The struggle is real! The agony of recruiting participants for empirical software engineering studies

Kashumi Madampe*, John Grundy*, Rashina Hoda
Monash University, Australia.
Email: firstname.lastname@monash.edu

Humphrey Obie
Telstra Health, Australia.
Email: humphreyobie@gmail.com

***Abstract*—Empirical software engineering researchers are committed to investigating better ways of improving software engineering practice. To do this, gathering data about software professionals' experiences and perspectives on the researched topic is crucial. However, in the last decade, our experience of recruiting software professionals as potential participants in our studies has met with fluctuating success. Therefore, we had to explore multiple possibilities when recruiting participants. This cost us more effort and time than anticipated. In this paper, we briefly discuss the challenges we faced when recruiting participants for empirical research in software engineering, the solutions we implemented to address these challenges, and a concise list of strategies for garnering the interest of potential participants.**

***Index Terms*—empirical studies, software engineering, participant recruitment, challenges**

I. Introduction

Empirical software engineering studies with industry participants are critical for advancing the discipline, including understanding key research and development (R&D) problems in industry settings, understanding work practices and tools, understanding how developers and stakeholders work together, and feeding into the software engineering (SE) research agenda. Many of these empirical studies focus on diverse human aspects and their impact on software engineering practice. Human aspects such as emotions [1], [2], [3], [4], [5], [6], personality [7], [8], [9], [10], [11], gender [12], [13], [14], and SE-related activities, such as software development [15], [16], [17], [18], requirements engineering [19], [4], [20], software testing [21], and agile practices [22], [23], [24], [17] continue to be popular areas for empirical research in SE.

However, it is extremely difficult to carry out experiments or quasi-experiments in industry settings to understand these human aspects and their software engineering impacts. More often, single or mixed-methods studies are usually employed, including surveys, interviews, focus groups, and observations. In the last decade, the number of hurdles faced when recruiting participants for such empirical research studies in SE has been high. In most cases, we faced difficulties in finding potential participants for the studies and having them participate to the expected level of time and quality input.

*The authors contributed equally to this work.

The key data collection methods we wanted to use included interviews and surveys, and to a lesser extent workplace observations, in our own studies. Given that we work in a constrained funding environment, we have limitations on the budget to support both our own research activity, personnel, and for participant recruitment. Our first attempt at participant recruitment in almost all of our studies has been finding voluntary participants. We advertised our studies on various social media platforms and shared them among personal contacts. If these methods failed to attract sufficient participants or sufficiently varied and representative participants, we tried recruiting participants via paid participant recruitment platforms. Each of these data collection methods, payment methods, location of the participant, and topic sensitivity, novelty and perceived practicality of the topic, have made participant recruitment challenging. In this paper, we discuss some key challenges, solutions, and ideas on how to further garner the interest of potential participants in empirical SE research.

The rest of this paper is structured as follows. First, we present the challenges associated with participant recruitment in Section II. Then we present some potential solutions for the identified challenges in Section III. Finally, we present how to maximise the recruitment potential in Section IV followed by the conclusion in Section V.

II. Challenges

A. C1. Recruitment challenges associated with data collection methods

Time-intensive data collection and analysis methods (qualitative methods) – such as interviews, focus groups, and observations – make participant recruitment more challenging than surveys (generally mixed-methods/quantitative). A key constraint is the amount of time the participants have to spend on participation, which is usually significantly more in the former methods than the latter. For example, an interview can usually take between 40–60 minutes, whereas a survey may be in the range of 15–30 minutes.

Interviews are also synchronous, i.e., the interviewer and interviewee must schedule a meeting, sometimes in very different time zones on a virtual platform. Data may need recording or live transcription, and post-collection transcribing, cleaning,

and potential follow-up checking with the participant. In contrast, most surveys are asynchronous in the participant's own time, and much of the analysis can be automated for the close-ended questions.

When our survey or interview studies were voluntary, the number of participants was generally much less than for paid participant recruitment. However, even if we paid the participants, if the studies took a longer time, the qualitative studies attracted fewer participants than mixed-methods/quantitative studies. The quality of answers could also be significantly impacted, including a greater number of invalid/incomplete responses in surveys.

B. C2. Recruitment challenges associated with rewarding participants

In the cases of many of our non-voluntary studies, "thank you" for participation methods – such as cash/gift card – were used. We noticed that the cash "thank you" offering research attracted more participants than the ones that offered gift cards. We also noticed that the choice of gift cards used also matters. For some participant groups, gift cards for groceries could be more important than a gift card to shop at Amazon or electronic stores.

We found significant overhead in obtaining and distributing gift cards, especially for larger studies. For example, buying 20 gift cards of different types, sending them to participants, and obtaining reimbursement from University finance systems is surprisingly challenging. Our systems also require the names and contact emails of all participants for tax reporting purposes. We found some participants, especially older ones, don't want a digital gift card but a physical one. Trying to buy, e.g., 10 physical gift cards proved challenging at some retailers. These also then had to be given to or mailed to participants.

C. C3. Recruitment challenges associated with paid participants

We have to obtain Human Subject Ethics Committee approval for all empirical studies. We also need to obtain finance approval for paid participant recruitment and for reward gifts. These can be time-consuming. We also found that our Ethics committee is uncomfortable with mixed participant recruitment, i.e., voluntary vs gift vs paid. Calculating a suitable pay rate for paid participant recruitment is important to reward participants sufficiently for their time, and ensure quality participation, but also remain within tight research budgets. We found some recruitment platforms have a large number of potential recruits, but their take-up of our survey invitations could be very slow vs other platforms with much quicker take-up and responses. For example, in our survey study which focused human-centric defects [25], we had quicker take-up and responses from Prolific than Amazon Mechanical Turk.

D. C4. Recruitment challenges associated with the territory

We have conducted empirical studies with participants from many countries. One such experience was that, in one of the interview-based studies, we first recruited participants from a country with a small land mass and population (*country A*). Then we tried to recruit participants for the same study from a country with a large land mass and population (*country B*). The number of participants who showed interest in participating in *A* was much higher than in *B*. Even though *A* is a much smaller country in size and population than *B*, the participant attraction we received was higher. The reason we assumed here was that as the SE network was more tightly knitted in *A* than in *B* (as the SE community is smaller in *A* and SE community is larger in *B* and spread across a larger geographical area).

We found paid participant recruitment from different countries to be challenging. We found a large number of participants are available from a couple of countries on most recruitment platforms, and on advertising a survey a large number will sign up from these countries. However, we found the time spent and the answers given to be very poor by almost all of these participants. We learned to specifically exclude such responses.

E. C5. Recruitment challenges associated with the sensitivity of the topic of the study

Some research topics are interesting to some potential participants, but might not be for others. For example, the participants for studies related to sensitive topics such as emotions in the workplace can have second thoughts before they decide to participate in the studies. Similarly, concerns about anonymity and ensuring their answers can not be traced back to them/their employer are very important for many participants. Some potential participants for our studies indicated, verbally or by email, that these are key concerns preventing their participation. Some empirical studies e.g. investigating mental health, workplace conflict, job security, various unpleasant work-related experiences, drug use, non-conformist opinions, or other challenging topics tend to present major recruitment, engagement and honesty challenges.

F. C6. Recruitment challenges associated with the novelty of the topic of the study

In one of our recent ongoing research projects, we came across an experience where the research topic was novel and emergent for potential participants. Some participants were not even aware of the existence of the emerging role of the 'AI product manager', even though they had been in the field for many years. This experience was further related to the geographical location as the role was a popular one in the USA but the SE community elsewhere had little visibility around the role and its key tasks. This made diverse participant recruitment much more challenging.

G. C7. Recruitment challenges associated with the practicality of the study

Some empirical software engineering studies can come across as very 'academic' and their value to software practitioners, the target audience for participants, be rather unclear. Sometimes this can be due to the language used by us as

researchers, the way the study is advertised, the lack of clear outcomes for practitioners from the study, or the lack of understanding of current practitioner interests and concerns.

III. Some Potential Solutions

A common solution that we have used for data collection method-related challenges (C1) was rewarding our participants appropriately, and reciprocating the effort they put into the study [3], [4], [5], [25], [20], [11]. We have also made an effort to make sure we chose the right "thank you" gift method appropriate for our participants (C2) [26], [27], [10]. We used feedback from our participants to guide rewards described in recruitment for future studies.

We tried several paid platforms for participant recruitment including Prolific [5], [4] and Amazon Mechanical Turk [3], to address our paid participant recruitment challenges (C3). The former has become our preferred paid participant platform [25], with more engaged participants, better participant filtering, generally higher quality survey answers, and easier-to-use quality assurance and participant payment support. However, it is required to exercise caution in recruitment and authenticating responses.

To address challenges associated with territory (C4), the only solution we identified was to focus on subsets of the community rather than the whole community. For example, rather than focusing on the SE community in an entire country *B*, we focused individually on city-focused SE communities [6]. In some studies we had a large number of participants from a single country, often the one where we have a large professional network, and often the country that the PhD student leading the study was from [19]. Using a paid participant recruitment platform, we also learned to do several batches of recruitment from selected parts of the world/specific countries [3], [5], [4], in order to get more diverse participant populations. This can, however, further complicate recruitment and finance processes.

Being extremely careful and empathetic when crafting study advertisements, recruitment statements, consent forms, and questionnaires has helped us to address challenges related to study topic sensitivity (C5) [5], [3]. Being very clear about anonymity, confidentiality, data protection, and publishing aggregated results from the analysis of the data are important in ensuring concerns about the topic are addressed. Similarly, it is critical for participants being interviewed to know their answers are decoupled from their identities before data analysis, recordings are highly protected and encrypted or destroyed after transcription, their feedback is not shared with work colleagues, and interviews are conducted with as much confidentiality as possible, especially for online interviews. While advertisements and explanatory statements for studies need to describe the study scope and purpose carefully, they also need to clearly address these issues.

In order to overcome recruitment challenges related to study topic novelty (C6) and practicality (C7), we have made sure to explain the topic clearly and pilot the study advertisements, explanatory statements and instruments before recruiting participants. We have had some success, but generally only if the practice related to the topic is being practiced by the participants. For example, as reported above with our experience based on an on going research related to the role of AI product manager.

Addressing practicality concerns (C7), better outlining how the results may impact participants in their role in the future, how practical guidelines, tools or practices may be identified, how best practices may be more widely shared and promoted, and how participants will be provided with an industry practitioner-friendly study summary - all may aid in recruitment and engagement [3]. Historical trends of providing these for other studies help build reputation and trust.

IV. Maximising Recruitment Potential

We outline below some ideas for garnering interest from industry participants in empirical software engineering studies. We outline things that may go wrong or right and the associated risks the recruitment methods might have.

A. Recruitment method: Social media: LinkedIn, Meetups, Twitter, Facebook

What can go right:

- Sharing on social media: Usually, if the post is catchy, the people in the network react (like/comment/share). Such user activities make the post visible to a wider audience.
- Reaction-rate according to our experience: LinkedIn $>$ Twitter $>$ Facebook $>$ Meetups (Dependency: The proximity of the social network).
- Asking Lab, Department, Faculty and/or industry partners to promote on their network can also greatly boost reach, and sometimes much better reach target audiences.
- Response (reply) rate according to our experience: Meetups $>$ LinkedIn $>$ Twitter $>$ Facebook.
- Meetups are limited to the geographical location. For example, if we post our advert in a Meetups group in our area, the number of potential participants who reach out would be greater.

What might go wrong:

- Messaging on social media: The probability of getting a response to a recruitment call via direct messaging potential participants is very low.
- A number of channels are protective about content and some empirical study topics are unwelcome. For example, many groups forbid advertising for study participants.

Associated risks:

- The chance of achieving the representative sample size is low within a limited time frame.
- Response rate is usually impossible to quantify.

Overcoming the risks:

1) Use hashtags.
2) Share the post on specific interest groups related to the research.

3) Keep on resharing the post – a few times at regular intervals. For example, twice a week.
4) Have the call for participation open for a longer period.
5) Consider participant rewards within budget limits.

However, executing #3 and #4 might not be viable in every case.

B. Recruitment method: Personal contacts: Ex-colleagues, ex-students

What can go right:

- Reaction rate: The reaction rate so far when we tried recruiting participants via professional contacts has been high. The rationale we identify in this case is that as the contacts are known to us, they try their best to support our study. Invitations to professional contacts can be customised and tuned to their known role, interests etc.
- Response rate: Usually, the response rate is high when the participants are personal contacts. Sometimes they have participated in prior empirical studies, and sometimes they are willing to promote in their organisations and via their own networks, giving the study higher practitioner relevance and visibility.

What might go wrong:

- Personal contacts might not respond.
- Personal contacts may feel overburdened.
- Personal contacts may lack time for multiple study participation.
- Diversity of participants may be self-limited.

Associated risks:

- Participants may be influenced by involvement in previous studies.
- Given the participants are personal contacts, the study might suffer from participation bias and/or response bias. This will pose a threat to the validity of the research findings.
- Offering a reward for participation may or may not be seen as 'appropriate' for a personal contact.

Overcoming the risks:

1) Make all surveys anonymous. If asking for contact details, e.g., interest in receiving study findings or interest in the post-survey interview, provide a different channel to provide contact details or remote contact details immediately from survey results and prior to analysis.
2) When collecting data where identity can no longer be anonymous (example: interviews, observations, focus groups), the research team member whose personal contact is the participant where possible avoids being involved in collecting the data from the participant and analyzing data from the participant.
3) Where feasible, use purposeful sampling to ensure diversity of, e.g., role, experience, country, gender etc. This may mean drawing on multiple research team member professional networks and inviting participation in tranches.
4) Think carefully about the need for any reward offered for participation and the nature of the reward. Ensure both Ethics and Finance approvals are in place before offering.

C. Recruitment method: Recruitment platforms: Amazon Mechanical Turk, Prolific etc.

What can go right:

- Reaction rate: Usually, the participants show interest when the study provides some sort of reward for them.
- Response rate: The higher the amount we pay the participants, the more responses we tend to get.

What might go wrong:

- The number of participants available across roles may greatly vary. We have had previous experience of attracting multiple participants without any effort vs no luck at all.
- Over-paying sets unrealistic expectations for future recruitment.
- Under-paying for time or effort spent reduces take-up, greatly increases incomplete, low-quality responses and is unethical.

Associated risks:

- Surveys: The quality of responses could be low. Since the main motive for participants to participate in the studies is money, the participants may fill out the questionnaires with inaccurate/incomplete data. Such responses will need to be discarded before analysis. Participants who may not match the target criteria may sign up especially if payment is relatively high. Some sub-groups seem to produce uniformly high (and low) quality responses.
- Interviews and observation studies: We have used paid recruitment much less frequently for these, using professional networks, social media distributed advertisements, and "thank you" gifts. Interviewee payments normally need to be significantly more than survey participants due to the time needed. Time zones are an issue due to worldwide recruitment and the need for synchronous interviewing.

Overcoming the risks:

1) Have attention-check questions at different points of the survey questionnaire to increase the quality of the research.
2) Pilot the survey and interview questions on 2-3 representative participants to accurately gauge the time needed.
3) Look for similar studies on paid platforms to get an idea of pay rate, participant selection criteria, advertisement approach, etc.
4) Use tranches of recruitment so that the selection criteria can be fine-tuned and be more purposeful in recruitment.

D. Recruitment method: Recruiting from organisations, i.e., from software companies

What can go right:

- Reaction and response: We have found that it is extremely difficult to recruit participants directly from software companies. The reaction from the companies is often negative, for a variety of reasons. This includes confidentiality and intellectual property (IP) concerns, time commitment, and differences of opinion on the usefulness of the study amongst the authorising personnel.
- If previous positive experiences have been established between the academic organisation and the software company, the academic researchers may find direct company-based recruitment to be viable.

What might go wrong:

- Low chance of recruiting participants even though the company encourages employees to participate. As the employees are occupied with their own work, they might not show any interest in participating in the studies.
- Some companies may already have their own empirical research projects, e.g., to boost productivity or team climate. This may make the employees opt out of external studies.

Associated risks:

- Since two organisations are involved (academic organisation, software company), many agreements, non-disclosure agreements (NDAs), or contracts could be involved before sharing about the study with potential participants. This may be very time-consuming, may result in no agreement being reached, and cause a loss of interest in the software company to allow their employees to participate in the studies.
- Multiple approvals may be needed for the company employees to participate in a project, e.g., participant and their manager. Our experience suggests many industry participants are not comfortable having their manager explicitly approve participation in a research project, for a variety of reasons.

Overcoming the risks:

1) Patience - this is needed from both organisations as internal processes of both organisations often take longer than expected.
2) The advertisement for the call for participation has to be crafted carefully to attract busy target potential participants quickly.
3) Some organisations may be very sensitive about IP and the time commitment of participants. This may make these organisations too difficult to recruit participants/obtain approvals.

V. Conclusion

We have conducted several surveys, interviews, focus groups and observation studies with industry participants in the past few years. Many of these studies have faced challenging recruitment issues. In this paper we have provided an outline of some of the key issues that we have faced. We have summarised some of the successful solutions, some possible advantages, limitations and risks of these approaches, and some risk mitigation strategies. We hope that these will be beneficial for the research community.

Acknowledgment

We thank all the many industry participants and their organisations for participation in our empirical studies.

Madampe and Grundy are supported by ARC Laureate Fellowship FL190100035.

References

[1] D. Girardi, F. Lanubile, N. Novielli, and A. Serebrenik, "Emotions and perceived productivity of software developers at the workplace," *IEEE Transactions on Software Engineering*, vol. 48, no. 9, pp. 3326–3341, 2022.

[2] D. Girardi, N. Novielli, D. Fucci, and F. Lanubile, "Recognizing developers' emotions while programming," in *2020 IEEE/ACM 42nd International Conference on Software Engineering (ICSE)*, pp. 666–677, 2020.

[3] K. Madampe, R. Hoda, and J. Grundy, "The emotional roller coaster of responding to requirements changes in software engineering," *IEEE Transactions on Software Engineering*, vol. 49, no. 3, pp. 1171–1187, 2022.

[4] K. Madampe, R. Hoda, and J. Grundy, "A framework for emotion-oriented requirements change handling in agile software engineering," *IEEE Transactions on Software Engineering*, 2023.

[5] K. Madampe, R. Hoda, and J. Grundy, "Supporting emotional intelligence, productivity and team goals while handling software requirements changes," *ACM Transactions on Software Engineering and Methodology*, 2023.

[6] K. Madampe, R. Hoda, and P. Singh, "Towards understanding emotional response to requirements changes in agile teams," in *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering: New Ideas and Emerging Results*, pp. 37–40, IEEE, 2020.

[7] S. D. Vishnubhotla and E. Mendes, "Examining the effect of software professionals' personality & additional capabilities on agile teams' climate," *Journal of Systems and Software*, vol. 214, p. 112054, 2024.

[8] S. D. Vishnubhotla, E. Mendes, and L. Lundberg, "Investigating the relationship between personalities and agile team climate of software professionals in a telecom company," *Information and Software Technology*, vol. 126, p. 106335, 2020.

[9] T. Kanij, R. Merkel, and J. Grundy, "An empirical investigation of personality traits of software testers," in *2015 IEEE/ACM 8th International Workshop on Cooperative and Human Aspects of Software Engineering*, pp. 1–7, IEEE, 2015.

[10] D. Hidellaarachchi, J. Grundy, R. Hoda, and I. Mueller, "The impact of personality on requirements engineering activities: A mixed-methods study," *Empirical Software Engineering*, vol. 29, no. 1, pp. 1–56, 2024.

[11] D. Hidellaarachchi, J. Grundy, R. Hoda, and I. Mueller, "What's personality got to do with it? a case study on the impact of personality on requirements engineering-related activities," 2024.

[12] D. Russo and K.-J. Stol, "Gender differences in personality traits of software engineers," *IEEE Transactions on Software Engineering*, vol. 48, no. 3, pp. 819–834, 2020.

[13] Y. Wang and D. Redmiles, "Implicit gender biases in professional software development: An empirical study," in *2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS)*, pp. 1–10, IEEE, 2019.

[14] T. Kanij, J. Grundy, J. McIntosh, A. Sarma, and G. Aniruddha, "A new approach towards ensuring gender inclusive se job advertisements," in *Proceedings of the 2022 ACM/IEEE 44th International Conference on Software Engineering: Software Engineering in Society*, pp. 1–11, 2022.

[15] J. L. Butler, T. Zimmermann, and C. Bird, "Objectives and key results in software teams: Challenges, opportunities and impact on development," in *Proceedings of the 46th International Conference on Software Engineering: Software Engineering in Practice*, pp. 358–368, 2024.

[16] C. Bird, N. Nagappan, P. Devanbu, H. Gall, and B. Murphy, "Does distributed development affect software quality? an empirical case study of windows vista," *Communications of the ACM*, vol. 52, no. 8, pp. 85–93, 2009.

[17] K. Madampe, R. Hoda, J. Grundy, and P. Singh, "Towards understanding technical responses to requirements changes in agile teams," in *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops*, pp. 153–156, IEEE, 2020.

[18] H. O. Obie, J. Ukwella, K. Madampe, J. Grundy, and M. Shahin, "Towards an understanding of developers' perceptions of transparency in software development: A preliminary study," in *2023 38th IEEE/ACM International Conference on Automated Software Engineering Workshops (ASEW)*, pp. 40–45, IEEE, 2023.

[19] K. Madampe, R. Hoda, and J. Grundy, "A faceted taxonomy of requirements changes in agile contexts," *IEEE Transactions on Software Engineering*, vol. 48, no. 10, pp. 3737–3752, 2021.

[20] D. Hidellaarachchi, J. Grundy, R. Hoda, and I. Mueller, "The influence of human aspects on requirements engineering-related activities: Software practitioners' perspective," *ACM Transactions on Software Engineering and Methodology*, vol. 32, no. 5, pp. 1–37, 2023.

[21] T. Kanij, R. Merkel, and J. Grundy, "A preliminary survey of factors affecting software testers," in *2014 23rd Australian Software Engineering Conference*, pp. 180–189, IEEE, 2014.

[22] M. Berntzen, V. Stray, N. B. Moe, and R. Hoda, "Responding to change over time: A longitudinal case study on changes in coordination mechanisms in large-scale agile," *Empirical Software Engineering*, vol. 28, no. 5, p. 114, 2023.

[23] V. Stray, D. I. Sjøberg, and T. Dybå, "The daily stand-up meeting: A grounded theory study," *Journal of Systems and Software*, vol. 114, pp. 101–124, 2016.

[24] E. Mendes, D. Viana, S. Datta Vishnubhotla, and L. Lundberg, "Realising individual and team capability in agile software development: A qualitative investigation," in *2018 44th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, pp. 183–190, 2018.

[25] V. Chauhan, C. Arora, H. Khalajzadeh, and J. Grundy, "How do software practitioners perceive human-centric defects?," *arXiv preprint arXiv:2402.02726*, 2024.

[26] M. H. Vu, J. Wyman, J. Grundy, and A. Madugalla, "Better understanding diverse end user website usage challenges with browser-based augmented reality approaches," in *International Conference on Evaluation of Novel Approaches to Software Engineering*, pp. 269–291, Springer, 2022.

[27] W. Wang, H. Khalajzadeh, J. Grundy, A. Madugalla, and H. O. Obie, "Adaptive user interfaces for software supporting chronic disease," in *Proceedings of the 46th International Conference on Software Engineering: Software Engineering in Society*, pp. 118–129, 2024.