Case Studies: Triangulation in Software Projects

Greg Hartman
gghartma@cs.cmu.edu
12 April 2005

Case studies are essential validation tools

- Software Engineering is concerned about **cost** and **timeliness** in addition to quality
  - There is often no empirical basis for decisions about these attributes
- These attributes are sensitive to the size of the project
  - Results from a 10K-line system may not apply to a 500K-line system
- Theories don’t reliably predict these attributes
- Traditional experiments are prohibitively expensive when applied to large projects

Case studies compliment experiments

- **Experiments**
  - Tend to be quantitative, but can be qualitative (surveys)
  - Small scale
  - Lab setting
  - Variables tightly controlled
  - Replication = identical results
  - Validity through construction, sample size
  - Causality and correlation

- **Case studies**
  - Often combine quantitative and qualitative measures
  - Large scale
  - Industrial setting
  - No control of variables
  - Replication = different results
  - Validity from multiple sources of data (triangulation)
  - Neither correlation or causality

Triangulation strengthens a case study

- A case study relies on multiple data sources
  - This sets it off from experience reports and examples, which tend to rely on a small number of sources
- Ideally, the data sources are chosen so they are subject to different uncontrolled variables
- If many sources support the hypothesis, our confidence improves
- Dissenting sources often point to interesting problems
  - These may motivate new experiments and studies
- Multiple case studies can also help
Examples of diverse data sources

- Tradition metrics
  - Lines of code
  - Time tracking
  - Defects
- Design documents
- Developer diaries
- Developer interviews
- Logging system use
- Customer feedback

Case studies are a summative technique

Example of triangulation

- Dingsøyr and Røyrvik, 2003
- Type of paper: experience report with multiple data sources
- Data sources employed
  - Interviews
  - Observation (logbooks, screenshots and pictures)
- Conflicting views of company structure
  - Observation of a flat structure, interviews indicated a hierarchy
- Conclusion (somewhat tangential for our purposes)
  - Use of tool seemed to depend on job function
  - No consensus on use of hierarchy to structure data in the tool

Hybrid techniques do exist (method)

- Salo & Abrahamsson, 2004
- Assumes an iterative development environment
- Inject ideas gleaned from last development cycle
- Look for changes in the metrics
- Paper type: improved method (for studying development)
- The authors correctly point out that the results still don’t generalize
- The paper contains many practical suggestions for running these studies
Concluding points

• Case studies can be valuable tools

• Don’t underestimate the complexity of case studies
  – Choose data sources carefully
  – Minimize the impact on participants
  – Analyze while you’re collecting data

• Look for and explain discrepancies

• Look for opportunities to create experiments to generalize the results