Qualitative Methods for CSCW: Challenges and Opportunities

Casey Fiesler
University of Colorado Boulder
Boulder, CO
casey.fiesler@colorado.edu

Jed R. Brubaker
University of Colorado Boulder
Boulder, CO
jed.brubaker@colorado.edu

Andrea Forte
Drexel University
Philadelphia, PA
aforte@drexel.edu

Shion Guha
Marquette University
Milwaukee, WI
shion.guha@marquette.edu

Nora McDonald
Drexel University
Philadelphia, PA
norakmcdonald@gmail.com

Michael Muller
IBM Research
Cambridge, MA
michael_muller@us.ibm.com

Abstract

Qualitative methods have long been an important component of CSCW research. However, it can be challenging to make qualitative work legible to a broader set of researchers, which is critical as mixed methods research becomes more common. Moreover, the shift towards larger scales of data and increasing calls for open data and more transparency pose new questions for qualitative methods in terms of data collection, analysis, reporting, and sharing. This workshop brings together researchers to discuss these challenges as well as new opportunities for qualitative methods, with goals to help build norms and best practices for (1) conducting qualitative research, (2) reporting that research, and (3) engaging and collaborating with CSCW researchers from other methodological traditions.

Author Keywords

Ethics; Grounded theory; Interviews; Methods; Qualitative methods; Reliability; Research; Thematic analysis; Transparency

Introduction

Qualitative methods account for a substantial proportion of research in the CSCW and social computing fields, but making this research legible to a broader set of researchers can be challenging when there is such a diversity of approaches.

Qualitative methods also open up a number of questions at the heart of CSCW—around issues like transparency, ethics, collaboration, and sharing—particularly in light of a growing trend towards mixed methods work [35] and increased calls for open data in our research community [9,22]. Conversations around these issues suggest that clarity and deliberation are needed to define local norms and help authors make methodological choices and communicate them with clarity and confidence.

Even the definition of qualitative inquiry is situated within a complex history that has led to off-shoots of different methodological traditions and distinct analytical paradigms (e.g., deductive, inductive, abductive, and others [13,28]). We define qualitative inquiry as an interpretive style of research in which researchers attempt to make sense of phenomena in their natural settings, in terms of the meaning that people bring to them [11].

CSCW is particularly notable for the way in which it draws on diverse research traditions, which continue to expand into more qualitative and interpretive study. However, at the same time, calls in our field for greater transparency and open data [22] have placed pressure on researchers not only to share their data, but also to be more precise in descriptions of methods, which has opened up ethical and practical questions when it comes to both practice and reporting.

Workshop Themes
Data Collection

Qualitative research encompasses disparate classical traditions of inquiry, epistemic paradigms, and modes of reasoning. In interdisciplinary fields like CSCW, researchers tend to draw from multiple approaches for both data collection and analysis.

Data collection methods are an important decision when designing a study, with multiple ways to approach the same questions. How do we decide which types of data are most appropriate [31], and how do we weigh different potential types of bias? Are we interested in observed versus reported behavior [21]? Logistically, what are the trade-offs of different collection methods for the same kind of data [12]? How do we create the least biased samples [10], and whose definitions of bias are we using [8]? How much data do we need to collect and how do we know when to stop collecting data [6,7,19]? When conducting mixed methods research, how do we decide on the sequence of data collection [20]? And how do we balance our research plans across all of these dimensions as well as additional value considerations?

Decisions around data collection also raise some of the most pressing ethical issues for qualitative research. For example, how do we consider issues of researcher positionality and power dynamics [32]? What are best practices for collecting data with human subjects, including informed consent [4]? What are the ethical implications of collecting data *without* consent (e.g., from social media) [15]? How do we navigate cultural and other collectivist rules about permissions to share data [26]? How do we weigh potential ethical concerns with concerns about sampling or rigor? The lack of norms regarding ethics for social media research specifically [34] has produced many conversations about this topic in recent years, including multiple workshops [16,17,36]. By contextualizing these concerns in a broad discussion of qualitative research methods trade-offs, we can further these conversations.

Organizers

Casey Fiesler and **Jed. R. Brubaker** are both Assistant Professors in the Department of Information Science at University of Colorado Boulder. **Andrea Forte** is an Associate Professor in the Department of Information Science at Drexel University. **Shion Guha** is an Assistant Professor in the Department of Computer Science at Marquette University. **Nora McDonald** is a PhD candidate in the College of Computing and Informatics at Drexel University. **Michael Muller** works as a research staff member at IBM Research in Cambridge. All organizers have been conducting research relevant to qualitative methods. Fiesler and Brubaker's ongoing research in this space is funded by NSF project #1764089.

Data Analysis

There are a number of formal qualitative analysis methods common to CSCW research--for example, grounded theory [7], content analysis [23], and thematic analysis [3], among others. An understanding of the origins, details, and differences among approaches can help researchers align analysis methods, questions, data, and goals for the study. One goal is to help researchers make informed decisions about their own work. Another goal is to determine how best to articulate these traditions so that readers can better understand how to evaluate a study's merits.

This discussion is particularly critical given current conversations around rigor in qualitative work. Multiple techniques exist for establishing reliability or validity, such as member checking, triangulation of multiple sources of data, theoretical sampling, constant comparison, relationality, and reflexivity in grounded theory [7,18,28]. Yet there are not clear standards for which of these are most appropriate in what circumstances. Additionally, use of quantitative measures such as inter-rater reliability [23] raise challenging questions when applied to qualitative data [14]. When is it appropriate to apply standards of quantitative research (often based on positivist traditions) to qualitative research (often based on interpretivist traditions), or when might this be actively harmful?

There are also a number of tools available for analysis, depending on the method. Researchers might use everything from post-it notes to spreadsheets to software such as Atlas.ti, NVivo, MaxQDA, or Dedoose. How do researchers choose among these options, and what are the trade-offs? How do we conduct research collaboratively, both with respect to process [24] and how we deal with disagreement [37]? Moreover, as the scale and depth of data we have access to increases, how can we potentially scale, improve and enhance qualitative analysis methods, or combine them with other methodological paradigms popular in CSCW research (e.g. statistical methods, machine learning, systems analysis, design fiction etc.) in complementary and useful ways [1,29]?

Data and Methods Reporting

Given the complexities of decisions around data collection and analysis as discussed above, how do we decide how much detail to include? How do we explain rigor for qualitative research for potential reviewers or readers who come from different methodological traditions? When is it appropriate to simply cite a method versus explaining exactly how the research was conducted?

In addition to explaining the methods themselves, there are also decisions about reporting data. When is it appropriate to quantify qualitative data (e.g., N number of participants said X)? How much detail about participants should be included? How do we anonymize data appropriately, including potentially obscuring or fabricating quotes for ethical reasons [25]? When might we need to carefully protect identities of participants--versus intentionally identifying them [5]? What are best practices for explaining these decisions?

Data Sharing

In recent years there have been increasing calls for transparency and openness within HCI research [9], and there has been clear movement along these lines for quantitative research [22]. Values of open science that are standardized in some other fields include transparency, openness, and reproducibility [30]. However, it is unclear which of, whether, and how these values might apply to qualitative research. Interpretive research can also pose challenges to replicability—for example,

since frameworks that investigate marginal perspectives and/or are critical of power structures may consider that agreement perpetuates these structures, and measurements of agreement as indicators of scholarly quality can overlook the potential impact of some arguments.

Moves towards more openness in scientific research have also led to publications requiring data sharing or stating strong values; for example, PLoS' statement that "we strongly believe that data should be freely available all the time without having to go through a gatekeeper" [2]. This is a movement that has reached the HCI and CSCW communities as well [9,22]. However, beyond the obvious ethical concerns with sharing data collected from human subjects research, qualitative and quantitative data differ in fundamental ways that are important to consider when assessing whether qualitative data should adhere to the same data sharing expectations as quantitative data [33]. Replicability in analysis, for example, is typically not possible in qualitative research without the analyst.

Is openness a value that qualitative researchers should take up, and if so, how could we do so while protecting the humans and preserving the cultures represented in our data? For example, there may be cases where permission to reference private stories may be given to a particular researcher, but not extended to others [26], or where there are questions about who should be credited with "authorship" of reports [27]. Are there ways to achieve transparency beyond sharing data?

This set of themes highlights the complexities of conducting qualitative research, and the importance of having more open conversations within our research community. In the tradition of recent workshops focused on research ethics that emphasize open conversation and norm-setting [16,17], our hope is that providing a venue for these important discussions will help researchers who are struggling with these challenges.

References

[1] Eric P.S. Baumer, David Mimno, Shion Guha, Emily Quan, and Geri K. Gay. 2017. Comparing grounded theory and topic modeling: Extreme divergence or unlikely convergence? *Journal of the Association for Information Science and Technology* 68, 6: 1397–1410.

[2] Theodora Bloom, Emma Ganley, and Margaret Winker. 2014. Data access for the open access literature: PLOS's data policy. *PLOS Biology*.

[3] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2: 77–101.

[4] Amy Bruckman. 2014. Research Ethics and HCI. In *Ways of Knowing in HCI*, Wendy A. Kellogg and Judith S. Olson (eds.). Springer, 449–468.

[5] Amy Bruckman, Kurt Luther, and Casey Fiesler. 2015. When Should We Use Real Names in Published Accounts of Internet Research? In *Digital Research Confidential*, Ezster Hargittai and Christian Sandvig (eds.). MIT Press, Cambridge, MA, 243–258.

[6] Kelly Caine. 2016. Local Standards for Sample Size at CHI. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)* 981–992.

[7] Kathy Charmaz. 2014. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis (2nd ed.).* SAGE Publications, London, UK.

[8] Ronald J. Chenail. 2011. Interviewing the Investigator: Strategies for Addressing Instrumentation and Researcher Bias in Qualitative Research. *The Qualitative Report* 16, 1.

[9]   Lewis L. Chuang and Ulrike Pfeil. 2018. Transparency and Openness Promotion Guidelines for HCI. In *Extended Abstracts on Human Factors in Computing Systems (CHI)*, 1–4.

[10] Imelda T. Coyne. 1997. Sampling in qualitative research. Purposeful and theoretical sampling; merging or clear boundaries? *Journal of Advanced Nursing* 26, 3: 623–630.

[11] Norman K. Denzin and Yvonna S. Lincoln. 2008. *Strategies of Qualitative Inquiry*. SAGE Publications.

[12] Jill P Dimond, Casey Fiesler, Betsy DiSalvo, Jon Pelc, and Amy Bruckman. 2012. Qualitative Data Collection Technologies: A Comparison of Instant Messaging, Email, and Phone. In *Proceedings of the ACM Conference on Supporting Group Work (GROUP)*.

[13] Paul Dourish. 2014. Reading and interpreting ethnography. In *Ways of Knowing in HCI*, Judith Olson and Wendy Kellogg (eds.). Springer.

[14] Robert Elliott, Constance T Fischer, and David L Rennie. 1999. Evolving guidelines for publication of qualitative research studies in psychology and related fields. *British Journal of Clinical Psychology* 38: 215–229.

[15] Casey Fiesler and Nicholas Proferes. 2018. "Participant" Perceptions of Twitter Research Ethics. *Social Media + Society* 4, 1.

[16] Casey Fiesler, Pamela Wisniewski, Jessica Pater, and Nazanin Andalibi. 2016. Ethics and Obligations for Studying Digital Communities. In *Companion to the ACM Conference on Supporting Group Work (GROUP)*

[17] Casey Fiesler, Alyson Young, Tamara Peyton, Amy S. Bruckman, Mary Gray, Jeff Hancock, and Wayne Lutters. 2015. Ethics for Studying Online Sociotechnical Systems in a Big Data World. In *Companion to the ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW)*, 289–292.

[18] Susan Gasson. 2004. Rigor in grounded theory: An interpretive prespective on generating theory from qualitative field studies. In *The Handbook of Information Systems Research*, Michael Whitman and Amy Woszczynski (eds.). IGI Global.

[19] Greg Guest, Arwen Bunce, and Laura Johnson. 2006. How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. *Field Methods* 18, 1: 59–82.

[20] Nataliya Ivankova, John W. Creswell, and Sheldon L. Stick. 2006. Using mixed-methods sequential explanatory design: From theory to practice. *Field Methods* 18, 1: 3–20.

[21] Carlos Jensen, Colin Potts, and Christian Jensen. 2005. Privacy and practices of internet users: Self-reports versus observed behavior. *International Journal of Human-Computer Studies* 63, 1–2: 203–227.

[22] Matthew Kay, Steve Haroz, Shion Guha, Pierre Dragicevic, and Chat Wacharamanotham. 2017. Moving Transparent Statistics Forward at CHI. In *Extended Abstracts on Human Factors in Computing Systems (CHI)*, 534–541.

[23] Klaus Krippendorff. 2004. *Content Analysis: An Introduction to Its Methodology*. SAGE Publications, Thousand Oaks, CA.

[24] K. M. MacQueen, E. McLellan, K. Kay, and B. Milstein. 1998. Codebook Development for Team-Based Qualitative Analysis. *Field Methods* 10, 2: 31–36.

[25] Annette Markham. 2012. Fabrication as ethical practice: Qualitative inquiry in ambiguous Internet contexts. *Information Communication and Society* 15, 3: 334–353.

[26] Keavy Martin. 2012. *Stories in a new skin: Approaches to Inuit literature*. University of Manitoba Press.

[27] Sophie McCall. 2011. *First person plural: Aboriginal storytelling and the ethics of collaborative authorship*. UBC Press.

[28] Michael Muller. 2014. Curiosity, creativity, and surprise as analytic tools: Grounded theory method. In *Ways of Knowing in HCI*, Judith Olson and Wendy Kellogg (eds.). Springer.

[29] Michael Muller, Shion Guha, Eric P.S. Baumer, David Mimno, and N. Sadat Shami. 2016. Machine Learning and Grounded Theory Method: Convergence, Divergence, and Combination. In *Proceedings of the ACM Conference on Supporting Group Work (GROUP)*, 3–8.

[30] Brian A. Nosek, George Alter, George C. Banks, Denny Borsboom, Sara D. Bowman, Steven J. Breckler, and Stuard Buck. 2015. Promoting an open research culture. *Science* 348, 6242: 1422–1425.

[31] Kathleen H. Pine and Max Liboiron. 2015. The Politics of Measurement and Action. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 3147–3156.

[32] Farhana Sultana. 2007. Reflexivity, positionality and participatory ethics: Negotiating fieldwork dilemmas in international fieldswork. *ACME: An International E-Journal for Critical Geographies* 6, 3: 374–385.

[33] Alexander C. Tsai, Brandon A. Kohrt, Lynn T. Matthews, Theresa S. Betancourt, Jooyoung K. Lee, Andrew V. Papachristos, Sheri D. Weiser, and Shari L. Dworkin. 2016. Promises and pitfalls of data sharing in qualitative research. *Social Science and Medicine* 169: 191–198.

[34] Jessica Vitak, Katie Shilton, and Z. Ashktorab. 2016. Beyond the Belmont Principles: Ethical Challenges, Practices, and Beliefs in the Online Data Research Community. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW)*, 941–953.

[35] James R. Wallace, Saba Oji, and Craig Anslow. 2017. Technologies, Methods, and Values: Changes in Empirical Research at CSCW 1990-2015. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW: 106: 1–18.

[36] Pamela J. Wisniewski, Jessica Vitak, Xinru Page, Bart Knijnenburg, Yang Wang, and Casey Fiesler. 2017. In Whose Best Interest? Exploring the Real, Potential, and Imagined Ethical Concerns in Privacy-Focused Agenda. In *Companion of the ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW Companion)*.

[37] Himanshu Zade, Margaret Drouhard, Bonnie Chinh, Lu Gan, and Cecilia Aragon. 2018. Conceptualizing Disagreement in Qualitative Coding. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 159:1–11.