

A Methodological Approach to Traditional and Large-Scale Qualitative Research

Susan Zickmund, PhD

Director, Qualitative, Evaluation And Stakeholder Engagement (Qual EASE) Research Services, CRHC Data Center

CHERP, VA Pittsburgh Healthcare System

Methodology Seminar

- Goal for the first part of the talk is to describe what is needed to accomplish traditional qualitative research
 - My goal here is to have a stand alone methodology lecture describing the basics of qualitative data collection and analysis in the health sciences

Methodology Seminar

- The second half of the talk I will exemplify an approach to large-scale qualitative methods by discussing the analytic approach to a current VA study: Disparities in Satisfaction with Care (DISC)

Organization of the Seminar

- Questions for Principal Investigators to use to help conceptualize qualitative and mixed methods Research
- Unique challenges and approaches to large-scale qualitative data

Part I: Questions for Principal Investigators to Use to Help Conceptualize Qualitative and Mixed Methods Research

Why, When, What and How?

A Word on Definitions: Qualitative and Mixed Methods Research

- All mixed methods research involves qualitative data
- Mixed methods:
 - utilizes multiple methods (e.g., intervention trials and in-depth interviews)
 - intentionally integrates or combines these methods

Why?

- It is important to have a clear rationale **why** qualitative research is being used within the larger study
 - This rationale needs to be explicitly stated within a proposal or article
 - Each layer of qualitative data collection needs to be justified

When?

- The Principal Investigator (PI) and qualitative expert should think about the entire project to determine **when** qualitative data collection and analysis should occur



Classic Mixed Methods Designs

- Convergent design (e.g. “to merge concurrent qualitative and quantitative data to address study aims”)
- Sequential design (e.g. “to have one dataset build on the results of the other”)

Mixed Methods Design

- The vast majority of mixed methods studies are **sequential**
- There are different types of data collected within a mixed methods study (e.g. qualitative, survey-based, chart review)
- The PI needs to decide **when** (e.g. during which year? During which phase?) the qualitative data collection will occur
 - During intervention development?
 - At the end of the intervention?
 - During dissemination of results?

What?

Types of Data Collection

- Typically, in clinical research, this involves a decision between:
 - Interviews
 - Focus groups



Rules of Thumb

- Whole courses have been written about this topic, but, practically, it is helpful to keep in mind when determining **what** type of data to collect...
 - Interviews typically produce the best data (in-depth, individual, uniquely tailored to capture specific information needed for a project)
 - Focus groups require less time (typically), fewer transcriptions, and less overall analytic work

Interviews

- Considered the gold standard in qualitative health research
- If distance is an issue (e.g. multi-site studies; participants are dispersed) then telephone-based approaches to data collection are best
- Less administrative structure is needed, but the analytic costs of transcription and coding are higher



Focus Groups

- Focus groups are face-to-face sessions which capture the discussion between participants as they discuss a given topic
 - Important to note that they require support at a site (e.g. room scheduling, delivery of food)
 - Are more difficult for multi-site studies with limited RA support
 - Are very difficult to use for collecting data from providers with busy schedules

Layers of Data Collection

- Different types of data collection can occur within a single study
- EX: With PCORI applications, early phase stakeholder engagement designed to tailor interventions *may* be best-served by using focus groups
- EX: Later phase experiences with an intervention *may* be best-served by using interviews

Qualitative Sample Size for Interview Studies

- Focus is on “thematic saturation” – the idea that once new themes no longer arise, data collection is complete
- For a single “cell” study (e.g. *views of Medicine fellows on their training*), saturation is likely achieved ~20-40 interviews
- For a study with multiple cells (e.g. *racial differences in urban versus rural diabetes patients*) sample size per cell is ~9-30 depending on overall size of the study and homogeneity of the cell
 - The more diversity, the larger the sample per cell must be

How?

Approaches to Qualitative Analysis

- Whole courses are written about this, but, at a minimum, this typically includes:
 - Qualitative methodology
 - Codebook development
 - Coding approach
 - Inter-coder reliability
 - Thematic analysis

Qualitative Methodology

- **All qualitative studies require a methodology**
 - A qualitative methodology is a specific philosophical approach to qualitative data that provides insight into how to collect, code and analyze the qualitative data
- A qualitative methodology should be clearly described and be an accepted approach

Qualitative Methodology

Frequently used qualitative methodologies in health services research include:

- Grounded Theory
- Ethnography
- Crabtree and Miller's Editing, Template, and **Quasi-Statistical approach** (developed in Family Medicine)

Codebook Development I: Open Coding

- Read/listen to the texts
- Use at least 20% of cases for initial open coding construction period
 - For a small study, may need 50-100%
- Goal is to record topics that come up frequently and that are important to the research question(s)

Codebook Development II: The Refinement Process

- Read the next 20% of the cases
- Keep (for now) all codes
- Note which ones:
 - Remain constant across cases (constant comparison)
 - Are important
 - Are new
- Repeat this step as needed for refinement

Codebook Development III: Define Codes

- Provide key definitions of codes, dates created (e.g. audit trail)
- Provide inclusion/exclusion criteria for codes and give clear and border examples

Satisfied	All codes under Satisfied are positive. (Things the subject likes about the VA)
▪ <u>General Satisfaction statement</u>	<p>Only use when they are needed to capture a new idea. E.g. “My nurses are personable and my doctor is good.”</p> <p>Do not use when no new idea would be captured by additional code. E.g. “My nurses are really personable. I have never had a problem with them.”</p> <p>Do not use for improvements question. E.g. “Nothing, I’m very satisfied” only code as Nothing; Keep doing what they’re doing. Do not use Very Satisfied</p> <p>Do not use if the subject is introducing or summing up a more specific comment by making a general statement. E.g. “They’re good people. They’re very professional and they’re very respectful, and they like their jobs.” Do not use Satisfied/Good</p> <p>“They say they’re going to do something and do it” Note: Does not specify how they do it/what it is that they’re doing</p>

Approach to Selecting Coders

- A PI needs to decide on:
 - The total number of coders (ideally 2)
 - Whether the coders include the PI, co-investigators or research assistants
 - The training process to train coders
 - The process for determining inter-coder reliability
 - Single coder (no inter-coder reliability)
 - Two coders with adjudication process
 - Two coders with adjudication process and kappa statistics

Interpreting Kappa Statistics and Reliability

- **0.00 = poor**
- **0.01-0.20 = slight**
- **0.21-0.40 = fair**
- **0.41-0.61 = moderate**
- **0.61-0.80 = substantial**
- **0.81-1.00 = almost perfect**
- **Rule of thumb: shoot for .70 and above**

Landis & Kock, Biometrics, 1977

Computer Data Management

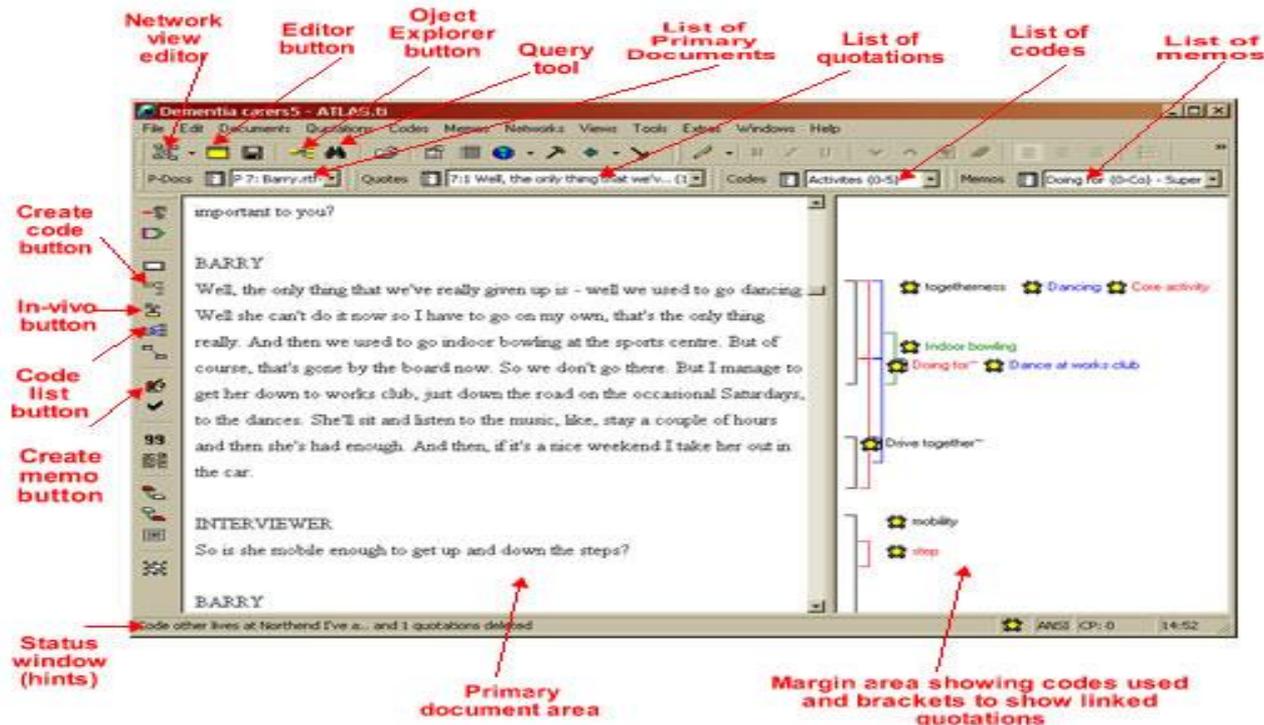
Software programs (Atlas.ti, Nudist) allow for computerized management of:

- Interview/focus group transcripts
- Codebooks
- Codes

Software enables a level of textual complexity not possible with notes alone

Traditional Data Management

Our Qualitative, Evaluation And Stakeholder Engagement (Qual EASE) Research Services in the Data Center uses Atlas.ti



Thematic Analysis

- The most typical way to code is to capture the major themes related to the research question
- This involves recording verbatim quotations categorized under a specific code
- A theme may be a single code (e.g. transportation barriers) or a combination of several codes (e.g. all barriers to adherence)
- Major themes are the ones included in an article

Part II: Approach to the Development of a Large-Scale Qualitative Study

A decorative graphic consisting of several horizontal lines of varying lengths and colors (teal, light blue, white) extending from the right side of the slide.

DISC Study Design

- Identify reasons for satisfaction or dissatisfaction with VA care based on race, ethnicity, and gender
- Original targeted N=1,350 from 25 VA Medical Centers
- Qualitative data: 1-hour interview
- Quantitative data: 5 surveys (e.g., Healthcare System Distrust, Race-Based and Gender-based Discrimination in Healthcare, and Cynicism) and health utilization data
- Focus on male/female gender by white, African American and Hispanic ethnicity
 - 6 cells per VA Medical Center
 - 9 participants per cell (lowest thematic saturation per cell)

DISC Study Sites and Interview Script Domains

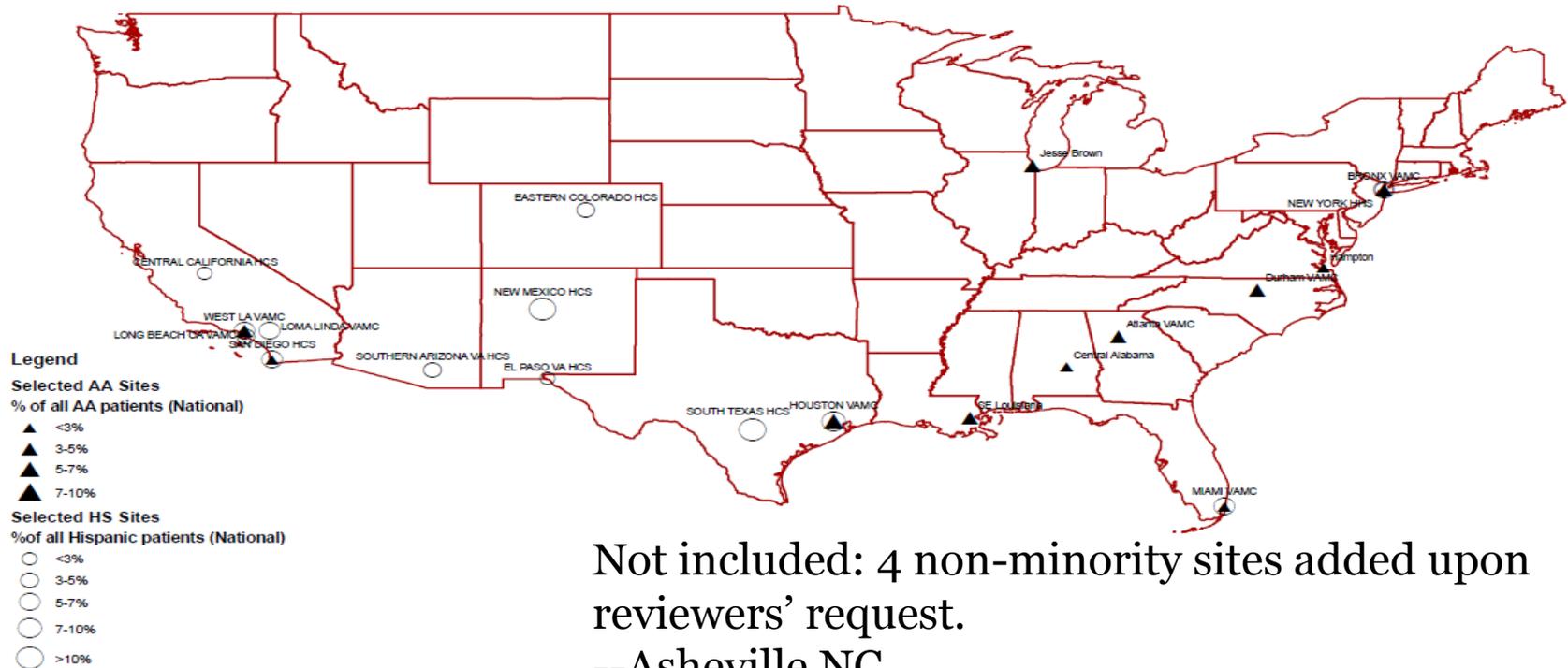


25 DISC VA Medical Center Sites

Facility Selection Criteria:

- Facilities with mostly high African American, high Hispanic, or moderately high numbers of both groups
- Geographic distribution across the country
- Additional 4 sites added at HSR&D reviewer request that served predominantly white Veterans

VA Facilities



Not included: 4 non-minority sites added upon reviewers' request.

--Asheville NC

--Grand Junction CO

--Prescott AZ

--Providence RI

DISC Domains

Entire Cohort

- Overall VA Care
- Outpatient VA Care
- Primary Care Provider--PCP
- Access
- Pharmacy
- Continuity/Coordination of Care
- Communication
- Respect
- Physical Facilities--Parent Station
- Cost

Subgroups

- Specialist
- Mental Health
- Pain Management
- Physical Facilities--CBOC
- Inpatient VA Care
- Women's Health

DISC Sample



DISC Study Enrolled Sample

Mailings Sent: 8,010

Eligible for Recruitment: 7,565

Interested in Learning More:
2,501

Eligible to Complete Interview:
1,928

Consented to Interview: 1,849

Consented to Voice Recording:
1,386

Completed Interviews: 1,207



Respondent Characteristics

Patient Characteristics	Study Sample (n=1207)
Gender	
Female	622 (51.5%)
Male	585 (48.5%)
Race/Ethnicity	
White	418 (34.6%)
African American	383 (31.7%)
Hispanic	390 (32.3%)
Other	16 (1.3%)
Age	
18-34	142 (11.8%)
35-44	140 (11.6%)
45-54	239 (19.8%)
55-64	340 (28.2%)
65-74	244 (20.2%)
75+	101 (8.4%)
VA User Only (Yes)	753 (62.4%)

DISC Qualitative Approach: Intercoder Reliability and Utterance Codebook



DISC Qualitative Codebook Development

- DISC requires a team of 7 coders to accomplish the needed coding
- The process of developing the codebook required input from all coders
- Initially two trained coders listened to ~200 interviews to develop the first draft of the codebook
- The coding team refinement with an additional ~100 interviews discussed during intensive meetings from August to November

Intercoder Reliability Process

- 7 coders operating in pairs
- Intercoder reliability conducted with cross coding between 3-4 teams
 - Inter-coder reliability with 20 per team for 5 rounds
 - a total of 300-350 interviews
- An adjudication process follows the coding of each interview within each team
- Large team meetings maintain consistency across pairs

Capturing Stories



How We Capture Stories

- Each domain (e.g. overall, access, specialist) has the same three elements
 - Five point Likert scale ranging from Very satisfied to Very dissatisfied
 - Each Likert scale is followed by the open-ended question: “What contributed to your rating for ____?”
 - Each domain ends with a final open-ended question: “What could the VA do to improve your satisfaction with ____?”

Satisfaction/Dissatisfaction and Improvement

Main Codes

Characteristics

Communication

Care

Access

Continuity/Coordination of Care

Facility

Cost

Non-Medical Services

Pharmacy

Dropdown Menu to Capture Employees Per Code

- **Unspecified/“They” (default)**
 - Employees/“Staff”
 - Doctor/Provider
 - Pharmacist
 - Receptionist/Clerk
 - Nurse
 - Scheduler
 - Student/Resident
 - Ancillary Staff
 - Other Veterans

Categories to Capture Services

- **Unspecified (default)**
- Specialty General
- Audiology
- Billing
- Cardiology
- Dental
- Dermatology
- Emergency Room
- Hepatology
- Inpatient
- Labs
- Mental Health
- Medical Records
- Optometry
- Outsourced contractor
- PCP
- Pharmacy
- Phone system
- VA Health Program
- VA system as a whole
- Women's Health

Dropdown Menu to Capture Location

- **Unspecified/Main DISC VA (default)**
- CBOC – Community-based outpatient clinic
- Other VA (Non-DISC site)



Coding Example

Location--
CBOC

Global
code/MST

Facility/women
do not have own
space

Employee--
unspecified

Services--
unspecified

- “The whole aspect of being a woman Veteran. For example, at the [__ VA Clinic] it’s kind of funny but they have a little room that says ‘Women Only, Veteran’s Healthcare’ and the whole thing about that came about because of MST and staff goes by there, left and right, whatever. I’ve even left suggestions in their box. And you go by there—I wouldn’t sit in there--there’s men in there. So you know, it’s kind of pointless. And they’re not aware.”

Computerized System to Capture Complex Coding

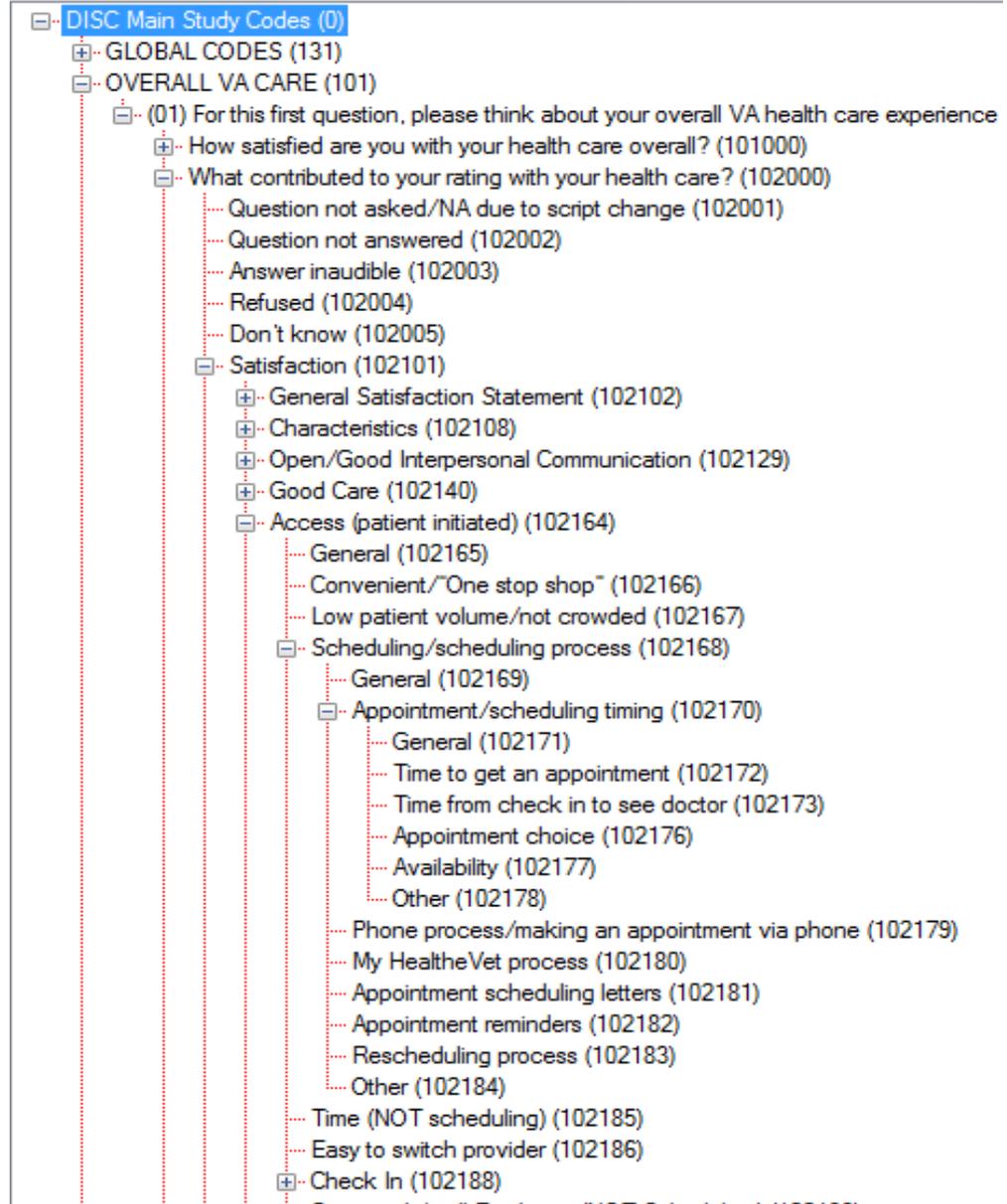
The screenshot displays a web-based application window titled "CodingEntry_pending_fe : Database (Access 2007 - 2010) - Micro". The interface includes a navigation pane on the left and a main content area. The main content area is divided into several sections:

- Coding Entry Header:** Includes a dropdown for "Select SampleID-StaffID" (value: 88888-01) and a button "View All Code Entries, by entry order".
- Form Fields:** "SampleID" (88888-01), "DateStarted:" (12/18/2014 8:09:34 AM), and "Coding Completed:" (DOE).
- Appt. Summary:** A text input field.
- Domain:** A grid of buttons for different care domains: 10-CONTINUITY OF CARE, 11-COMMUNICATION, 12-HUMANENESS (RESPECT), 13-PHYSICAL FACILITIES (PARENT SITE), 14-PHYSICAL FACILITIES (CLINIC/SUBS), 1-GLOBAL CODES, 2-OVERALL VA CARE, 3-OUTPATIENT VA CARE, 4-COMPETENCE (PRIMARY CARE PROVIDER), 5-COMPETENCE (SPECIALISTS), and 6.
- Question Section:** A heading "(01) For this first question, please think about your overall VA health care experience at [insert VA medical center/clinic]". Below it, "Individual Question(s):" lists several questions, including "4-How satisfied are you with your health care overall?" and "2-What contributed to your rating with your health care?".
- Filter Section:** A text input field for filtering code boxes and a "Do Not Filter/Clear Filter" button.
- Quote Section:** A "Quote" field with a dropdown menu showing various satisfaction statements and characteristics.

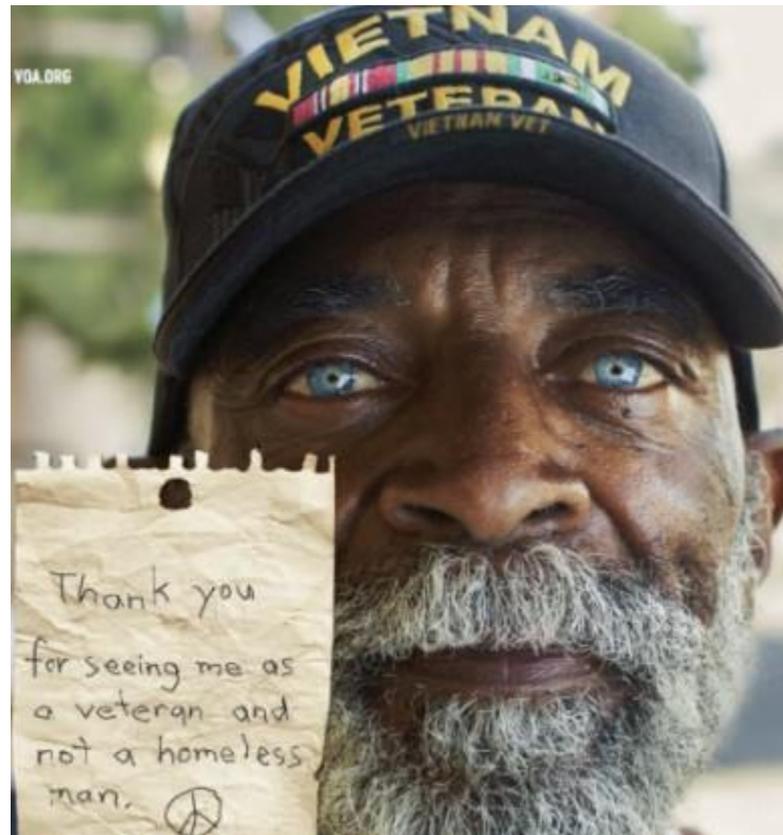
Example: Overall VA Care – “What contributed to your rating with your [Overall VA Care]?”

Coding Hierarchical Structure

- Codes stored as numeric variables
- Hierarchical structure maintained in metadata
 - Upward collapsing



In Conclusion



What We Can Learn from Large-Scale Qualitative?

- Size allows for complex analytic plans
- Can show the ratio of satisfaction/dissatisfaction themes per Likert rating (e.g. very satisfied, somewhat satisfied)
 - Can examine which qualitative themes predict Likert rating
- Can drill down to the type of employee or service most associated with positive or negative themes

VAPHS Study Co-I Team (It Takes a Village)

VA Pittsburgh Co-Investigators:

- Sonya Borrero, MS, MS
- Kelly Burkitt, PhD
- Shasha Gao, PhD
- Leslie Hausmann, PhD
- Scott Obrosky, MS
- Keri Rodriguez, PhD
- Roslyn Stone, PhD
- Galen Switzer, PhD
- Michael Fine, MD, MS (Senior)



Questions?

