

“Making a square fit into a circle”: Researchers’ experiences reusing qualitative data

Ayoung Yoon

University of North Carolina at Chapel Hill
216 Lenoir Dr. CB #3360
100 Manning Hall, Chapel Hill, NC 27599
ayyoon@email.unc.edu

ABSTRACT

Although researchers have shown their skepticism of qualitative data reuse due to the epistemological issues, interest in qualitative data reuse has continuously grown. Discussions regarding qualitative data sharing and reuse have also been very active, especially in Europe and Australia. Compared with efforts in other countries, the discussions do not seem to be as prominent in the United States, in spite of a long history of data-depositing and -curating practices and researchers’ reusing qualitative data in some disciplines. This research aims to explore qualitative researchers’ experiences reusing data in the field of social science in US, which have not been empirically addressed yet. The preliminary results from the in-depth interviews with qualitative researchers who have used secondary data are presented, along with the barriers or hindrances to reusing qualitative data and the keys to successful data reuse.

Keywords

Data reuse, qualitative data, data curation.

INTRODUCTION

In recent years, data sharing and data reuse in scientific research have been discussed with increasing frequency. This has occurred due to the revolution in the field of science known as data-intensive research and the growth of data in “big science.” While there is a well-established tradition of quantitative data reuse in social science, researchers have shown their skepticism of qualitative data reuse and do not work in a research-sharing culture that actively encourages data reuse among qualitative researchers and students (Corti, 2007). However, interest in

qualitative data reuse has continuously grown, and situations have changed markedly, particularly in Europe and Australia (Bishop, 2009; Corti et al., 2005). For instance, in the United Kingdom (UK), this burgeoning interest in qualitative data reuse resulted in the establishment of the Qualitative Data Archival Resource Centre (QUALIDATA, ESDS Qualidata since 2003, and now merged into UK Data Archive) at the University of Essex in 1994, as well as the publication of the Economic and Social Research Council (ESRC) Datasets Policy (1996), which requires a deposit of data as a condition of ESRC research funding (Corti & Backhouse, 2005). Compared with efforts in other countries, the discussions do not seem to be as prominent in the United States (US), in spite of a long history of data-depositing and -curating practices and researchers’ reusing qualitative data in some disciplines.

This research aims to explore qualitative researchers’ experiences reusing data in the field of social science in US, which have not been empirically addressed yet. The preliminary results from the in-depth interviews with qualitative researchers who have used secondary data are presented, along with the barriers or hindrances to reusing qualitative data and the keys to successful data reuse.

LITERATURE REVIEW

The benefits of data sharing and reuse, including qualitative data, have already been discussed by a number of researchers, for example, making “unmined” data available, validating existing results, generating new findings built on the work of others, avoiding duplications, and reducing burdens of research participants (Bishop, 2009; Borgman, 2010; Fry et al. 2009; Hey & Trefethen, 2003; Hey, Tansley, & Tolle, 2009). In addition, supporters of qualitative data reuse also argue that original researchers often do not have enough time to fully explore their data given the richness and the amount of qualitative data (Fielding & Fielding, 2000).

Despite the general consensus on the benefits of qualitative data reuse, researchers who expressed skepticism have discussed epistemological issues in reusing qualitative data, which is related to the nature of qualitative research and data. In qualitative research, researchers themselves become

{This is the space reserved for copyright notices.}

ASIST 2014, November 1-4, 2014, Seattle, WA, USA.

[Author Retains Copyright. Insert personal or institutional copyright notice here.]

the research instrument, and all results are filtered through their perception and understanding of the social situations (Dale et al., 1988). Because qualitative research is “insider activity” (Mauthner et al., 1998) and “knowledge about qualitative data is highly contextual and experience-dependent” (Niu & Hedstrom, 2008), qualitative research involves interpretation and subjectivities not concrete (or transportable) enough to be documented and reused (Broom et al., 2009). Ethical considerations are another concern for researchers, because consent forms used in qualitative research usually allow only the use of primary research, not secondary research (Stebe et al., 2010). Confidentiality and anonymity of participants are other major concerns (Bishop, 2009). This is related to the challenges of data reuse, as in some cases, a thorough anonymization of participants make data reusers difficult to understand the data (Carusi & Jirotko, 2009).

The major challenges of qualitative data reuse, i.e., whether all necessary contextual information can be documented and transferred, are often related to this distinctive nature of qualitative data and qualitative research. Parry & Mauthner (2005) said contextual information is not automatically accessible to data reusers, and recovery of context can only be partial. While documentations are one way to help retrieve contextual information about data for reuse (Parry & Mauthner, 2005), they are rarely sufficient (McCall & Appelbaum, 1991). In addition, access to the data may be limited due to confidentiality issues, and approval of the data collector may be required (McCall & Appelbaum, 1991).

Compared to practices in other countries like the UK, where qualitative data archives have established, qualitative data reuse is not yet a common research practice in the US, and there is no methodological guidance or publicly available catalogued qualitative research data (Corti, 2005). However, some qualitative researchers in some disciplines (e.g., education, sociology, public health, etc.) have been conducting secondary analysis. Investigating their experience is important to understand the current situation of qualitative data reuse and to promote data reuse culture by supporting researchers’ needs.

RESEARCH METHOD

In order to identify individuals who have qualitative data reuse experience, this study used data citation tracking from major databases (i.e., Sage journals, ERIC, etc.). Tracking data citation is a challenging process and may have limitations, but it is still the most effective way to identify data reusers. Data reusers were identified from a keyword search from each database, using the search term “secondary data” or “secondary analysis.” These searches were limited to journal publications and conference proceedings published in North America since 2004, following reverse chronological order (i.e., starting from the most recent ones). In case the authors’ contact information was not provided in the publications, an additional Google

search was conducted to collect the authors’ contact information (email addresses). For articles written by multiple authors, either the information of the corresponding author or the first author was collected. If the first author did not work directly with datasets, I asked the first author who would be the appropriate person for this research and contacted that person. A total of 40 researchers were identified and invited to this study, and 13 researchers responded (response rate 32.5%). 8 were interviewed at the time of this article’s writing, and more interviews were scheduled.

PARTICIPANTS’ CHARACTERISTICS

All participants were female, and all of them were researchers either in academic or research organizations (2 were associate or full processors; 4 were assistant professors; and 2 were researchers and consultants). Their ages varied from 20s to 50s. Their experience in qualitative research varied from 3 to 25 years, but all of them had only 1 or 2 experiences using qualitative data for secondary analysis.

INITIAL DATA ANALYSIS RESULTS

Initial results reported below are based on narrative interviews with 8 participants.

1. Prior connection to the original data and original investigators (who collected and owned data) was the condition for reuse.
 - a. All participating researchers got the data from an advisor or an advisor’s co-worker, from their colleagues, or the organizations that they had worked with previously.
2. Choosing data from someone they already know is an important part of their trust judgment of data, because qualitative data is the end product of original investigators’ worldview, research philosophy, and experiences.
 - a. “I think I would only do it with data sets from researchers that I actually knew, and knew well, and that they were aligned with my philosophy.” (QP06)
 - b. “It’s important to look at [...] is this someone who is strong in methods and who was experienced [in qualitative research]. Just because you can’t speak to that yourself, because you haven’t gone through the process.” (QP03)
3. The researchers usually relied on original investigators during the process of understanding data for reuse because:
 - a. The researchers believed that is the way to “really understand the original context of the data collection, and then [original researchers] could answer questions for me.” (QP06). She also said, “One of the key things that was important to me is that I have access to people

- who had been involved in that original study;” and
- b. The level of documentation was varied. Usually the researchers received project descriptions and interview transcripts or observation notes, but other information that could help to enhance data reusers’ understanding of data (e.g., PIs’ information, interviewers’ information, original investigators’ memos and notes, etc.) was provided in a few cases. Access to audiotapes was also limited due to confidentiality issues, although the researchers often found it useful to listen the actual audio.
4. Overall, the researchers felt the data reuse experience is “making a square fit into a circle in some ways” (QP01), as the data was not created for their intended research.
 - a. In one case, the research found that the original investigators and original study participants did not agree with her interpretation: “So I had some people that disagreed with my interpretation of what was going on and they made some demands that I could not meet. They wanted me to reinterpret a document entirely and that just wasn’t possible.” (QP05)
 5. Major challenges that the researchers found during their data reuse experiences included:
 - a. Missing information and context: typos in the transcripts and notes, abbreviations, study participants’ tones of voice, reasons why a certain probe was made during the interviews, etc.
 - b. Restrictions on data access
 - c. Different practice in different fields:
 - i. When researchers use data across disciplines, disciplinary differences present a challenge. For instance, one researcher in social science used data collected in the field of humanities, which uses the real names of participants, whereas social scientists emphasize privacy and confidentiality.
 - ii. “It wasn’t just the challenge of secondary data analysis, but it was that the ways... It was a disciplinary difference in how the data was collected and what was included in the set.” (QP05)
 - d. Publishing
 - i. Publishing was a major challenge to the qualitative researchers and they said there was not many room for publish secondary data analysis.
 - ii. “Quantitative secondary analysis happens all the time, but qualitative secondary analysis is very unique and I think it takes

a lot of over explaining. We ended up publishing the pieces, you know, but it took a while in order to pitch it in a way that it would be deemed acceptable.” (QP08)

CONCLUSION

The initial data analysis results present qualitative data reusers’ strong preference and needs for personal interaction with original investigators, mainly because of the nature of qualitative data, and in order to get all necessary contextual information. As the data were not collected for the data reusers’ research purposes, they ran into different challenges. Some could be solved by interacting with the original investigators, but some needed to be addressed in a bigger frame, as they were related to academic culture and practices.

These initial analysis results shed light on qualitative researchers’ lived experiences in data reuse, which have not been empirically investigated before. More in-depth analysis will be conducted as more data are collected. The final product of this research will provide implications for qualitative data curation.

REFERENCES

- Bishop, L. (2009). Ethical sharing and reuse of qualitative data. *Australian Journal of Social Issues*, 44(3), 1-15.
- Borgman, C. L. (2010). Research Data: Who will share what, with whom, when, and why? Presented at the China-North America Library Conference, Beijing. Retrieved from <http://works.bepress.com/borgman/238>
- Broom, A., Cheshire, L., & Emmison, M. (2009). Qualitative Researchers’ Understandings of Their Practice and the Implications for Data Archiving and Sharing. *Sociology*, 43(6), 1163–1180. doi:10.1177/0038038509345704
- Carusi, A., & Jirotko, M. (2009). From data archive to ethical labyrinth. *Qualitative Research*, 9(3), 285–298. doi:10.1177/1468794109105032
- Corti, L. (2005). Qualitative Archiving and Data Sharing: Extending the Reach and Impact of Qualitative Data. *IASSIST Quarterly*, 2005(Fall), 8-12.
- Corti, L. (2007). Re-using archived qualitative data – where, how, why? *Archival Science*, 7(1), 37–54. doi:10.1007/s10502-006-9038-y
- Corti, L., & Backhouse, G. (2005). Acquiring Qualitative Data for Secondary Analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 6(2). Retrieved from <http://www.qualitative-research.net/index.php/fqs/article/view/459>
- Corti, L., Witzel, A. & Bishop, L. (eds.) (2005). Secondary analysis of qualitative data. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 6(1). Retrieved from <http://www.qualitative-research.net/fqs/fqs-e/inhalt1-05-e.htm>.

- Dale, A., Arber, S. L., & Procter, M. (1988). *Doing Secondary Analysis*. London: Allen & Unwin Ltd. <http://onlinelibrary.wiley.com/doi/10.1002/0470867167.ch36/summary>
- Fielding, N.G. & Fielding, J. L. (2000). Resistance and Adaptation to Criminal Identity: Using Secondary Analysis to Evaluate Classic Studies of Crime and Deviance. *Sociology*, 34(4), 671–89. doi: 10.1177/S0038038500000419
- Fry, J., Lockyer, S., Oppenheim, C., Houghton, J. & Rasmussen, B. (2009). *Identifying Benefits arising from the Curation and Open Sharing of Research Data*. UK Higher Education and Research Institutes, November. Retrieved from <http://ie-repository.jisc.ac.uk/279/>.
- Hey, T., Tansley, S., & Tolle, K. (2009). *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Retrieved from <http://research.microsoft.com/en-us/collaboration/fourthparadigm/contents.aspx>.
- Hey, T., & Trefethen, A. (2003). The Data Deluge: An e-Science Perspective. In F. Berman, G. Fox, & T. Hey (Eds.), *Grid Computing* (pp. 809–824). John Wiley & Sons, Ltd. Retrieved from
- Mauthner, N. S., Parry, O., & Backett-Milburn, K. (1998). The Data are Out there, or are They? Implications for Archiving and Revisiting Qualitative Data. *Sociology*, 32(4), 733–745. doi:10.1177/0038038598032004006
- McCall, R. B., & Appelbaum, M. I. (1991). Some issues of conducting secondary analyses. *Developmental Psychology*, 27(6), 911–917. doi:10.1037/0012-1649.27.6.911
- Niu, J., & Hedstrom, M. (2008). Documentation Evaluation Model for Social Science Data. In *Proceedings of the Annual Meeting of the American Society for Information Science & Technology (ASIS&T)*. Columbus, OH.
- Parry, O., & Mauthner, N. (2005). Back to Basics: Who Re-uses Qualitative Data and Why? *Sociology*, 39(2), 337–342. doi:10.1177/0038038505050543
- Stebe, J., Hudales, J., & Kragelj, B. (2010). Archiving and Re-using Qualitative and Qualitative Longitudinal Data in Slovenia. *IASSIST Quarterly*, 2010(Spring/Summer/Fall/Winter), 50-59.

The columns on the last page should be of approximately equal length.