Freie Universität Berlin

When Grounded Theory Methodology is Not Enough

Additions for Video-Based Analyses
of Software Engineering
Process Phenomena

Franz Zieris
zieris@inf.fu-berlin.de

Qualitative Research in a Nutshell

The Qualitative Approach:

- **Naturalistic inquiry of a part of social reality** (rather than laboratory settings)
- **Open research design & purposeful sampling** (rather than fixed plan & random sampling)
- **Holistic perspective & rich data** (rather than simple cause-effect measures)
- **Develop/discover theories** (rather than test given theories)

based on [Patton, 2002] & [Flick et al., 2004]

Grounded Theory Methodology:

- **Theoretical Sampling**
- **Theory-Oriented Coding**
  - Open Coding (conceptual labels)
  - Axial Coding (interaction model)
  - Selective Coding (narrative, context)
- **Constant Comparison**
- **Memo Writing**
- **Non-Linear Process** of data collection, coding, and writing

[Strauss & Corbin, 1990]

as summarized by [Przyborski & Wohlrab-Sahr, 2014]

Motivation for Amending the GTM

- Basic properties and coding perspectives are useful, *but:*
- **Difficulties** of applying GTM
  - Some due to unspecified aspects
  - Others due to assumptions

- GTM alone is not enough → **Additions**

Our Research Case:

"Understand how Pair Programming (PP) works"

(Why? Meta-analysis of PP effects from controlled studies:
Mere tendencies, lot of unexplained variation)

- Next: Five research stages to illustrate problems and solutions

Data Collection: Interviews?

Data for GTM:

- In principle: *"All is data"* [Glaser, 2007]
- Actual: focus on interview transcripts

Problem:

- Interviews are not naturalistic
- Can practitioners explain their PP process in an interview?

Solution: Combine observations and interviews

- Primary: Record PP sessions (screen, audio, webcam)
  - Capture aspects which the subjects are not aware of
  - Repeatable in-depth analyses
  - Less biased than field notes
- Secondary: Reflective interview with pair afterwards
  - Capture subjects' perspective

Data Collection: "Smash & Grab"? [Dey, 1999]

Data Collection in GTM:

- Opportunistic, be open, adjust on site
- Save time to not need to come back

Idea: Visit company, record many sessions

Problem:

- Lack of context makes interpretation difficult
- PP for researcher's sake: not naturalistic
- What about one-off behaviors?

Solution:
Stay Around & Come Back

- Don't hurry to finish data collection
  - Stay at companies for longer than just main data collection
  - Water cooler discussions with developers
  - Understand company and team climate in which the PP sessions happen
- Involve participants in study
  - Return with results

Theoretical Sampling

Theoretical Sampling in GTM:

- When research need arises: collect additional data with special properties
- But: Purposeful sampling *"can also be difficult if you do not have unlimited access to sites, persons, or documents"* [Strauss & Corbin, 1990]

Problem:

- Building trust with a company takes time.
- Then: How to find a PP session with desired properties?

Solution:
Data Repository

- Over time: Build stock of reusable data (naturalistic, rich, with context information)
  - Repository **PP-ind**
  - Since 2007: 13 companies, 57 developers in 67 PP sessions, avg. 1:35 hours

  [Zieris & Prechelt, 2020b]
- Then: theoretical sampling from repository

Analysis: How to Code?

Going through Data in GTM:

- Open Coding: label data as to *"what it is"*
- Filter by (implicitly): theoretical sampling, selective coding, theoretical sensitivity

Problems:

- What am I looking at?
  - Industrial software development is complex, even more with two experts talking about it
- What am I even looking *for*?
- What *is* it that I see?

Solution:
Define a Perspective

1. **Filter**: In which regards do I expect the data to yield insights?
2. **Epistemology**: What kind of interpretations do I allow myself to make?
3. **Goal**: What kind of result do I aim for? (e.g. coding scheme or full theory)

[Salinger et al., 2008]

Analysis: Develop a Theory?

Goal of GTM:

- Integrated theory with saturated and fully grounded categories
  - from square one to dissertation

Problem:

- How to integrate work of more than one study and/or researcher?
- Is a full theory really necessary?
  - see ← *Define a Perspective*

Solution:
Reusable Concepts

- Develop low-level, generic-but-domain-specific concepts first (this takes time!)
  - Base Layer: ~70 well thought-out concepts, "atoms" of all PP processes [Salinger & Prechelt, 2013]
  - Groundwork for specialized PP topics (e.g. knowledge transfer, decision making)
- Reuse them in later studies when fit
  - Knowledge Transfer Episodes (ESEM '14)
  - Resynchronization Behavior (ICSE-SEIP '16)
  - Overall PP Session Dynamics (ICSE '20)

  [Zieris & Prechelt, 2014; 2016; 2020a]

Filling the Gaps in the Methodology

- Open Aspects in the GTM
  - How to perform naturalistic inquiry beyond interviews? →  Combine Observations & Interviews
  - Unclear role of the researcher →  Stay Around & Come Back

- Assumptions in the GTM
  - Access to data →  Maintain and Sample from Repository
  - Easy-to-understand data →  Define Perspective on Data
  - Self-contained research →  Work with Reusable Concepts

Thank you!

References

On Qualitative Research

[Strauss & Corbin, 1990] *Basics of Qualitative Research. Grounded Theory Procedure and Techniques* (Sage Publications)

[Dey, 1999] *Grounding Grounded Theory: Guidelines for Qualitative Inquiry* (Emerald Group Publishing)

[Glaser, 2007] *All Is Data* (Grounded Theory Review, Vol. 6, Issue 2)

[Patton, 2002] *Qualitative Research and Evaluation Methods* (3rd Edition, Sage Publications)

[Flick et al., 2004] *A Companion to Qualitative Research* (Sage Publications)

[Tracy, 2010] *Qualitative Quality: Eight "Big-Tent" Criteria for Excellent Qualitative Research* (Qualitative Inquiry, Vol. 16, Issue 10)

[Przyborski & Wohlrab-Sahr, 2014] *Qualitative Sozialforschung. Ein Arbeitsbuch* (4. Ausgabe, Oldenbourg Verlag)

Our Research

[Salinger et al., 2008] *A Coding Scheme Development Methodology using Grounded Theory for Qualitative Analysis of Pair Programming* (Human Technology: An Interdisciplinary Journal on Humans in ICT Environments, Vol. 4)

[Salinger & Prechelt, 2013] *Understanding Pair Programming: The Base Layer* (Books on Demand, Norderstedt)

[Zieris & Prechelt, 2014] *On Knowledge Transfer Skill in Pair Programming* (Proc. 8th ESEM 2014)

[Zieris & Prechelt, 2016] *Observations on Knowledge Transfer of Professional Software Developers During Pair Programming* (Proc. 38th ICSE 2016 Companion)

[Zieris & Prechelt, 2020a] *Explaining Pair Programming Session Dynamics from Knowledge Gaps* (Proc. 42nd ICSE 2020)

[Zieris & Prechelt, 2020b] *PP-ind: A Repository of Industrial Pair Programming Session Recordings* (arXiv:2002.03121 [cs.SE])

Images

 Icon "Qualitative research" by Template from the Noun Project
 Icon "Pair Programming" by Creative Stall from the Noun Project
 Icon "combine" by vigorn from the Noun Project
 Icon "Switch positions" by Gregor Cresnar from the Noun Project
 Icon "documentation" by lastspark from the Noun Project
 Icon "see" by Deivid Sáenz from the Noun Project
 Icon "decomposition" by Arthur Shlain from the Noun Project