University of Kentucky
UKnowledge

Information Science Faculty Publications | Information Science

9-2016

Advancing Research for Library and Information Science with Qualitative Secondary Analysis

Jenny Bossaller
University of Missouri

C. Sean Burns
University of Kentucky, sean.burns@uky.edu

Amy VanScoy
University at Buffalo

Follow this and additional works at: https://uknowledge.uky.edu/slis_facpub

 Part of the Library and Information Science Commons

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Repository Citation

Bossaller, Jenny; Burns, C. Sean; and VanScoy, Amy, "Advancing Research for Library and Information Science with Qualitative Secondary Analysis" (2016). *Information Science Faculty Publications*. 55.
https://uknowledge.uky.edu/slis_facpub/55

This Article is brought to you for free and open access by the Information Science at UKnowledge. It has been accepted for inclusion in Information Science Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

Advancing Research for Library and Information Science with Qualitative Secondary Analysis

Notes/Citation Information

Published in *Qualitative and Quantitative Methods in Libraries*, v. 5, p. 593-600.

© ISAST

Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the articles in *Qualitative and Quantitative Methods in Libraries* without asking prior permission from the publisher or the author.

This article is available at UKnowledge: https://uknowledge.uky.edu/slis_facpub/55

Advancing Research for Library and Information Science with Qualitative Secondary Analysis

Jenny Bossaller[1], C. Sean Burns[2], and Amy VanScoy[3]

[1]University of Missouri
[2]University of Kentucky
[3]University at Buffalo

Abstract: This paper presents the results of a study utilizing a seldom-used method in Library and Information Science (LIS): Qualitative Secondary Analysis. The data is drawn from two phenomenological studies about experiences of Reference and Information Services (RIS) librarians. We discuss how we repurposed the interview data in this study, and also the strengths, weaknesses, and wider applications of the method across LIS.

Keywords: Qualitative Methods, Qualitative Secondary Analysis, Reference and Information Services, sociology of work

1. Introduction

We would like to start with a problem: qualitative data collection is time-consuming and sometimes messy. Because it takes so much time to collect and process, the data is valuable (Glaser, 1963). Qualitative data often provides rich, interpretive, and descriptive information about people and their lives (Creswell, 2007). If there is a way to re-use data from qualitative research collection, should you? This paper describes how and when a researcher might want to reuse qualitative data. The paper builds on Johnston's (2014) QQML presentation, but is grounded in a recent study, which will be used to illustrate qualitative data re-use. Recommendations, based on experience and the literature, round out a discussion of the study.

Before we delve into the particulars of qualitative secondary analysis (QSA), though, we will establish that secondary analysis is more common with quantitative data. Big data problems, in particular, might offer ways to re-analyze and combine expansive data sets. Here, we focus on the United States because the authors are familiar with U.S. sets, although there are likely similar data sets available in many other countries. In the U.S., social science (e.g., American Community Survey) and educational data sets (e.g., National

Received: 18.3.2016 Accepted: 21.5.2016 ISSN 2241-1925
© ISAST

Assessment of Educational Progress), Health data (e.g., Medicare Part C and D Performance Data) are available for download from data.gov. There are other sources for large data sets from nonprofits, such as the Pew Research Center. R-dir (https://r-dir.com/reference/datasets.html) has a list of free datasets available for download. The possibilities for mining huge sets of data collected on personal devices (e.g., Fitbits and e-readers) looms large (presenting ethical challenges, as well). Such data sets can be combined and mined to produce new findings, revealing meaningful intersections between the sets. Smaller, institutional data sets, likewise, can be re-analyzed and combined, or layered within larger sets. This is quantitative secondary analysis.

Secondary analysis, then, is simply a way to explore data that was produced for another purpose. The data from a study that uses qualitative methods, such as interviews or focus groups, might produce extra evidence that was not fully explored in the original study, as well. The researchers might have purposefully collected extra data (because if you are already collecting data, why not throw in a few more questions?) or have done so inadvertently (because during unstructured or semi-structured interviews or focus groups, participants might take the conversation in directions that were not originally planned). The data, therefore, might offer insight into a new problem. The secondary analysis begins when a new question is applied to existing and possibly shared data sets. This re-use of data is convenient, but there are some precautions that researchers should take when conducting secondary analyses. Here we outline distinct benefits, drawbacks, and possible problems with any secondary analysis.

2. Secondary Analysis in LIS

A number of studies within LIS have re-used data; here, we focus on some that illustrate variety in LIS, emphasizing reasons for data re-use.

Undergraduate Use of the Library and Student Success: Whitmire (2001) combined data from components of the National Study of Student Learning (NSSL), which looked for the impact of various experiences on student learning and development, attitudes, and persistence. There were three instruments within the survey that looked at undergraduate college activities and experiences, learning outcomes, and background characteristics. One of the instruments, the Collegiate Assessment of Academic Proficiency (CAAP), was designed to measure critical thinking ability. By analyzing the data sets using several regressions, Whitmire was able to correlate background characteristics and college experiences that determined library use throughout students' undergraduate careers, including correlations with critical thinking. For instance, certain activities (e.g., being enrolled in classes that require term papers in the Junior year, or being an 'engaged writer') had a high correlation with library use; other activities (e.g., having an off-campus job, or taking more classes in the natural sciences or mathematics) had a negative or neutral correlation on library use.

Searching for Health Information: Kwon and Kim (2009) looked at use of libraries by patients with cancer. Cancer patients are potentially overwhelmed with information about their disease from the Internet. Libraries have the potential to help people find the most relevant and best information for health. Kwon and Kim used data from the Health Information National Trends Survey (HINTS), published by the National Cancer Institute (NCI) to assess how people sought information about cancer. As the authors point out, the NCI has access to a much wider swath of data about the general population than libraries, which mitigates a persistent problem that much library research is "largely restricted to the behaviors of either library patrons or people in a geographically limited location" (p. 199). The outsiders' perspective offered by a national survey can be helpful, but it can also be of somewhat limited value; that is, there might be a lot of 'noise' amidst the 'signal.' Again, supplementing the findings with qualitative data, or data that directly answers the question, would provide additional, valuable insights into the problem.

E-books versus Print: Libraries are investing huge sums of money into e-book collections. Zhang and Kudva (2014) asked: *Are e-books are replacing print*? The authors re-analyzed data from Pew Research Center's Internet & American Life Project's Reading Habits Survey by downloading the source survey data file from Pew. By weighting variables to ensure "a balanced and representative sample of the national population parameters for gender, age, education, race and ethnicity, U.S. Census region, population density, and telephone usage" (p. 1699), they were able to identify factors that contribute to e-book adoption in readers in the United States for various reader groups. For instance, they found that people who use the Internet more read more e-books; that in suburban communities, there was a larger proportion of readers who read both print and e-books; and that younger people were more likely to be e-book only readers. They concluded that e-book use is contextual and situational, and that it is unlikely that print will become obsolete in the near future.

Academic Library Value: Crawford (2015) set out to measure the value of academic libraries by analyzing the relationships between several readily available data sources on academic institutions and library expenditures, library use, and retention and graduation rates. A number of different tests were used to find correlations between the data; significantly, for libraries, he found that library expenses per FTE had "the highest correlation with…graduation rates and the second highest with the retention rate" (p 53). The relationship between library use and student success, though, was not pronounced (p. 55). While he could demonstrate some correlations, he also said that exclusive use of quantitative data was a limitation, and that "more qualitative measures would give depth to the findings" (p. 56). The study, thus, illustrated both a strength and a weakness in secondary quantitative analysis of large data sets.

Generally speaking, Secondary Qualitative Analysis does not seem to be as commonly used; we were, in fact, unable to find any research similar to ours

(though that does not mean that it does not exist). Practically speaking, it is difficult to search for research by method. However, there are some good reasons that secondary analysis of qualitative data is not commonly completed, especially since the data might be sensitive or include personally identifiable information. Currently, few data sets are publicly available for download (though Elman, Kapiszewski, and Vinuela, 2010, discuss some options for archiving as well as some sources for sets of qualitative data).

3. This study

The study described in this paper reused data from two phenomenological analyses of Reference and Information Services (RIS) librarians' work life (Burns and Bossaller, 2012; VanScoy, 2013). The studies had some similarities in both method and findings. Through discussions spurred by an intriguing yet vexing conference call for proposals (Canadian Association of Information Science 2015, "Time is of the Essence: Organizing People, Data, Information and Knowledge as Memory and Participation"), we realized that the concept of 'time' was present in our recently completed studies and suspected that analyzing the data for references to time might possibly offer an important, yet unexplored, framework for thinking about work life. While analyzing the data, we found that combining and re-examining the data resulted in findings that were conceptually and theoretically strong and that this was because the subject had come naturally from the two studies. That is, even though time was not directly investigated in the original studies, it was present in very important, if tacit ways, in the interviews.

The two original phenomenological studies sought to uncover how academic librarians experienced their lives through or during various aspects of their work. In one of the studies, Burns and Bossaller (2012) conducted nine in-depth interviews with academic reference librarians on the subject of information and communication technologies and librarianship. From seven emergent themes, they described synchronous and asynchronous communication patterns that disrupt work flow and librarians' ability to assist users. In the other study, VanScoy (2013) used interpretive phenomenological analysis (IPA) (VanScoy, 2015) to study how eight academic reference librarians make sense of their experiences doing reference and information services work. The themes that were identified in that study were: the importance of the user, variety and uncertainty, fully engaged practice, emotional connection, and sense of self as reference professional.

Each dataset was re-analyzed by one of the original researchers for words, phrases, or concepts relating to time. Interview participants from 'round one' of data collection certainly understood that excerpts from their interview data would be shared, but there was no discussion during the consent process about the possibility of other researchers gaining access to the interview transcriptions. Therefore, the nature of the original studies was unchanged (explorations of the experience of RIS work), but the researchers felt that it was important from an

ethical standpoint to maintain the privacy of the original transcripts and to share only excerpts related to time. This methodological choice ensured we maintained context (since we were re-analyzing our own data sets) while being sensitive about maintaining confidentiality of the participants' data.

In order to form themes, data that related to time were excerpted, shared, and analyzed thematically by all of the authors, and then were grouped and regrouped to develop the themes. We found three themes that describe time in relation to RIS work: *perceiving time as discrete and continuous*, the *consequences of time as a commodity*, and *framing narratives and identities*. The first theme describes the perception of time and how the professionals experience it as both continuous and bounded moments. The second theme focuses on how the common treatment of time as a commodity influences how we measure and value the time we spend on tasks, and how that affects our perception of whether time at work is positive or negative. The third theme was uncovered as we realized that participants related experience through stories that were framed through the lens of time. This contributed to their sense of self and personal identity by providing a way to generate meanings of their professional lives (e.g., how much time did they spend doing reference work versus administrative work?).

Qualitative Secondary Analysis proved to be a useful method for this study in that it provided a method for a rich, new exploration of two existing data sets. There were some problems that the researchers had to negotiate during the course of re-analysis, such as sharing data and creating meaningful codes with the diverse data sets. Negotiations about the method opened the researchers to new ways of thinking about qualitative data, and how it might be shared between researchers to discover the answers to arising practical or theoretical problems. We will turn to some of these lessons, answering to existing cautionary tales with secondary analysis of qualitative data.

4. Strengths and Weaknesses of the method

Glass (1976), writing about primary, secondary, and meta-analysis of research, said, "extracting knowledge from accumulated studies is a complex and important methodological problem" (p. 8). Qualitative research processes (collection, transcription, etc.) are time-intensive and expensive (Glaser, 1963). Researchers are sitting on piles of data, and re-using them can be quite meaningful, while simultaneously saving time and money. Secondary analysis, though, has been criticized because drawing conclusions from data collected for a different purpose might introduce problems with validity (Boslaugh, 2007).

One impediment to secondary analysis of qualitative data might be the intimacy of the data: qualitative data is often composed of interviews and focus groups. Gathering this data is naturally more personal, and relies on forming a relationship with the subject. It requires a certain degree of trust. "Most concern revolves around issues of harm, consent, deception, privacy, and confidentiality

of data" (Punch, 2005, p. 168). If the data is re-used for a purpose beyond that approved by the initial interviews, might it be a breach of trust? Punch points out that in, for instance, feminist research, there is a "standpoint epistemology" that "inhibits deception of the research 'subjects'; "the personal is related to the ethical, the moral, and the political standpoint" (p. 169). Action research demands that 'subjects' are seen as partners in the research process in a "constructivist paradigm that is based on avoidance of harm, fully informed consent, and the need for privacy and confidentiality" (p. 169 – 170). The subject is, therefore, not to be treated as a subject, but as an equal, which might imply that the data should be kept in confidence and not used for a purpose other than its original intentions.

Qualitative data collection techniques are time-consuming and potentially stressful for the participants themselves, as well. Beyond this, certain populations, especially minority populations, may be turned to frequently for research, adding yet another dimension to what is referred to as "invisible work" (Evans, 2007; Hart et. al., 2009). Participants may appreciate re-use of their data as a means of more efficiently making their voices heard.

Perhaps expense is one reason that qualitative data are used much less frequently in LIS research than quantitative methods; another reason, though, might be the feeling that quantitative data is more generalizable or valid. Although exact figures are difficult to determine, recent studies suggest that qualitative research use in LIS is still low: VanScoy and Fontana (in press) found that only 12% of the empirical research on RIS, a subset of LIS research, is qualitative; Hider and Pymm (2008) found that 19% of research studies in LIS used qualitative analysis. If we can re-use qualitative data, should we not? We assert that, yes, we should, but with caution. Advanced planning and careful documentation of research protocols (on the front end) is necessary, might make re-use much more manageable, and also increase validity of the secondary analysis.

Johnson (2014) pointed out that within LIS, secondary data-analysis is an under-used research technique: "In a time where the large amounts of data are being collected, compiled, and archived by researchers all over the world are now more easily accessible, the time has definitely come for secondary data analysis as a viable method for LIS research" (p. 626). Johnson emphasizes, though, that the researcher doing the secondary analysis must obtain all of the research processes and protocols used in the primary research. The secondary research should only make use of primary data if the creators "have a reputation for excellence in research integrity" (p. 623).

We assert that qualitative secondary analysis might be examined in a less traditionally systematic way; we coded the excerpts anew in order to re-examine interview data. However, we limited our re-analysis to our own data. With interview data, especially, the context is meaningful and influential on the data

that is gathered, so context cannot be dismissed. Because we were re-analyzing our own data, we were aware of the context and could sift through statements that were not meaningful.

Our study demonstrates the usefulness of secondary analysis for understanding and improving library practice. It capitalized on the similarities of the original two studies, creating a larger and richer dataset for the focus of the secondary analysis. It sought to uncover the meaning of a specific aspect of work: the way that RIS librarians experience time at work. The results of the study have both practical and theoretical implications.

5. Recommendations and Conclusions

Perhaps the most problematic aspect of our research was the researchers' ethical concerns and subsequent decision not to share original interview transcripts. Elman, Kapiszewski, and Vinuela (2010) propose a method for archival storage of qualitative data for widespread re-use. Archiving interview, focus group, and ethnographic data would allow more comparative research, as well as the "vertical integration of primary data [and] secondary analysis" (p. 24) leading to greater scholarly output. However, they also discuss some of the problems that we averted by not sharing data – namely, that interview participants (and participants in other forms of field research) are often collected under the premise that their data will remain anonymous. They propose "washing" the data in order to protect the participants, which might be impossible under some conditions. However, they also say that "disciplinary norms will need to change if sharing and reusing qualitative data – and producing scholarship relying…on secondary data analysis – are to become accepted practices" (p. 25).

During the research design process, researchers should consider the possibility of future re-use of data and design study procedures with this re-use in mind. This may include gaining specific consent from participants for data re-use and plans for archiving the data.

Scholars in the field should discuss what it means to "anonymize" qualitative data. Removing a participant's name, position title, and organization from the data will prevent, to some degree, those with access to the data from identifying a participant (though all precautions should be taken to remove any identifying information). But there may be words or actions that the participant feels are too intimate to share with other researchers, such as confessions of poor choices or tearful revelations. Researchers may want to have participants review the data and flag sections that they feel are too personal for re-use.

Within LIS, there is certainly room for expanding the use of Secondary Qualitative Analysis. We feel that this research approach is an especially important methodological advancement for library and information science studies and will support the future of librarianship by adding a greater understanding of both librarians' and library patrons' experiences.

References

Boslaugh, S., (2007). *Secondary Analysis for Public Health: A Practical Guide*, Cambridge, New York, NY,.

Burns, C.S. and Bossaller, J., (2012). Communication overload: a phenomenological inquiry into academic reference librarianship, *Journal of Documentation*, Vol. 68, No. 5, 597-617.

Creswell, J. W. (2007). *Qualitative Inquiry & Research Design: Choosing Among Five Approaches*, Sage Publications, Thousand Oaks, CA.

Crawford, G. A., (2015). The academic library and student retention and graduation: An exploratory study. *portal: Libraries & The Academy*, Vol. 15, No. 1, 41-57.

Elman, C., Kapiszewski, D., & Vinuela, L., (2010). Qualitative data archiving: Rewards and challenges. *PS: Political Science and Politics*, Vol. 43, No. 1, 23–27.

Evans, S. Y., (2007). Women of color in American higher education, *Thought & Action*, 131 – 138.

Glaser, B. G., (1963). Retreading research materials: The use of secondary analysis by the independent researcher, *The American Behavioral Scientist*, Vol. 6, No. 10, 11-14.

Glass, G. V., (1976). Primary, secondary, and meta-analysis of research. *Educational researcher*, Vol. 5, No. 10, 3-8.

Hart J.E., Grogan, M.A., Litt, J.A., and Worthington, R.O., (2009). Institutional diversity work as intellectual work at the University of Missouri-Columbia. In W.R. Brown-Glaude (Ed.), *Doing diversity in higher education: Faculty leaders share challenges and strategies*, Rutgers University Press, New Brunswick, 61-82.

Hider, P., and Pymm, B., (2008). Empirical research methods reported in high-profile LIS journal literature, *Library & Information Science Research*, Vol. 30, No. 2, 108-114.

Irwin, S. and Winterton, M., (2011). Debates in Qualitative Secondary Analysis: Critical reflections. Timescapes, *Working Paper Series* No. 4. Research Council UK: Economic & Social Research Council.

Johnston, M. P., (2014). Secondary Data Analysis: a method of which the time has come. *Qualitative and Quantitative Methods in Libraries (QQML)*, Vol. 3, 619-626.

Kwon, N., and Kim, K. ,(2009). Who goes to a library for cancer information in the e-health era? A secondary data analysis of the Health Information National Trends Survey (HINTS), *Library & Information Science Research*, Vol.31, No. 3, 192- 200.

Punch, K.F., (2005). *Introduction to Social Research – Quantitative & Qualitative Approaches*, Sage, London.

Vagle, M., (2014). *Crafting Phenomenological Research*, Left Coast Press, Walnut Creek, CA.

VanScoy, A. (2013), Fully engaged practice and emotional connection: aspects of the practitioner perspective of reference and information service, *Library & Information Science Research*, Vol. 35 No. 4, pp. 272-278.

VanScoy, A. and Evenstad, S.B., (2015), Interpretive phenomenological analysis for LIS research, *Journal of Documentation*, Vol. 71 No. 2, pp. 338-357.

VanScoy, A. and Fontana, C. (in press). How reference and information service is studied: Research approaches and methods. *Library & Information Science Research*.

Whitmire, E., (2001). The relationship between undergraduates' background characteristics and college experiences and their academic library use, *College & Research Libraries*, Vol.62, No. 6, 528-40.

Zhang, Y., and Kudva, S., (2014). E-books versus print books: Readers' choices and preferences across contexts, *Journal of The Association For Information Science & Technology*, Vol.65, No. 8, 1695-1706.