Quantitative Methods in Software Engineering, IIITDM Jabalpur, October, 2009

RESEARCH METHODS IN SE

Atul Gupta

Software Engineering (SE)

- A multi-disciplinary field
- We need to investigate
  - Which tools, techniques, and processes they use?,
  - What social and cognitive processes surrounding them?
  - How individual software engineer develop software?
  - How teams and organizations coordinate their efforts?

Research Methods

- Controlled Experiments (including Quasi-Experiments)
- Case Studies (both exploratory and confirmatory)
- Survey Research
- Ethnographies
- Action Research

Empirical Investigations

- Common tasks
  - Identify goals, questions and measures (metrics)
  - Choosing research method(s)
  - Planning, designing, and carry out investigations
  - Data collection,
  - Data Analysis
  - Results
  - Validity considerations
  - Conclusions

Two Motivating Examples

- To assess the effectiveness of a novel fisheye-view file navigator
- To assess how developers in industry use (or not) UML diagrams during software design

Research Questions

Kind of research questions

- *RQ. "Is a fisheye-view file navigator more efficient than the traditional view for file navigation?"*
- *Ambiguous !*
  - *Who?*
  - *File navigation?*
  - *Circumstances?*
  - *Efficiency?*
- *RQ. "how widely are UML diagrams used as collaborative shared artefacts during design?"*
- *Ambiguous !*

Defining the precise meaning of terms is Crucial !

Formulating exploratory Questions

- Existence Questions
  - Ex: “Does X exists?”
- Classification Questions
  - Ex: “What is X like?”
- Descriptive-Comparative Questions
  - Ex: “How does X differ from Y?”

Formulating Base-rate Questions

- Frequency and distribution questions
  - Ex: “How often does X occur?”, “What is an average amount of X?”
- Descriptive-Process questions
  - Ex: “How does X normally work?”
- Relationship questions
  - Ex: “Are X and Y related?”
  - Correlation and Causality

Formulating Base-rate Questions

- Causality Relationship questions
  - Causality questions -"Does X cause Y?"
  - Causality-Comparative questions - "Does X cause more Y than does Z?"
  - Causality-Comparative Interaction "Does X or Z cause more Y under one condition but not others?"

Choosing a Research Method

- Depends
  - available resources,
  - access to subjects,
  - opportunity to control
  - the variables of interest,
  - the skills of the researcher
  - how closely the method aligns with the question(s) that have been posed

Controlled Experiments

- Investigation of a testable hypothesis suggesting causal relationships
- Identification of study variables and instruments
  - Independent , dependent, block variables
  - Subjects, objects
- Design of experiments
- More of a quantitative than qualitative method
- Theory driven
- Validity consideration

Controlled Experiment

- Strengths
  - Theory driven
  - Exercise maximum degree of control over study variables
  - Repeatable
  - Better generalization is possible
  - A tool for SE education

Controlled Experiments

- Weaknesses or Limitations
  - Small Scope
  - Underlying assumptions may be wrong
  - Tends to be artificial

Case Studies

- "An empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident." [Yin 2002]
- Exploratory vs. confirmatory
- Selection of a case is a crucial step
  - Purposive sampling vs. random sampling
- Quantitative as well as qualitative

Case Studies

- Strengths
  - Offer in-depth understanding of a phenomenon in real situations
  - Situations where effects are expected to be wide ranging, or take a long time (e.g. weeks, months, years) to appear
  - situations where the context is expected to play a role in the phenomena

Case Studies

- Weaknesses
  - the data collection and analysis is more open to interpretation and researcher bias
  - Limited generalization

Surveys

- Characteristics of a broad population of individuals
- Use of interviews/questionnaires for data collection
- Selection of a *representative sample* from a well-defined *population*
  - More of a qualitative than quantitative

Surveys

- Strengths
  - Small to large scope
  - Generalizable

Surveys

- Weaknesses
  - control for sampling bias
  - Low response rate can be problematic
  - Inappropriate survey questions
  - Difficult to interpret or less insightful

Ethenographics

- Field studies
- Study a *community* of people
- focus on a broad technical community (e.g. Java programmers in general), or a small, closely knit community (e.g. a single development team)
- Researcher as an observer or a participant observer
- The biggest challenge is to perform detailed observation, data collection, and analysis

Action Research

- Solving a real-world problem while simultaneously studying it
- A long drawn method
- An iterative approach to problem solving
- Immaturity as an empirical method (tends to be ad hoc)
- An appealing framework for mixing research with professional activities

Mixed-Mode Approaches

- All methods have limitations
- Weaknesses of one method can be compensated for by the strengths of other methods
- Key decisions involve
  - the strategy for data collection
  - the sequence in which different methods are employed
  - quantitative and qualitative data
- Three general approaches
  - *Sequential explanatory strategy*
  - *Sequential exploratory strategy*
  - *Concurrent triangulation strategy*

Data Collection Techniques

- Separate lecture

Empirical Validity

- Construct validity:
  - the "right" (relevant and minimal) metrics, design
  - "Are we actually measuring what we intend to measure?"
- Internal validity:
  - the "right" data values.
  - "Does the data really follows from the experimental concepts?"
- Conclusion validity:
  - the right (appropriate) data analysis?
- External validity:
  - the "right" (representative) context
  - "Can the results of the experiment be generalized?"

Threats to validity

Examples:

- Violated assumptions of statistical tests
- Researchers may influence the results by looking for a specific outcome.
- If the group is very heterogeneous there is a risk that the variation due to individual differences is larger than due to the treatment.
- Incorrectly designed Experiment
- Effect of confounding factors
- Research bias
- …

Other Dangers

- Manipulating several explanatory variables simultaneously so the effect on the response variable cannot be interpreted
- Over-complex design
- Carryover effects or learning
- Insufficient subjects (lack of power)
- Unrealistic task/environment
- Unrepresentative subjects - usually students

Confounding Effects in SE Experiment

- Technique Learning Effect
- Object Learning Effect
- Boredom Effect
- Enthusiasm Effect
- Experience Effect
- Procedure Effect

Practical Considerations

- Relate to time, budget and personnel resources, and access to data
- Methods that are primarily qualitative
  - Designing good research question
  - Researching skills – Observation, recording social behavior, patience,
- Methods that are primarily quantitative
  - Require more significant time in the planning
  - Pilot studies
  - Researching Skills – critique, management, innovation

Conclusions

- An overview of research methods
- Each has its own strengths and weakness
- Use of mixing-methods
- Validity considerations

References

- Guide to Advanced Empirical Software Engineering, Forest Shull, Janice Singer, and Dag I.K. Sjøberg, Springer 2008