soomro et al. (2015) [19]	Studying relationship between personality traits, team climate and performance	Malaysia	NA	36
	Salleh et al. (2014) [20]	Exploring Agile methods adoption	Indonesia	NA	21
Quasi experiment	Su et al. (2016) [21]	Studying consumer's usage of software Architecture Documents (ADs) when performing information-seeking tasks	NZ	NA	16[a]
	Su et al. (2016) [21]	Studying consumer's usage of ADs when performing information-seeking tasks (using KaitoroCap)	NZ	NA	18[b]
	Kanij et al. (2013) [16]	An empirical study of the effects of personality on software testing	Australia	NA	48

[a]16 industry + 14 academic; [b]18 industry + 11 academic + 9 students; NA: not applicable.

3.3 Analysis Method

Through discussion and consensus, the five authors agreed that the main phases of research carried out in industrial contexts, common to all included studies, and exhibiting key challenges to empirical software engineering researchers are *recruitment*, *engagement*, and *feedback*. In order to systematically collate and analyse our experiences and to determine key insights from them, related to each phase, we performed a thematic analysis of the included studies. To do this we:

1. Designed a detailed data collection instrument in the form of an Excel spreadsheet as described above (key aspects summarized in Table 3).
2. Each study's main author identified key challenges that occurred based on their experiences, and outlined the successful aspects of the study. Where possible, the co-author(s) cross-checked the extracted information, adding further or revising as needed;
3. Each author then went through each study's extracted data and tagged the key methods employed, challenges identified, and where present, solutions used to address challenges;
4. The authors collectively cross-checked this tagging, merging and labeling sub-themes used to identify commonalities across several studies;
5. The authors grouped sub-themes to identify key "questions" and associated challenges asked when designing and conducting empirical studies with industry
6. The authors then identified challenge-related "solutions", as used in our studies. This included identifying studies that failed to properly address challenges or would have benefited from subsequent study lessons and approaches.

Fig. 1 shows the key challenges identified under each research phase.

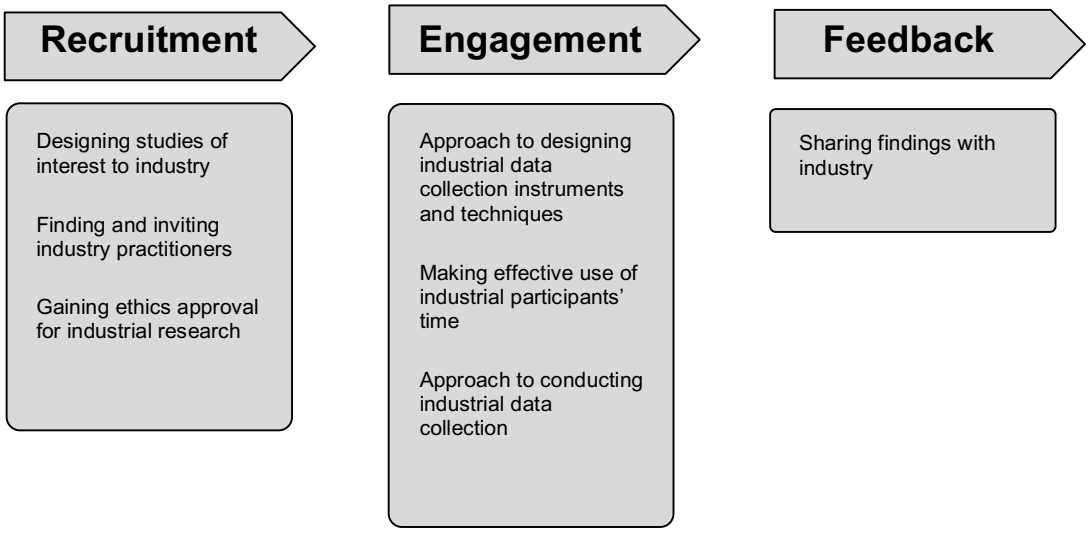

Fig. 1. *Challenges of research carried out in industrial contexts across three key phases: recruitment, engagement, and feedback*

A couple of examples of how challenges were derived from the analysis is shown here. A challenge encountered in study underpinning [21] was: *"As a new researcher, I knew only a few industry people with Software Architecture background making recruitment even harder."*

This challenge naturally fell under the recruitment theme. Since the challenge was related to finding participants, it was further categorized under *finding participants*.

Other entries described challenges with appropriate ways of inviting participants, for example this challenge faced in study underpinning [15]: *"We sent a long email to potential participants explaining details of the purpose of the survey, the expected outcome, benefits of the survey and the ethics approval details. This appeared not really helpful, using shorter, concise email text was rather effective."*

It was evident that long and detailed invitations were not effective. Thus, part of the recruitment problem was the type of invitations being sent out, i.e. a challenge with *inviting participants*. Given the intertwined nature of these two issues, the challenge of inviting participants was later combined with that of finding them (as described earlier) to produce the revised and final sub-theme '*finding and inviting industry participants*'.

3.4 Forming Challenge Themes and Questions and Identifying Examples

We also phrased the challenges using short, direct questions to better represent the crux of the challenge and improve relevance for readers. For example, *How and where can I find the right industry practitioners?* and *How do I invite them to become participants?* were associated with the theme *finding and inviting industry practitioners* as one of the first challenges most researchers face during the first phase of research carried out in industrial contexts, **recruitment**.

Similarly, a set of questions such as *How can I make the most effective use of industrial participants' time? How should I design my industrial data collection instruments and techniques?* were associated with the challenges faced in the second phase of research carried out in industrial contexts, **engagement**.

Finally, questions such as *What are the best mechanisms to share research findings with industry?* constituted the challenges faced in the last phase of research carried out in industrial contexts, **feedback**. Collectively, these make up the most pertinent questions we as industrial researchers faced during our various studies. Not all studies encountered every challenge. However, we included the challenges that were most common across multiple studies.

3.5 Forming Challenge Solutions and Identifying Examples

We then extracted identified solutions to each of the challenges as reported in our publications, using each reported study as an exemplar of applying the solution. Along with this we also collaboratively listed solutions to each of the challenge based on experience and recommendations from our own practice and those previously reported e.g. in [1][25][26][27]. Another interesting observation is that sometimes a researcher had faced a challenge but did not necessarily have a solution, while another researcher had not faced the same challenge on account of having a useful strategy to avoid the challenge. This helped formulate solutions to all challenges, representing our collective experiences. It also added to our collective knowledge and repertoire of research strategies.

Table 4 summarises the key recommendations based on the challenges that we extracted from our studies analysis. In the following three sections, we present and discuss the key challenges that we have encountered in conducting our empirical SE research studies involving industrial stakeholders; the solutions we have found useful in avoiding or overcoming these challenges; exemplars of these challenges, solutions, and lessons learned, and our overall recommendations for other researchers for their studies.

Table 4
Themes (challenges) extracted from our studies analysis along with recommendations for each phase of research carried out in industrial contexts (recruitment, engagement, and feedback) from our analysis

Industrial Research Phase	Challenges/Questions	Recommendations
RECRUITMENT	Designing studies of interest to industry *"Why aren't they interested in my really important software engineering research?"*	• Network with local practitioner community to identify their interest and refine research focus accordingly • Pilot study early to acquire practitioners' interest • Use practitioner feedback to guide future studies
	Finding and Inviting Industry Practitioners *"I don't know them and they don't know me"*	• Get genuinely involved with and contribute to the local practitioner community to build a strong reputation as a genuine researcher and contributor. • Approach managers and team coaches as they are critical source of access to recruit more individuals in their teams, and sometimes the full teams. • Approach online groups through moderators to improve authenticity. • Craft the call for participations (CfPs) carefully to avoid a 'spam effect' • Prepare a small invitation email with catchy slogan to attract participation • Hire enumerators where necessary to help recruit participants • Perform snowball sampling (or word of mouth

		references) to reach wider participation ● Begin recruiting 'at home' (locally) before venturing out internationally, through opportunities at conferences and events, to build local relationships
	Gaining Ethics Approval for research carried out in industrial contexts *"How do I increase industry's confidence that my research is being conducted correctly so that they are more likely to participate?"*	● State clearly the method used to ensure confidentiality, consent, anonymity, and data security in the participant information sheet to improve industry confidence ● Ensure an appropriate safety protocol is defined and agreed and ensure practitioner manager informed consent is obtained to comply with ethics regulations
ENGAGEMENT	Approach to designing industrial data collection instruments and techniques *"How do I design data collection instruments and techniques that promote industrial engagement?"*	● Perform pilot data collection and refine before approaching industrial participants to improve industrial relevance ● Design demographic surveys to capture basic details prior to the main data collection session to customize and make the most of face-to-face time (primarily applicable for interview-based and observational studies) ● Questions should be designed to achieve high clarity to help elicit useful responses (simple language, clear instructions and avoiding jargon) ● Surveys should aim for an appealing presentation, and adequate layout to improve completion rates ● Specialised data collection tools must be secure, reliable and accessible as well as professional-looking to attract and sustain industry interest
	Making effective use of industrial participants' time *"How do I make the most out of my industrial participant's time?"*	● Be flexible with meeting schedules to accommodate busy professionals ● Schedule one or two additional backup slots in case of schedule changes ● Schedule observations between/around interviews on site to utilize participant's time effectively ● Ask for the minimum data as needed to answer research questions to prevent participants feeling overwhelmed
	Approach to conducting industrial data collection *"How do I encourage a curious mindset and conducive environment for industrial data collection?"*	● Adopt a curious mindset, not an auditing approach to enable participants to be forthcoming ● Be flexible around participants' preference for recordings and be prepared to take extensive notes instead ● Avoid including team leaders or managers in interviews with subordinates so that they don't feel intimidated and can be confident of anonymity

FEEDBACK TO INDUSTRY	Sharing findings with Industry *"What mechanisms can I use to share research findings with my participants and other industry practitioners?"*	• Various industry-friendly formats and mediums should be employed to share findings with the industry, including short videos, posters, brief reports of main takeaways, and talks or presentations to industry at industry-focused events to meet industry standards and expectations • Results should be shared after all data collection at a given company is complete so as not to bias other participants from the same company.

4 RECRUITMENT

Empirical software engineering researchers need to design studies of interest to practitioners, locate suitable industry practitioners for their studies; recruit them to participate in their studies; and ensure their studies are carried out adhering to required ethical processes and policies, including data security.

4.1. Designing studies of interest to industry

Challenge "Why aren't they interested in my really important software engineering research?"

Arguably the most frustrating problem for empirical SE researchers in recruitment is a lack of industry practitioner interest in what the researchers perceive to be very important studies for the discipline [1][25]. In our experience, practitioners predominantly want to take part in studies of close interest to their job area, that may make a difference in their current workplace, that they feel are of practical value to the industry, and where there is cost-benefit for them personally i.e. the benefit they gain versus the demand on their time is positive. In contrast, researchers may perceive a much larger challenge or problem in software engineering practice that, while an empirical study would provide a useful research contribution, its short-term practical benefit is unclear.

Solutions

We found this to be a key issue in our own early surveys around tester personality. It is a very interesting research problem as to what impact tester personality has on tester performance, but the benefit to practitioners participating in very detailed personality analysis data collection was unclear to them, resulting in low take-up. In subsequent studies, we used much more lightweight data collection as well as much more clearly articulated immediate benefit to participants from our study. For example, in [14] we modified the survey design where participants no longer required to use our proposed performance appraisal form. Instead, the form was presented to the participants for review and then the feedback questionnaire was presented. The lightweight survey took around 15-20 minutes to complete. This produced much higher responses to the survey.

Establishing a strong and wide network with the practitioner community helps researchers to better understand their interests and needs. We drew on our considerable relationships with the Agile community [4] and tester community [15] as motivation for our studies in these areas. Furthermore, detailed interviews, observations or discussions with industry participants in our testing tools [18], agile teams [4], and software AD usage studies [21] all helped informing us of current and potential future

topics. We found that participation in industry-based discussion forums and events such as open spaces helps gauge industry interest in research topics, also born out by others [1].

Failure to engage practitioners deeply and to suitably refine research questions and study focus are common problems [1]. Piloting a study with a small number of practitioners has been demonstrated to be valuable in previous work in case study research with industry, where they recommend structuring case studies iteratively, taking early samples and pilot testing [1]. We failed to sufficiently pilot our early personality testing surveys, resulting in low engagement [14]. In contrast, we refined our worklog, tester appraisal and software AD usage studies, tasks and instruments, and GT interview questions from early feedback from such pilot studies [12, 17, 21]. This resulted in better engagement but also greater practitioner interest in the studies and their outcomes.

Empirical SE studies need to be framed with both good research questions and useful, tangible practitioner outcomes that will attract wide interest and participation. For example, our mobile app testing case study aimed to advance both a deeper understanding of current mobile app testing tool usage and limitations from a practitioner perspective and at the same time identify key app testing challenges and gaps that require better theoretical platforms and basic research projects [18]. This was also found in the case study research of [1], where they got early feedback from industry to identify whether they can commit to and are interested in the project. Our various GT studies were directly driven by industry interests as GT focuses on the most important concerns of the practitioners (e.g. [2-4]). We have found including open questions on surveys can also provide very useful suggestions and insights from practitioners for future studies that would be of real benefit to them [17].

Recommendations:

- *Network with local practitioner community to identify their interest and refine research focus accordingly.*
- *Pilot study early to acquire practitioners' interest*
- *Use practitioner feedback to guide future studies*

4.2 Finding and Inviting Industry Practitioners

Challenge "I don't know them and they don't know me"

An early challenge of research carried out in industrial contexts is to find and invite industry practitioners to become research participants [1]. As researchers starting out on an empirical software engineering study with industry, these are the first challenges encountered [25].

Solutions

A traditional approach to finding and inviting practitioners is to attend practitioner-based events and become genuinely involved in practitioner-based networks in the area of interest. [1] suggest to visit practitioner conferences and share the new methods, insights, and lessons learned to spark interest in new studies. We have chosen to attend local meet-up groups, conferences and networking events. These provide a range of access to potential participants [2-6, 9, 11, 17]. An interesting aspect to note here is that managers and team coaches not only participated in some of our studies, but also became a critical source of access to multiple individuals in their teams or indeed, full teams in some cases [2-9, 11].

A more modern approach to finding and inviting practitioners is the online approach. For example, we have used the social media channels such as LinkedIn user groups, discussion boards, twitter, and mailing lists [2-9, 11, 15]. However, some of our early experiences were disappointing where we had low take-up of surveys and experiment invitations despite using a wide range of social media channels [14][15]. This was part due to the tendency for these channels to have too many irrelevant advertisements including for participation in empirical SE research [25]. Another cautionary issue others have found [25] is to avoid a "spreading spree" through mailing lists, as it's impossible to control who will read the message and also unable to measure the response rate. We were criticised by some referees for our papers [14][15] for not being able to quote response rates when using social media recruitment channels.

The online approach is prone to such overuse and call for participants (CfPs) can easily be ignored as spam. Learning from our earlier experiences, one approach we have now been using extensively to mitigate this is to receive permissions from the moderator/administrator [12] and/or request them to send the CfP out [2-6, 9, 11]. The major benefits of this process was the more moderators we can convince, the larger number of people we could invite. This process of recruitment helped us to get a good number of participants in the short available time, for two of our follow-on surveys from the earlier less successful ones [14, 15].

Furthermore, investing effort in crafting the CfPs is needed to minimize this 'spam effect'. In some of our earlier studies we did not emphasise the academic nature of the study and our credentials, contributing to low participation [15][21]. From then on we have made researcher credentials including affiliation and contact information upfront and clear. We highlighted the academic nature of the research study, reinforcing in the email/forum posting heading, in the introduction text, and again at the end. A quality online profile for the researcher can further help establish authenticity. We avoided keywords in the heading that might even suggest at a commercial, spam-like posting or email, lest aggressive auto-filters categorise them as such. For example our tester effectiveness email subject was "*A survey of key factors affecting effectiveness of software testing professionals*", tester personality "*The Effect of Personality Traits on Effectiveness of Software Testing*", and tester tasks "*What Do Software Testers Spend Their Time Doing?*" [14, 17]. This experience is borne out by others [25], who highlight that good recruitment invitations can help boost participation and engagement.

Recruitment messages getting lost on busy social media feeds is highly likely. One very effective technique we learned from an industrial testing consultant was sending a short invitation email accompanied by a short "teaser" tweet to a large group of professionals to gain their attention and help attract participation [15]. The tweet was sent by our consultant collaborator from their industry account to their network. Our tweet was: "*Are developers from Mars and testers from Venus? Help Swinburne university researchers find out: <link of the survey>*". As a result, we received a very much higher number of responses in our survey using this approach compared with more traditionally worded and long invitation emails as used in our previous studies [17].

Representing multiple perspectives on a research topic may require researchers to elicit a range of software roles such as developers, testers, business analysts, as well as senior management [2-11]. This can be challenging. On the other hand, some studies focus on specific roles, e.g. testers [12, 15]. On some of our earlier studies we were not clear enough about the types of industry participants needed e.g. testing managers for performance assessment [13] and software architects [21]. We learned from these that we

must make sure that all such varied and specific participant requirements can be conveyed clearly in the call for participation (CfPs) and in the early survey or interview screening processes.

For such cases where specific and/or large numbers of participants are required, another strategy is to hire enumerators who have experience in recruiting participants [19]. An honorarium was paid to these enumerators for their effort in hiring the participants, similar to Sjoberg et al. [33]. Others have also used similar approaches, where they offered possible rewards e.g. raffles, payment, sharing results, to attract participation [25]. Snowball sampling is another approach that we have used [32], where participants invite more participants (of their network and who they think will be interested to participate) [2-9, 11, 21]. Finally, fellow researchers can also provide a rich source of participants for surveys, case studies and experiments [16, 21]. However, [25] do caution that uncontrolled forwarding invitations could extend the sample in an unforeseen way e.g. participants being disengaged or wholly unsuitable.

Some participants are more difficult to recruit than others and this needs to be planned for. For example, we often found senior, very experienced testers and senior software architects (often resources shared across teams) to be very challenging to recruit for our worklog and AD usage studies [12, 21]. Personal and professional contacts and their referrals had to be mobilized for recruiting such key targeted participants. Unfortunately we still got very low participation. On reflection, to find such challenging participants, we should have used our wider researcher contacts earlier to help find others by snowballing.

Researchers tend to begin ‘at home’, looking for industrial practitioners in their local communities, and indeed others have encouraged this [1]. However, when a paucity of local participants is encountered, researchers can explore international options. Our experience and other researchers [1] have shown that conferences and events provide a viable opportunity for international data collection. We have found that both these approaches can work well and compliment each other. Relying on only one or the other is likely to be limiting.

Recommendations:

- *Get genuinely involved with and contribute to the local practitioner community to build a strong reputation as a genuine researcher and contributor*
- *Approach managers and team coaches as they are critical source of access to recruit more individuals in their teams, and sometimes the full teams.*
- *Approach online groups through moderators to improve authenticity*
- *Craft the call for participations (CfPs) carefully to avoid a ‘spam effect’*
- *Prepare small invitation email with catchy slogan to attract participation*
- *Hire enumerators where necessary to help recruit participants*
- *Perform snowball sampling (or word of mouth references) to reach wider participation*
- *Begin recruiting ‘at home’ (locally) before venturing out internationally, through opportunities at conferences and events, to build local relationships.*

4.3 Gaining Ethics Approval for Research Carried Out in Industrial Contexts

Challenge "How do I increase industry's confidence that my research is being conducted correctly so that they are more likely to participate?"

Some universities and research institutes require ethics approval prior to starting investigation. Industry typically requires NDAs [1] and sometimes other in-advance legal agreements to protect their and their clients sensitive data and interests. Such approvals are particularly important when working with industry as they are outside the university context and may not agree or always comply with said and unsaid norms for engagement prevalent within university contexts, for example, issues surrounding confidentiality, consent, anonymity, data storage and security.

Solution

How industrial data, processes and other sensitive information is going to be collected, stored, used and potentially disclosed are critical areas for clarity and upfront agreement. We have learned that it is a must to explicitly state the way data in particular will be used and how it will be stored so that industry participants and our University ethics committees are well informed about the study procedure. This can be accomplished through a participant information sheet that explains the purpose of the study and emphasizes the voluntary and confidential nature of the study. For example, we provided the participants with industry-oriented information sheets and consent letters [2-9, 11,18, 21]. Others have emphasised the importance of preparing a Non-disclosure Agreement (NDA) to spell out what can be published/ to ensure transparency [1]. Additionally, we also made it clear that the data will be retained for a period of six years, after which they will be destroyed. This seemed to put participants at ease.

As empirical software engineering researchers working with industry often need to visit practitioners, an additional challenge for Ethics Committees is the safety of their researchers. Additionally, the clear consent of industry managers for participation of their staff is often required. In some of our earlier studies [12, 13] this was not explicitly required by our Ethics committees. However, we have more recently had approvals withheld pending us carefully describing safety protocols explaining how our researchers will conduct data collection and their safety during industrial interactions. We have also been required to explicitly show that practitioner managers have given explicit consent to interviews, experiments, and exemplar data usage in the data collection and related publications [18].

Recommendations:

- *State clearly the method used to ensure confidentiality, consent, anonymity, and data security in the participant information sheet to improve industry confidence*
- *Ensure an appropriate safety protocol is defined and agreed and ensure practitioner manager informed consent is obtained to comply with ethics regulations*

5 ENGAGEMENT

Stakeholder engagement can be defined as "*an iterative process of actively soliciting the knowledge, experience, judgement, and values of individuals selected to represent a broad range of direct interests in a particular issue, for the dual purposes of: creating a shared understanding; making relevant, transparent, and effective decisions*" [35]. Engagement therefore involves a bidirectional relationship between researchers and the stakeholders and this requires ongoing communication to maintain the

stakeholders' interests in the studies. We describe below challenges and recommendations for initiating and maintaining effective engagement with industry stakeholders based on our experiences.

5.1 Approach to designing industrial data collection instruments and techniques

Challenge "How do I design data collection instruments and techniques that promote industrial engagement?"

For survey research, a poorly designed survey can be a turnoff for potential participants [25]. This is particularly relevant for industrial participants to meet professional standards and to motivate their engagement. Size/length and time taken for the data collection can influence the participants' motivation to engage. This is particularly critical in research carried out in industrial contexts as industrial participants are often very busy professionals and the researcher may not get second chances [1]. Even after refining the data collection instruments and techniques through piloting, given the relative lack of time industrial participants generally have, it can be challenging to collect all information during the face-to-face data collection (e.g. interviews or observations) [1].

Solutions

Testing a study's data collection devices (e.g. recorders); estimating time taken (e.g. survey or interview) and refining accordingly if too long or too short; practicing our data collection techniques (e.g. interviewing style, pace, follow-up questioning, notes taking etc.) are essential preparatory tasks so the researcher is confident and can perform data collection effectively. Piloting allows researchers, especially novices, to test and practice all of these aspects, refine them and then apply them effectively with real industrial participants. We have used piloting as a data collection design technique for most of our studies [2,3,5,7,8,11,15,18,19,21]. This has also been emphasised as a critical step by others [1]. This has enabled us to refine our respective data collection instruments and techniques and maximise suitability for industrial practitioner engagement.

It is important to accurately collect data about participants to ensure appropriate population sampling and for study analysis and reporting. This also helps to keep the main data collection short and focused. Doing this meant that we not only shortened the time for the actual interviews [6-8, 11], but also the interview questions were able to be customized and better focused in light of the demographic contextual information collected prior to the interview [5, 11]. In our tester surveys we used a lightweight demographic capture [13, 14, 17], that helped us to confirm the representativeness of the sample.

Instruments need to be very accessible and easy to use for industry participants. The data collection instrument (e.g. the online survey, semi-structured interview guide) should be composed using simple language, and clear instructions with jargon avoided or explained [25]. For survey research a pleasant design (e.g. Google forms provides several templates to select from) adds to a favourable participant experience. The layout should be adequate, e.g. appropriately sized response boxes, such as a paragraph textbox instead of a one line answer space if expecting a long answer, and providing an opportunity to add information that may not fit elsewhere, e.g. a last question that say "do you have any other comments?". This last point is equally applicable to interviews. For example, in our personality surveys it was critical to enable participants to supply a large inventory of personality trait information and these needed careful explanation plus ease of selection [17]. These things might sound trivial to novice researchers causing them to be ignored. In fact, they ensure that a data collection instrument is

understandable and non-frustrating to the participants, especially since surveys typically do not enable opportunities for clarification.

Some studies need to collect specific artefacts or data and require purpose-built tools for the study. In our work this was the case for capturing tester worklogs [12], software testers' personality traits [17] and AD usage data [21]. We had to ensure our tools were secure, reliable and accessible, both in terms of the data collection instruments they represented, capture of informed consent, and protection and management of data [17,12,21]. In our tester studies [12][15][17] some of our custom-built tools lacked a professional look and feel which adversely impacted on participation and feedback. Our Ethics committee also commented on this. On reflection, we should have paid more attention to a professional, polished interface as well as the specialised data collection and analysis aspects.

Recommendations

- *Perform pilot data collection and refine before approaching industrial participants to improve industrial relevance*
- *Design demographic surveys to capture basic details prior to the main data collection session to customize and make the most of face-to-face time (primarily applicable for interview-based and observational studies).*
- *Questions should be designed to achieve high clarity to help elicit useful responses (simple language, clear instructions and avoiding jargon)*
- *Surveys should aim for an appealing presentation, and adequate layout to improve completion rates*
- *Specialised data collection tools must be secure, reliable and accessible as well as professional-looking to attract and sustain industry interest*

5.2 Effective use of industrial participants' time

Challenge "How do I make the most out of my industrial participant's time?"

Project managers and software professionals have busy schedules and are constantly under the pressure of project deadlines and careful use of industrial participants time is an important issue [1]. Therefore, a major challenge in obtaining industry participants' engagement in empirical SE studies is in effectively organizing the meetings with them and/or getting them to respond to surveys [25].

Solutions

Cancellation of meetings by busy industry participants are inevitable. In some cases, if we could not re-schedule the meetings based on their requests, we lost the participants [16, 18, 21]. One strategy to cater for cancellation and extension of meetings is to schedule one or two additional slots. It also required us to factor in flexibility in the meeting schedule to cater for the convenience of the industry participants. For instance, having a flexible visit schedule was critical in our observational studies at participating companies [2-6, 18]. Meeting schedules need to be flexible but cancellation and rescheduling of meetings should be controlled so that participants do not experience diary/scheduling fatigue.

Observations of workplace and practices can be difficult to arrange on their own. Conducting observations in between or around scheduled interviews, and requesting for these on the spot during interviews can yield good results and makes for effective use of the participants' time [2-4, 9].

Overly complex and time-consuming instruments can have a negative effect on industry participants, especially shown in other survey research [25]. In our tester performance appraisal survey and interview, we unfortunately built a very detailed instrument to capture rich feedback from testing team managers. Initial deployment with this showed it was far too long and detailed for our target audience. Learning from this experience, we reduced its size significantly while still obtaining the key data we needed in a subsequent study [14].

It is critical to get participants to provide the data that is essential for answering the research questions. In one of our quasi-experiments on AD usage [21], we collected a wide variety of AD usage data (e.g. annotation data such as ratings, comments, tags, and interaction data such as clicks on hyperlinks). This turned out to be a mistake as while we got a rich set of data from small number but not all participants, rendering the data insufficient for deriving meaningful findings. Hence, we should have minimized the things a participant should do or provide in performing an experiment in order to reduce (or effectively use) the participant time and to prevent participant from feeling overwhelmed.

Recommendations

- *Be flexible with meeting schedules to accommodate busy industry professionals*
- *Schedule one or two additional backup slots to use in case of schedule changes*
- *Schedule observations between/around interviews on site to utilize participant's time effectively*
- *Ask for the minimum data as needed to answer research questions to prevent participants feeling overwhelmed*

5.3 Approach to conducting industrial data collection

Challenges "How do I encourage a curious mindset and conducive environment for industrial data collection?"

One major challenge we faced when conducting data collection in many of our studies is our ability to build the right mindset when approaching the industry participants. Industry participants had some perceptions about academia and therefore approaching them with the correct mindset we deemed important to obtain their honest feedback. Similarly, the physical environment when collecting data from industry participants could potentially influence the level and quality of engagement during data collection session. The environment and the right mindset both are very important in making the industrial participants comfortable in providing their responses. As emphasised by [1], skills do matter and they highly recommend to get involvement of experienced colleagues.

Solutions

As researchers, we should be wary of coming across as experts judging or auditing practitioners on their practices. This can happen particularly as the researcher gains a reputation for being the area expert. As such, we should provide a safe, non-threatening environment in which industrial participants can open up

and share their honest perspectives and experiences during interviews or data collection session. For example, conduct meetings at their workplace or a relaxed venue such as a cafe, instead of in our academic offices.

A related issue is the use of audio/video recordings which not all participants may be comfortable with. In such cases, we have attempted to explain the rationale for using recording (e.g. allows the researcher to focus on the conversation as opposed to rushing to jot the notes down) and that has helped participants relaxed and allowed for recordings [2,3,9]. In other cases, where the participants are adamant not to have recordings, the researcher must be ready to take extensive hand-written notes. This has happened occasionally where we adapted to the participant's preference and took extensive notes instead [2,3].

Participants need to be put at ease and well-engaged during interviews. This could help reduce the Hawthorne effect, i.e. the effect of being observed [42]. A relaxed approach and demeanour (e.g. being humble, responsive, and a good listener) can help them be themselves when being interviewed and/or observed. It has been emphasised in other case study research with industrial participants that use of experienced colleagues is important to assist with this [1]. An experienced researcher when teams up with a novice researcher will also provide a valuable apprenticeship opportunity for the novice researcher [1]. We learned that having a team of two researchers helped us to back-up each other during meetings and better engage the industrial participants [21]. For example, another researcher can take over when one gets engrossed in the conversation or misses a critical point. Having an experienced researcher in data collection meetings can also increase the credibility of the research team and attract more senior industry participants' engagement.

An environment of openness during interviews is also necessary to foster. We learned that it is better not to include a team leader with their developers in the same focus group meeting as this reduces the freedom of the developers to share their opinions or insights due to the presence of team leader in the session [18]. Manager engagement, however may potentially help to keep participant attrition rates down [18].

Recommendations

- *Adopt a curious mindset, not an auditing approach to enable participants to be forthcoming*
- *Be flexible around participants' preference for recordings and be prepared to take extensive notes instead*
- *Avoid including team leaders or managers in interviews with subordinates so that they don't feel intimidated and can be confident of anonymity*

6 FEEDBACK TO INDUSTRY

One of the key motivations for Empirical SE studies is not just to better understand industry challenges and current practices, but to inform practitioners and where feasible, to assist them to modify practices based on emerging knowledge and trends.

Challenge "What mechanisms can I use to share research findings with my participants and other industry practitioners?"

While it is common for academics to report the research findings in venues like journals and conferences, these type of reporting might not necessarily be appropriate for industry consumption. For example, in academic journals, reporting the findings must be accompanied by a rigorous description of research methodology, detailed presentation of results and thorough discussion and implications of findings [43]. In reality, practitioners typically do not want to read detailed reports or academic papers as managers and practitioners probably do not necessarily have the time or interest for it [26][27]. Additionally, the research papers will likely include research specific jargon and be written in formal academic language that can alienate industry readers, who may be struggling to find the main takeaways for practice [1].

Hence, the reporting must be adapted to make it suitable for industry audience and consider the timeliness of dissemination [27]. Reports can be written to address decision makers' information needs to assist these industry stakeholders in making better judgement of the suitability of adopting the software technologies under study [48].

Solutions

There are numerous ways in which researchers can opt for providing feedback and sharing the research findings to practitioners. A researcher can choose to either make the findings available through informal/formal discussion (face-to-face), or make them accessible online via a project or institutional website. Based on our experiences in [2-5], we shared findings from the included studies through practitioner events and conferences such as the AgileNZ, Agile India, Agile NCR, and local Auckland-based Agile community meetups and we found that our industry participant greatly benefit from the sharing sessions.

Most previous studies of empirical studies in industry have emphasised the difficulty industry has to grasp concept of significance in experimental analysis and that researchers should focus on diagrams, stating key findings more clearly and discussing the impact of results in an industry context [27]. They offer various suggestions including informal presentations and blog posts over technical reports to peer-reviewed publications [1], and the use of simple visual representations e.g. histograms, boxplots [26]. In our performance appraisal study [15], we provided a summary of the survey outcomes to the participating managers who registered interest to receive the outcome during participation. We compiled a snapshot summary of results targeted to testing team manager audience to maximise value and minimise time to read. Additionally, we also provided access to the appraisal instrument with usage instructions, which are convenient for practitioners (testing team managers) who are willing to try and adopt it [15]. Similarly, after our preliminary survey [17] and personality study [14] were completed, we shared the results with the participants who wished to receive a copy of the outcome as an industry-oriented report.

Surveys, after data analysis, lend themselves to practitioner updates via one or more of the communication mediums discussed above. We have asked for practitioner contact details for those interested in receiving a summary of results in many of our studies, storing separately from anonymous raw data [14, 17]. We provided a practitioner-oriented summary of key results rather than an academic paper pre-print. It has been suggested previously to share first with the company where the study was conducted and that the whole report should be made available through a technical report and/or open platform [1]. It is also worth thinking about sharing results as a means to reward participants [25].

It is helpful to allow interested practitioners to easily opt-in (or out) of project updates. We created researcher and project web sites for several projects [12, 17] that allowed us to both provide background to participants during recruitment but also provide practitioner-oriented outcomes. Similarly, we also

shared videos of conference presentations through YouTube links on the project websites [2, 3]. In [5], we created a A4 sized poster that summarized the main findings and takeaway of the study in a visually-appealing and diagrammatic format. This was also made available to the participating company for their reference and potential use in helping in translating research findings to practice.

Participants need to feel some value add or gain from participating in studies. In our early tester personality profile work the detailed personality analysis was popular and we found that practitioners wanted their own personality profile. However this was very time-consuming to produce. In our later personality profile survey, we could generate such a profile and this provided to be a very popular side-benefit of doing our large scale personality assessment of developers and testers [17]. We presented this in a user friendly way immediately after they answered the 50 Five Factor Personality Trait inventory in our purpose-built survey tool. This proved an attractive reason for many participants to do the survey.

Regardless of the medium, the material used to present empirical SE study findings should offer a concise and visually appealing representation. A research summary should be written using clear and an understandable language and show useful conclusions and recommendations for practitioners. With regard to timing for feedback, it is recommended to share results after all data from a given company has been collected so as not to bias the participants [1]. Similarly, for surveys and experiments, require the study to be completed and analysed before any useful insights can be gained by the team, meaning any reporting to participants can only sensibly be done post-study.

Other empirical studies in an industrial context have suggested the appointment of a champion to to help communicate study purposes and industry impact potential [1]. Some cautionary feedback includes the experiences with industry based experiments [27], where even though the results convinced managers to adopt the strategies, developers disagreed.

Recommendation

- *Various industry-friendly formats and mediums should be employed to share findings with the industry, including short videos, posters, brief reports of main takeaways, and talks or presentations to industry at industry-focused events to meet industry standards and expectations*
- *Results should be shared after all data collection at a given company is complete so as not to bias other participants from the same company*

7 CONCLUSION

Industry stakeholders are key to all industrial empirical SE research. Based on our extensive body of empirical SE studies involving over 400 industry participants in grounded theory, case study, survey and quasi-experimental studies over the past decade, we presented the most common and pertinent challenges faced in three areas of industrial studies: recruitment, engagement, and feedback to industry. Despite the variety of research methods, procedures, instruments, and approaches used, most of these challenges applied to all studies. Based on our collective experiences, we presented a set of recommendations around study design, conduct and reporting to try and mitigate some of these challenges. These recommendations

can guide researchers working in close collaboration with industry stakeholders to make the most of their industrial SE research.

Key challenges in recruitment activities include designing studies of interest to industry, finding and inviting industry practitioners, and gaining ethics approval. We outlined a set of 13 recommendations based on our experiences in dealing with challenges faced when recruiting participants. In addressing the issue of lack of industry practitioner interest, we recommend researcher to network with local practitioner community to identify their interest; perform pilot study research questions early, and use practitioner feedback to guide future studies. In finding potential participants, we encourage researcher's involvement in the local practitioner community, approach managers and team coaches, approach online groups through moderators, careful crafting of the CfPs, prepare small invitation email with catchy slogan, hiring enumerators, use snowball sampling and begin recruiting at "home" before venturing out internationally. In gaining ethics approval, state clearly the method used to ensure confidentiality, consent, and data security to improve industry confidence. Additionally, ensure that safety protocol is defined and practitioner manager informed consent is obtained.

Engaging with stakeholders throughout the study period is important to obtain continuous participation. A significant number of challenges have been identified including the approach to designing data collection instruments and techniques, effective use of industrial participants' time, and the approach to conducting industrial data collection. We reported eleven (11) recommendations in addressing these challenges. When designing industrial data collection instruments and techniques, we suggest researcher to i) perform pilot data collection and refine, ii) design demographic surveys to capture basic details prior to main data collection session, iii) design the surveys to achieve high clarity, iv) ensure adequate layout and appealing presentation, and v) ensure security, reliability, and accessibility of specialised data collection tools. To ensure effective use of practitioners' time, we need to be flexible in meeting schedule, allocate one or two additional back-up slots, and obtain only necessary data to avoid participants from feeling overwhelmed. When conducting industrial data collection, it is important to adopt a curious mindset, not an auditing approach, be flexible around participant's preference for recordings, and avoid including team leaders in interview meetings to make participants feel at ease to provide responses.

Providing participant feedback on study outcomes and translating research results into industrial practice are as yet under-researched areas. We proposed some recommendations on the mechanisms to share findings with the industry, which include using industry-friendly formats and mediums such as short videos, posters, brief reports and talks or presentations at industry-focused events. The results should be shared after data collection is completed so as not to bias results. We believe a lot remains to be done in this area, particularly to effectively communicate the research outcomes, seeking industry feedback for possible refinements, and ideally to achieve the level where research findings can demonstrate significant impact with respect to improving industry practice.

Based on the data analysis of our collective set of challenges, the most frequently encountered challenge across all our studies by far was recruitment. Within recruitment, the most frequent sub-problem encountered was finding and inviting participants. Since this is likely also one of the first problems encountered in conducting empirical studies in industrial contexts, it is arguably the most significant.

The key challenge relating to engagement across most studies was careful use of industry participant's time. This manifested in cancelled meetings, cut-short meetings, failure to finish surveys, low response rates, poorly used time by inexperienced researchers, and limited open ended feedback. Piloting

instruments, reduction of survey questions to the minimum to answer research questions, scheduling additional meeting times up front, being flexible to industry demands and schedules, and using experienced researchers in helping design and run interviewing tasks were all successful solutions.

In our collective experience the approaches being used to providing feedback to industry can be improved. While we have made efforts in our various studies in the past as explained above, a lot remains to be done in this area. For example, researchers should not only disseminate results to industry but also actively seek their feedback leading to possible refinements. The most desirable, yet arguably very difficult to achieve, level of this feedback can be in the form of making impact to practice where research can demonstrate tangible improvements in practice.

We hope that our experiences in conducting empirical studies in industry will help guide researchers particularly in addressing the challenges that occur during recruitment, engagement, and translation of results into practice. We believe the recommendations can be applied to different types of empirical studies to improve the overall conduct of the study that involves industry practitioners.

Acknowledgements: The authors gratefully acknowledge the many funding agencies that have supported their empirical studies, including the Australian Research Council, Ministry of Business, Innovation and Employment, Foundation for Research, Science and Technology, Ministry of Higher Education Malaysia, BuildIT NZ, Agile Alliance, Software Process and Product Improvement for New Zealand Software Industry (UOAX0710), and their respective Universities. The studies would not be possible without the help of many of our colleagues and students, and the willing industry practitioners, their managers and organisations.

References

[1] D.M. Fernández, S. Wagner, Case studies in industries: What we have learnt, in: Proceedings of IEEE/ACM 4th International Workshop on Conducting Empirical Studies in Industry (CESI 2016), 2016, pp. 25 – 31.

[2] R. Hoda, J. Noble, S. Marshall, Organizing self-organizing teams, in: Proceedings of the IEEE International Conference on Software Engineering, Cape Town, 2010, pp. 285 - 294.

[3] R. Hoda, J. Noble, S. Marshall, Self-organizing roles on agile software development teams, IEEE Transactions on Software Engineering, 39(3) (2013) 422 - 444.

[4] R. Hoda, J. Noble, Becoming agile: A grounded theory of agile transitions in practice, in: Proceedings of the IEEE International Conference on Software Engineering, Buenos Aires, 2017.

[5] Y. Andriyani, R. Hoda, R. Amor, Reflection in agile retrospectives, Agile Processes in Software Engineering and Extreme Programming, Springer, 2017, pp. 3 – 19.

[6] R. Hoda, L.K. Murugesan, Multi-level agile project management challenges: A self-organizing team perspective, J. Syst. Softw. 117 (2016) 245-257.

[7] Z. Masood, R. Hoda, K. Blincoe, Exploring workflow mechanisms and task allocation strategies in agile software teams, in: Proceedings of the Agile Processes in Software Engineering and Extreme Programming (XP2017), Springer, 2017, pp. 267-273.

[8] Z. Masood, R. Hoda, K. Blincoe, Motivation for self-assignment: A case study of factors agile developers consider, in Proceedings of the Cooperative and Human Aspects of Software Engineering (CHASE), 2017.

[9] R. Hoda, J. Noble, S. Marshall, Impact of inadequate customer collaboration on self-organizing agile teams, Inf. Softw. Technol. 53 (2011) 521-534.

[10] Y. Shastri, R. Hoda, R., Amor, Does the "Project Manager" still exist in agile software development projects?, in: Proceedings of the 23rd Asia-Pacific Software Engineering Conference (APSEC), 2016.

[11] Y. Shastri, R. Hoda, R. Amor, Understanding the roles of the manager in agile project management, in: Proceedings of the Innovations in Software Engineering (ISEC), India, 2017.

[12] T. Kanij, R. Merkel, J. Grundy, An empirical study to review and revise job responsibilities of software testers, in: Proceedings of the IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC), 2014, pp. 89 – 92.

[13] T. Kanij, R. Merkel, J. Grundy, Performance assessment metrics for software testers, in: Proceedings of the 5th International Workshop on Collaborative and Human Aspects of Software Engineering (CHASE), 2012.

[14] T. Kanij, R. Merkel, J. Grundy, Performance appraisal of software testers, Inf. Softw. Technol. 56(5) (2014) 495 – 505.

[15] T. Kanij, R. Merkel, J. Grundy, A preliminary survey of factors affecting software testers, in: Proceedings of the Australasian Conference on Software Engineering (ASWEC 2014), 2014.

[16] T. Kanij, R. Merkel, J. Grundy, An empirical study of the effects of personality on software testing, in; Proceedings of the 26th International Conference on Software Engineering Education and Training (CSEE&T 2013), 2013, pp. 239-248.

[17] T. Kanij, R. Merkel, J. Grundy, An empirical investigation of personality traits of software testers, in: Proceedings of the 8th International Workshop on Collaborative and Human Aspects of Software Engineering (CHASE 2015), 2015.

[18] S. Zein, N. Salleh, J. Grundy, Mobile application testing in industrial contexts – An exploratory multiple case study, in: Proceedings of the SOMET 2015, Naples, Italy, 2015, pp. 30 – 41.

[19] B.S. Soomro, N. Salleh, A. Nordin, How personality traits are interrelated with team climate and team performance in software engineering? A Preliminary Study, in: Proceedings of 9th Malaysian Software Engineering Conference (MySEC2015), 2015.

[20] N. Salleh, E. Al-Kautsar, R. Hoda, A.L. Asnawi, A window into the emergence of agile software development landscape in Indonesia, Int'l Journal of Advances in Soft Computing and its Application (IJASCA), 6 (1) (2014).

[21] M.T. Su, J.G. Hosking, J. Grundy, E. Tempero, Usage-based chunking of software architecture information to assist information finding, J. Syst. Softw. 122 (2016) 215-238.

[22] S.L. Pfleeger, Experimental design and analysis in software engineering, Annals of Software Engineering, 1(2) (1995) 219 – 253.

[23] S. Jain, M.A. Babar, J. Fernandez, Conducting empirical studies in industry: Balancing rigor and relevance, in: Proceedings of the IEEE/ACM 1st International Workshop on Conducting Empirical Studies in Industry (CESI 2013), 2013, pp. 9 - 13.

[24] T. Kanij, R. Merkel, J. Grundy, Lessons learned from conducting industry surveys in software testing, in: Proceedings of the IEEE/ACM 1st International Workshop on Conducting Empirical Studies in Industry (CESI 2013), 2013, pp. 63 – 66.

[25] M. Torchiano, D.M. Fernández, G.H. Travassos, R.M. de Mello, Lessons learnt in conducting survey research, in: Proceedings of the IEEE International Conference on Software Engineering, Buenos Aires, 2017.

[26] S. Vegas, O. Dieste, N. Juristo, Difficulties in running experiments in the software industry: Experiences from the trenches, in: Proceedings of the IEEE/ACM 3rd International Workshop on Conducting Empirical Studies in Industry (CESI 2015), 2015, pp. 3 – 9.

[27] N. Juristo, Experiences conducting experiments in industry: The ESEIL FiDiPro Project, in: Proceedings of the IEEE/ACM 4th International Workshop on Conducting Empirical Studies in Industry (CESI 2016), 2016.

[28] O. Dieste, N. Juristo, M.D. Martinc, Software industry experiments: A systematic literature review, in: Proceedings of IEEE/ACM 1st International Workshop on Conducting Empirical Studies in Industry, CESI 2013, 2012, pp. 2 – 8.

[29] R.M. de Mello, G.H. Travassos, Surveys in software engineering: Identifying representative samples, in: Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM '16), Article No. 55, 2016.

[30] B.G. Glaser, Theoretical sensitivity: Advances in the methodology of grounded theory, Sociology Press Mill Valley, CA, 1978.
[31] W.G. Cochran, Sampling Techniques, 3rd Ed., John Wiley & Sons, 1977.
[32] B. Kitchenham, S.L. Pfleeger, Principles of survey research: Part 5: Populations and samples, ACM SIGSOFT Software Engineering Notes, 27(5) (2002) 17-20.
[33] D.I.K. Sjoberg, The industry is our lab - Organisation and conduct of empirical studies of software engineering in Simula, in: A. Tveito, A.M. Bruaset, O. Lysne (Eds.), Simula Research Laboratory, Springer, 2010, pp. 443 - 458.
[34] K. Kelley, B. Clark, V. Brown, J. Sitzia, Good practice in the conduct and reporting of survey research, Int'l Journal for Quality in Health Care, 15(3) (2003) 261 - 266.
[35] P.A. Deverka, D.C. Lavallee, P. J. Desai, L.C Esmail, S.D. Ramsey, D.L. Veenstra, S.R. Tunis, Stakeholder participation in comparative effectiveness research: defining a framework for effective engagement, Journal of Comparative Effectiveness Research, 1(2) (2012) 181 - 194.
[36] J.S. Lemon, The effect of reminder intervals on response rates for web surveys, in: M. Trotman et al. (Eds.), Survey and Statistical Computing IV The Impact of Technology on the Survey Process, Association for Survey Computing, 2007.
[37] E.J. Weyuker, Empirical software engineering research - The Good, The Bad, The Ugly, in: Proceedings of the 2011 International Symposium on Empirical Software Engineering and Measurement (ESEM'11), 2011, pp. 1 – 9.
[38] Segen's Medical Dictionary. S.v. "data collection instrument". Available: http://medical-dictionary.thefreedictionary.com/data+collection+instrument, (accessed 12.04.2017).
[39] J. L. Bastos, R. P. Duquia, D. A. González-Chica, J. M. Mesa, R. R. Bonamigo, Field work I: selecting the instrument for data collection, An Bras Dermatol., 89(6) (2014) 918–923.
[40] SurveyMonkey, Available: http://www.surveymonkey.com/, (accessed 19.03.2017).
[41] C.A.B. Warren, T.X. Karner, Discovering qualitative methods: Ethnography, interviews, documents and images, 3rd Ed. Oxford University Press, 2014.
[42] S. Coombs, I. Smith, The Hawthorne effect: Is it a help or hindrance in social science research?, Change: Transformations in Education, 6 (1) (2003) 97 – 111.
[43] J.M. Verner, J. Sampson, V. Tosic, N.A.A. Bakar, B.A. Kitchenham, Guidelines for industrially-based multiple case studies in software engineering, in: Proceedings of the Third International Conference on Research Challenges in Information Science (RCIS 2009), 2009.
[44] C. Wohlin, P. Runeson, M. Hst, M.C. Ohlsson, B. Regnell, A. Wesslen, Experimentation in Software Engineering: Springer Publishing Company, Incorporated, 2012.
[45] G. Cuddeback, E. Wilson, J.G. Orme, T. Combs-Orme, Detecting and statistically correcting sample selection bias, Journal of Social Service Research, 30(3) (2004) 19 - 33.
[46] S. Zein, N. Salleh, J. Grundy, A systematic mapping study of mobile application testing techniques, J. Syst. Softw. 117 (2016), 334 - 356.
[47] P.M. Wilson, M. Petticrew, M.K. Calnan, I. Nazareth, Disseminating research findings: What should researchers do? A systematic scoping review of conceptual frameworks, Implementation Science, 5 (91) (2010).
[48] A. Jedlitschka, N., Juristo, D. Rombach, Reporting experiments to satisfy professionals' information needs, Empirical Software Engineering, 19(6) (2014), 1921-1955.
[49] F. Shull, M.G. Mendonca, V. Basili, J. Carver, J.C. Maldonado, S. Fabbri, G.H. Travassos, M.C. Ferreira, Knowledge sharing issues in experimental software engineering, Empirical Software Engineering, 9 (1) (2004) 111 - 137.
[50] T. Gorschek, C. Wohlin, P. Garre, S. Larsson, A model for technology transfer in practice, IEEE Software, 23(6) (2006) 88 - 95.

[51] L. Blessing, W. Seering, Preparing for the transfer of research results to practice: Best Practice Heuristics, in: Impact of Design Research on Industrial Practice, A. Chakrabarti and U. Lindemann (Eds.), Springer International Publishing, 2016, pp. 3 - 21.

[52] B. Glaser and A. Strauss, The discovery of Grounded Theory. Adline, Chicago, 1967.

[53] R. Hoda, J. Noble and S. Marshall, Developing a grounded theory to explain the practices of self-organizing Agile teams. Empirical Software Engineering, 17(6), (2012), pp. 609-639.

[54] K. J. Stol, P. Ralph, & B. Fitzgerald, Grounded theory in software engineering research: a critical review and guidelines. In Software Engineering (ICSE), IEEE/ACM 38th International Conference on, (2016), pp. 120-131.

[55] P. Runeson, M. Host, A.W. Rainer, B. Regnell. Case Study Research in Software Engineering. Guidelines and Examples. Wiley, Hoboken (2012).

[56] R. K. Yin. *Case Study Research: Design and Methods*, 3rd edition. SAGE Publications, 2003.

[57] I. Benbasat, D. K. Goldstein, and M. Mead. The case research strategy in studies of information systems. *MIS Quarterly*, 11(3):369, 1987.

[58] A. Fink.: The Survey Handbook, 2nd edn. SAGE, Thousand Oaks/London (2003).

[59] C. Robson: Real World Research: A Resource for Social Scientists and Practitioners- Researchers, 2nd edn. Blackwell, Oxford/Madden (2002)

[60] B. A. Kitchenham and S. L. Pfleeger. Principles of survey research: parts 1 – 6. ACM SIGSOFT Software Engineering Notes, 26–28, 2001–2003.

[61] J.S. Molléri, K. Petersen, E. Mendes, Survey guidelines in software engineering: An annotated review, Proc. of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, 2016.

[62] T. D. Cook and D. T. Campbell. *Quasi-Experimentation. Design & Analysis Issues for Field Studies*. Houghton Mifflin Company, 1979.

[63] S. Bick, K. Spohrer, R. Hoda, A. Scheerer and A. Heinzl. *Coordination Challenges in Large-Scale Software Development: A Case Study of Planning Misalignment in Hybrid Settings*, IEEE Transactions on Software Engineering, 2017. 10.1109/TSE.2017.2730870