UNIVERSITY OF BRISTOL

University of Bristol - Explore Bristol Research
General rights

Inclusivity in TAS research: An example of EDI as RRI

Helen Smith [a,*], Arianna Manzini [a,b], Jonathan Ives [c]

[a] *Centre for Ethics in Medicine, Bristol Medical School, University of Bristol, Canynge Hall, 39 Whatley Road, BS8 2PS, Bristol, United Kingdom*
[b] *Radcliffe Department of Medicine, MRC Weatherall Institute of Molecular Medicine, University of Oxford, OX3 9DS, Oxford, United Kingdom*
[c] *Centre for Ethics in Medicine, Bristol Medical School, University of Bristol, United Kingdom*

A B S T R A C T

Responsible research and innovation (RRI) aims to achieve the goal of making research activities responsible and ethical. To those ends, it is vital for researchers to actively engage with equality, diversity and inclusion (EDI) which, if not attended to, may detrimentally affect both potential research participants and the research itself.

Our paper offers an account of our ongoing discussions surrounding the importance of EDI when designing our research, how we employed EDI to intentionally make our recruitment process more inclusive, and our ongoing planning to make all our research activities as inclusive, diverse, and accessible as possible.

There is no one-size-fits-all approach for adopting EDI principles in RRI, however, we posit that their consideration is essential for research communities who wish their work to represent the perspectives of those who will be affected by future novel technologies.

Introduction

Whilst being attentive to responsible research and innovation (RRI) is a clear and generic requirement for all researchers, the UKRI notes that there is "no prescriptive overarching checklist to embed responsible research and innovation in the research and innovation process" (UKRI, 2021). Instead, researchers are urged to reflect on whether their work is controversial, has significant ethical or moral components, or raises issues of trust or social acceptability (UKRI, 2021). In this process, a core requirement of the RRI framework is to engage a wide range of stakeholders at an early stage, so that different perspectives can contribute to the identification, and addressing, of the social and ethical concerns that may arise from research (Stilgoe, 2013). In this short paper, we reflect on our own attempts to embed RRI into our research, and how that led us to recognise the necessity of incorporating thinking about equality, diversity and inclusion (EDI) from day one. In particular, we reflect on how we have attempted to adjust for EDI concerns in our sampling and recruitment strategy.

EDI considerations in our study

Our research, funded by the UKRI Trustworthy Autonomous Systems (TAS) hub, is exploring how we should conceptualise and operationalise trustworthiness in the design of autonomous systems with evolving functionality (the ability for a system to autonomously change its process and adapt in response to unpredicted changes in the system or in the surrounding environment). Part of our ethics-focussed work package (ARET Study, 2021) involves speaking with stakeholders, including interviewing potential end users about how they think about trust in relation to technologies they use, and why they make decisions to trust a technology or not.

As ethicists ourselves (some of whom teach and publish on research ethics), we felt very comfortable thinking about ethical issues arising from our research. Thinking about informed consent, respecting autonomy, and avoiding harm (etc.) are, after all, our primary activities. As such, it took some of us somewhat off-guard when confronted with the possibility that our research practice could be unethical in a very specific way, unless we took steps to address it.

The issue arose when we were planning our protocol for sampling and recruiting participants into our qualitative study. We noted early in our work that novel technological developments are frequently framed as being designed to assist human activities, ranging from packing our shopping to helping us dress. We were already aware that we needed to include a wide range of people in our study, so that we could gather insight from a range of different perspectives. In addition, we recognised there might be important and specific viewpoints that we needed to actively sample for. For example, persons with disabilities may be more likely to engage with certain assistive technologies and have particular and important standpoints on trust and trustworthiness. The needs of the population a technology seeks to aid are as diverse as its members, and

our intention was to conduct research that was as diverse and inclusive as possible. In order to do this, we felt we needed to ask potential participants for certain pieces of demographic information at the point of contact, so that we could purposively sample for diversity and monitor to what extent we were achieving a diverse participant population. This is an accepted and non-controversial practice in qualitative research.

The difficulty arose when we were designing the initial contact form, which potential participants would be invited to complete to express an interest in the study, and at the same time provide some information about themselves so that we could purposively sample a diverse participant population. Our initial design asked standard questions about, for example, age group, ethnicity and gender, and also included a question about disability, as disabled people may very well be disproportionately high users of autonomous systems with evolving functionality in future applications. We explained clearly why we were asking for this information and what it would be used for. So far, so diverse.

However, one of our team members noted in these early drafts that questions about certain characteristics were being asked in a way that was restrictive to the point of potentially limiting inclusivity, thus being subtly exclusionary. In particular, our initial draft of our expression of interest form used pre-categorised ethnic groups. Each ethnic group's descriptions were broad, commonly used terms, and might well have been appropriate for many people filling in the form; for example: 'White', 'Black', 'Asian'. We initially provided the category 'Other' for people to self-describe if their ethnic category was absent from our list or if they preferred an alternative description. It was the 'Other' category that prompted our group's discussion about how 'othering' people fails to recognise the importance of their characteristics and may antagonise or alienate potential participants to the point that they do not engage with the research – and would serve to exclude them.

We discussed why we were using these categories, and the reasoning was twofold. First, broad categories saved us from having to identify sub-groups within larger ethnic groups, which would have raised the question of how and why those more specific categories had been identified, and who had made that choice. Second, we initially wanted to use broadly set categories to allow us to easily group people together for demographic reporting. The worry was that if we ended up with a large number of participants all describing their ethnicity slightly differently, we would have to report all of those differences, which would be unwieldy and not terribly useful. Alternatively, we may have ended up having to group them together into broader categories for reporting at a later stage, which rather defeats the object of not having pre-defined categories in the first place.

After careful consideration we concluded that, whilst it may be more convenient to use pre-defined categories, the convenience was relatively minor compared to the risk of putting off (and risking excluding) potential participants. We felt that participants needed to be free to express the fullness of their ethnic identity, and to withhold that option devalues their identity. If we devalue potential participants from the outset, we risk them losing interest in engaging in our activities, which would subsequently harm our aim of undertaking inclusive research. Even if one person was put off or did not engage as result of not feeling their ethnicity was valued, that would be one person too many. We only made a small change to rectify this: we are inviting potential participants to self-describe their ethnicity. In this way, we hope to be able to capture what participants wish to tell us, with as much precision as they wish to offer. We felt that this was a far more respectful way to ask people about their ethnicity, without asking them to fit into a pre-defined category that they may not identify with.

The same argument then seemed to apply to our enquiries about gender. Although our initial draft had a non-binary option, we acknowledged that people may have multiple and varied ways of describing their gender and being forced to select a description they do not identify with could limit inclusivity. We recognise that there are more than two genders and were, again, keen to avoid the 'Other' box.

Listing all the possible genders a person may identify with (just as with all the different possible ethnicity descriptors) would lead to a cumbersome and unwieldy form, and so, again, we are inviting our potential participants to self-describe their gender.

Regarding disability, we ask our potential participants if they have disabilities that might give them a particular insight into the research topic. Our question here serves to highlight their inclusion and allows us to favour their inclusion over that of others if we have too many responses to our initial call for participants. These discussions also encouraged us to consider the importance of ensuring accessibility with regards to research encounters, both in terms of physical access and flexibility in the mode of interviewing to suit different needs.

However, when looking at our inclusive methodology, we did identify a problem. In requesting participants to self-describe we risk generating large numbers of unique demographic descriptors, populated by potentially as few as one person, which would threaten anonymity if reported. For example, a person who identifies themselves as from a particular sub-section of an ethnic group, who is gender-non-binary, and lives with a specific disability, could potentially be identifiable as they may be one of very few people with those particular characteristics. The promise of anonymity is often central to the consent a participant gives, and that promise would be broken if demographic reporting of relatively unique individuals is not done carefully. We have not yet fully worked out how to manage this, but note that being attentive to it means we will be in position to do so. We suspect that the solution will be to report the demographic make-up of the sample in general, but not that of individuals. This could well create problems for reporting, as we will not be able to link a particular quotation to person with specified characteristics. We will have to wait and see to learn if this will be a problem (given the participants we get), but also how reviewers and editors will respond when we explain our reporting strategy.

Lessons learned

We started from the premise that *good research is curious about being inclusive*. Inspired by Nind's (2017) use of Aristotle's five virtues, we considered "what should I do now, in this situation, given these circumstances, facing this particular person, at this time" (Schwandt, 2007, p.243). Our aim is to recruit participants who are representative of the diverse population who will be affected by autonomous systems with evolving functionality.

We are still at the early stages of planning for our qualitative research, but these early discussions have led us to plan for other research-related activities, such as public engagement events, to be as inclusive as possible and so be able to encourage participation from a wide variety of potential stakeholders. Our considerations have so far included actions such as:

- Allowing for translation services where needed (languages which are spoken as well as signed).
- Captioning in our videos and/or providing transcripts of presentations.
- Clear and simple language use in all our written and audio/visual materials.
- When planning engagement activities, only considering venues which are readily accessible for all persons, including those with disabilities, both visible and invisible.
- Interviewing people either face-to-face or remotely with flexible timings to accommodate factors such as the participant's working lives, care requirements/commitments, or their reluctance to risk exposure to other people in the context of the COVID-19 pandemic.
- Creating space for diversity by inviting our participants to express their needs so that we may help to make their participation experience with us as barrier-free as we possibly can, thus promoting equality of access and inclusion.

Our EDI considerations have not been revolutionary, but that does not mean that our pursuit of improvements in inclusivity are any less important. These are all very small accommodations with the intention of proactively removing every barrier that we can to promote equitable access to engaging with our research, and a more personalised approach when interacting with stakeholders. Despite our efforts, we fully expect to fall short of what could potentially be achieved. We are aware that it is impossible to plan in advance for the needs of every potential participant, or to proactively attempt to make provisions for every person who could possibly engage with our work. For example, we do not know how many participants will have access needs, nor what those access needs would be. We welcome feedback from our participants and aim to meet their needs in the present, as well as the needs of others in future projects. In this way, we aim to remain teachable, thereby improving our inclusion, and, as a result, our research.

It might be asked whether there is actually any need for these endeavours at all, as the number of people disincentivised by our original initial contact form would be probably very small. We are also never likely to know if any of our efforts actually made any difference, and regardless of our planning, potential participants may choose not to engage with us for a wide variety of reasons. We are also aware that in order to judge whether our sample is diverse enough (allowing us to purposively recruit more people from groups that are underrepresented), we may ultimately need to group people together into the pre-defined categories that we had decided to set aside in favour of self-identification. For example, whilst we might have applicants who self-identify as White British, White Irish, White European, White South European, etc., we might need to group them together into 'White'. Our recruitment aims to ensure that our research benefits from the perspectives of ethnically, culturally, disability, and gender diverse populations, rather than our research representing multiple sub-categories of only one broad ethnic group and then incorrectly declaring that to be diverse. As such we have to make assumptions about what counts as *meaningful* diversity – and that conversation is ongoing.

The value of what we have done does not lie in having removed pre-defined categories from an initial contact form *per se*, but in having made a conscious effort to be interested in widening participation, in order to honour the ethical obligation to respect persons. Ultimately, by inviting people to describe their ethnicity and gender, as well as to declare their disability, we wish to treat people with respect and value them as individual persons, so that they may feel confident we will value their contribution to our project. In turn, our approach, which takes EDI considerations seriously, makes it more likely that our work will benefit from a more diverse participant population. In line with the requirements of the RRI framework, a more diverse population could help us anticipate and address potentially controversial implications that our work seeks to reveal, which would have otherwise been missed had a more homogenous group been recruited.

Final thoughts

EDI considerations are particularly important within TAS, as trust will not be achieved when the interests of the diverse populations which TAS seeks to serve are neglected. As such, EDI needs to be ongoing work that is established in all aspects of research. Whilst this paper is focussed on recruitment activities for our own work, there is a persistent broader requirement to ensure continued representation of diversity in all TAS work. This includes - but is not limited to - ensuring the representation of a variety of ages, cultures, genders, nationalities, geographical locations, (dis)abilities, and (non)faiths when performing consultation and other research and engagement activities. As a publicly funded endeavour, TAS research ought to be useful and (as appropriate) generalisable to those cross-sections of society that it would affect and/or seeks to influence. This can generate research outputs that are both informed by, and tailored to meet, the needs of those which it seeks to serve, and to reduce the risk of harm to those it may reach.

Inclusivity in research works across a spectrum, from accessing participation to accessing the findings, and we cannot expect people to take part in research if the end product is inaccessible to them. Whilst we have focussed on participation above, this is only part of the story. As such, we also call on TAS researchers to consider accessibility at the other end of the research spectrum, and ensure they publish their work using Open Access platforms. This measure a) allows for equality of access to publicly funded research by those both inside and outside of academia, thus enabling b) *all* persons to benefit from TAS research and to apply it to their own work and understandings.

EDI considerations in research are not a one-time tick box; they are an ever-ongoing effort to strive for inclusion of all who wish to contribute. If we are not attuned to this, whatever else we do, we can never be considered to be doing research on technical innovation in a responsible way.

Declaration of Competing Interest

Acknowledgements

References

ARET Study (2021). Retrieved from https://tasfunctionality.bristol.ac.uk/aret/Accessed: 15 December 2021.

Nind, M. (2017). The practical wisdom of inclusive research. *Qualitative Research, 17*(3), 278–288.

Schwandt, T. A. (2007). *The SAGE dictionary of qualitative inquiry* (3rd ed). Sage.

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy, 429*, 1568–1580.

UKRI (2021). Research integrity in healthcare technologies. Retrieved from https://www.ukri.org/councils/epsrc/guidance-for-applicants/what-to-include-in-your-proposal/health-technologies-impact-and-translation-toolkit/research-integrity-in-healthcare-technologies/responsible-research-and-innovation/Accessed: 13 November 2021.