



Sharing qualitative data & analysis. With whom and how widely?: A response to ‘Promises and pitfalls of data sharing in qualitative research’



Amber Wutich ^{a,*}, H. Russell Bernard ^b

^a Center for Global Health, Arizona State University, School of Human Evolution and Social Change, Arizona State University, P.O. Box 872402, Tempe, AZ 85287-2402, USA

^b Institute for Social Science Research, Arizona State University, Department of Anthropology, University of Florida, P.O. Box 874602, Tempe, AZ 85287-4602, USA

ARTICLE INFO

Article history:

Received 27 September 2016

Accepted 28 September 2016

Available online 30 September 2016

Tsai and colleagues respond to an important challenge: access by the growing number of scholars who use qualitative data and do qualitative analysis to top-tier journals that have data-sharing requirements. The authors offer guidance on how those data-sharing requirements could be modified to accommodate the needs of these colleagues—a welcome addition to a long-standing, at times controversial, and ongoing conversation about sharing of qualitative data.

Our response is offered in the spirit of advancing this conversation. We make two points: (1) Qualitative data and qualitative analysis are different and should not be combined into a single term, “qualitative research.” (2) We can leverage the lessons from already successful efforts to archive and share qualitative data.

1. Qualitative data and qualitative analysis are different—and they require different procedures for sharing

Tsai and colleagues conclude that “verification is likely to be impossible in the setting of qualitative research,” but they also describe how an “audit trail” could be assembled to capture the analytic stages in qualitative research. The problem here, in our view, is the combining of qualitative data and qualitative analysis into one phrase, “qualitative research.” Qualitative data—whether

words or images, sounds or artifacts—can be analyzed quantitatively or qualitatively (Bernard et al., 2016) and several traditions of social science research that rely on qualitative data have fully or nearly fully developed documentable steps.

Classical content analysis, for instance, has clear rules for segmenting texts, assessing inter-rater reliability, and coding systematically (Krippendorff, 2012). Word-based analysis, which encompasses machine-driven counts of words or word clusters, also allows for statistical verification of hypotheses derived from qualitative data (Ignatow and Mihalcea, 2016). Semantic network analysis involves the production of matrices summarizing relationships within a set of texts across codes or words (Handwerker and Borgatti, 2014). This latter field is burgeoning with the development of programs that read natural language texts and produce themes. In all of these traditions, partially or entirely quantitative analysis of qualitative data can be reproduced, replicated, and verified. This is one reason we advocate distinguishing between procedures for sharing qualitative data and qualitative analysis.

While the problem of documenting some qualitative analyses is easily solved, in many cases it is not. Most of the concerns that Tsai and colleagues raise involve the documentation of analyses that involve many inductive inferences—inferences that can, indeed, be difficult to document. There are, however, several well-developed literatures on documenting these inference steps. Grounded theory, for instance, uses memo-writing to document initial coding (e.g., in-vivo), later-stage theoretical coding (e.g., axial coding), and

* Corresponding author.

E-mail address: amber.wutich@asu.edu (A. Wutich).

theoretical saturation (Charmaz, 2014). Discourse analysis, narrative analysis, and schema analysis also have conventions for the documentation of key analytic inferences. Cutting across these traditions, it is possible to identify a core set of analytic steps that are widely shared: lists of codes and subcodes; detailed definitions of major codes; memo-writing to document analytic thinking; and the production of coded data segments (“coding queries,” in the authors’ terms).

As Tsai and colleagues suggest, it is possible to move closer to a standard set of expectations for reporting on most forms of qualitative analysis and, based on such reports, for reproducing major thematic discoveries. When shared, data protocols should be clear about what kinds of analyses the researchers believe can be accomplished: validation of a particular study or avenues to novel secondary studies (Hammersley, 1997).

2. We can leverage the lessons from already successful efforts to archive and share qualitative data

Tsai and colleagues focus on sharing textual data derived from field notes and interview transcripts. Qualitative data, however, take many forms—film, audio, photographs, sketches, artwork, and the like—and ethnographers have shared all of these materials from the earliest days of anthropological research (Cliggett, 2013; Kemper and Royce, 2002; Pelto, 2015; Silverman and Parezo, 1992). The National Anthropological Archives house the field notes of many anthropologists, as well as over 8 million feet of film and over a million photographs, collected over the course of 150 years (Leopold, 2008, 2010; Smithsonian Institution, 2016). And the Council for the Preservation of Anthropological Records (copar.org) lists over 70 major ethnographic archives all around the world.

The UK’s Data Service (<https://www.ukdataservice.ac.uk/>) has developed a web-based data sharing mechanism that includes survey data, census data, and qualitative and mixed methods data. Similar archives are being developed in other countries across Europe (Rasmussen, 2010/2011). The Inter-University Consortium for Political and Social Research (ICPSR) began in 1962 with 22 universities as members. Today, the ICPSR has over 750 members and curates over 250,000 files, including files of qualitative data. ICPSR “accepts replication datasets for researchers who need to publish their raw data in relation to a journal article” (<https://www.icpsr.umich.edu/icpsrweb/>).

Tsai and colleagues are right to highlight the risks in sharing primary qualitative data, particularly when sensitive health and medical data are involved. The risks of harm to respondents (Grinyer, 2009), to research communities (Karnieli-Miller et al., 2009), and even to the researcher (Irwin, 2013) are not trivial. Still, a major lesson from existing efforts to archive and share qualitative data is that the risk to research participants is not all that different from the risk in sharing quantitative data—and that the risk can be managed. For example, some of ICPSR’s data sets, like one on “Wife Abuse among Vietnamese Immigrants to the U.S.” (ICPSR 35247) say explicitly that “one or more files in this study are not available for download due to special restrictions.” In those cases, there is a link to apply online for access to the data. This

approach is how sensitive data are protected in the Tsimane’ Amazonian Panel Study (TAPS <http://heller.brandeis.edu/sustainable-international-development/tsimane/>) and the now 50-year and still going Wisconsin Longitudinal Study (<http://www.ssc.wisc.edu/wlsresearch/data/>).

Tsai and colleagues recommend that any journal requiring the release of qualitative data should institute an appeal process; we emphatically support this. Beyond this, however, we also emphasize the need for a “stakeholder approach” (Neale and Bishop, 2012) in discussing with research communities the nature and appropriateness of data sharing. Such a process should recognize that some communities may have an urgent need to raise awareness of the problems they face and will desire speedy and comprehensive data sharing to address this. We are fortunate to have an extensive literature on the ethics of sharing qualitative data to guide us (Bishop, 2005, 2009; Miller et al., 2012; Parry and Mauthner, 2004; Punzalan and Caswell, 2016).

References

- Bernard, H.R., Wutich, A., Ryan, G.W., 2016. *Analyzing Qualitative Data: Systematic Approaches*. SAGE Publications.
- Bishop, L., 2005. Protecting respondents and enabling data sharing: reply to Parry and Mauthner. *Sociology* 39 (2), 333–336.
- Bishop, L., 2009. Ethical sharing and reuse of qualitative data. *Aust. J. Soc. Issues* 44 (3), 255.
- Charmaz, K., 2014. *Constructing Grounded Theory*. SAGE Publications.
- Cliggett, L., 2013. Qualitative data archiving in the digital age: strategies for data preservation and sharing. *Qual. Rep.* 18 (24), 1–11.
- Grinyer, A., 2009. The ethics of the secondary analysis and further use of qualitative data. *Soc. Res. Update* 56 (4).
- Hammersley, M., 1997. Qualitative data archiving: some reflections on its prospects and problems. *Sociology* 31 (1), 131–142.
- Handwerker, W.P., Borgatti, S.P., 2014. Reasoning with Numbers. *Handbook of Methods in Cultural Anthropology*. Rowman & Littlefield, pp. 519–532.
- Ignatow, G., Mihalcea, R., 2016. *Text Mining*. SAGE Publications.
- Irwin, S., 2013. Qualitative secondary data analysis: ethics, epistemology and context. *Prog. Dev. Stud.* 13 (4), 295–306.
- Karnieli-Miller, O., Strier, R., Pessach, L., 2009. Power relations in qualitative research. *Qual. Health Res.* 19 (2), 279–289.
- Kemper, R.V., Royce, A.P. (Eds.), 2002. *Chronicling Cultures: Long-term Field Research in Anthropology*. Rowman Altamira.
- Krippendorff, K., 2012. *Content Analysis*. SAGE Publications.
- Leopold, R., 2008. *The Second Life of Ethnographic Fieldnotes*, vol. 32. *Ateliers d’anthropologie*.
- Leopold, R., 2010. Ethics of Exhibiting Culturally Sensitive Materials Online. http://anthropology.si.edu/naa/home/culturally_sensitive.html (accessed 08.09.16.).
- Miller, T., Birch, M., Mauthner, M., Jessop, J. (Eds.), 2012. *Ethics in Qualitative Research*. SAGE Publications.
- Neale, B., Bishop, L., 2012. *The Ethics of Archiving and Re-using Qualitative Longitudinal Data: a Stakeholder Approach*. *Timescapes Methods Guide Series*, Guide No. 18.
- Parry, O., Mauthner, N.S., 2004. Whose data are they anyway? Practical, legal and ethical issues in archiving qualitative research data. *Sociology* 38 (1), 139–152.
- Pelto, P.J., 2015. What is so new about mixed methods? *Qual. Health Res.* 25 (6), 734–745.
- Punzalan, R.L., Caswell, M., 2016. Critical directions for archival approaches to social justice. *Libr. Q.* 86 (1), 25–42.
- Rasmussen, K.B. (Ed.), 2010/2011. *Qualitative and Qualitative Longitudinal Resources in Europe. Mapping the Field and Exploring Strategies for Development*. *IASSIST Quarterly*. Vol. 34 (3 & 4), Vol. 35 (1 & 2).
- Silverman, S., Parezo, N.J. (Eds.), 1992. *Preserving the Anthropological Record*. Wenner-Gren Foundation for Anthropological Research.
- Smithsonian Institution, 2016. National Anthropological Archives. <http://anthropology.si.edu/naa/home/naahome.html> (accessed 08.09.16.).