28th International Conference
on Software Engineering

Tutorial F2

Case Studies for Software Engineers

Steve Easterbrook University of Toronto

Jorge Aranda University of Toronto

This tutorial was originally developed with:

Susan Elliott Sim University of California, Irvine
Dewayne E. Perry The University of Texas at Austin

28th International Conference
on Software Engineering

1. Empirical Methods in Software Engineering

Many methods available:

- Controlled Experiments
- Case Studies
- Surveys
- Ethnographies
- Artifact/Archive Analysis ("mining"!)
- Action Research
- Simulations
- Benchmarks

Controlled Experiments

experimental investigation of a testable hypothesis, in which conditions are set up to isolate the variables of interest ("independent variables") and test how they affect certain measurable outcomes (the "dependent variables")

- good for
  - quantitative analysis of benefits of a particular tool/technique
  - (demonstrating how scientific we are!)
- limitations
  - hard to apply if you cannot simulate the right conditions in the lab
  - limited confidence that the laboratory setup reflects the real situation
  - ignores contextual factors (e.g. social/organizational/political factors)
  - extremely time-consuming!

See:

Pfleeger, S.L.; Experimental design and analysis in software engineering. *Annals of Software Engineering* 1, 219-253. 1995

Case Studies

"A technique for detailed exploratory investigations, both prospectively and retrospectively, that attempt to understand and explain phenomenon or test theories, using primarily qualitative analysis"

- good for
  - Answering detailed how and why questions
  - Gaining deep insights into chains of cause and effect
  - Testing theories in complex settings where there is little control over the variables
- limitations
  - Hard to find appropriate case studies
  - Hard to quantify findings

See:

Flyvbjerg, B.; Five Misunderstandings about Case Study Research. Qualitative Inquiry 12 (2) 219-245, April 2006

Ethnographies

Interpretive, in-depth studies in which the researcher immerses herself in a social group under study to understand phenomena though the meanings that people assign to them

- Good for:
  - Understanding the intertwining of context and meaning
  - Explaining cultures and practices around tool use
- Limitations:
  - No generalization, as context is critical
  - Little support for theory building

See:

Klein, H. K.; Myers, M. D.; A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems. MIS Quarterly 23(1) 67-93. March 1999.

Action Research

"research and practice intertwine and shape one another. The researcher mixes research and intervention and involves organizational members as participants in and shapers of the research objectives"

- good for
  - any domain where you cannot isolate {variables, cause from effect, …}
  - ensuring research goals are relevant
  - When effecting a change is as important as discovering new knowledge
- limitations
  - hard to build generalizations (abstractionism vs. contextualism)
  - won't satisfy the positivists!

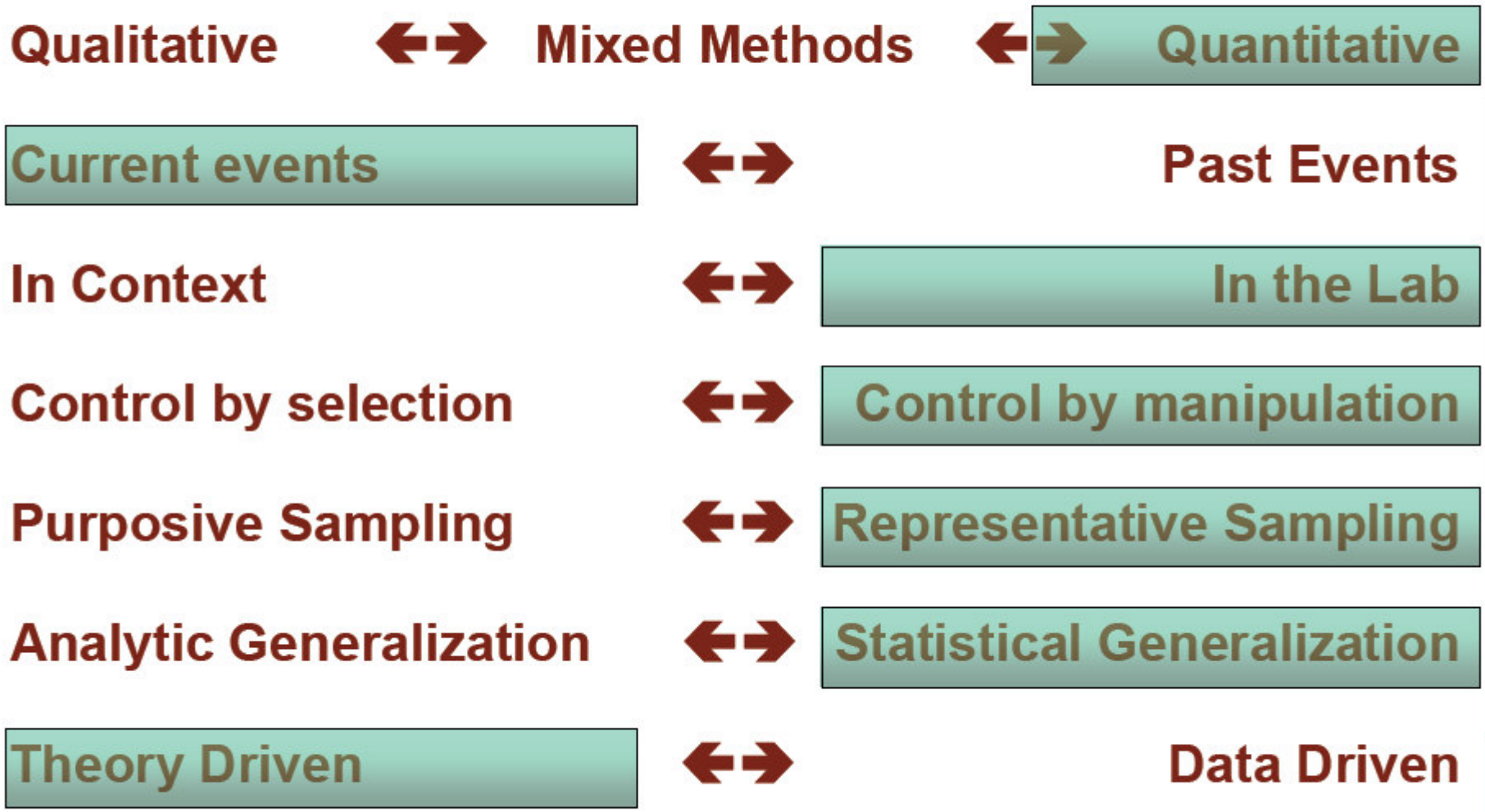
See:

Lau, F; Towards a framework for action research in information systems studies. Information Technology and People 12 (2) 148-175. 1999.

Case Studies

Qualitative	↔ Mixed Methods ↔	Quantitative
Current events	↔	Past Events
In Context	↔	In the Lab
Control by selection	↔	Control by manipulation
Purposive Sampling	↔	Representative Sampling
Analytic Generalization	↔	Statistical Generalization
Theory Driven	↔	Data Driven

Experiments



28th International Conference
on Software Engineering

2. Basic Elements of a Case Study

When should you use a case study?

- When you can't control the variables
- When there are many more variables than data points
- When you cannot separate phenomena from context
  - Phenomena that don't occur in a lab setting
  - E.g. large scale, complex software projects
  - Effects can be wide-ranging.
  - Effects can take a long time to appear (weeks, months, years!)
- When the context is important
  - E.g. When you need to know how context affects the phenomena
- When you need to know whether your theory applies to a specific real world setting

When should you use a case study?

- When you can't control the variables
- When there are many more variab… …nts
- When you cannot separat… …context
  - Phenomena that don't …
  - E.g. large scale …
  - Effects can …
  - Effect… …appear (weeks, months, years!)
- …s important
  …need to know how context affects the phenomena
- …you need to know whether your theory applies to a specific real world setting

When the interest is on the case (bounded system) not on the functions

Why conduct a case study?

- To gain a deep understanding of a phenomenon
  - Example: To understand the capability of a new tool
  - Example: To identify factors affecting communication in code inspections
  - Example: To characterize the process of coming up to speed on a project
- Objective of Investigation
  - Exploration- To find what's out there
  - Characterization- To more fully describe
  - Validation- To find out whether a theory/hypothesis is true
- Subject of Investigation
  - An intervention, e.g. tool, technique, method, approach to design, implementation, or organizational structure
  - An existing thing or process, e.g. a team, releases, defects

Misuses of the term "Case Study"

- Not a case history
  - In medicine and law, patients or clients are "cases." Hence sometimes they refer to a review of interesting instance(s) as a "case study".
- Not an exemplar
  - Not a report of something interesting that was tried on a toy problem
- Not an experience report
  - Retrospective report on an experience (typically, industrial) with lessons learned
- Not a quasi-experiment with small n
  - Weaker form of experiment with a small sample size
  - Uses a different logic for designing the study and for generalizing from results

How can I tell it's a case study?

- Has research questions set out from the beginning of the study
- Data is collected in a planned and consistent manner
- Inferences are made from the data to answer the research questions
- Produces an explanation, description, or causal analysis of a phenomenon
  - Can also be exploratory
- Threats to validity are addressed in a systematic way

Has a well defined case (bounded system)

Parts of a Case Study Research Design

- A research design is a "blueprint" for a study
  - Deals more with the logic of the study than the logistics
  - Plan for moving from questions to answers
  - Ensures data is collected and analyzed to produce an answer to the initial research question
  - (Analogy: research design is like a system design)
- Five parts of a case study research design
  1. Research questions
  2. Propositions (if any)
  3. Unit(s) of analysis
  4. Logic linking the data to the propositions
  5. Criteria for interpreting the findings

Practical

1. Critical reading of Runeson's guidelines, comparing it with Merrian's and Easterbrook's approach.
2. Read Flyvbjerg's article
3. Answer the question:
   - Should my research use case study as a method and why?