

# Working with big amounts of qualitative data: What you might want to do with it

Data sets from qualitative research studies can be stored in digital archives and available for reuse, such as the UK Data Archive: <http://www.data-archive.ac.uk/>. These archives present the possibility of merging interview transcripts, fieldnotes, etc. drawn from multiple small-scale archived qualitative projects sharing the same broad topic into one big data set, and conducting secondary data analysis.

Merging together several small qualitative studies into one large data set can be seen as a form of 'big qualitative data' akin to quantitative 'big data' and needing computational processing to work with the information. Critics point out that bigger volumes of data do not offer 'better' or more accurate forms of knowledge. Biases within big data need to be acknowledged, and attention given to the context of its creation and use, otherwise context and its associated meaning is lost, misunderstood, or left unexplained. The 'context' of big data relates to the social and cultural circumstances in which it was generated; how it was sourced and cleaned; and the approaches taken to analysis.

Although convinced of the value of small-scale qualitative enquiry in advancing knowledge, we are also convinced of the benefits of assembling and working with a large amount of qualitative data drawn from different data sets. These include:

## 1. Comparison

Merging data from multiple qualitative studies offers opportunities for new research questions to be addressed making comparative use of differences between the studies. This comparison can be guided by metadata about, for example, the focus of the studies, sample characteristics, geographical and temporal location, and so on. Some assemblages enable comparison across disciplinary differences. Researchers are able to ask questions that could not be answered by individual projects.

## 2. Generalisation

Merging data across projects may also enhance the possibilities of generalising from the data. Multiplying the number of small unrepresentative samples will never add up to a representative sample, but increasing the diversity of samples and total number of research participants may strengthen claims about understanding how social processes work.

### SOME USEFUL RESOURCES:

There are discussions of working across several data sets and large amounts of qualitative data on the BigQual website, in a series of guest blog posts, including:

- Anna Tarrant on working across two qualitative data sets  
<http://bigqlr.ncrm.ac.uk/2016/03/09/assessing-the-feasibility-of-secondary-analysis-within-and-across-two-qualitative-longitudinal-datasets-reflections-from-the-mplc-study/>
- Bren Neale on working across assemblages of archived materials  
<http://bigqlr.ncrm.ac.uk/2017/12/04/guest-post-13-prof-bren-neale-research-data-as-documents-of-life/>
- A range of references to do with BigQual can be found at: <http://bigqlr.ncrm.ac.uk/resources-3/>

### See also:

Boyd, D. and Crawford, K. (2012) Critical reflections for big data, *Information, Communication & Society*, 15(5), 662-679.

Mason, .J (2007) 'Re-using' qualitative data: The merits of an investigative epistemology, *Sociological Research Online*, 12(3): <http://www.socresonline.org.uk/12/3/3.html>.

Tarrant, A. (2016) Getting out of the swamp? Methodological reflections on using qualitative secondary analysis to develop research design, *International Journal of Social Research Methodology*, 20(6), 599-611.