Using Qualitative Methods in Empirical Studies of Software Engineering

Carolyn Seaman

University of Maryland Baltimore County

Fraunhofer USA Center for Empirical Software Engineering Maryland

ESELAW 2009

12 November 2009

São Carlos-SP, Brazil

UMBC

AN HONORS UNIVERSITY IN MARYLAND

Definitions

- Qualitative data - data in the form of text and pictures, not numbers
- Qualitative analysis – analysis of qualitative data in order to discover trends, patterns, and generalizations
- Grounded theory – theory formed bottom-up from the (usually qualitative) data
- Rich data – data that includes a lot of explanatory and context information

© Carolyn Seaman, 2009

UMBC
AN HONORS UNIVERSITY IN MARYLAND

Why Qualitative Methods?

- Problem: Difficult to answer complex SE questions with a purely quantitative approach because
  - Working with human subjects
  - Typically have small sample sizes
  - Experiments are expensive to run
  - Need some support for a hypothesis before investing effort in full experiment
- Solution: Use a qualitative approach that includes a quantitative aspect

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Types of results

A qualitative study will result in:

- Propositions tied to a trail of "evidence"
- Well-grounded hypotheses
- Complex findings that incorporate the messiness of the phenomenon under study
- Explanations
- Areas for future study

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Types of Research Questions

Qualitative methods are most appropriate when:

- Subject of study involves human behavior
- No concrete hypotheses
- Variables hard to define or quantify
- Little previous work
- Quantitative results may be hard to interpret

© Carolyn Seaman, 2009

UMBC
AN HONORS UNIVERSITY IN MARYLAND

Advantages to Researchers

- Richer results
- Results more explanatory
- Closer to sources of data
- Avoid errors in interpretation

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Advantages to Practitioners

- Richer, more relevant results
- Terminology of results
- More part of the research process
- Opportunity to clarify and explain findings

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Overview of Techniques

Data Collection

- Prior Ethnography
- **Participant Observation**
- **Interviewing**
- Surveys
- Document Analysis

Data Analysis

- Coding
- **Constant Comparison Method**
- Cross-case analysis
- Member checking
- Auditing

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Participant Observation

Definition: non-covert direct observation of phenomenon

Example: Observation of code inspection meetings

- collected both qualitative and quantitative data
- did not participate in the inspection
- used data forms as well as field notes

© Carolyn Seaman, 2009

UMBC
AN HONORS UNIVERSITY IN MARYLAND

Observation Data Form

Inspection Data Form

Class(es) inspected Inspection date: Time:

Author:

Moderator:

Reviewers:

Name	Responsibility	Preparation time	Present

Amount of code inspected:

Complexity of classes:

Discussion codes:

D = Defects Q = Questions C = Classgen defect U = Unresolved issues G/D = Global defects G/Q = Global questions P = Process issues A = Administrative issues

M = Miscellaneous discussion

Time logged (in minutes):

D_____ Q_____ C_____ U_____ G/D_____ G/Q_____ P_____ A_____ M_____

© Carolyn Seaman, 2009

UMBC
AN HONORS UNIVERSITY IN MARYLAND

Field Notes Example

The "step" function is a very important but complicated function. [Reviewer1] did not have time to review it in detail, but [Author] said he really wanted someone to go over it carefully, so [Reviewer1] said she would later.

There was a 4-minute discussion of testing for proper default values. This is a problem because often the code is such that there is no way to tell what a particular variable was initialized to. [Reviewer2] said "I have no way to see initial value". This was a global discussion, relevant to many classes, including [Reviewer2]'s evidently.

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Interviewing

- Interviews are good for getting
  - opinions
  - feelings
  - goals
  - procedures (both formal and informal)
- not facts

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Standard Interview Formats

- **Structured** (standardized)
  - Tightly scripted, almost verbal questionnaire
  - Replicable, but lacks richness
  - Analyze like questionnaire
  - *"How many times a day do you access the internet?*

    [0, 1-5, 5-10, 10-15, 15+]"

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Standard Interview Formats

- **Unstructured**
  (Open/Informal/Conversational)
  - Guided by a very scant script.
  - Rich, but not replicable.
  - Difficult to be systematic, problem of coverage.
  - Minimize interviewer effects, preserves interviewee point of view.
  - Interviewee led, interviewer probes.
  - *"Please, tell me about your internet usage..."*

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Standard Interview Formats

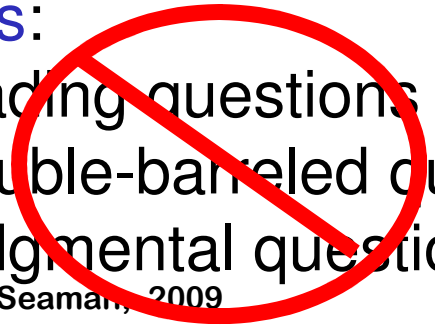
- **Semi-structured**
  - Guided by a script (interview guide), but interesting issues can be explored in more depth.
  - Good balance between richness and replicability.
  - Mixed analysis techniques.
  - *"In a typical day, how often do you use the internet?"*

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Interview questions

- Closed
  - Predetermined answer format (e.g. Yes/No)
  - Easier to analyze
- Open
  - No predetermined answer format
  - More complete response
- Combination
  - Closed, with opportunity to elaborate
- Probes
- Pitfalls:
  - leading questions
  - double-barreled questions
  - judgmental questions



UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Interview Guide

- A script for use by interviewer only
- “Wish list” vs. structured
- Flow/direction to interview
- Required topics
- Transitions between topic areas
- Important for replicability
- Wording and sequence are critical

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Interview Design Considerations

- Context switching
- Flow between open and closed questions
- "**Shape**" of interview
- Most important stuff first
- Wording

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Interview Shapes

- Funnel
  - Begin with open, gradually become more closed
  - Good if you're not sure what you're going to get
- Pyramid
  - Begin with closed, gradually become more open
  - Good with nervous interviewees
- Hour glass
  - Begin with open, gradually become more closed, then open up again at end to pick up things you might have missed
  - Good if you know what you want, but suspect there are important things you don't know about yet

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Example Interview Guide

Interview Guide 2a: In-depth project interviews
Who: Developers on [Project1], [Project2], [Project3]
Subjects covered: general opinions of GSS processes and products
Duration: 60-90 minutes

What do you like about the current process using GSS?

What do you dislike about the current process using GSS?

Do you depend on any other groups, either for information or help with GSS, or for work to be done related to GSS?

What do you like about the applications resulting from using GSS?

What do you dislike about the applications resulting from using GSS?

Have there been any problems with the interface between GSS and other COTS products?

What do you see as the top risks associated with the use of GSS? How would you mitigate these risks?

UMBC
AN HONORS UNIVERSITY IN MARYLAND

Interviewing Pointers

- give clues about the level of detail you want
- establish rapport, but be subject neutral
- avoid jargon, esp. academese
- dispel any notion of the "right" answer
- play the novice when appropriate
- probe, but do not lead
- always be aware of your biases
- be sensitive to their work (environment/schedule)
- no more than 60 minutes
- let interviewee know next steps
- end with "anything else I should know?"
- say Thank you!

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Recording of interviews

- Audiorecording
- Notetaking
- Scribing

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Audiorecording

- Best memory mechanism
- Full transcription or just verbatim quotes
- Still take notes
  - Tapes fail, digital files are deleted
  - Does not record all aspects (esp. context / facial expressions)
- Required consent
  - Always ask first.
  - Do NOT hide recorder, keep it visible at all times.
  - Give the option to turn it off at any point.

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Notetaking

- **Very** hard to take notes and interview at the same time
- There are some super-researchers who can do it
- Inevitably results in incomplete notes
- Slows down the interview
- Sometimes inevitable

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Scribing

- Partner-based interviewing
- Advantages of a single contact vs. trading-off
- Can share roles (interviewer/scribe)
  - BOTH take notes, though to different degree
- Group debrief: what did you get/miss?
- Synchronize notes: overlap and emphasis
- Clarify while it is still in your head

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Writing up the interview

ASAP!!!!

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Interview Notes

- Write it up immediately
- Descriptive vs. reflective notes
- Use Observer's Comments
  - Impressions, state of mind, assumptions, notes to self
- How detailed?
  - Verbatim transcript
    - only possible with audiorecording
    - Extremely labor-intensive
  - Summaries with major points quoted
    - OK, but use LOTS of quotes
  - Start closer to verbatim at the beginning of a study

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Interviewing Exercise

- **Background**:
  - The National Federation of Makers of Feijoada (FNFF) is concerned that the national consumption of feijoada is declining due to decreasing quality of feijoada.
  - So they have asked us to interview the top feijoada chefs in the country (as determined by regional competitions)
  - The goal is to find out the secrets to master feijoada making, so that it can start to be taught in elementary schools.

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Interviewing Exercise

- Three versions of the interview guide
- I will be the interviewer
- You will be the interviewees
  - So take a moment to think of your favorite feijoada chef

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

1. How often do you make feijoada and how long does it take you?
2. What do you think makes your feijoada the best?
3. Of course, you always wash your hands thoroughly before you start, right?
4. Do you add the sausage near the beginning or near the end of the cooking?
5. What kind of pot do you use?

3. How long does it take to make feijoada?
4. What are the ingredients you use?
5. What do you think makes your feijoada the best?

- Switching from topic to topic
- Switching between open and closed

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Constant Comparison Method

- Qualitative analysis method
- Meant to generate grounded theory
- Operates on a set of field notes
- Basic process:
  - coding
  - grouping
  - writing field memo
  - forming hypotheses
- Repeated periodically in parallel with data collection

© Carolyn Seaman, 2009

UMBC
AN HONORS UNIVERSITY IN MARYLAND

What's a Code?

- A label
- A concept
- A topic
- A category
- A relationship
- A theme

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

What's Coding?

- Open coding - assigning codes to pieces of textual data
  - Coded "chunks" can overlap
  - One chunk can have several codes
- Axial coding - grouping, categorizing, combining coded chunks
- Selective coding - making sense of it

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Open Coding

What's here? What are the pieces?

- Identification/discovery of concepts
- Classification (labeling of phenomena)
- Abstraction (this is part of that)
- Comparative analysis (this is different from that)
- Categorization (organization, grouping)
- Value-neutral, at least initially
  - "complexity" not "high complexity" or "low complexity"

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Open Coding Process

- Preparing for coding
  - Read the data
  - Read background material and research design
  - Create pre-formed codes, if applicable
- Coding by hand
  - Document markup (colored pens, etc.)
  - Photocopy, scissors, and envelopes
  - MS Word comments
  - Excel
- Coding tools – NVivo, Atlas TI
- Coding scheme
  - *Pre formed* or *post formed* codes
  - Constant iteration
  - Structure develops over time

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Open Coding Exercise

- Background:
  - Study of the role of documentation in software maintenance
  - Interviews with experienced software maintainers in several organizations
- Process:
  - I'll show you an example
  - Then you'll try it – code one excerpt with one code
  - Find a partner – compare your codings
  - I'll show you my coding of the excerpt

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Coding Scheme

Respondent Background
Information Gathering
Transition to maintenance
T
C
C
P
M
C
L
I
H
Quality of Process
Great Quotes

Human Sources of Information

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Open Coding and Quantification

- One form of coding
- Objective is to derive quantitative data from qualitative data for future statistical analysis
- Usually involves counting
  - How many subjects said…?
  - How many times did subjects use the term …?
  - How many times did …?
- Or timing
  - How long did subjects spend doing…?
  - How long did it take to …?
- Inevitably loses richness
- Often seems a little like missing the point
  - What's the point of collecting rich data when you're just going to condense it down to numbers?
- But often is an effective and necessary way to reduce the size of the data

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Inspection Data Form

Class(es) inspected: Inspection date: Time:
Author:
Moderator:
Reviewers:

Name	Responsibility	Preparation time	Present

Amount of code inspected:
Complexity of classes:

Discussion codes:

D **Defects**
Reviewer raises a question or concern and it is determined that it is a defect which the author must fix; time recorded may include discussion of the solution

Q **Questions**
Reviewer asks a question, but it is not determined to be a defect.

C **Classgen defect**
Reviewer raises a defect caused by classgen; author must fix it, but it is recognized as a problem to eventually be solved by classgen

U **Unresolved issues**
Discussion of an issue which cannot be resolved; someone else not at the meeting must be consulted (put name of person to be consulted in () beside the code); this includes unresolved classgen issues. It also includes issues which the author has to investigate more before resolving.

G/D **Global defects**
Discussion of global issues, e.g. standard practices, checking for null pointers, which results in a defect being logged (does not include classgen defects)

G/Q **Global questions**
Same as above, but no defect is logged

P **Process issues**
General discussion and questions about the inspection process itself, including how to fill out forms, the order to consider material in, etc., but not the actual execution of these tasks.

A **Administrative issues**
Includes recording prep time, arranging rework, announcing which products are being inspected, silence while people look through their printouts, filling out forms.

M **Miscellaneous discussion**

Time logged (in minutes):

D_____ Q_____ C_____ U_____ G/D_____ G/Q_____ P_____ A_____ M_____

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn S

freshly painted room – smells + is hot
just had a task meeting – 39 classes needed in 6 weeks
SM: "This is a nightmare, and it's going to get worse."
– started 30 minutes late because of meeting

Class(es) inspected: ANI.3, EV3, EV3.1 **Date:** 3/15/96 **Time:** 2:00 **Page** 1 **of** 2

Time	Participants	Code	Notes
30	SM	A	get started; SM having problems finding right files
31	AP → RK	G/D	0 change to null – actually several different small defects
33	AP, SM, MI	Q	don't change now, wait for TB3
34	AP	D	couts
35	MI	G/D	
36	MI	Q	"good thorough test plan" – some FYIs – not standard format – do for next TB – other style – don't take time now
38	MI → RK, SM	Q	MI went through everything she did – no defects – showed RK+SM something on paper – don't change for now
40	RK → M	Q	re DB filename
41	SM	A	nothing on category
42	SM → RK	G/D	null instead of 0 – had trouble finding it
44	SM → RK, MI	D	ParameterError exception – trying to figure out where it's thrown
46	SM → RK	U	similar to above "This leads me to my BIG QUESTION" –SM
47	SM → RK, MI	U (RK+SM)	RK catching error that will never happen MI: you're making it a lot more complex than you need to – too much error checking discussion of meanings of various parameters – MI: "action item for the 2 of you to battle out"
53	RK → SM	Q	why is certain error generated by classgen?
55	RK → MI, SM	Q	clarification
56	SM → RK, MI	D	ParameterError – handle differently from the way classgen does it

"Ken already gave you his stuff, correct?" – SM to RK
yes

I'm having a hard time concentrating

CA standing up by door

EV3

EV3.1

43 CA leaves

printout very small – hard to read
– SM + AP talking off their gtb's

47 CA comes back

MI gave RK marked up copy of test plan

Lots of time for everyone trying to find right place in printout – small print is a factor

© Caro

UMBC
AN HONORS UNIVERSITY IN MARYLAND

Axial Coding

How are things related?

- Initial process of reassembling
- Relationships among categories and codes
- Structure (why?)
- Process (how?)
- Explanations not causal prediction

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Selective Coding

How does it all fit together?

- Also called **sense making**
- Relationships among relationships
- Theory construction
- The central category
- Storyline memos
- Role of literature
- Write, write, write!!!
- **Field Memos**

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Field Memos

- The "*single most powerful analytical tool*" for qualitative researchers
- Simply, a piece of writing
- Maybe will later become part of a report, maybe will be thrown out
- Summarizes and synthesizes:
  - A proposition
  - An open question
  - A chain of evidence and logic
  - The complexity of a concept
  - Rich description
- Version control and organization

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Judging Validity

- Validity of methods
  - Triangulation
  - Documentation
  - Contradictory evidence
- Weight of evidence
  - How much is enough?
  - Variety as well as quantity of evidence

UMBC
AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009

Using Qualitative and Quantitative Methods Together

- Qualitative and quantitative methods best used in combination
- Can simply be used in parallel to address the same research questions
- There are other strategies to better exploit the strengths and weaknesses of the methods

© Carolyn Seaman, 2009

UMBC
AN HONORS UNIVERSITY IN MARYLAND

Example Design 1: Statistical Hypothesis Testing with Follow-up Interviews

- Classic design – often done without fully exploiting the interview data
- Example scenario:
  - Blocked subject-project experiment to evaluate a new testing technique
  - Statistical results show that technique is more effective on some applications than on others
  - Qualitative results show why

© Carolyn Seaman, 2009

UMBC
AN HONORS UNIVERSITY IN MARYLAND

Example Design 2: **Using Grounded Theory to Identify Variables**

- Want to evaluate a new technique, but not sure what the evaluation criteria should be
- Example scenario:
  - Evaluating a collaborative design process
  - Use participant observation of design meetings to generate hypotheses about properties of the resulting designs
  - Grounded hypotheses are used to design a quantitative evaluation of the resulting designs

© Carolyn Seaman, 2009

UMBC
AN HONORS UNIVERSITY IN MARYLAND

Example Design 3: Using Prior Investigation to Operationalize Variables

- Relevant variables are known, but the range and types of values is difficult to specify
- Example scenario:
  - Want to study the relationship between developer experience and types of defects
  - First use interviews to identify the range of developer experience (in its complexity) and a taxonomy of defect types
  - Quantitative study then is much more effective when using this operationalization

© Carolyn Seaman, 2009

UMBC
AN HONORS UNIVERSITY IN MARYLAND

Conclusions

- Empirical software engineering researchers are addressing more and more complex research questions that have increasingly human elements
- Qualitative methods, usually in combination with quantitative methods, can be helpful in handling this complexity
- Qualitative methods are both flexible and rigorous
- Qualitative analysis provides richer, more relevant, and more explanatory results
- The most effective research designs combine qualitative and quantitative methods

© Carolyn Seaman, 2009

UMBC
AN HONORS UNIVERSITY IN MARYLAND

Bibliography

- Day, I. *Qualitative data analysis: A user-friendly guide for social scientists.* New York: Routledge. 1993.
- Eisenhardt, Kathleen M. "Building theories from case study research." Academy of Management Review, 14:4, 1989.
- Gilgun, J.F., "Definitions, Methodologies, and Methods in Qualitative Family Research." *Qualitative Methods in Family Research*. Sage, 1992. pp. 22-29
- Glaser, Barney G. and Anselm L.Strauss. The Discovery of Grounded Theory: Strategies for Qualitative Research. de Gruyter, 1967.
- Jick, Todd D. "Mixing qualitative and quantitative methods: Triangulation in action." Administrative Science Quarterly, 24:4, 1979.
- Judd, Charles M., Eliot R. Smith, and Louise H. Kidder. Research Methods in Social Relations. Harcourt, Brace, Jovanovich, 1991.
- Miles, Matthew B. and A. Michael Huberman. Qualitative Data Analysis: An Expanded Sourcebook. Sage Publications, 1994.

UMBC

AN HONORS UNIVERSITY IN MARYLAND

© Carolyn Seaman, 2009