# Scientific Research

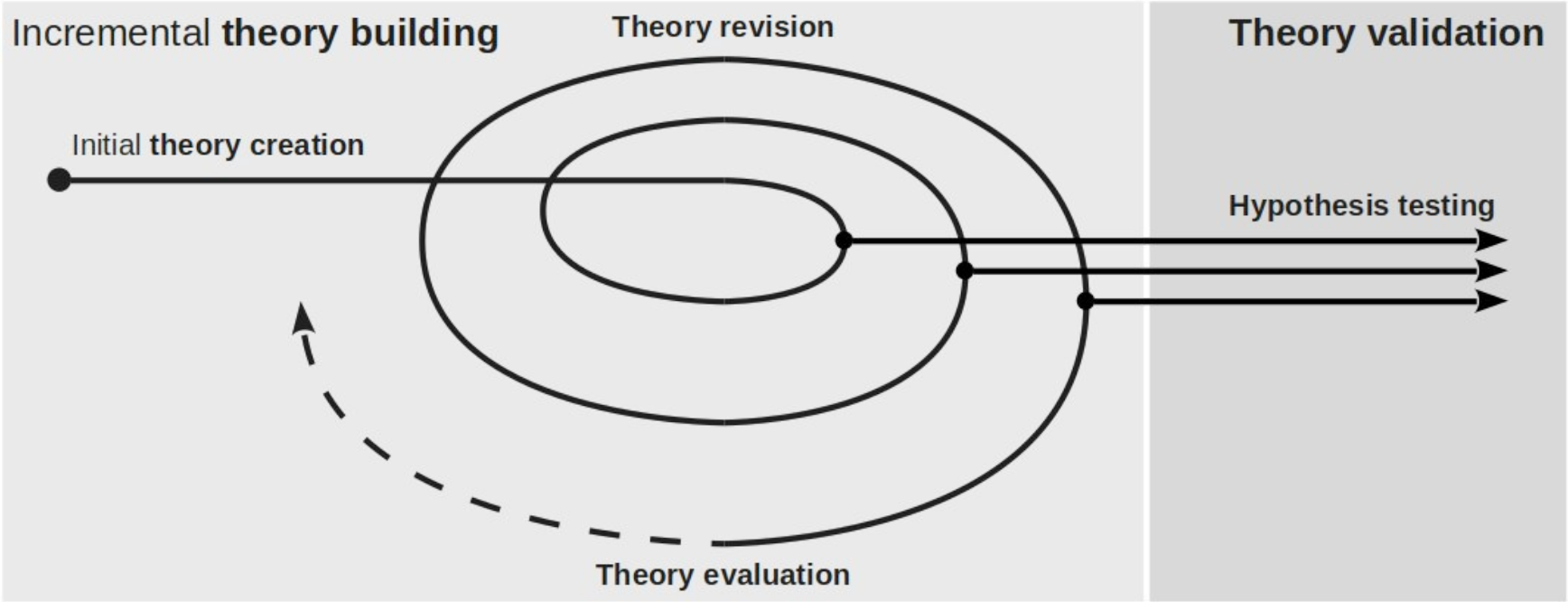
Dirk Riehle, Univ. Erlangen

**NYT B02**

# Agenda

1. Research questions
2. The research process
3. Research methodologies
4. Research methods and practices
5. Research quality criteria

# Theory Building and Validation (Recap)



# 1. Research Questions

# Research Question and Hypothesis

A **research question** is

- The question of interest to be answered by research

A **research hypothesis** is

- A research question that can only be answered with yes or no, true or false

Research questions should answerable and non-trivial

- The answer can be big (a whole theory) or small (yes/no)

# Hilbert's 23 Problems [1]

| Problem | Brief explanation | Status | Year Solved |
|---|---|---|---|
| 1st | The continuum hypothesis (that is, there is no set whose cardinality is strictly between that of the integers and that of the real numbers) | Proven to be impossible to prove or disprove within the Zermelo–Fraenkel set theory with or without the Axiom of Choice (provided the Zermelo–Fraenkel set theory with or without the Axiom of Choice is consistent, i.e., contains no two theorems such that one is a negation of the other). There is no consensus on whether this is a solution to the problem. | 1963 |
| 2nd | Prove that the axioms of arithmetic are consistent. | There is no consensus on whether results of Gödel and Gentzen give a solution to the problem as stated by Hilbert. Gödel's second incompleteness theorem, proved in 1931, shows that no proof of its consistency can be carried out within arithmetic itself. Gentzen proved in 1936 that the consistency of arithmetic follows from the well-foundedness of the ordinal $\varepsilon_0$. | 1936? |
| 3rd | Given any two polyhedra of equal volume, is it always possible to cut the first into finitely many polyhedral pieces which can be reassembled to yield the second? | Resolved. Result: no, proved using Dehn invariants. | 1900 |
| 4th | Construct all metrics where lines are geodesics. | Too vague to be stated resolved or not.[n 1] | – |
| 5th | Are continuous groups automatically differential groups? | Resolved by Andrew Gleason, depending on how the original statement is interpreted. If, however, it is understood as an equivalent of the Hilbert–Smith conjecture, it is still unsolved. | 1953? |

[1] See https://en.wikipedia.org/wiki/Hilbert%27s_problems

# Example Research Questions [1]

| | People | Practices | Artifacts |
|---|---|---|---|
| **Method of Analysis** | How to measure communication effectiveness? | How to track bug fixes in software development? | How to measure effort spent on a piece code? |
| **Method of Development** | How to distribute roles in an agile development team? | How to perform code reviews? | How does agile development ensure high quality? |
| **Evaluation of Instance** | What skills of Linus Torvalds are key for the Linux kernel? | What are some new types of code refactorings? | How patchy is the Apache web server? |
| **Generalization or Characterization** | What does it take to be a good programmer? | What's common to all open source processes? | How does the the open source commit distribution look like? |
| **Feasibility or Exploration** | Do energy drinks help programmers write code faster? | Is “Test First” a viable principle? | What could a power plant security hole do to safety? |

[1] Derived from Shaw (2003): Writing good software engineering research papers.

# Related Work

**Related work** is

- (All) existing scientific work related to your research question

Related means

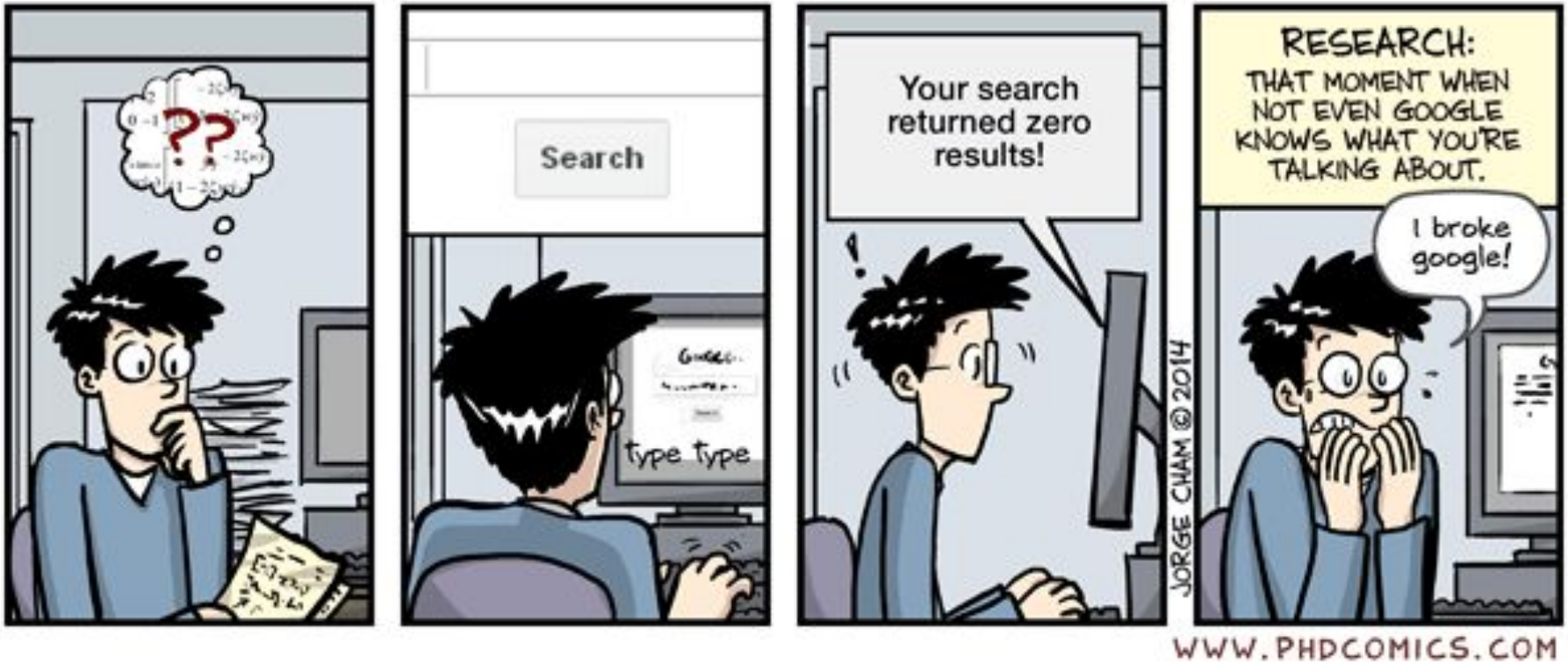
- Someone else worked on the same or similar research question
- The work takes an alternative or contradicting starting point
- Your proposed work requires and builds on this (related) work

Searching for and identifying

- **Related work is likely to change your research question**

A literature review is a common form of searching for related work

# One Stopping Criterion [1]




[1] See https://phdcomics.com

# Research Result

A **scientific research result** is

- Any piece of information resulting from scientific research
- Where scientific research is work following proper methods

The **answer** to a research question is

- A scientific result that answers a corresponding research question

A scientific result should be

- verifiable,
- reproducible, and
- independent of the researcher

# Paper

A **research paper** is a (science-idiomatic) term for

- A research article

It may or may not yet have passed peer review

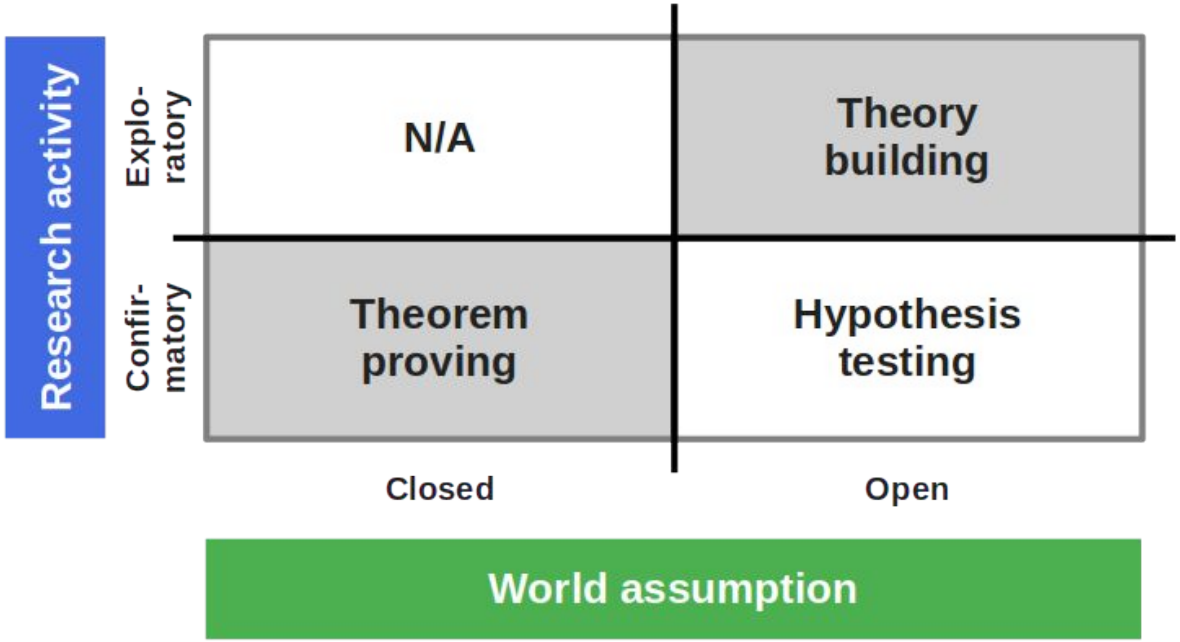
# Contribution

A **contribution** is a (science-idiomatic) term for

- A publishable / published research result

To be found in the contributions section of a research paper

# Open vs. Closed World Assumption



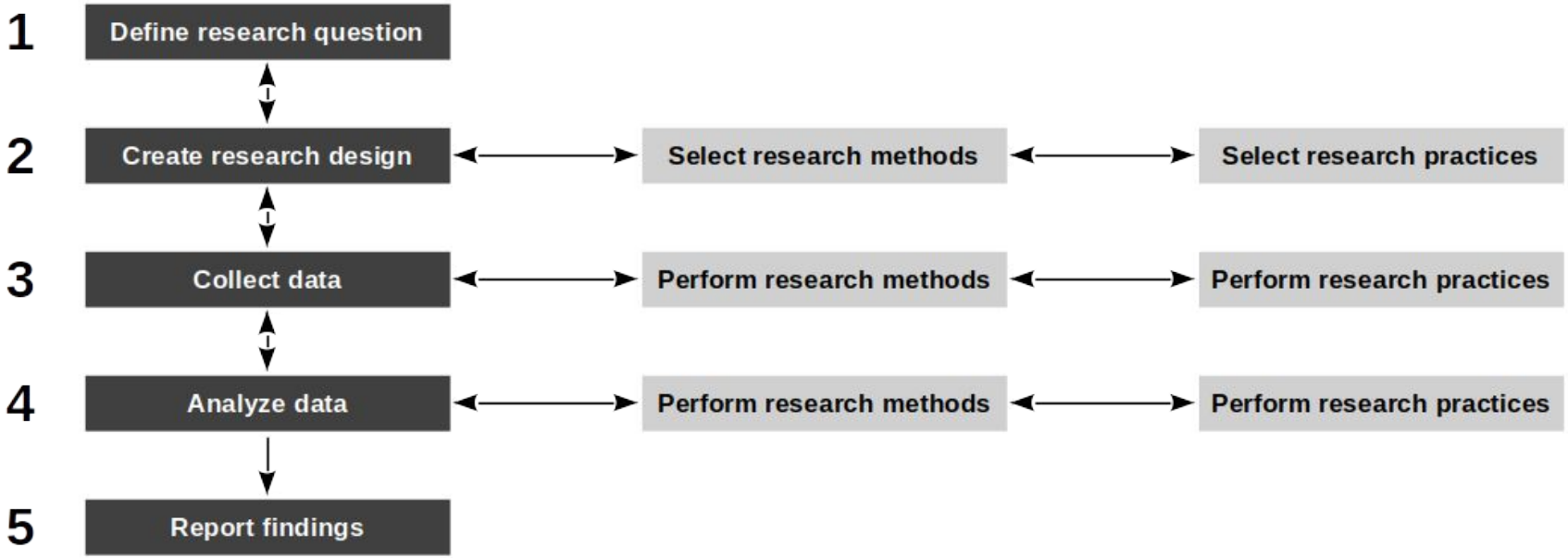
# Is This So? Why / Why Not?

**“In theory, practice and theory are the same. In practice, they are not.” [1]**

[1] Attributed to Albert Einstein

# 2. The Research Process

# Generic (Idealized) Research Process



# Research Methodology vs. Design vs. Method vs. Practice

A **research methodology** is

- A start-to-finish framework for performing theory building

A **research design** is

- A process description for answering a research question

A **research method** is

- A method for answering a type of research question

A **research practice** is

- A way of doing something with a defined outcome

# Research Projects

A **research project**

- Is a project that contributes to scientific progress
- Typically answers a (large or small) question
- As a project has a start and an end date

Research projects can be built from parts

- A principal investigator may define a whole project
- Graduate researchers may work on subprojects
- Thesis students may work on one question

# (Almost) All Research is Iterative

The scientific process is

- Iterative and incremental in how it builds out theory

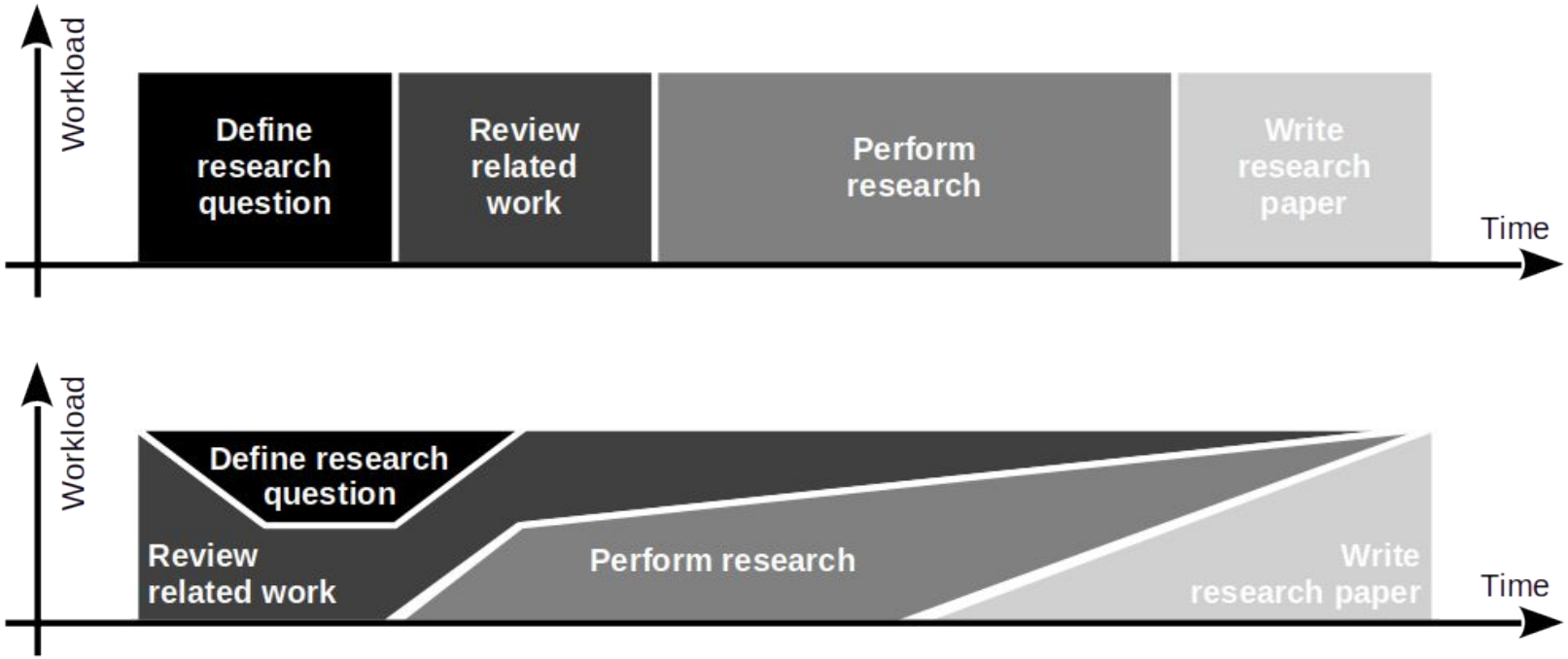
Its results presentation should

- Follow a linear logical structure usually corresponding to the research design

Beginners mistake when writing a thesis

- Focusing your process steps, not your results

# How Time is Spent in Research



# Research Protocol and Process Tracking

Document your research design and proposed process before you start

- Register this so-called research protocol with independent resource

Track data, process, and progress of actual work in lab book (audit trail)

- Discuss plan vs. execution in discussion and limitations section

# 3. Research Methodologies

# Research Methodology

A **research methodology** is a

- Start-to-finish framework for theory building
- Constraint system for research designs
- Pattern for research designs

# Example Research Methodology

The **qualitative survey** [1] according to Jansen (2010)

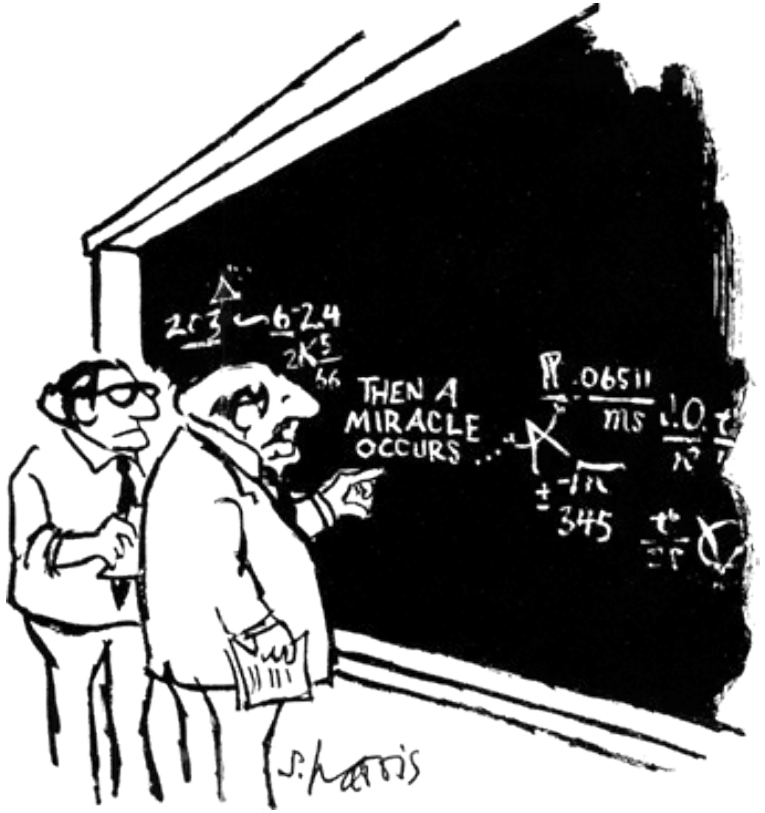
1. Write research protocol [3]
2. Build sampling model [2]
3. Sample for theory building [2]
4. Perform interviews [2]
5. Analyze transcriptions [2]
6. Determine saturation [3]
7. Iterate or conclude

[1] A.k.a. interview study (when restricted to interviews)
[2] A research method
[3] A research practice

# Some (Categories of) Research Methodologies

- Systematic survey
- Qualitative survey
- Action research
- Case study research
- Grounded theory
- Ethnographies

# “I think you should be more explicit in step 2” [1]




[1] See http://www.sciencecartoonsplus.com/

# Research Design

A **research design** is

- A structured plan to answer a research question
- Created by applying a research methodology

# Example Research Design Inner Source Challenges

Research question

- What are challenges of inner source collaboration in software development?

Possible research approach

- Qualitative survey according to Jansen (2010)

Resulting research design

1. Model units of analysis (companies, employees) of interest
2. Identify units of analysis using sampling model
3. Contact and interview experts
4. Analyze interviews using thematic analysis according to Braun & Clarke (2012)
5. Determine saturation and iterate if not satisfied

# The Marshmallow Experiment [1]

[1] See https://www.youtube.com/watch?v=Yo4WF3cSd9Q

# 4. Research Methods and Practices

# Research Method

A **research method** is

- A prescriptive procedure to answer a research question
- Which collects and analyses one set of data

Research method vs. **research methodology**

- A methodology provides a framework while a method provides a prescription
- It ain’t a methodology if it ain’t theory building

# Example Research Methods

- Braun & Clarke (2012): Thematic analysis
- Wohlin et al. (2012): Controlled experiments
- Fagerholm & Fritz (2020): Biometric measurements

# Research Practice

A **research practice** is a

- An established procedure that when employed provides a defined result
- That is known to practitioners of science but has not been codified yet

Research practice vs. **research method**

- A research practice may not have a name or formal definition
- A research practice may be a prototype of a method

# Example Research Practices

Example practices **with a name**

- Forward snowballing (an article search strategy)
- Polar sampling (a purposive sampling strategy)
- Peer debriefing (a quality assurance practice)

Example practices **without a name**

- How to derive a code name
- How to report on a Student’s t-test results
- How to codify a theory using best practices handbooks

# Research Methods (Expanded)

A **research method** is

- A prescriptive procedure
  - For collecting and analyzing one type of data
  - To answer a research question

An **established** research method has been validated

- It has been shown to deliver what it promises
- Methods should have their own quality assurance criteria

# Two Categories of Research Methods

A **qualitative research method** is a

- Research method that
  - Collects and analyses qualitative data
  - Creates theoretical insights

A **quantitative research method** is a

- Research method that
  - Collects and analyses quantitative data
  - Lets researchers draw descriptive or statistical conclusions

# Qualitative vs. Quantitative Data (Expanded)

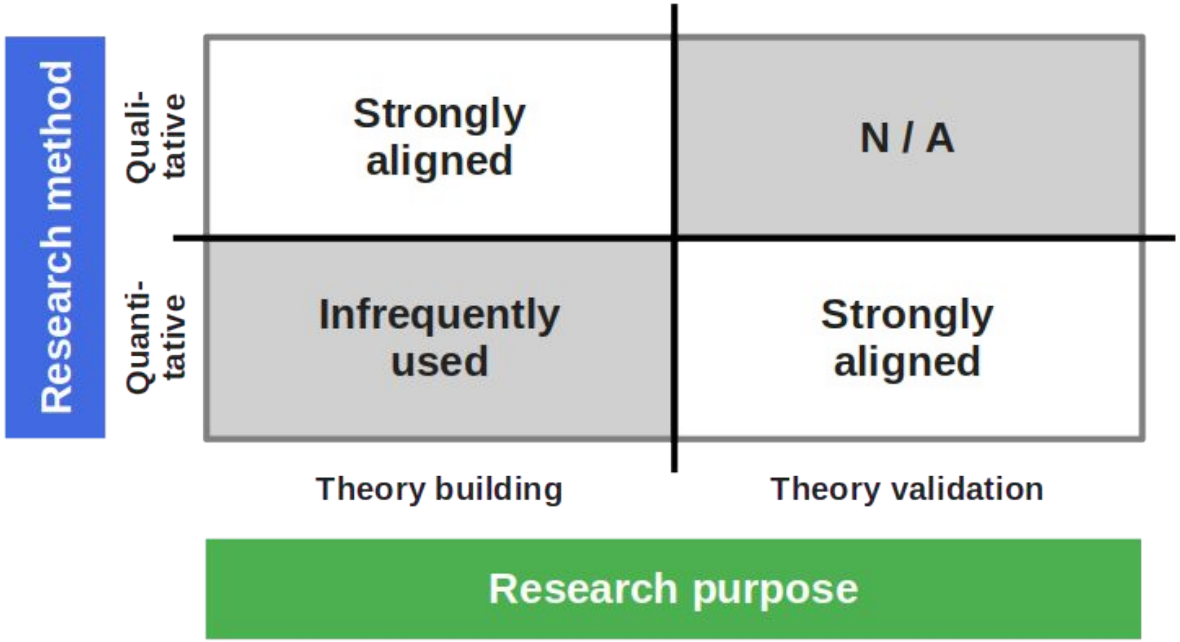
**Qualitative data** is data that

- Usually is not numeric
- Requires interpretation
- Is used in theory building
- Denies statistical generalizations

**Quantitative data** is data that

- Usually is numeric (various scales)
- Has objective (formal) definition
- Is used for theory validation
- Serves statistical generalization

# Qualitative / Quantitative vs. Theory Building / Validation



# “Soft” vs. “Hard”

**Qualitative research**

- Deals with qualitative data
- Feedback takes time
- Feedback is not binary

**Quantitative research**

- Deals with quantitative data
- Feedback is fast
- Feedback is precise / binary

# Should You Invent Your Own Methods?

Never invent your own methods

- Rather use established research methods

There will always be exceptions, but think hard if your situation is one

# 5. Research Quality Criteria

# Quality Criteria for Research Methods [1]

| Intuition | Qualitative research | Quantitative research |
|---|---|---|
| Truth value | Credibility | Internal validity |
| Applicability | Transferability | External validity |
| Consistency | Dependability | Reliability |
| Neutrality | Confirmability | Objectivity |

[1] See Guba & Lincoln (1982): Naturalistic inquiry.

# Truth Value (Category)

Truth value (framed as question)

- To what degree can we establish confidence in the “truth” of findings?

Credibility (naturalism)

- The degree of confidence in the truth of the findings

Internal validity (rationalism)

- The degree of confidence in a cause and effect relationship

# Applicability (Category)

Applicability (framed as question)

- To what degree can we determine the applicability of findings in other contexts?

Transferability (naturalism)

- The degree to which findings can be transferred to another context

External validity (rationalism)

- The degree to which findings can be generalized beyond the study

# Consistency (Category)

Consistency (framed as question)

- To which degree can the findings be consistently repeated?

Dependability (naturalism)

- The degree to which findings are stable over time

Reliability (rationalism)

- The degree to which findings can be recreated under the same conditions

# Neutrality (Category)

Neutrality (framed as question)

- To which degree are the findings independent of the researcher?

Confirmability (naturalism)

- The degree to which the findings can be confirmed by other researchers

Objectivity (rationalism)

- The degree to which the findings can be repeated by other researchers

# Summary

1. Research questions
2. The research process
3. Research methodologies
4. Research methods and practices
5. Research quality criteria

# Thank you! Any questions?

dirk.riehle@fau.de – https://oss.cs.fau.de

dirk@riehle.org – https://dirkriehle.com – @dirkriehle

# Legal Notices

License

- Licensed under the CC BY 4.0 International license

Copyright

- © 2012, 2023 Dirk Riehle, some rights reserved