

In an era of open science, how is ‘transparency’ shifting within qualitative research

openwashedata.

Marc Kalina

Global Health Engineering, ETH Zurich

2023-10-02

Note: This blog post is a re-published version from the [Global Health Engineering website](#). For attribution, please see the [Citation information](#) at the end of this post.

Qualitative research is inherently not reproducible or replicable, a fact those with a more positivist view of knowledge creation find hard to digest. Reproducibility means obtaining consistent results using the same protocol and data, while replicability entails obtaining the same results using the same methods but with different data or a different sample. When using interpretive qualitative methods (which rely on the co-creation of knowledge between researcher and respondent), both the study conditions and the results obtained can hinge on a variety of factors, not least of all, the positionality of the individual researcher. Simply put: when you ask someone a question, you are not guaranteed the same response I might receive.

Therefore, within qualitative research it is more common to speak of transparency, rather than reproducibility or replicability, when evaluating the scientific rigour of a given piece of work. And, while the broader scientific community has moved towards increased openness with data and analysis, there has been a growing debate within qualitative research communities on how much transparency is necessary or possible within current praxis. Sebastian Karcher¹ of the [Qualitative Data Repository](#) has helpfully distinguished between three types of transparency most relevant in qualitative research.

1. **Production transparency:** Information on how a study generates or collects data. This can include an interview schedule, and the types of methodological detail that would flesh out a good methods section, such as sample, subject recruitment, exclusion criteria, etc.
2. **Analytic transparency:** Information on how data is prepared and analysed.
3. **Data access:** Data availability, including access conditions.

Of these forms, the first two have become increasingly *de rigueur*, with detailed methods sections becoming expected, and information on analysis, including codebooks, becoming increasingly common as supplementary submissions to articles. However, data transparency, including open use and access for qualitative material, especially within the social sciences, has lagged far behind other disciplines where open access is increasingly becoming the standard. There are practical reasons for this, because, as Sebastian Karcher notes, transparency within

¹This is a great blog post by Open Working summing up a presentation by Sebastian Karcher at TU Delft on the limits of reproducibility within qualitative research: <https://openworking.wordpress.com/2019/02/11/what-does-reproducibility-mean-for-qualitative-research/>

qualitative research is not a straightforward path, and is often fraught with ethical considerations and practical barriers. What are we doing at Global Health Engineering (GHE) to move this discussion forward and become more transparent in our own work?

At GHE we have made a pledge to embrace the full spectrum of transparency: to make all of our data, either quantitative or qualitative, open access under permissive licenses, and to publicly document how it has been produced and analysed². This is, of course, important, so that the rigour of our research practice can be clearly evaluated, and so that our data can be accessible to other researchers, who may favour other modes of analysis, allowing the words of our respondents to resonate further and make an impact beyond what we may have foreseen.

Through our partnership with the [openwashdata](#) team we recently published [our first open access qualitative dataset](#): a series of 61 semi-structured interviews with biogas plant owners in southern Malawi [1]. Moving forward, we intend to publish all of our work in similar fashion, but as a qualitative researcher going through this process for the first time, what did I learn from the experience?

First, making our data open access necessitates a change in ethical practice, particularly in how we obtain consent and inform participants on how their responses will be utilised. Will participants be as willing to speak freely if they know their responses will be made public, albeit anonymously? Second, the preparation of qualitative material for open access is an important step– but is time and resource intensive. Even with an in-house ETH data steward who was able to automate some of these processes, the anonymisation and packaging of transcripts is time consuming. Although we were able to devote the time and resources towards publishing the data package, I am concerned about the feasibility for scholars in less-resourced institutional contexts of giving their work the same platform as standards for transparency shift.

Nonetheless, we are pleased with the results of our first effort, and invite engagement and feedback from the greater scientific community. Moving forward, what do we want to improve? First, we would like to be more timely with our outputs, with data packages being published soon after they are collected, ensuring their broader usefulness and to support our own published analyses. Second, to date, we have only published transcripts in English. In the future we will also publish transcripts in the language in which the data collection was conducted, so that the un-translated voices of our respondents can stand on their own. Finally, we intend to develop data management guidelines and open source tools to support data preparation processes so that scholars from low-resource contexts can pursue data publishing without regard to material constraints.

The future is hard to predict, but we are hopeful it will be more transparent.

Bibliography

- [1] L. Schöbitz, S. C. Loos, M. Kalina, J. O. Ogwang, J. Kwangulero, and E. Tilley, “Biogasout-comesmalawi: Data for 61 Semi-Structured Interviews with Biogas Owners in Malawi.” Aug. 2023. doi: [10.5281/zenodo.8289621](https://doi.org/10.5281/zenodo.8289621).

²Learn more about how we do this in this interview with Dr. Julian Dederke of the ETH Library: <https://ethz.ch/staffnet/en/news-and-events/internal-news/archive/2023/03/interview-eine-vision-fuer-open-science-and-data-stewardship-an-der-eth-zuerich.html>