SAGE researchmethods cases

Quantitative Coding of Qualitative Data Gathered in a Lab Experiment

Abstract

This case outlines some of the processes involved in writing a PhD thesis on a sensitive topic, using forensic administrative data, in an Indigenous population. Issues raised include the author's initial concerns about the project, a discussion of using forensic administrative data sourced from police information management systems, dealing with one-to-many relationships between variables, missing data on a main variable, and working with data relating to Indigenous populations. The case concludes with a discussion of the author's emotional response to the topic, with advice on how to manage vicarious trauma that may be caused by doing research on child sexual abuse.

Learning Outcomes

By the end of this case, students should be able to

- Describe the benefits of using mixed-method approaches in research
- Describe the steps for developing a coding scheme
- Employ strategies for quantitative coding of qualitative data
- Identify the resources required for completing quantitative coding of qualitative data

Project Context

The legal standard for sexual harassment and sex discrimination cases requires evidence that the behavior was unwanted by the target, that is, that a "reasonable person" would react assertively if she experienced behavior that was inappropriate (EEOC, 1990). Although women predict that they will actively respond to sexism, they rarely assertively confront sexist perpetrators when they experience sexist behavior (Swim & Hyers, 1999; Woodzicka & LaFrance, 2001). This suggests a disconnect between predicted and actual behavior in response to sexism. We tested whether the divide is due to the different goals that women may have in mind when imagining versus experiencing discrimination. When women predict their responses to sexism, they may imagine that their primary goal will be to be respected and to restore that respect via confrontation. When women are actually in a sexist situation, the fundamental need to belong (Baumeister & Leary, 1995) may override the goal to be respected. This is because when faced with sexism in the workplace context, a woman may pursue the goal to be liked by the perpetrator because a job, evaluation, or promotion may depend on the perpetrator's opinion of her. The fear of backlash in the moment may therefore stifle confrontation even if the behavior was clearly inappropriate and sexist. In Mallett and Melchiori (2014), we hypothesized that women would only pursue an assertive response to sexism if the goal to be respected outweighed the goal to be liked.

We were interested in documenting women's actual responses to sexism. Therefore, we wanted to capture their real-time responses to a sexist situation. We decided to use a mixed-method approach and collect both qualitative, open-ended responses, in the form of instant-messaging conversation responses, and quantitative, closed-ended responses, in the form of Likert scale ratings. Mixed-methods approaches allow researchers to gain rich information from qualitative data while maintaining tight control of closed-ended information from quantitative data. Mixed-methods approaches allow for "triangulation"—the use of different methods, sources of data, and/or perspectives to focus on a particular research question. Triangulation allows researchers to view the research question from multiple perspectives and gain insights into the problem at hand from different vantage points. If multiple methods yield similar results, we can be confident that our conclusions are sound.

One reason we collected qualitative data was to reveal what women actually do when faced with sexist behavior. Recording the language used by women is an important reason for collecting qualitative data. Doing so allowed us to either support or contradict the expectation that a reasonable person always confronts sexism. Collecting qualitative data in an experimental paradigm also allows us to experimentally manipulate aspects of the situation and capitalize on the power of random assignment to account for individual differences in sensitivity and response preferences.

There are several methods for collecting qualitative information from participants. One may ask people to recall a certain type of experience using focus groups, conduct one-on-one interviews, or ask people to keep a daily diary of a certain type of experience (Hyers, Swim, & Mallett, 2006). Each method provides a unique piece of information and increases our ability to understand complex behavior such as responses to discrimination. Collecting real-time responses has some advantages relative to other methods. Focus groups and interviews rely on the ability to retrospectively recall an experience and often result in people retrieving memorable but perhaps extreme examples of a behavior (Swim, Pearson, & Johnston, 2007). Moreover, memories of one's responses to discrimination may be affected by attempts to cope with the stressful nature of the experience. For example, memories may become distorted by our attempts to make sense of what happened (Reis & Wheeler, 1991). Daily diaries are a great way to capture information from daily life including the frequency and nuances of a behavior, but we sacrifice internal validity for external validity.

We conducted laboratory experiments to test our hypothesis. A detailed discussion of our high-impact, high-psychological realism methodology can be found in a previous SAGE Research Methods Case (Melchiori & Mallett, in press). In this case, we will discuss the coded open-

ended data published in Mallett and Melchiori (2014). The two studies in the published article are similar in their methodology. We focus on Study 1 because it has a more straightforward design but includes the exact same qualitative, open-ended data coding as Study 2.

Research Practicalities

A team of five or six undergraduate research assistants collected data on the studies each semester. Over the course of the experiment, female participants interacted with a confederate posing as a male university staff member over an online instant message platform. This alleged staff member would ask the participant either sexist or surprising interview questions during the online conversation, and the participant would have the opportunity respond to each question via instant massage. These text message responses made up the open-ended, qualitative data that we then quantitatively coded.

Our full research lab of 12 undergraduate research assistants completed quantitative coding of our qualitative data each semester. The first few weeks of one semester were dedicated to refining our coding scheme. We based our coding scheme, in part, on what past research had coded for, relying on Woodzicka and LaFrance's (2001) coding as a starting point. We added codes and examples of anchor points as we refined the coding scheme via discussion at lab meetings. Once the coding scheme was in place, we gave the research assistants a weekly coding assignment to complete. Two research assistants coded each response independently. Coders were partially naïve (i.e., "blind" or unaware) to condition. We did not tell coders what condition the participants were in, but sometimes the nature of the response to a question revealed the condition (e.g., if the response mentioned a "boyfriend," it was obviously a response to one of the sexist questions and not the surprising questions control condition). Coders only coded one of the three responses of each participant, therefore if the condition was revealed in one response, the coder for the next response would potentially still be naïve to condition. Moreover, coders were naïve to the exact hypotheses regarding the relation between independent and dependent variables.

Coding qualitative data is a time- and human resource-intensive process. There were 154 participants in Study 1, and three responses for each participant, for a total of 462 responses we needed to code *twice*. For each weekly assignment, coders would independently code approximately 20 responses to a single question (there were three questions that each participant responded to in our study). Each weekly coding assignment took approximately 1 to 2 hr for each of the 10 to 12 research assistants to complete. Each research assistant completed approximately 5 hr of coding for the study, for a total of 100 hr of labor to complete the coding for Study 1. We also dedicated 20 to 30 min of each weekly research lab meeting to

a coding "check-in" to increase and maintain coder reliability. Note that time for coding was in addition to the 154 hr required to collect the data. Thus, simply compiling the data for this single study took more than 250 hr.

Method in Action

Study Design

Following Woodzicka and LaFrance's (2001) interview paradigm, we brought women into the lab one at a time to participate in study about online job interviews for telecommuters. After she was greeted by the experimenter, each participant was seated at a computer station in a small room, and the research assistant was seated at a computer station in a small adjoining room. All study materials were presented via computer.

We told participants that a 32-year-old male campus employee would conduct the mock interview for a research assistant position. In reality, the research assistant conducted the chat from another room. All participants received three control questions ("What is your year in school?" "What is your major or intended major?" "Why are you interested in this research position?") and were given time to respond in the chat. After receiving their response, we randomly assigned half the participants to view the sexist questions ("Do you have a boyfriend?" "Do people find you desirable?" "Do you think it is important for women to wear bras to work?") and assigned the other half to view the surprising, control questions ("Do you have a best friend?" "Do people find you morbid?" "Do you think it is important for people to believe in God?"). These questions were taken directly from Woodzicka and LaFrance's (2001) study. Participants were given time to respond in the chat. After sending their responses, the research assistant entered the room and announced the interview portion of the study was over due to an issue with the internet connection.

We then assessed the assertiveness of participant responses via self-report and coding. For self-reported assertiveness, participants rated how much they used six assertive response strategies based on Woodzicka and LaFrance (2001) (e.g., "Asked why the question was asked" "Told the interviewer it was none of his business"; 1 *not at all* to 7 *very much*) in response to each question. For coded assertiveness, two independent coders were trained by the researchers to identify each of the assertive response strategies as well as whether the question was simply answered with a yes or no response. We also asked participants to report their goal preference by answering the question "If you had to choose between being liked and being respected by the interviewer, which do you regard as more important?" with 1 = *most important to be liked*, 4 = *equally important*, 7 = *most important to be respected*.

Preparing a Coding Scheme

After the data were collected, we began to draft our coding scheme (see Figure 1 for the final coding scheme). We started by including the types of responses found in Woodzicka and LaFrance (2001). This included stating whether participants simply answered the question, asked why the question was asked, refused to answer, stated the question was irrelevant to the position, told the interviewer it was "none of his business," and stopped the interview.

Figure 1. Coding Scheme.

			Question #: (circle one)	1	2	3
PID_____	RA Initials_____					

	NO	YES
Did the participant provide a yes or no answer _to the interviewer_? In other words, did the participant <u>convey to the interviewer</u> that she has a boyfriend, thinks other people find her morbid, thinks women should wear bras, etc.?	NO	YES

	NO	YES
Q1: Did the participant say "yes" <u>with little elaboration</u>?	NO	YES

↳ If YES, <u>skip to The Vibe</u>. If NO:

↓

	NO	YES
Q2: Did the participant say "no" <u>with little elaboration</u>?	NO	YES

↳ If YES, <u>skip to The Vibe</u>. If NO:

↓

	NO	YES
Q3: Did the participant refuse to answer the question <u>with little elaboration</u>?	NO	YES

↳ If YES, <u>skip to The Vibe</u>. If NO:

↓

If you answered NO to the three items above and the participant gave a more elaborate answer to the question, please use the checklist below. Put a 1 in the box for her first response, a 2 for her second response, etc. Important: only code <u>actions/behaviors</u>. Do <u>not</u> code for thoughts or feelings with this section (save those judgments for The Vibe.)

A	Say she didn't know how to respond	E	State that she refused to answer	I	Report the interviewer to his supervisor	M	Try to be funny
B	Provide a clarification	F	Laugh	J	Leave the interview	N	State she did not understand a specific word
C	Say it is irrelevant to position	G	Say something rude	K	Was polite, kind, nice, etc.	O	Provided a rationale for her response
D	Answer the question	H	Say it is none of the interviewer's business	L	Say the question was inappropriate		

	NO	YES
Did the participant ask a question in her response?	NO	YES

The Vibe

OVERALL…

	not at all	slightly	moderately	extremely
…how offended was the respondent?	1	2	3	4

	very negative	slightly negative	slightly positive	very positive
…what was the affective tone of the response?	1	2	3	4

Next we browsed through the qualitative data, noting any other thematic trends. Based on this brief review, we decided to include a few additional behaviors in our coding scheme (e.g., clarified the question or explained their answer). If participants gave more than a simple "yes," "no," or refusal to answer the question, coders were allowed to select more than one behavior code (see Figure 1). Initially, we asked coders to mark which response was given first, second, and so on if more than one selection was made. Our thought was that we would rather begin with richer data coding to allow for more flexibility for exploratory data analysis, even though we did not hypothesize an order effect for the different responses. During data analysis we elected to ignore response order and recoded the data so that each behavior was coded as either present or absent.

The final section of our coding scheme (see Figure 1) we referred to as "The Vibe." We wanted participants to engage in minimal intention-guessing for the behavioral codes. Therefore, we asked coders to save their judgments about the intentions, thoughts, and feelings of the participant responses for the last two questions in the coding scheme: "Overall, how offended was the respondent?" and "Overall, what was the affective tone of the response?"

Coding Training

We used the coding training process to refine our coding scheme and to boost interrater reliability by providing coders with definitions, examples, and answers to common questions. Before training, we made copies of the coding scheme draft and selected several responses from our data set to use as examples during the weekly lab meeting. We made an effort to select a broad array of answers, from simple one-word or one-sentence replies to replies that spanned multiple lines of text.

After the research assistants arrived, we walked them through the coding scheme (Figure 1) then used a computer and projector to display example participant responses to the lab room. All lab members, including undergraduate research assistants, graduate students, and faculty, coded each example response according to the current scheme and participated in the subsequent lab discussion. We encouraged the coders to make note of any questions they had while coding. After coding each response, we went through the coding scheme as a lab group. We carefully noted points of confusion, any additional themes that emerged, or other structural changes that we could make to the coding scheme to make the process more clear. For example, as a result of training, we added codes such as "State she did not understand a specific word" after it became clear that many students did not know the definition of "morbid" and added the code "try to be funny" after a discussion about how to interpret "haha" or "lol" responses. We elected to include these codes, even though we did not have a priori

hypotheses for them, for two reasons. First, having additional codes would give us a richer data set with which to conduct exploratory analyses, should we so choose. Given that data coding is time- and human resource-intensive, it made sense to gather the additional data at the same time as our initial coding. Second, including these additional codes helped our coders refine their inclusion criteria for our main behaviors of interest.

As a result of the training process and lab discussion, we also decided to create a "coding cheat sheet" (see Figure 2). Here, we provided a one page, double-sided list of codes, coding directions, and example responses to further guide our coders as they completed their coding outside of the lab meeting. Coders could refer to this list of standard examples if they encountered a tricky response on their own. This increased reliability as coders were referring to the same set of definitions for each code.

Figure 2. Coding Cheat Sheet.

<u>YES with little elaboration</u>: **(one line, < 15 words)**
    "Yes, I do."
    "Yes, I have a few best friends from home."
    "I would say that I think women should be required to wear bras in the workplace."

<u>NO with little elaboration</u>: **(one line, < 15 words)**
    "Nope, currently single"
    "I don't think so!"
    "I would say that I did not have a best friend."

<u>REFUSE TO ANSWER with little elaboration</u>: **(one line, < 15 words)**
    "I would not answer that question."
    "I think I would ask him to go on to the next question."

Coding Process

We created a coding assignment list that include each coder's name, initials, and the participant identification numbers (PIDs) that the coders would be responsible for coding each week. We gave the coders this assignment document and an Excel spreadsheet that included the PIDs and responses that they were responsible for coding. These documents were stored on a secure university network drive; coders could complete their assignments on any university computer, but were instructed to only code when they were in a private space such as their residence hall or the research lab. To ensure confidentiality, they were also instructed not to discuss any of the coding with friends who were not in the lab. All documents that they created or completed while coding were also stored on the secure university network drive.

Coders were required to fill out a coding scheme sheet (see Figure 1) for each response either electronically or by hand. After completing their weekly assignment, we asked coders to then enter their responses into an Excel spreadsheet with the help of a data entry guide that we created (see Figure 3). Coders were instructed to email us their completed Excel files one day before our scheduled lab meeting. Coders also either emailed their digitally completed coding schemes or brought their hand-completed coding sheets with them to the lab meeting.

Figure 3. Data Entry Sheet.

Variable	Question	Value
PID	Participant identification number	[same as PID #]
RA	RA Initials	[type your initials]
Question#	Question number	[1, 2, 3]
Answer	Did the participant provide a yes or no answer to the interviewer? In other words, did the participant convey to the interviewer that she has a boyfriend, thinks other people find her morbid, thinks women should wear bras, etc.?	0 = no 1 = yes
Q1	Did the participant say "yes" with little elaboration?	
Q2	Did the participant say "no" with little elaboration?	
Q3	Did the participant refuse to answer the question with little elaboration?	
RespA	Say she didn't know how she would respond	0 = did not make response 1 = first response 2 = second response 3 = third response 4 = fourth response 5 = fifth response 6 = sixth response 7 = seventh response 8 = eighth response 9 = ninth response 10 = tenth response
RespB	Provide a clarification	
RespC	Say it is irrelevant to position	
RespD	State that she would answer the question	
RespE	State that she would refuse to answer	
RespF	Say she would laugh	
RespG	Say something rude	
RespH	Say it is none of the interviewer's business	
RespI	Report the interviewer to his supervisor	
RespJ	Leave the interview	
RespK	State that she would be nice, polite, kind, etc.	
RespL	Say the question was inappropriate	
RespM	Try to be funny	
RespN	State she did not understand a specific word	
RespO	Provide a rationale for her response	
AskQuestion	Did the participant ask a question in her response?	0 = no 1 = yes
Vibe1	…how offended was the respondent?	1 = not at all 2 = slightly 3 = moderately 4 = extremely
Vibe2	…what is the affective tone of the response?	1 = very negative 2 = slightly negative 3 = slightly positive 4 = very positive

Before the lab meeting, we combined independent coder files of the same PIDs. We then flagged responses where the coders had drastically different codes (e.g., a "Vibe" question where one coder marked a 1 and the other coder a 3; when one coder marked two totally different behaviors from the other coder). We printed out these 5-6 sets of flagged documents and brought them with us to the lab meeting. Coders sat next to their coding partners during the meeting for a reliability check-in. We instructed coders to review the flagged discrepancies. We instructed the coders to discuss why their codes were discrepant (e.g., one person read the wrong response). If, as a result of their discussion, one or both coders came to a new realization about a code or realized there was a typo, the coders were allowed to change their codes on the sheet. We emphasized that it was perfectly fine to agree to disagree and leave the flagged codes as they were.

After each coding pair finished their discussion and made desired changes, we met again as a group to discuss any points of clarification or coding guidelines that any of the pairs had now set for themselves. We included this final discussion to ensure that any new coding guideline set by one coding pair was also followed or discussed among all coding pairs. We then collected the coding documents and made the updates to the digital coding file, always being sure to retain the file that was created before the coding check-in. The process was completed each week there was a coding assignment for the lab group. By the end of the coding process we had 90% to 100% agreement across codes and coders.

To analyze our finalized coded data, we summed across answers to the three questions in each condition (sexist and surprising) to create a variable for each behavior that ranged from 0 to 3. We created one variable to represent compliant responses by computing the average of "simply answer" and "explain." We created another variable to represent the assertiveness of responses by averaging the remaining behaviors of interest we identified from past research.

Conclusions and Practical Lessons Learned

Based upon our coded data, we were able to conclude that both coded responses and self-reported ratings showed that women were more assertive following sexist than merely surprising questions. Furthermore, we found a significant mediation effect such that women in the sexist (vs surprising) questions condition reported greater pursuit of a respect goal (vs a liking goal), which in turn was associated with an increase in coded and self-reported response assertiveness (Mallett & Melchiori, 2014).

One caveat is that we captured only one version of how people respond to discrimination. Our conclusions must be considered within the context of our study. Our participants were in an

experimental lab setting where the costs of responding assertively are lower than they are in the real world. Therefore, what we captured may be more assertive than what we would see in a workplace context. Alternatively, the norms for compliance and politeness in an experiment mimic the workplace. As such, we may have come close to recreating a workplace environment. In both cases, the response is public—meaning that the women know someone else will read their thoughts.

Our research assistants had an intensive learning experience via their involvement in this project. Although less than half of the research assistants were personally responsible for collecting the data, all of our undergraduate assistants helped design and implement the coding scheme. They saw how we were able to form a preliminary test of our idea and modify the coding scheme to accommodate unanticipated patterns of response. They used what they learned from classes on research methods, social psychology, and the psychology of prejudice to help us refine our constructs and map examples onto codes. Many research assistants report that this experiential learning made them feel like they were part of a team and that they had a home in the research lab. They recognized that we valued their contributions and relied on them to complete the project.

Research assistants acquired a valuable skill set from their participation in creating and using the coding scheme. This skill helped at least three of our research assistants get jobs and internships after graduation. By reading women's responses to the sexist and surprising questions, research assistants also gained a better understanding of the variability in responses to sexism. They were able to see firsthand that women almost never assertively responded. They were frequently shocked by how many of our participants ignored the inappropriate comments or answered the questions, sometimes with elaborate detail.

Interestingly, several of our research assistants reported that before their involvement in conducting this research they did not consider themselves to be a feminist. However, after seeing how hundreds of women responded to sexism, they now identified as feminist. Moreover, they were more willing to engage in conversations with their peers and family because they had a better understanding of the responses they could use during such a conversation. Thus, the rich data we gathered with our mixed-methods design not only lead to interesting and important research findings, but also lead to personal growth within our research lab.

Exercises and Discussion Questions

1. What is one benefit of using a mixed-methods approach to data collection?

2.What is one limitation of relying on retrospective recall and self-report of past behavior?

3.What resources are necessary for completing quantitative coding of qualitative data?

4.What are the steps to creating and refining a coding scheme?

5.Why should you include more detail in the coding scheme than that which is included in your study hypothesis?

Further Reading

Tashakkori, A., & **Teddlie, C.** (Eds.). (2010). *The SAGE handbook of mixed methods in social & behavioral research*. Thousand Oaks, CA: SAGE.

Willig, C., & **Stainton-Rogers, W.** (Eds.). (2017). *The SAGE handbook of qualitative research in psychology*. Thousand Oaks, CA: SAGE.

References

Baumeister, R. F., & **Leary, M. R.** (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117, 497–529.

EEOC. (1990). *Policy guidance on current issues of sexual harassment*. Retrieved from https://www.eeoc.gov/policy/docs/currentissues.html

Hyers, L. L.Swim, J. K., & **Mallett, R. K.** (2006). The personal is political: Using daily diaries to examine everyday gender-related experiences. In **S. N. Hesse-Biber** & **P. Lina Leavy** (Eds.), *Emergent methods in social research* (pp. 313–335). Thousand Oaks, CA: SAGE.

Mallett, R. K., & **Melchiori, K. J.** (2014). Goal preference shapes confrontations of sexism. *Personality and Social Psychology Bulletin*, 40, 646–656.

Melchiori, K. J., & **Mallett, R. K.** (2018). Examining responses to sexism using a high-impact, high-psychological realism lab study. *SAGE Research Methods Cases*. http://dx.doi.org/10.4135/9781526439031

Reis, H. T., & **Wheeler, L.** (1991). Studying social interaction with the Rochester Interaction Record. In **M. Zanna** (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 269–318). San Diego, CA: Academic Press.

Swim, J. K., & **Hyers, L. L.** (1999). Excuse me—What did you just say? Women's public and private responses to sexist remarks. *Journal of Experimental Social Psychology*, 35, 68–88.

Swim, J. K.Pearson, N. B., & **Johnston, K. E.** (2007). Daily encounters with heterosexism: A week in the life of lesbian, gay, and bisexual individuals. *Journal of Homosexuality*, 53, 31–48.

Woodzicka, J. A., & **LaFrance, M.** (2001). Real versus imagined gender harassment. *Journal of Social Issues*, 57, 15–30.

Quantitative Coding of Qualitative Data Gathered in a Lab Experiment