Sharing Qualitative Research Data for Secondary Analysis: Why, How, and With Whom?

Workshop at ASEE
Sunday June 25, 2023  1 - 3:30 pm

This project is supported by the USA National Science Foundation Grant 8060753.
Outline (for facilitators with timings):

**Understanding what SDA might encompass**

- **Part I: Understanding SDA**
  - 1:05 - 1:10 pm (5 min): Overview of SDA and findings from the project to date (Jenni) - including opportunity for participation that will be available at end of this workshop (include discussion of remaining funds and possibility of supporting additional projects - email us if you’re interested in moving forward)
  - 1:10 - 1:30 (20 min): Examples: Stories of SDA from pilot projects (10 min for each project: Zhenya, Tyler; lean on positionality)
  - 1:30 - 1:45 (15 min): Reflection & Observation (Susan)

**To SDA or Not to SDA? Exploring Possibilities**

- **Part II: To SDA or Not To SDA? Exploring Possibilities**
  - 1:45 - 2:05 (20 min): Exploring your own data - guide participants through the process of describing their data sets for others. (Rachel)
  - 2:05 - 2:20 (15 min): Exploring other people’s data - guide participants through generating potential SDA research questions for other people’s data. (includes time for break, which can also spill into the next activity) (Marie)
  - 2:20 - 2:35 (15 min): Reconsidering your own data - guide participants through reflecting on if and why (or why not) they would consider allowing any of the potential SDA research questions. (Lisa)

**The SHARE Framework: Trying it out**

- **Part III: The SHARE Framework: Trying It Out** (Holly & Lisa)
  - 2:35 - 2:40 (5 min): Brief explanation of the SHARE Framework
  - 2:40 - 3:00 (25 min): “Rapid Prototyping” - participants will work in groups to talk through the details of 1 selected SDA project using the SHARE Framework. Interactive, one principle of a time
  - 3:00 - 3:10 (10 min): Lessons Learned - group discussion about Parts II and III to consider factors that support or inhibit SDA.

**Now What?**

- **Part IV: Now What? (Jenni)**
  - 3:10 - 3:20 (10 min): Wrap-Up - mention opportunity to obtain funding,
TODAY’S AGENDA

Understanding what SDA might encompass

To SDA or Not to SDA? Exploring Possibilities

The SHARE Framework: Trying it out

Now What?
Today’s Facilitators

Holly Matusovitch, VA Tech
Jennifer Case, VA Tech
Marie Paretti, VA Tech
Lisa Benson, Clemson
Rachel Kajfez, Ohio State
Yevgeniya (Zhenya) V. Zastavker, Olin
Susan Lord, U San Diego
Tyler Young, Ohio State
Other Team Members

David Delaine, Ohio State

Shawn Jordan, Arizona State
Part I

Understanding what SDA might encompass

To SDA or Not to SDA? Exploring Possibilities

The SHARE Framework: Trying it out

Now What?
Part 1: Understanding SDA
The starting point for this grant (Jenni Case):

Secondary data analysis entails:
“the analysis of a single dataset that has been previously collected and analyzed by a different researcher/ set of researchers”

Secondary data sets are found on data repositories, such as:
- ICPSR, Institute for Social Research, University of Michigan (https://www.icpsr.umich.edu/web/pages/)
- The Qualitative Data Repository (https://qdr.syr.edu/)

Secondary data analysis has advantages:
- Resource savings (time/money)
- More complete use of data/better stewardship of data
Outcomes from our first workshop:

A broadened vision of secondary data analysis:

1. Analysis of multiple datasets
   a. Mixing and merging datasets on related topics
2. “Data as product”
   a. Thinking of the data itself (not the analysis) as a product
3. Data collected with broad data sharing in mind (i.e., Design Thinking Research Symposium led by Robin Adams)
   a. Share only with selected researchers
   b. Share on repository

Secondary data analysis challenges:

1. Careful thinking about process and cultural aspects
   a. Investment and rewards
   b. Best practices (IRB and other in-process ethical considerations)
2. Potential gatekeeping around who has access to or gets invited to engage in secondary analysis
Questions you hope we will talk about in this workshop?
PROJECT 1
(Rachel Kajfez and Zhenya Zastavker)

SDA as a Research and a Training Tool in Engineering Education Research:

First-Year Engineering Experiences
our story
who we are

Rachel Kajfez, Ohio State

Yevgeniya (Zhenya) V. Zastavker, Olin
who we are

Who are we?

- Olin College of Engineering
- Both currently engineering students
- Used to an education with COVID

[Images of two individuals]
Project 1: Kajfez & Zastavker

Using SDA to train UG students in EER

- **Dataset** – interviews with engineering students at one institution about their experience, longitudinal

- **Challenges being explored:**
  - Researchers are undergraduate students from another institution without grad program – does not have IRB
  - Undergraduates analyzing data from other undergraduates

- **Work in progress** – data prepping, students doing summer work, report-out at end of summer
original project

Collaborative Research: Understanding Engineering Pathways and Their Impact on Community and Identity

- Funded by NSF
- Study Sites: 2 main institutions with 2 regional campuses
- Study Population: 36 UG students
- Data Sources: interviews over three years (not all participants), survey responses, focus groups
- Dissemination: 2 journal articles, conference posters, WIPs, and a workshop
- Challenge: while findings were strong, the dataset offered more unexplored opportunities
sda project: opportunities

○ Coming together of 2 scholars from 2 radically different institutions in a co-creative way
○ Training of 2 engineering undergraduate students in qualitative methods research
sda project: outcomes

- Preliminary findings: verification and emergence of new themes through quasi-thematic and grounded theory analyses:
  - Both institutions: Impostor syndrome, What is engineering, Who is an engineer
  - Institution 1: Transition to college
  - Institution 2: Weed-out culture, Teaming difficulties, Support networks, Gender

- Emergent questions from SDA process:
  - For senior scholars: reimagining scholar identity & stepping into new story
  - For junior scholars: meaning-making about their scholar identity
  - For team: weaving a new story of individual and group identities as an SDA process and product
sda project: benefits & challenges

● Benefits:
  ○ Training of undergraduate students on deep qualitative investigation
  ○ Further exploration of data
  ○ Development of new scholarly networks
  ○ Support of existing findings and exploration of new findings

● Challenges:
  ○ Providing context without findings
  ○ Learning data for which context is not as present
  ○ Development of new processes
PROJECT 2
(E. Tyler Young, David Delaine, & Shawn Jordan)

SDA as a Research and a Training Tool in Engineering Education Research:

Diné Sovereignty
Project 2: Diné Sovereignty

**Purpose**

- Describe how Diné (Navajo) engineers’ understandings of tribal sovereignty (and their identities as tribal citizens) mediate their perception of engineering and their academic or work pursuits
- Pilot a framework for operationalizing tribal sovereignty in a PhD student’s dissertation work

**Problem**

- Collecting data in American Indian communities is fraught with difficulties:
  - Historic abuse of these communities by outside researchers
  - Can be intrusive and time-consuming for participants with little benefit offered in return
  - Tribes are sovereign nations - an additional level of necessary oversight is needed for legal/ethical reasons
- Novice researchers (e.g., PhD students) may be overwhelmed by the complexities of qualitative data collection

**Solution:** Conduct a secondary thematic analysis on an existing dataset consisting of de-identified transcripts generated through interviews with Diné engineers
Project 2 Original Dataset: Navajo Engineers

- Interview transcripts with Navajo engineering professionals
- Dataset previously used to analyze intersections between Navajo culture and engineering design practice
- 20 participants
- Navajo / Half Navajo
- Male (12) / Female (8)
- Raised on / off of Navajo Nation
- Live off of Navajo Nation
- Work on / off of Navajo Nation
- Variety of engineering disciplines
Challenges and Adaptations (Project 2)

**CHALLENGE**

The consent process from the original study did not include a clause for secondary analysis.

**ADAPTATION**

PI on original study contacted participants to obtain updated consent.

Data collected from the student population in the original study was excluded from SDA, as it was deemed too difficult to reestablish contact with these individuals to obtain updated consent.
Challenges and Adaptations (Project 2)

CHALLENGE
The new investigators only have access to the de-identified interview transcripts from the original study

ADAPTATION
The original PI remained active in the SDA, providing insights and context from the original study when necessary
Challenges and Adaptations (Project 2)

CHALLENGE
The new lead investigator (PhD student) is not familiar with the Diné (Navajo) tribal context

ADAPTATION
Add an additional Diné investigator
Highlighted Outcomes (Project 2)

- Expanded preliminary code tree
  - (6 => 40+ codes/subcodes)
- Draft interview protocol for dissertation work
- 100+ training hours for PhD student in qualitative analysis methods
Discussion

(Susan Lord)

What questions did this discussion raise for you?

What questions did it answer for you?

Challenges and opportunities?

Take a few minutes to share with your neighbors and then we’ll share all together.
Part II

Understanding what SDA might encompass

To SDA or Not to SDA? Exploring Possibilities

The SHARE Framework: Trying it out

Now What?
Part 2: To SDA or Not to SDA
Your Data (Rachel Kajfez)

For those with data...
Please create a post-it

Project topic or title (can be shorthand)

Who are the participants?

How was the data collected?

Purpose of the data

Status of current analysis (if any?)

For those without data...
What else would you want to know?

Brainstorm a list of 3-5 questions you could ask of each post-it

Shame in Eng. Ed

Focus: deep, psychological experiences of shame

9 x interviews with White Males
7 x interviews with URM students
Analyzed all interviews using IPA

Focus: Social construction of expectations

10 x focus groups with 2-5 participants
(all White male and diverse)
Analyzed focus groups using thematic analysis/grounded theory
Exploring Other People’s Data - A Working Break (Marie Paretti)

Your tasks for the next 15 minutes:

- Refresh and renew yourself as needed!
- Grab a handful of post-it notes.
- Visit at least two posters and use the post-its to respond to the following:
  - What questions do you have for the authors? What *else* would you want to know about the data before engaging in SDA?
  - What energizes or excites you about this data set? What draws your interest?
  - What new research questions and/or frameworks and/or methods would you bring to the data in a collaborative SDA? Be creative here!
Reconsidering your own data  
(Lisa Benson)

Think of sharing data as a form of “technology transfer.” So many benefits! You want to share, but…

● How do you protect your intellectual property?
● What do new users need to know?
● If the original project is really complex, how do you preserve that complexity while sharing?
● Will you have any control over how others use your data?

To avoid conflict or negative outcomes:

● Document expectations about collaborations, credit, and co-authorship
● Carefully consider who can access the data, how it might be used, and how well it aligns (if at all) with the original project
Another approach: SDA within a graduate research methods course

What was shared?
- De-identified data from an NSF-funded project
data (n = 527) collected at two time points (Years 3 and
4) from students in a single undergraduate engineering program:
Survey data, interview data collected in Year 4 from three
students
- Background information (anonymized) on the project
- Survey key (definitions of 9 constructs plus demographic info)
- Interview data collected in Year 4 from three students

What did it take?
- Data cleaning: 10+ hours (and counting)
- Documentation: 4 hours
- Developing and launching course project: 3 hours
- Student meetings: 8 hours
- Developing feasible research questions
- Interpreting findings
- Troubleshooting data analysis

What was learned?
- Positive learning experience for students (“using real data”)
- Anonymized data set with documentation available for SDA for
  education or research purposes → IM and BI for NSF project
- SDA for education purposes is a good first step in sharing data
- Anonymized data set with documentation available for SDA for
  education or research purposes → IM and BI for NSF project
- SDA for education purposes is a good first step in sharing data
Part III

Understanding what SDA might encompass

To SDA or Not to SDA? Exploring Possibilities

The SHARE Framework: Trying it out

Now What?
SHARE

(Lisa Benson & Holly Matusovich)

- Stewarding collaborative relationships
- Honoring context of data
- Aligning questions and data
- Responsibly reusing data
- Expanding capacity and ownership
## SHARE principles

| **Stewarding collaborative relationships** | The deeply contextualized nature of qualitative data often requires data stewards to facilitate collaborations between researchers originating and seeking to use data, ensuring that data are anonymized. Data originators could potentially serve as collaborators on projects or co-authors on publications to ensure that the original context is preserved and recognized. |
| **Honoring context of data** | Those originating and analyzing the data need to consider where, how, when, and by whom data are collected while maintaining necessary anonymity of participants. |
| **Aligning questions and data** | As with all research studies, research questions should be aligned with data collection and analysis methods. SDA can be approached collaboratively where a researcher with specific research questions works with a data steward to determine whether a dataset is rich enough to help answer their questions, or approached inductively where secondary analyzers review samples of the data in search of potential research questions that could be answered, seeking to fill gaps in existing literature. |
| **Responsibly reusing data** | Ethics and trust are critical to any data sharing and analysis project. It is imperative to protect the identities of participants, respect their lived experiences and conduct research that has the potential to benefit the original participants or the population they represent. In sharing data, it is also important to develop a trusting relationship between the data originator and secondary analyzer that acknowledges the vulnerability involved with sharing a data set. |
| **Expanding capacity and ownership** | Sharing data can fulfill the need to acknowledge diverse approaches to capability development and build capacity of the research community by bringing new researchers into the process without requiring them to collect their own data. SDA can also broaden ownership of data so that others can own and shepherd it as well. |
Lessons learned from SDA for educational purposes

- Documentation and communication are key:
  - Document what is in the data set, background of the original project, original data collection protocols, timing of data collection, cross-referenced and/or matched data sources, etc.
  - Two-way communication - ask and answer questions with those you are sharing with
- Be prepared to put time into data cleaning, documentation and communication
- Carefully consider how others will access and use your data

How could SHARE help guide the process of prepping a data set and introducing it to secondary analyzers?
## Using SHARE principles to guide SDA

<table>
<thead>
<tr>
<th>Principle</th>
<th>Explanation</th>
<th>Example reflection question</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stewarding collaborative relationships</strong></td>
<td>The deeply contextualized nature of qualitative data often requires data stewards to facilitate collaborations between researchers originating and seeking to use data, ensuring that data are anonymized. Data originators could potentially serve as collaborators on projects or co-authors on publications to ensure that the original context is preserved and recognized.</td>
<td>Do secondary analysts understand the richness of the context and the importance of maintaining participant anonymity?</td>
</tr>
</tbody>
</table>

Envision sharing a data set; what other questions would you have related to stewarding collaborative relationships between the data originator and secondary analysts?
<table>
<thead>
<tr>
<th>Principle</th>
<th>Explanation</th>
<th>Example reflection question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Honoring context of data</td>
<td>Those originating and analyzing the data need to consider where, how, when, and by whom data are collected while maintaining necessary anonymity of participants.</td>
<td>Has data been carefully de-identified such that the nature of the context is kept intact while stripping out specific details that would allow participants to be identified?</td>
</tr>
</tbody>
</table>

Envision sharing a data set; what other questions would you have related to honoring the context of the data that should be considered by the data originator and secondary analysts?
**Principle**  
**Explanation**  
**Example reflection question**

<table>
<thead>
<tr>
<th>Principle</th>
<th>Explanation</th>
<th>Example reflection question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aligning questions and data</td>
<td>As with all research studies, research questions should be aligned with data collection and analysis methods. SDA can be approached collaboratively where a researcher with specific research questions works with a data steward to determine whether a dataset is rich enough to help answer their questions, or approached inductively where secondary analyzers review samples of the data in search of potential research questions that could be answered, seeking to fill gaps in existing literature.</td>
<td>Is the dataset content appropriate for exploring the phenomenon within the proposed SDA?</td>
</tr>
</tbody>
</table>

Envision sharing a data set; what other questions would you have related to aligning research questions and the data that should be considered by the data originator and secondary analysts?
<table>
<thead>
<tr>
<th>Principle</th>
<th>Explanation</th>
<th>Example reflection question</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Responsibly reusing data</strong></td>
<td>Ethics and trust are critical to any data sharing and analysis project. It is imperative to protect the identities of participants, respect their lived experiences and conduct research that has the potential to benefit the original participants or the population they represent. In sharing data, it is also important to develop a trusting relationship between the data originator and secondary analyzer that acknowledges the vulnerability involved with sharing a data set.</td>
<td>Do the potential outcomes of the planned SDA meaningfully expand on those of the original research?</td>
</tr>
</tbody>
</table>

Envision sharing a data set; what other questions would you have related to responsibly reusing data that should be considered by the data originator and secondary analysts?
Principle | Explanation | Example reflection question
---|---|---
**Expanding capacity and ownership** | Sharing data can fulfill the need to acknowledge diverse approaches to capability development and build capacity of the research community by bringing new researchers into the process without requiring them to collect their own data. SDA can also broaden ownership of data so that others can own and shepherd it as well. | Are the limitations of the secondary analyzers identified?

Envision sharing a data set; what other questions would you have related to expanding capacity and ownership of the data that should be considered by the data originator and secondary analysts?
Gallery Walk

Think about:

- What interests you about these datasets?
- What questions/concerns do you have about using these datasets?
- (Assume the IRB issues are under control)
- Who might you want to engage with in further conversation around SHARE principles?

**does this go here or earlier when Marie discussion other people’s data?**
### Work through SHARE for one dataset on the wall

<table>
<thead>
<tr>
<th><strong>Stewarding collaborative relationships</strong></th>
<th>The deeply contextualized nature of qualitative data often requires data stewards to facilitate collaborations between researchers originating and seeking to use data, ensuring that data are anonymized. Data originators could potentially serve as collaborators on projects or co-authors on publications to ensure that the original context is preserved and recognized.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Honoring context of data</strong></td>
<td>Those originating and analyzing the data need to consider where, how, when, and by whom data are collected while maintaining necessary anonymity of participants.</td>
</tr>
<tr>
<td><strong>Aligning questions and data</strong></td>
<td>As with all research studies, research questions should be aligned with data collection and analysis methods. SDA can be approached collaboratively where a researcher with specific research questions works with a data steward to determine whether a dataset is rich enough to help answer their questions, or approached inductively where secondary analyzers review samples of the data in search of potential research questions that could be answered, seeking to fill gaps in existing literature.</td>
</tr>
<tr>
<td><strong>Responsibly reusing data</strong></td>
<td>Ethics and trust are critical to any data sharing and analysis project. It is imperative to protect the identities of participants, respect their lived experiences and conduct research that has the potential to benefit the original participants or the population they represent. In sharing data, it is also important to develop a trusting relationship between the data originator and secondary analyzer that acknowledges the vulnerability involved with sharing a data set.</td>
</tr>
<tr>
<td><strong>Expanding capacity and ownership</strong></td>
<td>Sharing data can fulfill the need to acknowledge diverse approaches to capability development and build capacity of the research community by bringing new researchers into the process without requiring them to collect their own data. SDA can also broaden ownership of data so that others can own and shepherd it as well.</td>
</tr>
</tbody>
</table>
Lessons Learned (Lisa & Holly)
Part IV

Understanding what SDA might encompass

To SDA or Not to SDA? Exploring Possibilities

The SHARE Framework: Trying it out

Now What?
What is next?

(Jenni Case)

- Come see our poster at the NSF grantees session on Wednesday 9:15-11:45 am “Lessons Learned doing Secondary Data Analysis in Engineering Education Research (EER)”
- Participate directly in SDA (jencase@vt.edu).
- Brainstorm: How do we increase accessibility to SDA?
Plan for today:

Part I: Understanding SDA
- 5 min: Overview of SDA and findings from the project to date (Jenni) - including opportunity for participation that will be available at end of workshop (include discussion of remaining funds and the possibility of supporting additional projects - email us if you’re interested in moving kind of thing)
- 5 min: Reflection & Observation
- 20 min: Practical Examples: Stories of SDA from two pilot projects (10 minutes for each project: Zhenya, Tyler; both lean on their positionalities)
- 15 min: Reflection & Observation

Part II: To SDA or Not To SDA? Exploring Possibilities
- 20 min: Exploring your own data - participants will be guided through the process of describing their data sets for others. (Rachel)
- 15 min: Exploring other people’s data - participants will be guided through generating potential SDA research questions for other people’s data (includes time for break, which can also spill into the next activity) (Marie)
- 15 min: Reconsidering your own data - participants will be guided through reflecting on if and why (or why not) they would consider allowing any of the potential SDA research questions. (Lisa)

Part III: The SHARE Framework: Trying It Out
- 5 min: Brief explanation of the SHARE Framework (Holly & Lisa)
- 25 min: “Rapid Prototyping” - participants will work in groups to talk through the details of one selected SDA project using the SHARE Framework. Interactive, one principle of a time
- 10 min: Lessons Learned - group discussion about Parts II and III to consider factors that support or inhibit SDA.

Part IV: Now What? (Jenni)
- 10 min: Data Sharing Mechanisms and Ways. Uses a commodity kit to explain what to make SDA accessible (except for problem...