Informed consent and participant recruitment in studies of software practice

Conference or Workshop Item

Informed consent and participant recruitment in studies of software practice

Helen Sharp
School of Computing and Communications
The Open University
Milton Keynes, UK
helen.sharp@open.ac.uk

Tamara Lopez
School of Computing and Communications
The Open University
Milton Keynes, UK
tamara.lopez@open.ac.uk

Michel Wermelinger
School of Computing and Communications
The Open University
Milton Keynes, UK
michel.wermelinger@open.ac.uk

***Abstract*—All empirical software engineering studies require informed consent from participants. In this paper we explore issues we have faced around the need for informed consent for qualitative studies in organisational settings, and the effect they have on participant recruitment and engagement. These issues are: the "chicken and egg" of study design and data collection; the nature of data being collected; the difference between requirements gathering and research data collection; benefits and coercion; and participant reluctance and uncertainty.**

Keywords—qualitative studies, workplace studies, ethics

I. INTRODUCTION

Establishing an ethical and professional relationship between researchers and their participants is of paramount importance. This ensures trust, mutual co-operation and ultimately a more valuable set of research results. Part of this process involves making sure that participants understand the purpose of the data gathering and how the data will be used, and that they feel comfortable participating without coercion or undue pressure. This requirement is often fulfilled via an informed consent form, which must be signed or somehow acknowledged by the participant. For example, this information may be included at the start of an online survey where participants are asked to check a box to confirm agreement that their data may be used for the stated research; in a face-to-face interview this may take the form of a sheet of paper that interviewees are asked to sign. In many countries, studies involving human participants need to be approved by ethics committees, and before such approval can be given the study must be designed and suitable informed consent provisions stated. These committees often do not include members from the software engineering discipline.

This paper discusses issues around informed consent and ethical approvals which, in our experience, have impacted on participant engagement and recruitment. The studies concerned are qualitative, in situ studies of software practitioners and usually involve interviews, observation, and artefact analysis. Access to participants is usually gained through a gatekeeper within the organisation, although individual participants can choose to take part or not. The studies we draw on span many years and their purpose has varied from process issues [4, 6] to remote working within teams [1], and from security questions [5] to UX concerns [7]. Often the end goal is to improve support for software development including new tools and processes. Some of these issues are echoed in literature pertaining to ethnographic studies more broadly [2] but they have not been discussed in the context of software organisations.

To illustrate the issues, we begin with an account of an ongoing project with a large commercial company. This account illustrates the five issues which are described in more detail in the next section: (1) the "chicken and egg" of data collection; (2) the distinction between personal data or beliefs and "facts"; (3) the belief that data collection is no more than requirements gathering; (4) concerns around coercion; and (5) participant reluctance.

A. An example study

The purpose of the overall project is to enhance an existing research prototype for use by the company in their mainstream software production. An agreement between the university and the company has been signed and this includes NDA (non-disclosure) provisions. The research design has three stages: an initial phase to understand current working practices; a technically-focused phase to develop and enhance the existing prototype; and an evaluation phase in which developers will use the tool in their daily work. The first and third phases rely on participant engagement from developers within the company; we are currently focusing on the first phase.

In order to design data gathering in detail, we must meet with gatekeepers in the organisation so they can indicate the kind of engagement we can expect (e.g. interview, observation, or survey) and identify potential participants. In these meetings we have covered a number of areas including the prototype and its goals, and existing workflows where the modified tool may be used. Through these conversations, various types of information have been exchanged. For example, facts about the development process such as lists of tools currently in use by the teams and their interface details such as what does a green flag and a red flag mean, but also individual perspectives such as how workflow processes are managed in practice. This latter data set may be viewed as personal data rather than facts [issue 2], and according to our university's ethics process it cannot be used in our research because there is no informed consent with individuals in place [issue 1]. Moreover our gatekeepers have observed that no informed consent is needed because we are collecting requirements for the new tool rather than seeking personal data [issue 3]. Initial discussions suggest that the introduction of formal consent procedures may result in participant reluctance [issue 5] because of this confusion with requirements gathering, but we know that we have already heard personal viewpoints which are relevant to our research, and these need to be protected by informed consent. The primary benefit of the research from our point of view is the integration of our prototype into the company's workflows, and the primary benefit from the organisation's point of view is improved development quality. Our ethics committee have raised concerns that participants may be coerced into taking part because the main benefit is to the employer and not to the individual [issue 4].

II. Issues affecting participant recruitment and engagement

The following issues related to informed consent and ethical approvals have arisen in multiple projects. It is difficult to determine whether any of these issues have directly prevented any studies from going ahead, but they have shaped relationships and hence study designs. Each organisational engagement requires negotiation and discussion, including matters of access, data protection and informed consent:

- *The "chicken and egg" of study design and data collection*. When engaging with an organisation for research purposes there are several stages. The earlier stages may involve a range of meetings and interviews to explore the context and to decide on participant access and detailed data collection design. Without a detailed study design it is difficult to obtain approval from the ethics committee and so during the course of these earlier meetings, informed consent is usually not in place. Discussions in these earlier stages often bleeds into data collection, but without informed consent in place that information cannot be used as research data. In past projects we have overcome this through member checking and verification activities once informed consent is in place.
- *The different types of data being collected.* In particular, does informed consent cover all types of information? For example do "facts" (except personal details) need to be covered by informed consent? In the project described above some of the information gathered relates to what various interface symbols mean. Is this research data and hence require informed consent? In our observational studies it is common to take photographs of work being undertaken in context. This may lead to individuals being included in a photograph, e.g. of a large open plan environment, who are not directly involved in the research project. In this context getting informed consent from everyone in the photograph would have been impractical, yet our strict ethics committee rules preclude explicit use of that photograph as research data, or in reports to illustrate the context of our participants' work. The notion of personal data being "information that relates to an identified or identifiable individual" [3] is problematic in qualitative studies where, arguably, even program code can be traced back to the originator.
- *The difference between requirements gathering and research data collection*. Software developers are familiar with requirements gathering and the need to collect information through interviews and other interactions. Indeed, user-centred design emphasises the importance of understanding users and their context as part of product development. So even when tool development is not the central goal of a study, data collection appears similar to requirements gathering. Informed consent between individuals providing and collecting data is not usually required when the goal is requirements gathering for tool development. This can lead to uncertainty for potential participants and gatekeepers: why is consent needed for sharing our requirements?
- *Benefit and coercion*. Participants need an incentive to take part in a study, i.e. there needs to be a perceived benefit to them. In the studies considered by this paper, the benefit from taking part is more easily perceived as a benefit to the organisation rather than to the individual. In addition participants are approached through an organisational gatekeeper, and the distinction between benefit and coercion (where the organisation's persuasion of individuals to take part is seen as pressure) can become confused. In a previous project, our gatekeepers were very open about this: they introduced the research project to the whole company and asked individuals to volunteer directly to us. While the gatekeepers supported involvement, the decision was explicitly left to individuals. In other studies we have used a combination of gatekeeper introductions and direct contact from the researchers, but participants usually want to know that research sessions will run in the organisation's time rather than their own and that requires management support.
- *Participant reluctance and uncertainty*. Any of the above issues can affect participant recruitment and engagement. If the informed consent process is seen as too heavy-handed then participants may be nervous. In many cases we have not used written informed consent but have audio-recorded verbal agreement to informed consent statements on the advice of our gatekeepers. Having an NDA in place can help with participant recruitment and engagement because participants feel less concerned about sharing confidential organisational information. However, informed consent is even more important if opinions about confidential information are shared. If the value of the project is tool development and hence data collection is primarily requirements gathering, then why is informed consent required? These concerns can be resolved through discussion, but we have found that balancing expectations between our collaborators' cultures, research ethics and study design considerations can be challenging.

III. Workshop contribution

The need to treat participants with appropriate respect, to protect them from any potential harm and to secure their data from unauthorised access is paramount. To do this requires appropriate informed consent but in practice, it is not always straightforward. This short paper presents participant recruitment and engagement issues related to informed consent in qualitative studies of software practice. These issues have arisen in studies taking place within an organisational context that require gatekeeper access to professional software engineers. Although issues around consent have not prevented research from continuing, they have caused overhead and delay, and have persisted across different studies and contexts, and over a significant timespan. At the workshop we would like to share these experiences, and discuss others' approaches and suggestions for how best to tackle issues around consent: how best to negotiate with collaborators, how best to account for consent in the ethics approval process, and how to work with committees who are not generally drawn from the software engineering domain.

Acknowledgement

The work was partially supported by UKRI/EPSRC (EP/R013144/1 and EP/T017465/1), NCSC, and ABC Ltd.

REFERENCES

[1] Deshpande, A., Sharp, H., Barroca, L. and Gregory, A.J. (2016) 'Remote Working and Collaboration in Agile Teams', in *Proceedings of International Conference on Information Systems*, Dublin, 11-14 December 2016, pp 4543-4559

[2] Hammersley, M. and Atkinson, P. (2019) *Ethnography: Principles in Practice*, Routledge

[3] ICO (2022) "What is personal data" retrieved from https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/key-definitions/what-is-personal-data/ 10th March 2022

[4] Karvonen, T., Sharp, H., and Barroca, L. (2018) 'Enterprise agility: why is transformation so hard?', in *Proceedings of XP2018*.

[5] Lopez, T., Sharp, H., Tun, T.T., Bandara, A., Levine, M., and Nuseibeh, B. (2019) 'Hopefully We Are Mostly Secure': Views on Secure Code in Professional Practice', in *Proceedings of CHASE 2019*, workshop at *ICSE 2019*, doi: 10.1109/CHASE.2019.00023.

[6] Sharp, H. and Robinson, H. (2008) 'Collaboration and Co-ordination in mature eXtreme Programming teams' *International Journal of Human-Computer Studies*, **66,** 506-518.

[7] Zaina, L.A.M., Sharp, H. and Barroca, L. (2021) 'UX information in the daily work of an agile team: A distributed cognition analysis', *International Journal of Human-Computer Studies*, 147 (March), 102574.