



NRC-CNRC

From **Discovery**
to **Innovation...**

Evaluating tools in ASE (or Everything I've learned in 20+ years taught in 1.5 hours)

Janice Singer

janice.singer@nrc-cnrc.gc.ca

National Research Council Canada



National Research
Council Canada

Conseil national
de recherches Canada

Canada

FIRST THINGS FIRST: Biases in Research (About Me)

- Know your researcher and their biases – makes a difference in interpreting results...
- Undergrad CS, Ph.D. Experimental Psychology
- Currently researcher in software engineering
- Started by asking the question “What do software engineers do on a daily basis, and how can we help them do so more effectively, productively, enjoyably, quicker, etc?” Still asking and answering the same question
- 25 years experience in conducting empirical investigations using both qualitative and quantitative methods
- STILL LEARNING!!!!

- So...1.5 hours, pretty impossible task, but maybe enough to get you started thinking about some of the issues you may encounter and how you can structure your thinking to start evaluating.

The six steps (more or less) – Outline of Talk

1. Specify/operationalize question
2. Design study
3. Gather your resources and do your paperwork
4. Collect data
5. Analyse data
6. Interpret findings with respect to Step 1

Questions

When to do each of the steps?

- Although presented serially
 - Actually can and in many cases should happen simultaneously
 - Especially in qualitative paradigms, research progresses and informs itself as the basic findings and ideas emerge from the data

The six steps (more or less)

- 1. Specify/operationalize question**
2. Design study
3. Gather your resources and do your paperwork
4. Collect data
5. Analyse data
6. Interpret findings with respect to Step 1

Step 1: Specify/operationalize question

- Before you can begin your research, you need to know what you are researching. That is, what is your question
- The more specific the question, the easier it will be to answer, although not all questions can be very specific

Types of questions

- Descriptive questions
 - Describe or classify a wide variety of situations
- Correlational Questions
 - Determine whether two events are related
- Causality Questions
 - Determine whether one event is caused by another event

Descriptive Questions

- Think of yourself as a reporter trying to write a story about your tool
 - Who, what, where, when, why, how + the special existence question
- Generally much richer data set
- Useful for understanding context of use and generating hypotheses

Examples of descriptive questions

- Why is my tool not being used?
- What do SEs do now without my tool?
- Where is my tool being used (both virtually and physically)?
- When is my tool helpful?
- Who is using my tool?
- How does my tool improve processes?
- Is my tool used?

Correlational questions

- Correlation does not equal causality
- It simply shows that when one event occurs, another event is also likely to occur
- On average, tall girls weigh more than short girls
- Can not know direction of causation
- Additionally, there could be another variable controlling both weight and height, therefore, we cannot know if one is causing the other.

Examples of Correlational Questions

- Is my tool used during debugging?
- Is better documentation related to better quality code?
- Does the length of documentation relate to better answers?
- Does documentation accuracy relate to use?

Causality questions

- Shows that one event causes another event to happen
- If you possess Gene X, then you are short
- Causality is not bi-directional
- If you are short, you do not necessarily possess Gene X

Examples of Causality Questions

- Does my tool improve productivity?
- Is code written with my tool less buggy?
- Are the processes engendered by my tool more efficient?

Operationalize question

- Need to decide how you're going to measure the construct that you are interested in.
- For instance, what does 'quality' mean? Could mean:
 - Fewer bugs
 - Less maintainence
 - More beautiful as judged by experts

Pitfalls to question asking

- Asking too vague a question
- A relatively specific question is necessary to get the research going
- Remember my question – what do Software Engineers do – pretty vague – too much data – too hard to analyse. Now looking at more specific question – how do they navigate during debugging?
- Lack of/and improper operationalization
 - Operationalization does not measure what you say it does

The six steps (more or less)

1. Specify/operationalize question
- 2. Design study**
3. Gather your resources and do your paperwork
4. Collect data
5. Analyse data
6. Interpret findings with respect to Step 1

S2: Design Study

- First question: Do I need humans to answer my question?
 - Some questions are impossible to answer without human input, but some will be. Examples include:
 - speed of algorithm
 - In some cases, accuracy of algorithm
 - It's easier to do a study without humans, if you can manage it.
- But I'm not an expert in this area, so...
- The rest of this mini-tutorial will focus on the human side – both direct and indirect contact!

So now we know we're studying humans

- Do you use qualitative or quantitative methods?
- Before answering, we need to take a short digression into philosophy...
- Research methods do not come value free. Deciding to use a particular method means that you are taking a stance on some important issues with respect to how knowledge is created

A grossly simplified distinction between qual and quant methods

- Knowledge claims surrounding Positivism:
 - Objectivity is possible
 - Reductionism can aid in understanding complex phenomena
 - Focus on causal determination, prediction, and statistical generalization
- Knowledge claims surrounding Qualitative research:
 - Knowledge is socially constructed
 - Can't understand knowledge without understanding context
 - Focus on illumination, understanding, and analogical generalization to similar situations

Mixed Methods approach

- Relatively new addition to empirical approaches
- Mixture of quantitative and qualitative methods
- Use both to offset shortcomings of each
- Pragmatist
- Use methods appropriate for answering research questions
 - Sequentially
 - Concurrently

Back to research design

- Decide which method you will use
 - Quantitative
 - Causality questions
 - Correlational questions
 - Qualitative
 - Descriptive questions
 - Correlational questions
 - Mixed-Methods
 - Both types of questions

Some Methods associated with Qualitative Research

- Ethnography
- Participant Observation
- Action Research
- Case Study
- Grounded Theory
- Historical Research/Content Analysis

Methods associated with Quantitative Research

- Controlled Experiments
- Simulations

Pitfalls to study design

- Study design does not match question
- Using the wrong method to answer question

The six steps (more or less)

1. Specify/operationalize question
2. Design study
- 3. Gather your resources and do your paperwork**
4. Collect data
5. Analyse data
6. Interpret findings with respect to Step 1

Gather your resources

- Study site
 - Naturalistic
 - Lab study
- Subject group
 - Experts
 - Students
 - Professionals
- Research assistants
- Research experts
- Materials

Do your paperwork

- Research ethics approval
 - Depending on where you live, may be necessary
 - If required, do it early, can take some time for approval
- Even if REB approval not required, makes sense to follow regulations
 - Protection from liability
 - The right thing to do
 - Relationship with study population
- Study Description and Consent form
 - Talks about purpose of study
 - Harms
 - Benefits
 - Statement that withdrawal is possible without penalty
 - Confidentiality
 - Actual consent

Pitfalls to resources

- Not getting the right resources
- Not getting enough resources
- Not realizing the scope of the problem
- Failing to get ethics approval

The six steps (more or less)

1. Specify/operationalize question
2. Design study
3. Gather your resources and do your paperwork
- 4. Collect data**
5. Analyse data
6. Interpret findings with respect to Step 1

Data Collection

- Match data collection to methods used and questions asked
- Match data collection methods to resources
- Think about flexibility required
- Think about reliability required
- What type of process are you interested in
 - Cognitive (why) vs. behavioural (how)

Quantitative Data Collection

- Done for experimental research
 - manipulate some variables and then measure the effects of this manipulation on other variables;
- Define/Operationalize dependent and independent variables
 - Independent variable – the variable manipulated by the experimenter
 - Dependent variable – what the experimenter is measuring
 - $DV = f(IV)$
- Collect measures for each dependent variable based on values on independent variable
- Between versus Within Subject design

Qualitative Data Collection

- Generally collecting textual or visual data
- Need some way to collect data
 - Notes
 - Audiotape
 - Videotape
 - Palm Pilot
- Need some way to transcribe data
 - 10 times to transcribe
 - Usually need some expertise in field for accurate transcription

First Degree Techniques – Direct access to subjects

Technique	Used by researchers when their goal is to understand:	Volume of data	Also used by software engineers for:
First Order Techniques			
Brainstorming and Focus Groups	Ideas and general background about the process and product, general opinions (also useful to enhance participant rapport)	Small	Requirements gathering, project planning
Surveys/Questionnaires/Interviews	General information (including opinions) about process, product, personal knowledge etc.	Large	Requirements and evaluation
Conceptual modeling	Mental models of product or process	Small	Requirements
Work Diaries	Time spent or frequency of certain tasks (rough approximation, over days or weeks)	Medium	
Think-aloud sessions	Mental models, goals, rationale and patterns of activities	Small	UI evaluation
Shadowing and Observation	Time spent or frequency of tasks (intermittent over relatively short periods), patterns of activities, some goals and rationale	Small	Advanced approaches to use case or task analysis
Participant observation (joining the team)	Deep understanding, goals and rationale for actions, time spent or frequency over a long period	Medium	

2nd Degree Techniques – access to subject environment as they work

Technique	Used by researchers when their goal is to understand:	Volume of data	Also used by software engineers for:
Instrumenting systems	Software usage over a long period, for many participants	Large	Software usage analysis
Fly in the wall	Time spent intermittently in one location, patterns of activities (particularly collaboration)	Medium	

3rd degree techniques – access to work artifacts

Technique	Used by researchers when their goal is to understand:	Volume of data	Also used by software engineers for:
Analysis of work databases	Long-term patterns relating to software evolution, faults etc.	Large	Metrics gathering
Analysis of tool use logs	Details of tool usage	Large	
Documentation analysis	Design and documentation practices, general understanding	Medium	Reverse engineering
Static and dynamic analysis	Design and programming practices, general understanding	Large	Program comprehension and metrics

Pitfalls to data collection

- Gathering the wrong data
- Gathering too much data
- Not gathering enough data

The six steps (more or less)

1. Specify/operationalize question
2. Design study
3. Gather your resources and do your paperwork
4. Collect data
- 5. Analyse data**
6. Interpret findings with respect to Step 1

Parametric Statistics

- Hypothesis testing – null hypothesis
- Used in experimental settings
 - t-test
 - Correlation
 - Based on Linear Model
 - ANOVA
 - MANOVA
 - Regression
- Results in some p-value
 - Basically says how likely is that?
- MAJOR PITFALL – multiple tests without correcting p

Nonparametric Statistics

- Less sensitive than parametric stats
- Used when you cannot use parametric statistics
 - Small sample
 - Non-normal distribution
 - Measurements that are not at least interval (i.e., nominal or ordinal)
- Generally speaking there is a non-parametric stat equivalent for all parametric stats.
 - Tests of differences between groups
 - Tests of differences between variables within group
 - Test of relations
 - Descriptive Stats

Qualitative Analysis

- Variety of approaches to Qualitative Analysis
- Usually based on Research Approach
 - Certain type of analysis done for Grounded theory
 - Different analysis done for Ethnographic research
- Generally, but not always
 - Code the data according to some scheme
 - Refine categorization
 - Re-categorize data
 - Validate categorization
 - Use findings to tell story

Triangulation

- Using a variety of approaches to provide complementary evidence to hypothesis/question/theory
- Just qualitative
- Just quantitative
- Qualitative and quantitative

Pitfalls to data analysis

- Choosing the wrong analysis strategy
- Question operationalization incorrect
- For Quantitative
 - Invalidating assumptions of statistical test
 - Misunderstanding the meaning of statistical significance
- For Qualitative
 - Categorization scheme does not reflect questions asked

The six steps (more or less)

1. Specify/operationalize question
2. Design study
3. Gather your resources and do your paperwork
4. Collect data
5. Analyse data

6. Interpret findings with respect to Step 1

Interpret findings wrt Step 1

Research is an iterative process!

- We shouldn't be testing tools, but testing theories
- Whether we state them explicitly or not, each of the tools/processes we create have embodied in them particular theories about how software development occurs
- We should try to make these theories explicit
- Then as our work develops, we are developing a corpora of ideas that are related to the theories, as opposed to just another tool, or just another process
- System approach

Therefore...

- We need to go back to our original question and see whether we've answered it adequately
- If not, why not
 - Failed to reject null hypothesis
 - Tool isn't ready for human consumption
 - Did we answer a different question
 - If so, which question did we answer and why did our question get changed
 - BTW, this isn't necessarily bad, and happens all the time in qualitative research
- Tie the question to a theory, and look to the theory for explanatory power
- Tie the theory to previous/other work to see how you are arguing for/or against similar constructs

Software

- For Qual analysis
 - Nvivo
 - Qualrus
 - Atlas-ti
 - Noldus
 - Ovo
 - Transana (open source)
- For Quant analysis
 - SPSS
 - SAS
 - Statistica

Useful References

- Creswell, J. (2003). *Research Design: Quantitative, Qualitative, and Mixed Methods Approaches*. Sage Publications.
- Denzin, N.K. & Lincoln Y.S. (1998). *Collecting and Interpreting Qualitative Materials*. Sage Publications Inc., Thousand Oaks California
- Lethbridge, T., Sim, S., Singer, J.(2005). Studying Software Engineers: Data Collection Techniques for Software Field Studies, *Empirical Software Engineering*, 10(3), pp. 311-341.
- Miles, M., Huberman, M. (1994). *Qualitative Data Analysis: 2nd Edition*. Sage Publications.
- Shull, F., Singer, J., Sjoberg, D. (2007??). *Advanced topics in empirical software engineering – to be published by Springer*.
- Seaman, C. (1999). Qualitative Methods in Empirical Studies of Software Engineering. *IEEE TSE*, 25(4): 557-572 (1999) Siegel, S., Castellan, N.J. (1988). *Nonparametric statistics for the behavioral sciences*, 2nd edition. McGraw Hill.
- Singer, J, Vinson, N (2002). Ethical Issues in Empirical Studies of Software Engineering. *IEEE TSE*, 28(12): 1171-1180 (2002)
- Wohlin, C., Runeson, P., Host, M., Ohlsson, M., Regnell, B., Wesslen, A. (2000). *Experimentation in software engineering: An introduction*. Kluwer Academic Publishers

For a specific recommendation for a paper or book explicating a research method, please contact me.

Questions?

**Email me for copies of slides:
janice.singer@nrc-cnrc.gc.ca**