

Journal Pre-proof

Embedding equality, diversity and inclusion in usability testing: recommendations and a research agenda

Sophie Rutter , Efpraxia Zamani , Jo McKenna-Aspell ,
Yuhua Wang

PII: S1071-5819(24)00062-4
DOI: <https://doi.org/10.1016/j.ijhcs.2024.103278>
Reference: YIJHC 103278



To appear in: *International Journal of Human-Computer Studies*

Received date: 22 September 2023
Revised date: 5 March 2024
Accepted date: 16 April 2024

Please cite this article as: Sophie Rutter , Efpraxia Zamani , Jo McKenna-Aspell , Yuhua Wang , Embedding equality, diversity and inclusion in usability testing: recommendations and a research agenda, *International Journal of Human-Computer Studies* (2024), doi: <https://doi.org/10.1016/j.ijhcs.2024.103278>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2024 Published by Elsevier Ltd.

Highlights

- Diversity in usability research participants is critical to ensure resulting technologies will be fit for purpose
- Giving participants a sense of value, trust and agency are likely universal principles for embedding equality, diversity and inclusion in research
- Organisations should facilitate an open and sharing culture where researchers can learn from mistakes
- Participants and researchers need to share expertise and perspectives to facilitate equality, diversity and inclusion

Journal Pre-proof

Title page

Embedding equality, diversity and inclusion in usability testing: recommendations and a research agenda

Authors:

*Sophie Rutter¹ s.rutter@sheffield.ac.uk,

Efpraxia Zamani^{1,2}, efpraxia.zamani@durham.ac.uk

Jo McKenna-Aspell¹. jmckenna-aspell1@sheffield.ac.uk

Yuhua Wang^{1,3} yuhua.wang@mmu.ac.uk

*Corresponding author

¹ Information School, University of Sheffield, The Wave, 2 Whitham Road, Sheffield, S10 2AH, U.K.

Present address

² Durham University Business School, Durham University, Mill Hill Lane, Durham, DH1 3LB, U.K.

³ Business School, Manchester Metropolitan University, Lyceum Pl, Manchester M15 6BY, U.K.

Abstract

Technologies support our everyday lives, and to ensure that people are not routinely excluded they must be *usable* by the wider population. However, technologies are not commonly tested with participants from a range of backgrounds. This paper reports on interviews and roundtable discussions with people whose identities can be underrepresented in usability testing and usability researchers to discuss how equality, diversity and inclusion (EDI) can be embedded in usability testing.

Key findings include (1) when people participate in research they need a sense of value, trust and agency, and (2) challenges for researchers for embedding EDI in usability testing include organisational pressures, stakeholder culture, getting guidance and recruiting who you need. Recommendations are made to researchers, and to the organisations that employ them. Additionally, we propose a research agenda for a community of users, creators of services and products, usability researchers, and all those advocating for EDI in usability research.

Keywords: Usability; EDI; Inclusive research; Diverse participation; Recruitment; Evaluation

1. Introduction

Increasingly, technologies are used in our everyday lives to support leisure, study and work (Yoo, 2010) and are an important driver of equality and empowerment (Johnstone, 2007). To

ensure that people are not routinely excluded from participating in these activities and opportunities, these technologies must be *usable* by the largest possible range of users, including those from minoritised and marginalised social groups.

Usability is defined in ISO 9241-11 as the “extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use” (ISO, 2018). Usability testing is crucial for improving the design of technologies (Tindale and Dimitri, 2022) and helps the development team understand user behaviour, and identify design constraints as well as design solutions (Moran, 2019). For users, it can reduce frustration by making technologies easier to learn, and their use more efficient, effective and satisfying. The importance of designing systems that are easy for **all** to use has long been recognised (Meiselwitz et al., 2009; Shneiderman, 2000) and increasingly usability is concerned with designing for “culturally diverse users, including underserved and underprivileged user groups, in the increasingly globalized world” (Acharya, 2022, p. n.p.). However, usability tests are still not usually conducted with a diverse demographic of users (W3C WAI, 2016), possibly because it is a resource intensive endeavour (van den Berg et al., 2023). Further, requirements elicitation and engineering techniques, which are closely related and relevant to usability testing, have often been criticised for failing to consider “marginalized social perspectives” and for being reductionist, further exacerbating exclusion and exclusionary practices (Raza, 2021, p. 1). This means that the needs of many actual and potential users can be omitted in the design and evaluation of technologies. The increased concerns for social justice and user empowerment (e.g. Harihareswara, 2015; Walls, 2016) have led to renewed calls for wider participation in usability testing (Twidale et al., 2021), including participation from marginalised groups (Acharya, 2022).

This paper contributes to the call for usability testing with users from diverse backgrounds. While there are studies on, and guidance for, conducting inclusive usability studies, these studies are usually tailored to a specific group of test participants, most notably for a particular age or disability. Increasingly, researchers are recognising that identities are complex, and there are limitations in only considering one aspect of a person’s identity, such as only gender or only disability (Sum et al., 2022). There is a missed opportunity to holistically consider how to inclusively accommodate research participants when they take part in usability tests without focusing on a particular characteristic such as age, gender or disability. Explicitly factoring in the diverse and complex identities of potential users during usability testing is of significant importance for representation and inclusivity purposes, and for reducing instances of stereotyping and biases, as studies from cognate disciplines have shown (Bohnert et al., 2021; Colliver, 2020). Our research questions are

- RQ1: How should researchers include people from diverse backgrounds in usability studies?
- RQ2: What are the challenges for researchers when embedding equality, diversity and inclusion in usability testing?

To answer these research questions, we conducted a multiphase study that brought people together from different backgrounds to discuss equality, diversity and inclusion (EDI) in usability testing, focusing on people whose identities can be underrepresented in usability testing, and usability researchers. Through the analysis of interviews, roundtable discussions and participant feedback, we provide practical recommendations for usability researchers as

well as for organisations that do usability research. These recommendations are intended to be universal and inclusive, and are not focused on any specific participant identity / identities. We also propose a research agenda for a community of technology users, creators of technologies, usability researchers, and all those advocating for EDI in usability research.

The paper is structured as follows. First, we provide an overview of relevant literature on EDI and usability testing. Then we describe our methodology, followed by the presentation of our findings. As part of the discussion, we propose recommendations for usability researchers, and the organisations that employ them, as well as a research agenda for the community. We conclude the paper by highlighting the novelty and significance of our research and summarising the contributions.

2. Related work

We review related work to consider the importance of EDI in user research and more specifically in usability testing. We also identify current guidance for conducting inclusive usability tests as well as what is missing from this guidance.

2.1 Examining usability through an equality, diversity and inclusion lens

In this study, we focus explicitly on EDI by leveraging the U.K. Equality Act 2010. Equality means that everyone is allowed equitable opportunities to achieve their full potential regardless of their protected personal characteristics. Diversity refers to respecting and valuing differences in abilities, ages, faith, cultural backgrounds, beliefs, sexual orientation and gender identity, whereby everyone irrespective of their personal characteristics is valued and supported to grow (European Institute for Gender Equality, 2023). Inclusion denotes everyone should feel they belong and that they are able to participate in full.

Issues of EDI have drawn considerable interest from researchers across a range of sub-disciplines and research areas such as ICT4D (Potnis and Gala, 2020), RRI (Smith et al., 2022), HCI (Mack and McDonnell, 2021; Schlesinger et al., 2017; Strohmayer et al., 2018), health (Imison et al., 2022; Tindale and Dimitri, 2022; Wheeler et al., 2022), government services (UK Government, 2017) and disability studies (Mankoff et al., 2010). These researchers have identified that there is a vital need to recognise the vulnerabilities of participants and the impediments to participating in research (Imison et al., 2022; Potnis and Gala, 2020; UK Government, 2017) and the need for participatory approaches that involve users at all stages (Imison et al., 2022; Mack and McDonnell, 2021; Mankoff et al., 2010; Tindale and Dimitri, 2022; Wheeler et al., 2022). There is also concern for how best to represent participant identities, and the need for intersectional frameworks (Strohmayer et al., 2018), richly describing participant identities (Schlesinger et al., 2017), as well as allowing participants to describe their own identities (Smith et al., 2022).

We consider that examining usability through the EDI lens can offer insights and improved outcomes for minoritised and marginalised social groups. Current research offers substantial

evidence with regards to the positive organisational, societal and business impacts of a diverse workforce and an inclusive society, and demonstrates how these can be embraced for providing equitable opportunities for all (Gagnon et al., 2022). More specifically, it has been shown that more inclusive approaches, especially within research contexts, can support interpersonal trust and more fruitful long-term collaborations with stakeholders (Tindale and Dimitri, 2022); in turn this can lead to technology designs that can be more acceptable by end users.

2.2 Embedding EDI in usability

We next discuss why embedding EDI in usability testing is particularly important. There are two key approaches to evaluating usability: inspection methods with usability experts (evaluating against predefined guidelines and heuristics) and empirical methods with users (Fernandez et al., 2011). The latter is of interest to this study. Ideally, those who already use the technology or are likely future users are recruited for the evaluation (Moran, 2019). These users are then observed while completing predefined tasks that are realistic and what the user may do in real-life (Moran, 2019). This focus on real users completing real tasks means that it is particularly important that participants in usability studies should be representative of all users and all tasks. However, diverse recruitment is considered resource intensive (Fernandez et al., 2011). Furthermore, researchers may lack motivation to diversify recruitment as a widely accepted usability recommendation is that only three to five users are necessary to identify most usability problems (Nielsen, 2000).

To improve the usability of technologies, testing in different contexts with a diverse range of participants is crucial for two reasons. Firstly, as a task-based approach is taken for usability testing with technologies tested against the specific goals of using the technology, increasing the diversity of tasks and contexts increases usability for a wider range of uses. Secondly, the wider the range of people who test a technology, the higher the likelihood that the technology is usable by everyone. For example, it is generally recognised that improving usability for disabled people improves usability for all (Rubin and Chisnell, 2008), and can increase understanding of how to design inclusive technologies (Elmore et al., 2014).

Three other features of the empirical usability test also make it an interesting case for investigating EDI in research. Firstly, users are observed by one or more persons, either remotely or in the same room. This could be intimidating for some users, particularly if the predefined tasks cannot be completed or completed easily; this is a reasonable probability given that it is the usability of the technology that is being tested (Rubin and Chisnell, 2008). Secondly, in a typical usability test users are asked to think-aloud while completing tasks or the facilitator may ask the user questions (Nielsen, 2012a). This adds a cognitive load for the user. An alternative approach, retrospective review, involves the facilitator playing back a recording of the tasks for the user to comment on (Rubin and Chisnell, 2008). This retrospective technique is not ideal as it takes longer to facilitate. Moreover, with the benefit of hindsight participants may justify and give alternative explanations for their behaviour (Rubin and Chisnell, 2008). Thirdly, usability testing can take place in (a) purposefully designed labs, (b) in natural settings, or (c) remotely online. The appropriate location for a

usability test will depend on the study design and the needs of the recruited participants (Cornet et al., 2020; Petrie et al., 2006; Rubin and Chisnell, 2008).

2.3 Usability guidance

How to conduct usability tests is well-established with generic guidelines readily available for set-up and planning, recruiting participants, conducting the tests, briefing observers, analysing data and reporting results (e.g. Nielsen, 2012b; Rubin and Chisnell, 2008). In addition, there is guidance on conducting inclusive usability studies on topics similar to those in the more generic guidelines (such as set-up and planning), this time tailored to a specific group of test participants (Table 3 in the appendix provides a summary). Overall, there are many thoughtful suggestions on how to accommodate the needs of the participant group under consideration, as well as specific characteristics of individuals (Darin et al., 2022). However, there are four notable gaps.

The first gap is the lack of diversity in user groups. Guidelines are either specific to a particular disability or age (e.g. Caliz et al., 2017; Hanna et al., 1997) combination of a particular disability and a particular age (e.g. Korte et al., 2015), or to disabilities in general (e.g. Henry, 2007). We were unable to find guidelines for how to conduct usability studies with people from other marginalised groups. This matters because some concerns could be missed. For example, while a disabled person may be concerned about access (e.g. *Can I get to the study?*), someone who identifies as LGBTQ+ may be concerned about representation (e.g. *Does the study speak to me?*).

The second gap is in the representation of participants. Whilst it was recognised that disabilities are difficult to categorise (e.g. Henry, 2007) and that the focus should be on fulfilling the usability objectives rather than on recruiting a specific disability (Van Der Geest, 2006) the limitations of only considering one aspect of a person's identity (such as only gender or only disability) were not discussed. A usability test might recruit a small group of people with a particular need or characteristic to fulfil the requirements of the usability test. However, those participants could equally have diverse and intersectional identities which are not highlighted during the recruitment process. Therefore, embedding broader EDI practices into usability testing is crucial. Furthermore, a more holistic guide could also be useful in accommodating other marginalised groups in usability tests where guidance is lacking (the first gap). This seems feasible as although the guidance is written for participants with seemingly different needs, there is also a lot of common ground such as the need for collaborations and partnerships (Marsh, 2019; Pernice and Nielsen, 2012; Razak et al., 2010), and the importance of participatory approaches (Cornet et al., 2020; Hutter and Lawrence, 2018). Furthermore, much of the guidance, such as establishing relationships and helping participants feel comfortable, is good practice and relevant for the wider population.

The third gap is in hearing from people who have never taken part in usability studies. The guidance is mostly derived from the experience of researchers based on the studies they have undertaken. It is also important to gain first-hand insights from those who do not take part in usability tests.

The fourth gap is the lack of comprehensive guidelines to support researchers with the challenges they face. Rather, researchers give guidance on how to improve the experience for participants and rarely explore how to improve the situation for themselves.

Through this study, we hope to close the first three gaps by investigating how people from diverse backgrounds would like researchers to include them in usability studies (RQ1). We intentionally do not focus on any specific characteristic of a person. Instead we consider identities to be complex (Sum et al., 2022) and aim to produce recommendations that could be considered general good practice for any participant and their intersectional identity. On the basis of the fourth gap, we investigate the challenges for researchers when striving to embed EDI in usability testing (RQ2).

Examining, first separately and then in unison, the perspectives of usability participants from diverse backgrounds and of researchers on usability testing and EDI is a purposeful choice. First, it allows us to give voice to underrepresented research participants, and highlight the importance and value of having these voices heard (RQ1). Second, it allows us to directly influence and inform usability testing in practice while considering the challenges researchers are often faced with (RQ2).

3. Methodology

3.1 Research design overview

Project partners from two charities, LGBT Foundation and Disability Rights UK, and three organisations that do user research, Ideasmiths, Paper and Sheffield Digital, supported this study. Through these partnerships we were able to gain preliminary insights relevant to our research questions and access to potential study participants.

Across a three phase study (interviews, roundtable discussions and requests for feedback), we spoke with usability researchers, and people from different backgrounds focusing on people whose identities are underrepresented in usability testing (figure 1).

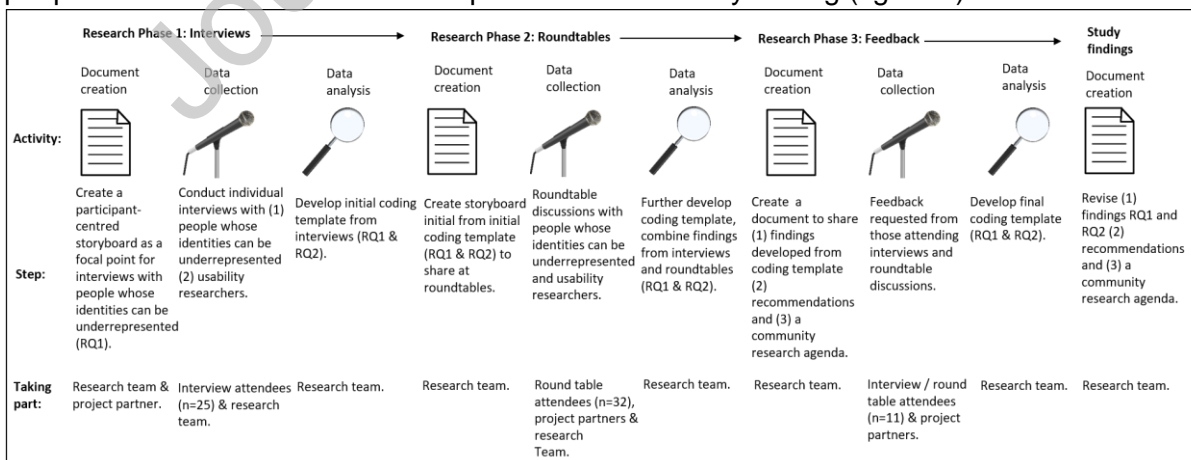


Figure 1: Overview of the research design.

In each of the three phases, we used an object-centred interview technique whereby an object is shared with participants to prompt discussion (Opiniano, 2021). In this study, the objects were storyboards and coding templates, offered in a choice of formats to accommodate participants' needs. The use of objects was important, as our participants had not previously taken part in usability studies. The storyboards (figure 4) and coding templates (figures 2, 3, 5 & 6) provided an accessible focal point for discussions.

The data were analysed using template analysis, a type of thematic analysis (Brooks and King, 2016). Central to template analysis is the development of a coding template (see Figures 2, 3, 5 and 6 for the chain of evidence). With template analysis there is flexibility in how codes are developed: researchers may use *a priori* themes as well as emergent themes (Brooks and King, 2016). This flexibility was important in this project where the data was collected over three phases, and the codes were trialled and developed with study participants.

Throughout the project we were mindful of the need to be flexible and to offer different routes to participation. The interviews were conducted per participant preference: synchronously using either Zoom and Google Meet, or asynchronously using Google Forms. All three roundtables took place in Zoom, alternative platforms were offered but not taken up. Two of our participants reported that they would have preferred to meet in-person but this was not possible as the study took place during a period of Covid uncertainty. Regardless of Covid, it is likely that at least some of our participants would prefer to meet online so that they did not need to travel, and this would account for those who might be providing care, or experience mobility issues and other obstacles, but still wishing for their needs to be considered and their voices heard.

For the comfort of our participants, we did not record the interviews and roundtables. Instead, two members of the research team attended each interview and took notes. Similarly, three or four members of the research team attended each roundtable. Our participants reported appreciating the relaxed atmosphere this engendered knowing that what they said could, if requested, be kept in confidence. To ensure the confidentiality of our participants all quotes are anonymised.

To reduce repetition and make it easier to follow, in the Results section, we aggregate the findings of the three phases under each of the two research questions. So that the reader can match the participant quotes to the research phase, we have assigned each participant a three-part code: (1) a unique number; (2) whether they are commenting as someone whose identity is routinely underrepresented in usability testing (UU) or as a usability researcher (UR) or as a project partner (PP); and (3) when they made their comment, at an interview (I), roundtable (RT), or in feedback (F). For example, P5-UU-I is participant number 5, whose identity is underrepresented in usability testing and made this comment at an individual interview. Additionally, at the roundtables some participants contributed anonymously using Google Docs.

3.2 Recruitment strategy

To identify how researchers can include people from diverse backgrounds in usability studies (RQ1) we focused on recruiting participants whose identities can be underrepresented, and to identify the challenges of embedding EDI in usability testing (RQ2) we recruited usability researchers. Where participants attended interviews and / or roundtables in their own time (i.e. not representing an organisation), they were compensated with a £25 high street shopping voucher of their choice.

Next, we describe our recruitment strategy for the two groups of participants. Under Phases of Research, we document how many people we recruited for each phase.

3.2.1 People whose identities can be underrepresented in usability studies

People from diverse backgrounds whose identities can be underrepresented in usability testing could include many people from different backgrounds, for instance, people who are digitally excluded, people from different socio-economic statuses, disabled people, people for whom English is a second language, people from different ethnic minorities, people who identify as LGBTQ+ and people with different levels of education. In our study, there were no inclusion / exclusion criteria as we were committed to diverse recruitment. However, as our project partners, Disability Rights UK and LGBT Foundation, helped us with recruitment this meant that many of our participants identified as disabled and/or LGBTQ+. Nevertheless with the exception of those who are digitally excluded, our participants had varied backgrounds (including from different socio-economic statuses, from different ethnic minorities, and those who speak English as a second language), very often characterised by intersectionalities (e.g., an older lesbian with mobility impairment). Furthermore, by working with these two groups, key EDI concerns such as access (e.g. *Can I get to the study?*) and representation (e.g. *Does the study speak to me?*) would likely be covered.

Participants were recruited with a combination of convenience and snowball sampling (Etikan, 2016) using the research team's social media networks and our project partners' networks and charity newsletters. We did not explicitly ask participants their demographic details (such as gender or age) as asking for this information out of context would not help us to understand which, if any, part of a participant's identity would influence their experiences of taking part in usability tests. Instead, we asked, "Are there issues that you might face that are particular to why you signed up for this study?" From the responses we received, it became clear that our participants' identities are indeed complex and not reducible to single categories (for example, P8-UU-I described themselves as autistic, dyslexic, hard of hearing and ambidextrous) thereby confirming the appropriateness of our intersectional approach.

At the interviews, we asked our participants if they had taken part in a usability test – none had. Many of our participants had, however, taken part in interviews and focus groups in other research projects. This meant that they were broadly aware of the processes for taking part in research but they were not familiar with the specific processes for usability testing.

A note on terminology. There is no universally accepted approach when writing about different identities. Therefore, to avoid unintended offence, we have chosen to emulate the person-centred approach advocated by our research partner, Disability Rights UK (n.d.), and the UK government (Disability Unit, 2021). As a result, we use the terms *disabled people* and *people identifying as LGBTQ+*.

3.2.2 Usability researchers

Participants were recruited using the research team's social media networks and our project partners' networks. A convenience sampling approach (Etikan, 2016) was necessary as we needed to access people who had experience of running usability tests. This was the only criteria for participation. It should be noted that some of our usability researchers are also disabled and/or identify as LGBTQ+, and this may have influenced their decision to participate in this study.

3.3 Phases of research

In total, we completed three phases of data collection.

- In phase 1, we interviewed people whose identities can be underrepresented in usability testing, and usability researchers.
- In phase 2, we presented and discussed the phase 1 findings at three roundtable meetings attended by people whose identities can be underrepresented in usability testing and usability researchers. These discussions formed part of our data for further analysis.
- In phase 3, we requested feedback on the findings including provisional recommendations and an agenda, from all those who participated in the interviews and roundtables.

We next describe these phases in more detail, including how many participants were recruited for each phase (summarised in Table 1).

	Individual interviews (phase 1)	Roundtables (phase 2)			Feedback (phase 3)
	Mar-Apr 2022	29 April 2022	6 May 2022	9 May 2022	Nov 2022
People whose identities can be underrepresented in usability tests	17	7	5	6	9
Usability researchers	8	3	4	7	2
Project partners	n/a	3	2	1	2

Table 1: Number of participants in each research phase

3.3.1 Phase 1: Interview study

In phase 1, we interviewed seventeen people whose identities can be underrepresented in usability testing to gain an initial understanding of how research participants would like to be included in usability studies (RQ1). We also interviewed 8 researchers to gain an initial understanding of the challenges for researchers (RQ2).

In preparation for participants who had no experience of taking part in usability tests, we worked with Paper (project partner) to create a participant-centred storyboard to illustrate the experience of undertaking a usability test (see Appendix Figure 4). This included what we considered the key decisions a participant would need to take during a usability test. The storyboard was shared with participants during their interviews and we asked them, "What are your preferred approaches to participating in research?" As the participants we recruited had experienced generic research but not usability testing, the storyboard was a useful interview object.

To understand the challenges for researchers we did not develop a storyboard as those interviewed were already familiar with usability testing. We asked them to describe the steps in the usability testing process and "identify which, if any, inhibit you from recruiting people from diverse backgrounds".

At the end of phase 1, we produced a coding template for each of the two research questions (see Appendix Figures 5 and 6). For RQ1, the third author revised the initial storyboard template based on a preliminary analysis of five interviews to more clearly foreground usability. With template analysis, the initial templates are built on a subset of the data (Brooks and King, 2016) and this allowed us to produce templates in the short time between phase 1 interviews and phase 2 roundtables. The template was later further refined by all authors with an analysis of all the interviews. For RQ2, the first author developed an initial coding template from the interviews, which was again later reviewed and refined by all authors.

3.3.2 Phase 2: Roundtable discussions

In phase 2, we arranged three roundtable discussions that brought together people whose identities can be underrepresented with usability researchers, and our project partners (see Table 1). Sixteen of those interviewed in phase 1 also agreed to take part in the roundtables. In addition, we recruited a further sixteen participants.

We used the phase 1 coding templates as the focus of the discussions. Based on feedback we received in phase 1, we shared the preliminary findings in advance as some participants said they like to process information in their own time. At the roundtables, we gave a further overview of the findings before breaking into smaller groups of three to seven participants to continue the discussion. Each roundtable group was mixed to ensure they included individuals from each type of participant group: individuals whose identities can be underrepresented in usability testing, usability researchers, and project partner representatives. These discussions were then fed back to the whole group.

At the end of phase 2, the first author analysed the roundtable discussions and combined this with the analysis of the individual interviews (phase 1) to produce two refined coding templates (figure 2 and figure 3 in Results). These templates were then reviewed by the other three authors, and a set of recommendations were developed from each of the two coding templates, as well as an overarching research agenda.

3.3.3 Phase 3: Feedback

In phase 3, we arranged for feedback with our study participants. Not only is this an important part of a co-production approach, our findings show that participants want to know what happens to research after their participation. We sent out the coding templates, a write up of the findings, the recommendations and the research agenda to all those who participated in the interviews and roundtables to garner feedback (table 1). No new people were recruited for this phase. Eleven participants and two project partners provided feedback. Three email addresses were no longer valid and one participant asked not to receive any further emails.

All other feedback was very positive, particularly about the clarity with which findings were communicated and the importance of the recommendations.

“This captures the discussions I was involved in and the key elements from my own perspective as a participant in the study. I also like the diagram illustrating this. Breakdowns of each area are clear and concise and illustrate key points well. The recommendations are concise and feel achievable/tangible. I would hope they are more likely to be adopted given this. It’s really positive to see the proposal for the development of a community and actions. This formed a valuable part of the focus group [roundtable] of discussions with diverse representations and [I] can see the value this was [and would] add as an ongoing community. (P15-UU-F)

“I have read over this and I feel like it is an accurate representation of the focus groups [roundtables] and interviews. The recommendations are also very good.” (P4-UU-F)

“The triumvirate of value, trust and agency cannot be understated. I think this is the essence of meaningful co-production and active participation in research, and in many other areas of engagement.” (P22-UR-F)

Two participants suggested that there was not enough emphasis on the importance of language (we subsequently strengthened this in the results), one participant reminded us that it can be difficult to know one’s value (feedback added in the results) and one participant would like to see EDI more generally adopted in organisational procedures. All other comments were related to phrasing to make the points clearer and stronger. No changes were made to the coding templates.

3.4 Ethics

Informed consent was obtained from all study participants and for confidentiality their data is anonymised. This research received approval (reference number 044992) from the University of Sheffield on 1 March 2022.

4. Results

4.1 Research question 1: How should researchers include people from diverse backgrounds in usability studies?

Based on our analysis, usability participants from diverse backgrounds would like a sense of value, trust and agency. Namely, participants wish (1) to take part in research that is of value to them and that they are valued for doing; (2) to be able to trust the study and the research team; and (3) to be in control and make their own decisions while participating in such studies. This is illustrated in Figure 2 and discussed in more detail below.

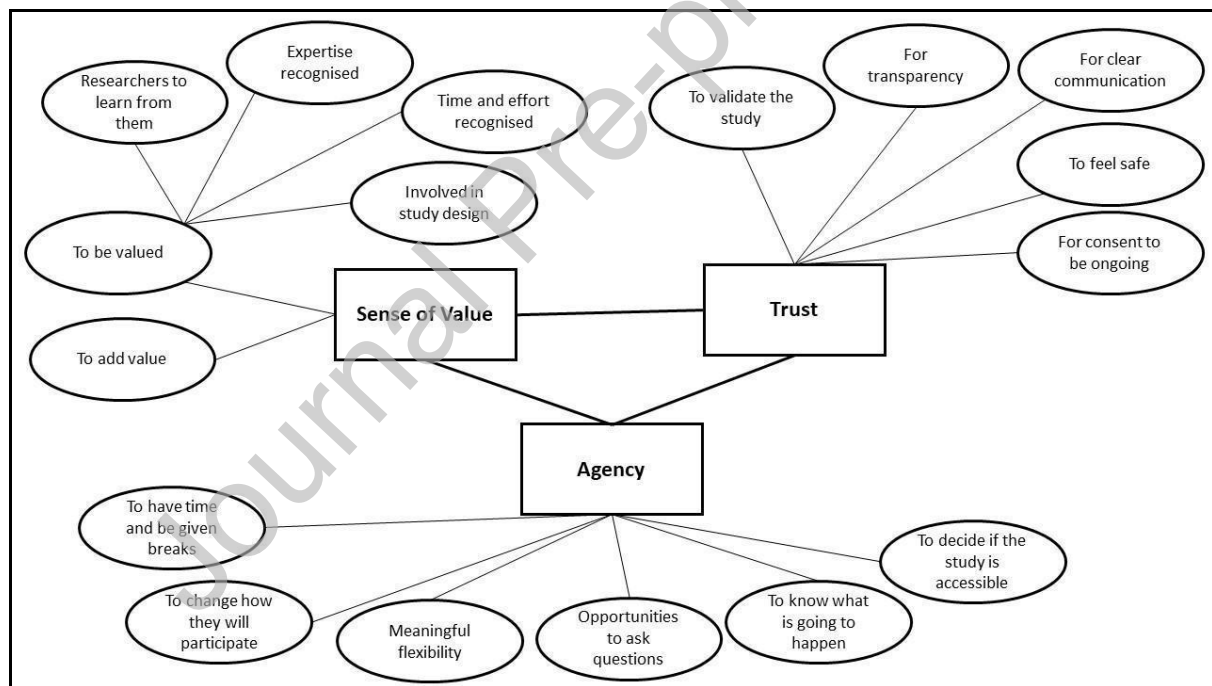


Figure 2: How researchers should include people from diverse backgrounds in usability studies. Rectangles indicate primary themes and ovals the sub-themes.

4.1.1 Sense of Value: Usability participants from diverse backgrounds want to add value to a study and for their contributions and time valued by researchers.

First, several participants indicated that when it comes to them participating in usability testing, what they consider important is that their contributions are acknowledged and valued. This might take different forms, for example being financially reimbursed for their

time and labour, whereby participants would like researchers to recognise their expertise because they are “expert[s] by experience” (P12-UU-I). Although participants in our study described taking part for altruistic reasons, many felt they “should not give their time for nothing” (P10-UU-I) and that they should be compensated for their time and expenses:

“It is not make or break for me. Useful for telling me how I will feel with people - knowing that someone is valuing my time from the off. More likely they will listen too and the environment is comfortable.” (P17-UU-I)

“You definitely need payment for travelling, some will travel for a good lunch. I don’t care about lunch.” (P11-UU-I)

However, it is not strictly about the financial reimbursement but rather about acknowledging and valuing the full effort and time it takes to participate in research. Research participants have wider lives, and, for them, participating in research is more than just the time taken during the usability test. Instead, this is likely to affect what a person can or cannot do that day and even beyond the specific day due to the “overall expenditure of energy, either cognitively or physically” (P15-UU-I):

“A lot of people assume only one hour out of your day but the time it takes to plan can be really time consuming.” (P17-UU-I)

“I have to think from beginning to end. I can end up with stress and anxiety. If you cancel at the last minute, I’ve not got anything else to do.” (P10-UU-I)

Secondly, participants highlighted the significance of researchers acknowledging explicitly the importance of having lived experience in relation to the investigated phenomenon. With regards to usability research in particular, because accessibility and usability are important facets of the everyday life of our participants, they expressed being keen in taking part in usability research, because they consider that doing so can add value to such studies and therefore make a difference to them and others:

“I think of the purpose of the study and have I got anything to offer it.” (P12-UU-I)

“Making sure that people like me and my disability need to be recognised and therefore this hopefully will influence XYZ in the future.” (P11-UU-I)

In other cases, however, it might be difficult to recognise if and what value one’s lived experience might add to a particular study. Especially those unfamiliar with research processes would need to understand this aspect better, highlighting the significance of the research team communicating clearly whether and how lived experience can enrich a study:

“It is difficult sometimes for people to recognise the full/specific value and expertise that people bring if it [taking part in research] is not an experience they are familiar with.” (P15-UU-F)

Third and most importantly, participants highlighted the importance of co-designing and co-producing research, which goes beyond contributing solely to the usability test. Many noted that being involved in setting the focus of the study and drafting and shaping usability tasks

could help researchers learn from them, their lived experience and their experience of participating in the study, which further contributes towards a sense of adding value and being valued:

“Participants should devise the questions ... instead of a standard set of questions.” (P12-UU-I)

“Researchers should include a short survey after the study: *genuinely, how did you find it?*” (P19-UR-RT3)

“We give our time, if we don’t hear anything back, you realise that giving your time is a waste of time in the end.” (PT-UU-RT2)

4.1.2 Trust: Usability participants from diverse backgrounds need to be able to trust the study and the research team. Trust needs to be established from the start of the study and maintained throughout.

Taking part in research oftentimes suggests that research participants will need to engage with the researchers, with other participants and interact with technologies and systems, either in-person or through online and hybrid means. In addition, it entails them sharing their experiences. Within this context, our participants indicated that the above necessitates a certain level of trust in terms of the study’s processes and outputs, spanning aspects of recruitment, informed consent, as well as the study’s motivation.

Recruitment techniques and decisions are critical for successful usability research, especially when it comes to minoritised individuals and hard to reach voices. Our findings suggest that being transparent about why someone is being recruited is highly valued by potential participants and it helps them to feel at ease. Several of our participants indicated that it is important to know why they were or were not selected for a study.

“I need an explanation on why I am chosen.” (P5-UU-I)

“The researcher emailed me [for another type of study] - really pleased you are interested, would love to work with you but here is why we can’t. Clearly explained. Seems genuine. It means you don’t take it personally.” (P17-UU-I)

Following the first contact with a prospective participant, researchers typically need to secure informed consent, a key principle in research, which ensures that a participant has sufficient information about a study and a good understanding around the risks and implications of making a decision of participating in it. All too often, however, participants may consent to taking part in a study at first, but while actually participating in it, they may still not know much about it and they need to be informed they can withdraw at any point, and exactly how they can do this:

“Going over the information sheet and consent form again. Give me the opportunity to say “yes, I’ve changed my mind”.” (P4-UU-I)

Even so, knowing why they have been chosen and granting consent is not sufficient. Before taking part in research, participants explained they wish to be aware of the purpose of the study and the research team, so as to be able to validate these for themselves. Those with research participation experience noted that in the past they either used the information provided to them or they relied on the (trusted) person who recommended the study to them:

“Overall ethos of why the study is done.” (P17-UU-I)

“I am doing this [study] because I know [name of person] because she mentioned this to me and I trust her.” (P8-UU-I).

In relation to this, participants noted that having clear and sufficient information is key to establishing trust. While participants may have lived experience on the studied phenomenon, typically they are less familiar with research processes; therefore, such information needs to be easy to understand, accessible, addresses language barriers and helps participants feel at ease:

“Key points aren’t obscured in fancy language. Helping people feel comfortable as though they are on a level playing field.” (P17-UU-I)

“Have everything that is said in a way that I understand it and it is accessible to me. And for this to be checked.” (P12-UU-I).

“Give key information first, then links to more information.” (P5-UU-I)

This includes the way relevant documentation is structured (for example ensuring that important information is easily accessed without participants feeling overloaded), and offered in alternative formats (for example during usability tests offering both written and oral instructions):

“I am conscious that one of the things with my long term health condition ... I am not sure how well I’d retain instructions.” (P15-UU-I)

Like all research studies, usability testing does not always go as planned and procedures may change mid way. Participants ideally would prefer that researchers stick to the agreed procedure, or at the very least to consider their needs while conducting the research:

“If you want a break. We’ve run over. They say “we’ll keep on going”.” (P16-UU-I)

In relation to the actual conduct of the research, another theme that emerged quite strongly from our findings highlights the importance of physical and psychological safety in relation to establishing trust. Participants require that they will feel safe when they participate in usability testing and expect that researchers will have considered the test environment and that they will provide them with information on what support is available:

“The environment matters: the room, its setting, the layout. If it’s friendly. If there is water, an open window, and so on.” (P4-UU-I)

“What’s going to be talked about? Is it triggering? Do they offer you any support with that?” (P16-UU-I)

In other words, participants expect and need researchers to protect participants’ safety and sadly several participants who had previously taken part in other types of research studies reported shocking experiences:

“Someone who is a moderator to remove an aggressor.” (P1-UU-I)

4.1.3 Agency: Usability participants want to be in control and have the agency to make their own decisions.

The third major theme that emerged from our study was that of agency: participants’ desire and need to make their own choices, and to express themselves freely during the research. Such choices can include decisions regarding the logistics of participating in the research, e.g., arriving, navigating oneself and leaving the venue. Participants wish to have enough information to inform their decision making and also have control over relevant decisions:

“Don’t arrange taxi times for me. I might want to go for a walk after completing the usability test.” (P10-UU-I)

“From the start I’d want to know. I find it hard going into something without knowing what is going to happen. How will I get to the place? Is it difficult for me to get to? Will it cost me anything?” (P16-UU-I)

Agency also relates to meaningful flexibility and the ways in which someone is able to participate in usability studies. A lack of flexibility can rob participants of their agency and can have a negative impact in terms of their energy for engaging in usability tasks. In contrast, having options whereby the “test [can be] tailored to the participant” (P5-UU-I) means participants feel less pressured and more involved in the usability test. Some choices that emerged from our findings include:

- whether participants can participate in-person or remotely;
- when to participate;
- which technology to use (Zoom, Google Meet, and also screen readers and assistive technology);
- how to communicate and receive information (website, emails, telephone);
- what data is collected (video, audio, other);
- how data is collected during the tests (think-aloud, retrospective review);
- how personal data is used, stored and shared;
- compensation formats (vouchers, cash, salary).

Agency also extends to participants being able to revisit their choices in how they might want to participate, even at short notice, as their needs can change depending on the day:

“On a day-to-day basis it [my needs] can change. It is not a blanket, *I need this*.” (P1-UU-I)

Participants may need time to “figure things out” (P6-UU-I), to fill in forms and other documents, and take breaks if and when needed, without feeling that they are delaying the researcher:

“I have absences and brief lapses of concentration. It is much worse if attention is shifted. Need to focus on one thing.” (P7-UU-I)

“If I am taking a lot of breaks, will I be kicked out for the next person when the time is up? It is just to have the awareness that things might take longer.” (P5-UU-I)

Having enough information to support their decision making is a core part of being able to exercise one’s agency, especially in terms of accessibility because different individuals will have different requirements depending on context, circumstances and other parameters. Having enough information can significantly help making an informed decision as to whether a study is accessible or not:

“Seemingly simple details can be really key for me and for others it can be crucial.” (P15-UU-I)

“The perception is always about the physical. But mental illness, the invisible ones, are not being understood.” (P5-UU-I)

Indeed, as one participant said, the more information that is provided up-front “the more agency you have” (P17-UU-I). Equally, however, having up-front information helps participants to feel reassured about what has already been “taken into consideration and that, therefore other aspects might be” (P15-UU-I), and provides encouragement especially to those who do not typically feel able to express their needs and desires, or those who may not know what they might need:

“[It is good for] people with anxiety, to be less anxious, not facing the unknown unknowns. It reassures confidence.” (P4-UU-I)

“Some people aren’t confident to say what they need because of past experiences of being knocked back or wondering, am I being reasonable.” (P3-UU-I)

“Giving a list of examples gives me permission to ask and an idea of the scale of adjustments I could ask about; giving a tick list is not so helpful.” (anonymous/RT3)

During our conversations with participants, several ideas emerged in terms of how such information can be provided and what it should cover. For example, “step by step walkthroughs” (P8-UU-I) help participants to know what to expect in a usability study and receiving information in advance lets participants “process it in [their] own time” (P15-UU-I). This is perhaps particularly important for usability studies where users are observed and engage in tasks that, by the nature of the test, may not be accomplished.

When the information provided is comprehensive, participants may not need to ask many questions as what they need to know has already been covered:

“I do like them to anticipate questions people have. Fill in the gaps for me.” (P16-UU-I).

That said, it is important for participants to “feel comfortable asking the questions” (P17-UU-I), starting from the moment they view the recruitment advert and continuing throughout the study: “It’s good to be able to ask questions, but it is good to do this also as you have started doing the tasks” (P11-UU-I), especially because this will allow for the results of the test to be more useful and easier to interpret.

4.2 Research Question 2: What are the challenges for researchers when embedding equality, diversity and inclusion in usability testing?

Organisational pressures, stakeholder culture, getting guidance and recruiting who you need are challenges for usability researchers. This is illustrated in Figure 3 and discussed in more detail below.

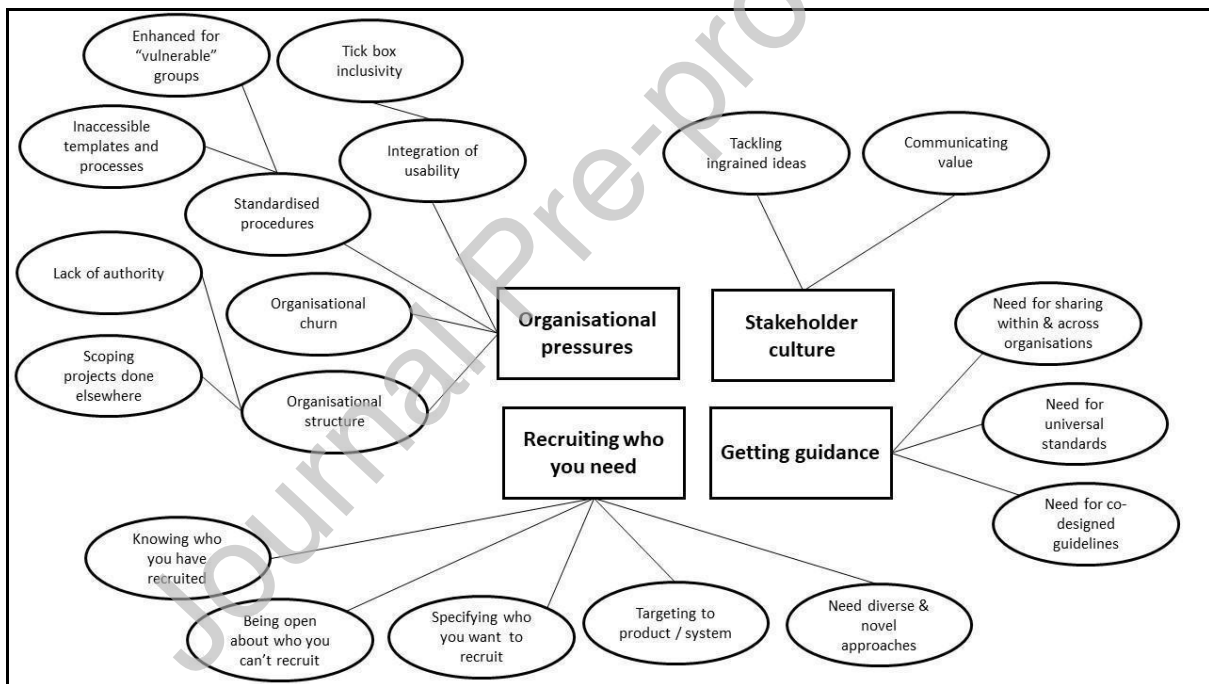


Figure 3: The challenges for researchers when embedding equality, diversity and inclusion in usability testing. Rectangles indicate primary themes and ovals the sub-themes.

4.2.1 Organisational pressures

Researchers are often working within organisations in which they are part of a larger hierarchical structure. Additionally, research teams are often not involved in scoping projects. Researchers find they are not in a strong position to advocate for (1) money to compensate participants and organisations such as charities that support recruitment, and (2) time to broaden recruitment and take a more flexible approach to the research design. Furthermore, when issues, particularly those relating to EDI, are found during the usability

test, researchers do not always have the authority to bring about design changes, and so it can be frustrating as they “bridge the information that the users give and the actions the organisation can take” (P37-UR-I).

“Our capacity in terms of staff, time and budget doesn't currently allow for the widest range of diversity engagement within user testing ... We select volunteers based on chosen criteria and their availability. We have no capacity to do more than five [participants].” (P25-UR-I)

Researchers are affected by changes in organisational structure, leadership, and direction, as well as changes to the broader environment. Moreover, many researchers are on short, fixed-term contracts and / or move between projects. This means that knowledge is lost and it is difficult to cultivate long term relationships. Furthermore, some organisations do not have the capacity to undertake all their testing in-house and employ external consultants.

“Constant changes to the health landscape can mean constant changes to the research.” (P21-UR-I)

“If my fixed term contract was renewed and I was assigned to another digital project which required user testing, I could contact the [names LGBT+ groups] to help us.” (P25-UR-RT1)

Legislation can mean that researchers need to follow rigid and standardised procedures such as gaining ethical approval, complying with GDPR and research governance. These procedures are designed to safeguard participants. However, as they often need to be signed off early in the research design, it can be difficult and time consuming to adapt research in response to participant requests. For participants defined as vulnerable, these procedures may be enhanced. Additionally, there may be organisational requirements to use templates and follow processes (such as compensation) that may not be accessible to everyone.

“With its rapid expansion [usability testing] lots of people aren't totally ignorant but they underestimate this process ... and are falling behind in understanding these wider issues such as GDPR, informed consent and so on.” (P33-PP-RT1)

Usability testing and user research are often integrated into the project at the end. Researchers are concerned that EDI should not be an add-on to tick a box. All our researchers thought it important that inclusivity is fully built-in and would like to include users from various backgrounds in all aspects of the research, including the setting of the research questions, the design of the usability tasks and what is tested.

“[Usability testing and EDI approaches should be] baked in from the start. Not tacked on at the end - we're giving 1% of the project's budget / time / staff to do all this stuff. It needs to be there from the beginning to be iterative.” (P25-UR-RT1)

“It becomes a bit easy for organisations, within structures that they've got, to carry on saying 'we've done some research on participants therefore it is participant-centred'.

And actually a bigger challenge would be to say, can you make things participant-led?" (P37-UR-RT2)

"Often people have an idea in mind of what should be developed before involving the user groups." (P24-UR-RT1)

4.2.2 Stakeholder culture

In some organisations there is a culture of embedded EDI, in other organisations staff have "ingrained ideas" and do not recognise the importance and value of EDI (P37-RI-I). Researchers stressed the advantages of working in a place which is diverse and inclusive. Some researchers are already working in such environments while others are not as fortunate.

"Having a diverse and inclusive team of colleagues involved in all we do is essential as is the company leadership and culture from the top down." (P36-UR-I)

"[Org] is not a diverse organisation. No one with a disability in management." (P37-UR-I)

Stakeholders tend to focus on what has been legislated and are therefore more concerned with accessibility rather than usability. Researchers stressed the importance of gaining client and senior colleague buy-in to ensure there is enough budget, time and staff. Several researchers reported the difficulties of trying to communicate value and would like evidence of (1) how usability improves products, systems and services; (2) how EDI can save and make money; and (3) how EDI helps organisations fulfil their policies and legal obligations. Researchers are finding that some progress is being made but this varies according to who is being recruited as some participant groups get more attention than others.

"Communicating the importance of usability testing to people who hold the budgets is really difficult ... we need to raise the importance and value of user-research and user-experience in large organisations." (P25-UR-I)

"[Organisation] is just starting to get to grips with physical disability and neurodivergence in their usability testing - everything else is a plus or a bonus." P25-UR-I)

4.2.3 Getting guidance

Researchers would like more guidance on how to recruit and conduct usability tests with a broader range of participants with different needs. Some researchers reported that their organisations have developed best practice guidelines but these are specific to a single organisation or sector, whereas researchers would like principles and guidance that can be used universally (i.e. in different organisations / sectors and for participants from a range of backgrounds). Universal guidance could help alleviate researcher concerns that they are taking an ethical approach, safeguarding their participants and following best practice. It was also thought important to recognise that researchers will make mistakes; these mistakes should be shared so others can learn from these and avoid repeating them in the future.

“We update that training as often as we can. Checking for gendered vocabulary, simpler verbs, ableism, etc. ... Alternative formats policy - leaflets, PDF. And for those who require different formats - digital or personal - to help them access information. Visual impairment, hearing loss, cognitive / learning disabilities. E.g. large print, braille, audio described, captioned, easy read (lots of pictures and simple sentences). Image data bank to use inclusive images of people who do have disabilities.” (P22-UR-I)

“Expanding it to other disciplines means working with other people with different knowledge and expertise. Involving user-centred or patient-centred work - has real potential to give you a lot of really good ideas.” (P17-UU-RT3)

Researchers would like greater engagement with users, and valued the roundtables for bringing together researchers and people from different backgrounds. It is important that any guidance should be co-designed with users.

“From a researcher's perspective, I'd like to see more events like this [the roundtable] to connect people who work in the area with people with all that lived experience ... the ability to speak to people who have worked through these problems already and potentially solved them and share.” (P21-UR-RT1)

“Always co-design public-facing information. It's not just about terminology; it's about making sense of the world differently and wanting different information from what researchers think people need”. (Anonymous-RT2)

4.2.4 Recruiting who you need

A key discussion point at all three roundtables was how to reach potential participants. Researchers described using one or more of the following methods.

- Recruiting participants within their or their client's organisation user base: it is often easier to recruit some people than others; for example, it is easier to recruit people where there is an established long term relationship.
- Using charities, voluntary and faith groups to tap into their networks: this was considered “the most effective way to connect with harder to reach groups” (P21-UR-RT1) but does require a budget as these organisations need to be recompensed for their time.
- Using recruitment agencies: this requires a budget but an advantage is that the recruitment partner likely understands the need for EDI recruitment and actively signs up a broad and inclusive range of users to their database (P36-UR-I).
- Using social media and reaching out to people's networks: this can take time and it helps if there are already established relationships.
- Going into spaces already used by potential participants, such as cafés.
- Finally, if researchers felt they did not have the time to recruit participants, they may not test with users and rely on knowledge of best practice instead.

Although researchers described using a variety of techniques to recruit participants, they felt that new approaches are needed for recruiting some groups such as those who are digitally excluded and those who do not usually volunteer to take part in research.

“The challenge we do face is doing testing on digital solutions with people who are digitally excluded - their voices are important to include in the design process, and possible with interviews, but trickier to show a digital solution, particularly during the pandemic.” (P44-UR-I)

“People who put themselves forward for research are still a certain kind of person. So you end up speaking to a lot of people who are really quite articulate, confident, they have the time to speak to you, they have the information to do so. And there’s a whole group of people who are really disengaged and really hard to reach, literally, who never put themselves onto a database. Would never volunteer for these kinds of things. And they might have some of the greatest needs and the greatest challenges.” (P42-UR-RT3)

Researchers are actively seeking solutions to what they see as recruitment challenges. One researcher reported their organisation is creating a full time position to support diverse recruitment. A second researcher felt that there is often a misperception that meeting the needs of a diverse range of people, particularly disabled people, is “difficult, expensive, intensive” (P15-UU-RT3) when they argue people are generally quite resourceful and tend to know their own needs really well. A third researcher reported that while it does take time, they are still able to take a flexible approach to participation.

“We do prepare users carefully with information, calls, testing of any technology that might be used in the usability session. This will be tailored to the user groups, different prep and help is needed for different categories. This is not a blocker but needs time.” (P44-UR-I)

Researchers also spoke about the importance of recruiting who you need rather than recruiting for diversity. It was thought that recruitment should be targeted to the technological product or system, and should not become a quota filling exercise. This means that sometimes researchers need to recruit a broad range of users with different characteristics, at other times they need to recruit specific niche groups.

“What we sometimes get wrong is that we don’t look at who actually needs to be involved to do, in this case, usability testing. There’s not much point in involving people from communities who are never going to use something just for the sake of it. If it’s not relevant to them.” (P19-UR-RT3)

“Depending on the nature of the work, our clients sometimes want us to recruit very specific groups of users, for example those with visual impairments who use assistive settings or hardware to test a website, or those from the LGBTQ+ community to advise us on inclusive wording and options around sex and gender on a medical app.” (P36-UR-I)

Researchers are also concerned about how they specify who they would like to recruit and how to capture participant identities that are meaningful to the usability test. Furthermore, as most disabilities are invisible and there is “a lack of definition or centralised database to describe disability - it’s so different for everyone” (P28-UR-RT2). Several researchers thought it important not to prescribe labels so participants can describe themselves how they choose, or alternatively describe the impact of their disability.

“Remove all these tick boxes because they don’t reflect people. It can be hard to describe, even for the potential participants themselves.” (P26-UU-RT2)

“It’s often more useful for us to better understand the impact that your condition/disability has on your day to day life, or your use of technology, or mobility etc rather than any label/checkbox.” (P35-UR-RT3)

On the other hand, researchers can find it useful to be highly specific about who they are looking to recruit. Clear communication about why particular people are being recruited, and why such personal information is collected is imperative as “it can seem inappropriate” (P33-PP-RT1). Researchers are also aware that it is not always possible or even necessary to identify what in the participant’s background and experience is informing the usability test. Furthermore, participant identities are not always captured because of the complexities of storing this data.

“Think intersectionality, not individual protected characteristics which often end up being a tick box exercise.” (Anonymous-RT2)

“That’s one of the issues I’ve found with recording this stuff. A lot of clients don’t want to go there because I point out to them that GDPR issues go up a step because you’re now collecting protected characteristics. And quite often, I’ve worked with some clients who didn’t have the processes in place to adequately record that.” (P33-PP-RT1)

Researchers also need to be open about who they have not been able to recruit.

“Researchers should call out who you’re missing. That’s a big step for some organisations to actually call that out and actually make that public. I think some organisations feel like if they say we don’t have enough of this community coming forward... they don’t have the confidence to even take that step. But I think that’s part of it, to acknowledge who you’re missing from a database, who you’re missing from a sign up. And then this is part of making proactive steps to find those people. ... Or to address the issues of why those people aren’t coming forward.” (P15-UU-RT3)

5. Discussion and recommendations

In this section, we discuss the study findings in relation to the research literature and make recommendations to usability researchers, and the organisations that employ them, as well as providing a research agenda for the community.

5.1 Sense of value, trust and agency are likely universal principles

Although none of our participants had previously taken part in usability tests, they had taken part in other types of research studies. Depressingly, several of our participants described previous negative experiences of participating in research as the motivation for taking part in this research. Furthermore, much of what our participants describe as needing should be standard research practice, for example, ongoing checking of consent. We can only conclude that we need to improve research practices and we need to more clearly communicate our practices to participants.

With regards to how our participants wanted to take part in usability testing, there was little common ground. Each of our participants has different needs and changing needs that affect where they can participate (in-person or remotely), the process of the usability test (think-aloud or retrospective review), forms of compensation (voucher or cash) and so on. However, common to all our participants was the need for a sense of value, trust and agency. We recommend that researchers are guided by three key principles: (1) they should ensure participants feel a sense of value, (2) they need to establish trust, and (3) they need to enable agency so that participants feel in control and are able to make their own decisions. For each of the three principles we make the following recommendations (Table 2). These recommendations can and should be adapted to respond to each study's objectives and the participants included in the usability test. As such, these recommendations are complementary to the guidance offered to particular groups of participants in Table 3 (in the Appendix).

	When		
	Set up & planning	During session	Post-session
SENSE OF VALUE			
Explain how your study adds value and could benefit participants and or the wider community.	✓		
Recognise participants' expertise. Make clear the value the participant adds to the study.	✓	✓	✓
Recognise the full time and effort of participation: offer compensation and pay expenses.	✓		✓
Include participants in the study design, including setting study objectives and usability tasks.	✓		
Learn from your participants about the experience of taking part in your study.			✓
TRUST			

Enable participants to validate both the study and research team by providing comprehensive information upfront.	✓		
Be transparent in everything you do. For example, explain why participants are or are not selected	✓	✓	✓
Use simple and clear language for all communication. Carefully structure documents and other information sources.	✓	✓	✓
Keep participants safe. Inform them about the available support and your safeguarding procedures. Check that participants are comfortable during and after the study. Be ready to intervene if a user feels unsafe.	✓	✓	✓
Keep checking for consent and do not change the goal posts.	✓	✓	✓
AGENCY			
Enable participants to decide if the study is accessible by providing upfront information on practicalities and what can be adjusted. Give users the opportunity to say what they would like.	✓	✓	✓
Use simple and clear language for all communication. Carefully structure documents and other information sources.	✓	✓	✓
Tell participants in advance what is going to happen, so they have time to plan and that there are no surprises.	✓		
Provide opportunities for questions.	✓	✓	✓
Be flexible, offer meaningful choices and do this from the start.	✓	✓	✓
Allow participants to change how they participate throughout the study.	✓	✓	✓
Design studies so that participants can take their time. Factor in breaks	✓	✓	

Table 2: Recommendations for researchers wanting to embed EDI in usability research

We suggest that the three principles of sense of value, trust and agency are likely universal principles and could be considered general good practice when including people whose identities are routinely underrepresented in usability studies as well as for including the wider population. We believe if usability researchers strive to meet these recommendations, their practice will be inclusive for all participants, no matter their identity. For instance, ensuring participants have agency means that you are letting them make decisions and lead the way.

If usability researchers take the time to build trust, a participant can feel comfortable to express their needs so accommodations can be made.

5.2 Adopting the cultural values of learning organisations

Our researcher participants considered learning from each other and learning from each other's mistakes important for embedding EDI in usability research. Such learning can be both hindered and encouraged as a result of the organisational culture (Davies and Nutley, 2000). Based on our findings, it emerges that learning, including from mistakes, is significant for an organisation to solve problems or launch new technologies, otherwise the organisation is likely to repeat old practices that do not lead to improvements (Garvin, 1993). We therefore posit that, to create a sense of value, trust and agency within usability testing, it is important that the organisation adopts a learning orientation. A learning orientation allows organisations to achieve improvements and paradigm shifts (Alerasoul et al., 2022); this should be fostered among usability research organisations, as suggested by the usability researchers participating in our study. Such an orientation entails adopting the cultural values that underpin learning organisations including celebrating success, valuing change and innovation, tolerating mistakes, believing in and trusting people, recognising tacit knowledge, being open and outward looking (Davies and Nutley, 2000) and supporting diversity in knowledge (López-Cabrales et al., 2011). While such learning is led by individuals, learning takes place at different levels (improvements, paradigm shifts, learning to learn) and an organisation can mobilise learning and maximise its potential (Davies and Nutley, 2000). We therefore propose the following recommendations, which have been compiled with our participants.

- Recognise the strategic importance of recruiting users from a range of backgrounds.
- Include usability teams in the scoping and initial development of projects.
- Seek to develop, train and maintain diverse research teams.
- Allow time and flexibility for following organisational and legal procedures.
- Give researchers time and resources to develop novel approaches to reaching and recruiting participants.
- Include compensation for charities, recruitment agencies and partners, and participants in funding bids.
- Ensure an EDI approach to usability and user testing is iterative and embedded into every stage of the research and design process.
- Involve users in all stages of the design and development process.
- Facilitate an open and sharing culture. Encourage sharing of best practice within and across organisations.
- Engender a culture where it is acceptable to admit to making mistakes and not getting it right.

5.3 Moving from underrepresentation to co-production

The importance of co-production and participatory approaches (Cornet et al., 2020; Hutter and Lawrence, 2018) was strongly supported in our research. Those who feel underrepresented in usability studies want to be included in all aspects of usability research including the setting of the research questions, methods and tasks. At the same time,

researchers realise the value this would add and want participants to have that involvement. However, as with other research areas, co-production is constrained by organisations and their management practices (Paylor and McKevitt, 2019). Nonetheless, to fully embed EDI in usability testing, we propose a research agenda for a community of technology users, creators of technologies, usability researchers, and all those advocating for EDI in usability research. This community should embrace co-production through the following research agenda.

1. Advocate for, and increase the visibility of, inclusive usability testing.
2. Ensure usability research is a positive experience in which users feel safe, comfortable and valued.
3. Develop a diverse community of users and researchers. The larger and more diverse the community, the more that can be achieved.
4. Value and respect everyone by recognising everyone's expertise, whatever their role and interest.
5. Create a knowledge bank to share best practices for embedding EDI in usability. Sharing knowledge would surface more research and prevent duplication of effort.
6. Be open to learning from mistakes and sharing what has gone wrong.
7. Collect and share evidence of how EDI has improved products, services and systems.

5.4 Limitations and future work

Naturally, our participant sample was a subset of a wider population that is relevant to the research questions. This means that our sampling is not representative of all the possibly underrepresented social groups and therefore we make no claim that our recommendations are applicable to all. Furthermore, recruitment was facilitated by our project partners and therefore focused on disabled people and people who identify as LGBTQ+. However, with the exception of those who are digitally excluded, our participants had varied backgrounds (including from different socio-economic statuses, from different ethnic minorities, and those who speak English as a second language), very often characterised by intersectionalities (e.g., an older lesbian with mobility impairment). Future research could usefully test the recommendations with a wider subset of underrepresented participants as well as those who are well-represented.

The findings of this study are based on our participants' experiences and their opinions collected during interviews and roundtable discussions. As many of our participants had little actual experience of participating in usability tests, it would be helpful in future work to run observations of usability tests conducted with this study's recommendations. It would also be useful in future work to investigate whether five users is sufficient to identify the majority of usability problems (Nielsen, 2000) when a diverse pool of participants is recruited. It is also recognised that some of the recommendations made to researchers and organisations may be culturally and structurally difficult to implement (such as allowing participants flexibility when there are standardised protocols to follow and there are expectations that research should be reproducible). How this tension could be resolved could also be helpfully investigated in future work.

The interpretations of our findings have inescapably been influenced through our own lived experience of conducting research studies either as PIs and Cols or as participants. This lived experience also incorporates aspects of our intersecting identities, where we, as research participants in the past, felt that our needs and interests are not addressed nor represented.

Only 11 participants and 2 project partners provided feedback in phase 3. This is likely because there was a six month gap between participation and the request for feedback; when we requested feedback immediately post roundtables 20 out of 35 responded. Through our website and social media channels we have continued to engage with many of our participants. It would be helpful in future research to establish ways to maintain connections with participants when not collecting data.

6. Conclusion

In this project, we brought together people whose identities can be underrepresented in usability testing with people who do usability studies to increase our knowledge and understanding of how we can embed equality, diversity and inclusion in usability testing.

From this research, we contribute practical recommendations to both researchers and organisations that conduct usability research. Our recommendations to researchers wanting to embed EDI in usability research are made from the point of view of those whose identities can be underrepresented in usability testing. This sets them apart from guidelines made from are experiences of usability researchers informed by their interactions with participants. Furthermore, previous guidance is siloed in studies that only represent one aspect of a participant's identity, yet identities are complex and what might be affecting usability may be difficult to isolate to a singular characteristic. By focusing on universal principles (sense of value, trust and agency), we believe that our recommendations, while pertinent to our participants, are applicable to all and can be considered as general good practice.

Our recommendations to organisations are important as there is a lack of guidance for organisations on how to support researchers wanting to embed EDI in usability testing. Previous guidance has focused on how researchers can improve the situation for test participants rather than for themselves. What is apparent from our study, is that for researchers to embed EDI in usability testing they also need the support of the organisations that employ them.

Taken together these recommendations (to researchers and to the organisations that employ them) could help in addressing inequalities that result from people being routinely excluded from research. Finally, to take the discussion forward there is a need to share expertise and perspectives. To this end, we contribute a research agenda for a community of technology users, creators of technologies, usability researchers, and all those advocating for EDI in usability research.

Acknowledgements

The authors would like to thank all the participants and our project partners who so generously shared their time and contributed to the recommendations and agenda.

Funding source: this work was supported by Research England QR-Policy Support Funding through the University of Sheffield.

Rights retention statement

For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

References

- Acharya, K.R., 2022. Promoting Social Justice Through Usability in Technical Communication: An Integrative Literature Review. *Tech. Commun.* 69, 6–26.
- Alerasoul, S.A., Afeltra, G., Hakala, H., Minelli, E., Strozzi, F., 2022. Organisational learning, learning organisation, and learning orientation: An integrative review and framework. *Hum. Resour. Manag. Rev.* 32, 100854. <https://doi.org/10.1016/j.hrmr.2021.100854>
- Bohnert, C.A., Combs, R.M., Noonan, E.J., Weathers, A.E., Weingartner, L.A., 2021. Gender Minorities in Simulation: A Mixed Methods Study of Medical School Standardized Patient Programs in the United States and Canada. *Simul. Healthc.* 16, e151–e158. <https://doi.org/10.1097/SIH.0000000000000532>
- Brooks, J., King, N., 2016. Template Analysis for Business and Management Students.
- Caliz, D., Martínez, L., Caliz, R., 2017. Helping People with Down Syndrome Through an Usability Testing Guide Proposal for Mobile Applications. *Int. J. Comput. Sci. Eng. Inf. Technol.* 7, 1–19. <https://doi.org/10.5121/ijcseit.2017.7201>
- Colliver, B., 2020. Representation of Lgbtq Communities in the Grand Theft Auto Series. *Video Games, Crime and Next-Gen Deviance: Reorienting the Debate* 131–149. <https://doi.org/10.1108/978-1-83867-447-220201008>
- Cornet, V.P., Toscos, T., Bolchini, D., Ghahari, R.R., Ahmed, R., Daley, C., Holden, R.J. (2020). 2020. Untold stories in user-centered design of mobile health: Practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. *JMIR mHealth uHealth* 8.
- Craven, J., Booth, H., 2006. Putting awareness into practice: Practical steps for conducting usability tests. *Libr. Rev.* 55, 179–194. <https://doi.org/10.1108/00242530610655984>
- Darin, T., Andrade, R., Sánchez, J., 2022. Usability evaluation of multimodal interactive virtual environments for learners who are blind: An empirical investigation. *Int. J. Hum. Comput. Stud.* 158. <https://doi.org/10.1016/j.ijhcs.2021.102732>
- Davies, H.T.O., Nutley, S.M., 2000. Developing learning organisations in the new NHS. *Br. Med. J.* 320, 998–1001. <https://doi.org/10.1136/bmj.320.7240.998>
- Disability Rights UK., n.d. About us [WWW Document]. Disabil. Rights UK. URL <https://www.disabilityrightsuk.org/about-us/>
- Disability Unit, 2021. Inclusive language: words to use and avoid when writing about disability. [WWW Document]. Gov.UK. URL <https://www.gov.uk/government/publications/inclusive-communication/inclusive-language-words-to-use-and-avoid-when-writing-about-disability>
- Elmore, K., Meloncon, L., James, C.A., 2014. Embracing interdependence: Technology developers, autistic users, and technical communicators., in: *Rhetorical Accessibility:*

- At the Intersection of Technical Communication and Disability Studies. Routledge., pp. 15–38.
- Etikan, I., 2016. Comparison of Convenience Sampling and Purposive Sampling. *Am. J. Theor. Appl. Stat.* 5, 1. <https://doi.org/10.11648/j.ajtas.20160501.11>
- Fernandez, A., Insfran, E., Abrahão, S., 2011. Usability evaluation methods for the web: A systematic mapping study. *Inf. Softw. Technol.* 53, 789–817. <https://doi.org/10.1016/j.infsof.2011.02.007>
- Gagnon, S., Augustin, T., Cukier, W., 2022. Interplay for change in equality, diversity and inclusion studies. *Hum. Relations* 75, 1327–1353. <https://doi.org/10.1177/00187267211002239>
- Garvin, D.A., 1993. Building a learning organization. *Harv. Bus. Rev.* 71, 78–91.
- Hanna, L., Ridsden, K., Alexander, K., 1997. Guidelines for usability testing with children. *interactions* 4, 9–14. <https://doi.org/10.1145/264044.264045>
- Harihareswara, S., 2015. User experience is a social justice issue. [WWW Document]. Code4Lib J. 28.
- Henry, S.L., 2007. Just Ask: Integrating Accessibility Throughout Design. [WWW Document]. Just Ask. URL www.uiAccess.com/JustAsk/
- Hutter, L., Lawrence, H.M., 2018. Promoting inclusive and accessible design in usability testing. *Commun. Des. Q.* 6, 21–30. <https://doi.org/10.1145/3282665.3282668>
- Imison, C., Kaur, M., Dawson, S., 2022. Supporting equity and tackling inequality: how can NIHR promote inclusion in public partnerships? An agenda for action.
- ISO, 2018. Ergonomics of human-system interaction - Part 11: Usability: Definitions and concepts.
- Johnstone, J., 2007. Technology as empowerment: A capability approach to computer ethics. *Ethics Inf. Technol.* 9, 73–87. <https://doi.org/10.1007/s10676-006-9127-x>
- Joyce, A., 2019. Usability Testing with Minors: 16 Tips. [WWW Document]. NN/g Nielsen Norman Group. URL <https://www.nngroup.com/articles/usability-testing-minors/>
- Korte, J.L., Potter, L.E., Nielsen, S.H., 2015. An experience in requirements prototyping with young Deaf children. *J. Usability Stud.* 10, 195–214.
- López-Cabral, Á., Real, J.C., Valle, R., 2011. Relationships between human resource management practices and organizational learning capability: The mediating role of human capital. *Pers. Rev.* 40, 344–363. <https://doi.org/10.1108/00483481111118658>
- Mack, K., McDonnell, E., 2021. What do we mean by accessibility research? a literature survey of accessibility papers in chi and assets from 1994 to 2019. *Conf. Hum. Factors Comput. Syst. - Proc.* <https://doi.org/10.1145/3411764.3445412>
- Mankoff, J., Hayes, G.R., Kasnitz, D., 2010. Disability studies as a source of critical inquiry for the field of assistive technology. *ASSETS'10 - Proc. 12th Int. ACM SIGACCESS Conf. Comput. Access.* 3–10. <https://doi.org/10.1145/1878803.1878807>
- Marsh, S., 2019. Designing and researching with people with aphasia. User research in government. [WWW Document]. User Res. Gov. URL <https://userresearch.blog.gov.uk/2019/08/30/designing-and-researching-with-people-with-aphasia/>
- Meiselwitz, G., Wentz, B., Lazar, J., 2009. Universal usability: Past, present, and future, *Foundations and Trends in Human-Computer Interaction.* <https://doi.org/10.1561/11000000029>
- Mihoc, A., 2020. Conducting remote research with people with access needs. [WWW Document]. User Res. Gov. URL <https://userresearch.blog.gov.uk/2020/05/20/conducting-remote-research-with-people-with-access-needs/>
- Moran, K., 2019. Usability Testing 101. NN/g Nielsen Norman Group. [WWW Document]. URL Nielsen Norman Group logoNielsen Norman Group
- Nielsen, J., 2012a. Thinking aloud: The #1 usability tool. [WWW Document]. NN/g Nielsen Norman Group. URL <https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/>
- Nielsen, J., 2012b. Usability 101: Introduction to Usability. [WWW Document]. NN/g Nielsen

- Norman Group. URL <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>
- Nielsen, J., 2000. Why you only need to test with 5 users. [WWW Document]. NN/g Nielsen Norman Group. URL <https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>
- Opiniano, M.J.M., 2021. Object-centered interviews in mixed methods: Yielding the emotions of overseas migrant households in family financial socialization. *Methods Psychol.* 5, 100073. <https://doi.org/10.1016/j.metip.2021.100073>
- Paylor, J., McKeivitt, C., 2019. The Possibilities and Limits of “Co-producing” Research. *Front. Sociol.* 4, 1–5. <https://doi.org/10.3389/fsoc.2019.00023>
- Pernice, K., Nielsen, J., 2012. How to Conduct Usability Studies for Accessibility Methodology Guidelines for Testing Websites and Intranets With People Who Use Assistive Technology.
- Petrie, H., Hamilton, F., King, N., Pavan, P., 2006. Remote usability evaluations with disabled people. *Conf. Hum. Factors Comput. Syst. - Proc.* 2, 1133–1141. <https://doi.org/10.1145/1124772.1124942>
- Potnis, D., Gala, B., 2020. Best practices for conducting fieldwork with marginalized communities. *Inf. Process. Manag.* 57, 102144. <https://doi.org/10.1016/j.ipm.2019.102144>
- Raza, S.A., 2021. Managing ethical requirements elicitation of complex socio-technical systems with critical systems thinking: A case of course-timetabling project. *Technol. Soc.* 66, 101626. <https://doi.org/10.1016/j.techsoc.2021.101626>
- Razak, F.H.A., Hafit, H., Sedi, N., Zubaidi, N.A., Haron, H., 2010. Usability testing with children: Laboratory vs field studies. *Proc. - 2010 Int. Conf. User Sci. Eng. i-USER 2010* 104–109. <https://doi.org/10.1109/IUSER.2010.5716733>
- Rubin, J., Chisnell, D., 2008. Handbook of usability testing: How to plan, design, and conduct effective tests, 2nd ed. Wiley Publisher.
- Schlesinger, A., Edwards, W.K., Grinter, R.E., 2017. Intersectional HCI: Engaging identity through gender, race, and class. *Conf. Hum. Factors Comput. Syst. - Proc.* 2017-May, 5412–5427. <https://doi.org/10.1145/3025453.3025766>
- Shneiderman, B., 2000. Universal usability. *Commun. ACM* 43, 84–91. <https://doi.org/10.1145/332833.332843>
- Smith, H., Manzini, A., Ives, J., 2022. Inclusivity in TAS research: An example of EDI as RRI. *J. Responsible Technol.* 12, 100048. <https://doi.org/10.1016/j.jrt.2022.100048>
- Strohmayr, A., Bellini, R., Meissner, J., Alabdulqader, E., Toombs, A., Finnigan, S.M., Balaam, M., 2018. CHiversity : Implications for equality, diversity, and inclusion campaigns. *Conf. Hum. Factors Comput. Syst. - Proc.* 2018-April, 1–10. <https://doi.org/10.1145/3170427.3188396>
- Sum, C.M., Alharbi, R., Spektor, F., Bennett, C.L., Harrington, C.N., Spiel, K., Williams, R.M., 2022. Dreaming Disability Justice in HCI. *Conf. Hum. Factors Comput. Syst. - Proc.* <https://doi.org/10.1145/3491101.3503731>
- Tindale, W.B., Dimitri, P., 2022. MedTech innovation across the life course—the importance of users and usability. *J. Med. Eng. Technol.* 46, 427–432. <https://doi.org/10.1080/03091902.2022.2091173>
- Tornblad, M.K., Jones, K.S., Inan, F.A., 2019. Individuals with Autism spectrum disorder: A review of human factors considerations and usability testing guidelines. *Proc. Hum. Factors Ergon. Soc.* 2019-Octob, 1575–1579. <https://doi.org/10.1177/1071181319631423>
- Twidale, M.B., Nichols, D.M., Lueg, C.P., 2021. Everyone everywhere: A distributed and embedded paradigm for usability. *J. Assoc. Inf. Sci. Technol.* 72, 1272–1284. <https://doi.org/10.1002/asi.24465>
- UK Government, 2017. Running research sessions with disabled people. [WWW Document]. Gov.uk Serv. Man. URL <https://www.gov.uk/service-manual/user-research/running-research-sessions-with-people-with-disabilities>
- van den Berg, A.C., Giest, S., Kraaij, W., 2023. Assessing inclusivity in online platforms

- through usability evaluation with Google Analytics. *Policy and Internet* 15, 55–77. <https://doi.org/10.1002/poi3.328>
- Van Der Geest, T., 2006. Conducting usability studies with users who are elderly or have disabilities. *Tech. Commun.* 53, 23–31.
- W3C WAI, 2016. Accessibility, Usability, and Inclusion. [WWW Document]. W3C WAI. URL <https://www.w3.org/WAI/fundamentals/accessibility-usability-inclusion/>
- Walls, D.M., 2016. User experience in social justice contexts. *SIGDOC 2016 - 34th ACM Int. Conf. Des. Commun.* <https://doi.org/10.1145/2987592.2987604>
- Wheeler, G., Mills, N., Ankeny, U., Howsley, P., Bartlett, C., Elphick, H., Dimitri, P., 2022. Meaningful involvement of children and young people in health technology development. *J. Med. Eng. Technol.* 46, 462–471. <https://doi.org/10.1080/03091902.2022.2089252>
- Williams, P., 2006. Developing methods to evaluate web usability with people with learning difficulties. *Br. J. Spec. Educ.* 33, 173–179. <https://doi.org/10.1111/j.1467-8578.2006.00436.x>
- Wood, R., Dixon, E., Elsayed-Ali, S., Shokeen, E., Lazar, A., Lazar, J., 2021. Investigating Best Practices for Remote Summative Usability Testing with People with Mild to Moderate Dementia. *ACM Trans. Access. Comput.* 14, 1–26. <https://doi.org/10.1145/3460942>
- Yoo, Y., 2010. Qiarterjy. *MIS Q.* 34, 213–231.

Appendix

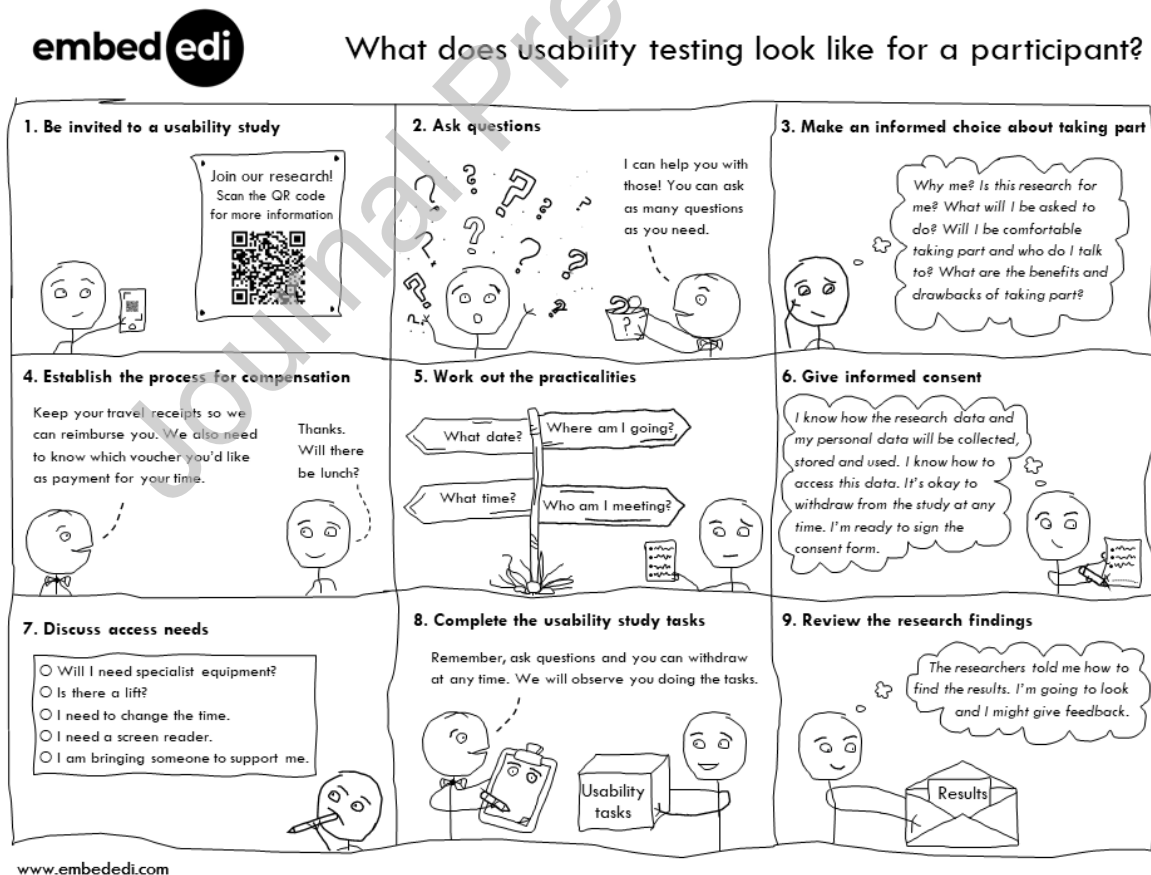


Figure 4: Storyboard of what does usability testing look like for a participant (published on project website)

Agency

People who have **lived experience** of a particular context can describe it more effectively than those who do not. If researchers and product designers **"talked to enough people"**, they'd have worked out what they actually want to do". Participants want meaningful **flexibility** when contributing to usability studies – not to have options dictated to them. A lack of flexibility robs them of their **agency** and impacts their energy for engaging in usability tasks. Whereas having **options** which are **"tailored"** means participants **feel less pressured** and more **"involved"** in the usability test.

Participant perspective

Usability studies should be **participant-centred**. This helps to create **space** for participants to express their viewpoint and communicate their needs. Participant-centred studies lead to **comfortable experiences** with **"warm and open conversation"**. Conversely, participants addressed a potentially unhealthy **power dynamic** that can lead to them feeling **"anxious"**, patronised or pushed. Participants do not want to feel **"judged"** if they cannot explain their choices in a usability test or **"embarrassed"** if they get something **"wrong"** whilst being observed.

Information transparency

Transparent information throughout the process is highly valued by participants. It helps them to **feel at ease** when **communicating their needs**. Receiving information and details about predefined usability tasks **in advance** lets participants **"process them in their own time"**. **"Checklists"** and **"step by step walkthroughs"** help participants to know what to expect in a usability study. Using **"plain" language** means participants are on a **"level playing field"**.

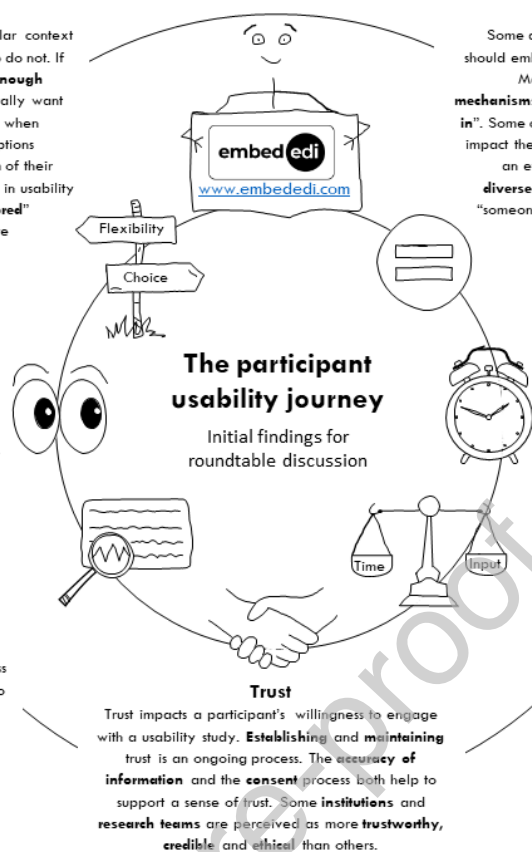


Figure 5: Coding template: The participant usability journey (published on project website)

The user-researcher perspective

Initial findings for roundtable discussion



www.embededi.com

Top force: intention

The desire

Strong desire to carry out research that is equitable, inclusive and diverse.

Second force: the broader situation

Time pressure

- * deadlines

Financial pressure

- * lack of resource

Organisational flux

- * staff change position / role
- * staff / agency staff on short term contracts
- * staff and project turnover
- * changes in organisational structure, leadership, direction and role

Third force: doing research

Rigid and standardised procedure

- * difficult and time consuming to adapt
- * expectation to use templates / procedures which may not be accessible for everyone
- * extra safeguarding procedures for 'vulnerable' participants
- * expected to create usability tasks before recruiting participants

Tackling organisational culture

- * ingrained ideas about diversity and inclusion
- * helping clients understand good EDI practice

Working within organisational structure

- * no authority to push ideas
- * deadlines / delivery decided elsewhere
- * usability testing is an afterthought / at the end of the project

Working within a defined project

- * tight deadline
- * budget
- * scope decided by clients and stakeholders
- * need to commit to approaches early – like ethical approval before recruitment

Getting guidance

- * siloed teams – knowledge is spread not centralised
- * lack of / unable to access guidance
- * lack of feedback / dialogue between researchers and stakeholders
- * creating usability tasks which are representative of diverse usage

Needing research credibility

- * expectation of a single, predefined experiment with consistent procedures
- * expectation of controlling for differences

Recruiting participants

- * existing channels (databases) lacking diversity
- * working with voluntary / faith groups but not able to develop long-term relationships
- * project deadline pressures
- * recruiting people who can fit available slots (see research space)

Getting access to research space

- * lack of accessible, physical space
- * limited availability of physical space
- * needing the right equipment to support participants with usability tasks
- * organisational restriction on communication technology options

Researcher worries

- * ethical approaches, including compensation
- * not knowing people's needs
- * safeguarding participants and the research team

Fourth force: the reality

Working within restraints

It is difficult to take a flexible approach. Do the best I can within the constraints in which I am operating.

Figure 6: Coding template: The user-researcher perspective (published on project website)

Authors	Key findings / guidance	Test participant
Cáliz, Martínez and Cáliz (2017)	Guide that includes task design, giving instruction, design the session and pilot test, recruiting participants, analysis and reporting of data.	People with Down's syndrome
Cornet et al (2020)	Guidance for researchers of the advantages and disadvantages of labs and in-the-wild testing, the need to adapt methods and be flexible, as well as balancing the number of concurrent evaluation methods to reduce participant cognitive load.	Older adults
Craven and Booth (2006)	A checklist for preparing for usability tests including deciding on the objectives of the test, who and how many to recruit.	Disabled people
Darin, Andrade and Sánchez (2022)	Guidance (planning, conducting and reporting) on choosing usability evaluation measures appropriate for test participants characteristics and evaluation goals.	Learners who are blind
Hanna et al. (1997)	Detailed guidance on setting up and planning the usability test, introducing the test and establishing a relationship, and what to do during and after the test.	With children
Henry (2007)	Detailed guidance on what to ask in a screener, and how to plan, prepare, conduct and report usability tests.	Disabled people
Hutter and Lawrence (2018)	Guidelines for inclusive practice focusing mostly on pre-session engagement with people who are deaf and ASL interpreters so that their perspectives are incorporated, the research team are knowledgeable, and sessions can be adapted as needed.	People who are deaf
Joyce (2019)	Guidance on how to recruit, design studies and facilitate sessions including advice on age appropriateness of incentives, language, tasks and study environment as well as how to dress.	People under 18
Korte et al. (2015)	Tips include minimising distractions, helping children feel comfortable, and offering alternative activities and communication channels.	Young deaf children
Marsh (2019)	Practical tips for preparing sessions, supporting participants and analysing results including the importance of collaboration, communicating	People with aphasia

	clearly, making the session comfortable, and taking things slowly.	
Mihoc (2020)	Tips on how to respond to challenges such as recruitment, data collection failures, and team building. Further tips including getting to know participants, tailoring sessions, being responsive to participants needs, and using preferred devices and assistive technology,	People with access needs
Pernice and Nielsen (2012)	Detailed guidance for online testing including designing the study, study location, using assistive technology, recruiting participants, as well as sample documents.	People who use assistive technology
Petrie et al. (2006)	Principles on when and how to choose between in-person and remote evaluations.	Disabled people
Razak et al. (2010)	Guidance on conducting usability studies in a laboratory or in the field.	Children
Tornblad et al. (2019)	Guidance on preparing and conducting tests, including ensuring the system being tested has base level accessibility, considering the comfort of participants, and ways to support participation.	Autistic people
Van der Geest (2006)	Guidelines on what to consider before, during and after a usability test including who to sample when, Invitations to participate, drawing a varied sample, designing tasks, measuring success, communicating with participants, test location and materials, getting consent and reporting results.	Elderly people or disabled people
Williams (2006)	Findings include the need to make tasks interesting and meaningful, ensuring the tasks can be understood, the need for assistive devices, and the need for supporters.	People with a learning disability
Wood et al. (2021)	Detailed guidance on planning and logistics, conducting remote tests and evaluating results. In addition, lessons learned are reported including supporting agency, reducing anxiety and reducing fatigue.	People with mild to moderate dementia

Table 3: Summary of guidance for conducting inclusive usability tests

Author contributions

Sophie Rutter: Conceptualisation, Methodology, Formal analysis, Investigation, Writing - Original draft, Visualisation, Supervision, Project administration, Funding acquisition

Efpraxia Zamani: Conceptualisation, Methodology, Investigation, Writing - Review & Editing, Supervision, Funding acquisition

Joanna-Louise McKenna-Aspell: Methodology, Formal analysis, Investigation, Resources, Writing - Review & Editing, Visualisation

Yuhua Wang: Methodology, Formal analysis, Investigation, Writing - Review & Editing

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: