

How to write and evaluate RE research papers

Roel Wieringa

University of Twente

<http://ewi.utwente.nl/~roelw/>

An updated introduction to paper classification and evaluation criteria

- Wieringa, R.J. and Maiden, N.A.M. and Mead, N. and Rolland, C. (2006) [Requirements engineering paper classification and evaluation criteria: A proposal and a discussion.](#) *Requirements Engineering*, 11 (1). pp. 102-107.
- Wieringa, R.J. and Heerkens, J.M.G. and Regnell, B. (2009) [How to write and read a scientific evaluation paper.](#) (Invited) In: *Proceedings of the 17th International Requirements Engineering Conference*, pp. 361-364. IEEE Computer Society.

Table of contents

A large, multi-layered burger is the central focus, served on a metal tray. The burger features a sesame seed bun, two beef patties, melted yellow cheese, pickles, red onions, tomatoes, and lettuce. It is topped with a generous amount of fried onion rings and a drizzle of sauce. A side of golden french fries is arranged around the base of the burger. In the background, a glass bottle and a plate with cutlery are visible but out of focus.

Introduction

Contents

Conclusion

The problem, part 1

- You design a wonderful technique that does what no one has been able to do before.
- It takes 15 LNCS pages to describe your research goal, describe and illustrate the technique, and compare it analytically to related techniques.
- You submit it to conference X.
- Rejected!
 - “The example is just an illustration, we want to see an empirical validation.”

The problem, part 2

- You implement a prototype of a novel, experimental technique that you published last year, and test it.
- It takes 15 LNCS pages to describe your the technique, what is currently known about it, the test setup, the measurements, your analysis, threats to validity, and the implications for practice and research.
- You submit it to conference X.
- Rejected!
 - “You published about this technique last year already. It is not novel.”

What is the problem?

- A design paper is evaluated as an empirical paper. Wrong!
- An empirical paper is evaluated as a design paper. Wrong!
- This way, you cannot win.

- Solution:
 - Make paper classifications,
 - And make their criteria explicit.

Paper classification schema: REFSQ 16

- **Full research papers** (up to 15 pages), including literature reviews, evaluation research, solution proposals, and validation research.
- **Experience reports** (up to 15 pages) describing positive and negative experiences.
- **Vision papers** (up to 6 pages) stating where the research in the field should be heading towards.
- **Problem statements** (up to 6 pages) describing open issues of practical or theoretical nature.
- **Research previews** (up to 6 pages) reporting research results at an early stage.

Paper classification schema: REFSQ 16

Empirical research

New design

- **Full research papers** (up to 15 pages), including literature reviews, evaluation research, solution proposals, and validation research.
- **Experience reports** (up to 15 pages) negative experiences. **How do you know which paper you are writing ...**
- **Vision papers** (up to 6 pages) stating where the research in the field should be heading towards. **or reading?**
- **Problem statements** (up to 6 pages) describing open issues of practical or theoretical nature. **By which criteria to evaluate?**
- **Research previews** (up to 6 pages) reporting research results at an early stage.

The problem

- For writers:
 - Under which category to submit your paper?
- For PC members:
 - How to know what paper you are reading?
- For all of you:
 - By which criteria to evaluate it?

Our 2006 paper

- Wieringa, R.J. and Maiden, N.A.M. and Mead, N. and Rolland, C. (2006) [Requirements engineering paper classification and evaluation criteria: A proposal and a discussion.](#) Requirements Engineering, 11 (1). pp. 102-107.

The core of that paper: the engineering cycle

The engineering cycle

- Problem investigation
- Treatment design
- Treatment validation
- Treatment implementation
- Implementation evaluation

An example

- *Why are our IS projects late?*
- *New effort estimation technique*
- *Test with experts; do pilot project*
- *Roll out in the organization*
- *Have estimations now improved?*

A model of rational change

But first: the top-level distinction

Journalist, detective, researcher

Knowledge questions

- What is the case?

• *When?*
Is it true?
• *Where?*

- How did it happen?
- Who was involved?
- How much?
- How many?

• ...

- Why did it happen?

Ask
for
facts

Asks
for
explanations
(Theories)

Engineer, architect, problem-solver

Design problems

- How to treat this problem?

• *How to use this tool?*
How does it work?
• *How to achieve this goal?*

- How to improve the business?
- How to learn this technique?
- How to test this program?

• ...

How to do something

Example: You

Knowledge questions

- What is the case?
- When?
- **What is the performance of my program?**
- How much?
- How many?

Facts

- ... **Why this performance?**
- Why did it happen?

Explanations
(Theories)

Design problems

- How to treat this problem?
- **How to improve it?**
- How to achieve this goal?
- **Redesign of architecture, or replacement of component**
- How to test this program?
- ...

How to do something

Example: Design science researcher: technical design validation

Knowledge questions

- What is the case?
- When?
- **Does the artifact work in the real world?**
- Who is involved?
- How much?
- How many?

Facts

- ...
- **Why?**
- Why did it happen?

Explanations
(Theories)

Design problems

- How to treat this problem?
- How to use this tool?
- **How to improve it?**
- How to achieve this goal?
- How to improve the business?
- How to learn this technique?
- How to test this program?
- ...

How to do something

The top-level distinction: evaluation criteria

Knowledge questions

- What is the case?
- When?
- Where?
- How did it happen?
- Who was involved?
- How much?
- How many?
- ...
- Why did it happen?

Is it true?

Facts

**Explanations
(Theories)**

Design problems

- How to treat this problem?
- How to use this tool?
- How to achieve this goal?
- How to improve the business?
- How to learn this technique?
- How to test this program?
- ...

How does it work?

How to do something

Truth, uncertainty and fallibility.

- Propositions about the real world are true or false.
- We cannot be certain about the truth of a proposition.
- Our truth claims are fallible.
- Scientific papers acknowledge this and reflect on the extent and limits of the support for their conclusions (“`validity”).

The top-level distinction: evaluation criteria for finite minds

Knowledge questions

- What is the case?
- When?
- Where?
- How did it happen?
- Who was involved?
- How much?
- How many?
- ...
- Why did it happen?

Is it well-justified?

Facts

**Explanations
(Theories)**

Design problems

- How to treat this problem?
- How to use this tool?
- How to achieve this goal?
- **How does it work?**
- How to improve the business?
- How to learn this technique?
- How to test this program?
- ...

How to do something

The top-level distinction: evaluation criteria for finite minds

Knowledge questions

- What is the case?
- When?
- Where?
- How did it happen?
- Who was involved?
- How much?
- How many?
- Why did it happen?

Is it well-justified?

Sound research methodology?

Clear presentation?

Facts

**Explanations
(Theories)**

Design problems

- How to treat this problem?
- How to use this tool?
- How to achieve this goal?
- How to improve the business?
- How to learn this technique?
- How to test this program?
- ...

How does it work?

How to do something

The top-level distinction: evaluation criteria for finite minds

Knowledge questions

- What is the case?
- When?
- Where?
- How did it happen?
- Who was involved?
- How much?
- How many?
- Why did it happen?

Is it well-justified?

Sound methodology?

Clear presentation?

Facts

**Explanations
(Theories)**

Design problems

- **How does it work?**
- How to use this tool?
- **Which artifact?**
 - **Components and their capabilities.**
 - **Interactions among components.**
- **How applied?**
- ...

How to do something

Summary and preview

Knowledge questions

- Ask for facts and explanations (theories)
- Answers evaluated on truth
- Truth claims must be well-justified
- This requires sound empirical research method
- Describe the research setup, data, and inferences from the data

Design problems

- Ask how to do something
- Answered by a treatment in which an artifact is applied to the problem
- Treatment evaluated by utility
- Artifact described by its architecture and internal interactions

Take a breath



← Design problems

← Knowledge questions

Back to the engineering cycle

The engineering cycle

- Problem investigation
- Treatment design
- Treatment validation
- Treatment implementation
- Implementation evaluation

An example

- *Why are our IS projects late?*
- *New effort estimation technique*
- *Test with experts; do pilot project*
- *Roll out in the organization*
- *Have estimations now improved?*

The engineering cycle

The engineering cycle



- Problem investigation
- Treatment design
- Treatment validation
- Treatment implementation
- Implementation evaluation

Solves a design problem.

- Researchers solve general problems;
- Consultants & engineers solve single problems.

Knowledge questions in the engineering cycle

The engineering cycle

- Problem investigation →
 - Treatment design
 - Treatment validation
 - Treatment implementation
 - Implementation evaluation
- **Knowledge questions:**
 - Stakeholders, goals, problematic phenomena

Design problem in the engineering cycle

The engineering cycle

- Problem investigation
- Treatment design
- Treatment validation
- Treatment implementation
- Implementation evaluation



Design problem

- Artifact components, their capabilities and interactions;
- Mechanisms: How does the artifact respond to events and conditions in the problem context?

Knowledge question in the engineering cycle

The engineering cycle

- Problem investigation
- Treatment design
- Treatment validation
- Treatment implementation
- Implementation evaluation



Knowledge question

Would this work if applied to a problem context?

- Effects?
- Trade-offs (different designs)?
- Sensitivity (different contexts)?
- Requirements satisfaction?

Implementation is transfer to the problem context

The engineering cycle

- Problem investigation → Real-world problem
- Treatment design
- Treatment validation
- Treatment implementation → **Technology transfer:**
Production, marketing,
- Implementation evaluation → **Knowledge question**
 - Stakeholders, effects, goal contribution?

The engineering cycle

- Problem investigation
- Treatment design
- Treatment validation
- Treatment implementation
- Implementation evaluation

Example

- *IS projects are always late*
- *New effort estimation technique*
- *Test in pilot projects*
- *Documentation, training, sales*
- *Customer satisfaction survey*

By “implementation” I mean: transfer to the problem context.

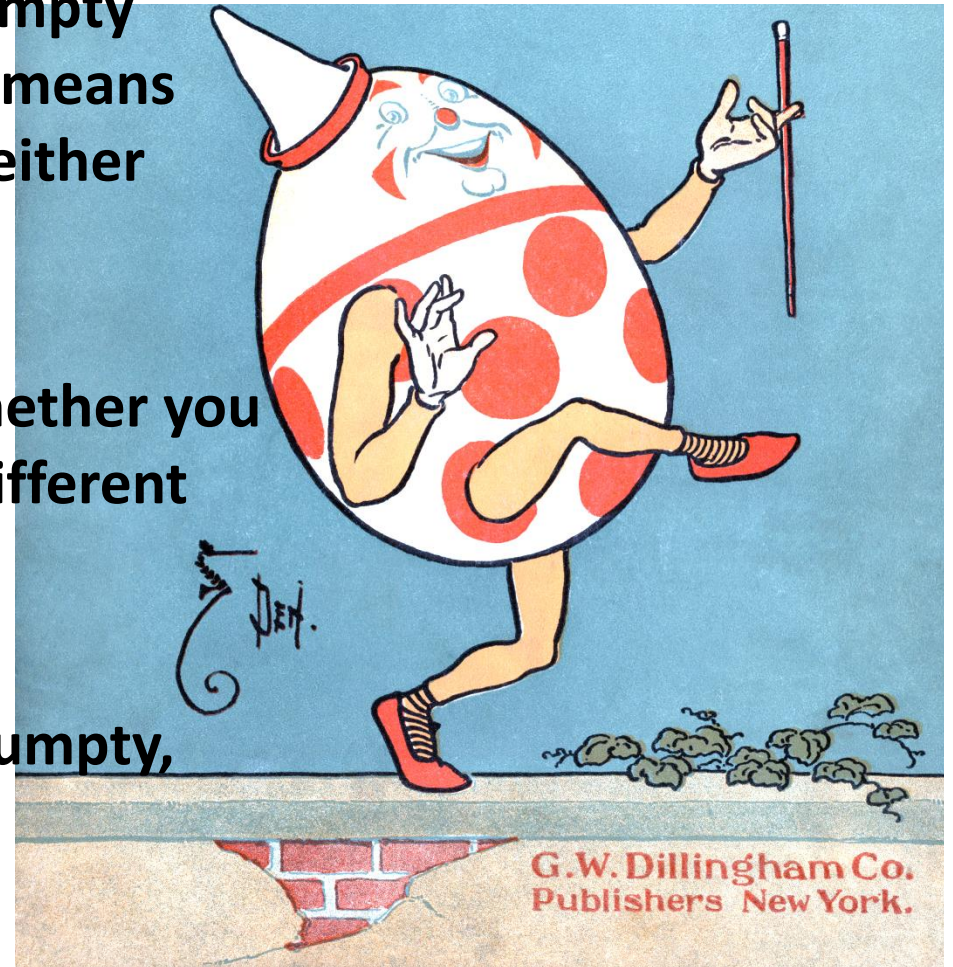
- What this means depends on what I think is the problem context.
- I think the problem context is in the **real** world.
 - Constructing a prototype is implementing a solution to **another** problem, namely to test a design idea.

Humpty Dumpty

"When I use a word," Humpty Dumpty said, in rather a scornful tone, "it means just what I choose it to mean—neither more nor less."

"The question is," said Alice, "whether you *can* make words mean so many different things."

"The question is," said Humpty Dumpty, "who is the boss—that's all."



Design science research

The engineering cycle

- Problem investigation
 - Treatment design
 - Treatment validation
 - Treatment implementation
 - Implementation evaluation
- Knowledge question
 - Design problem
 - Knowledge question
 - Knowledge question

Typical design research project: Design cycle

Take a deeper breath



← Design problems, design cycle

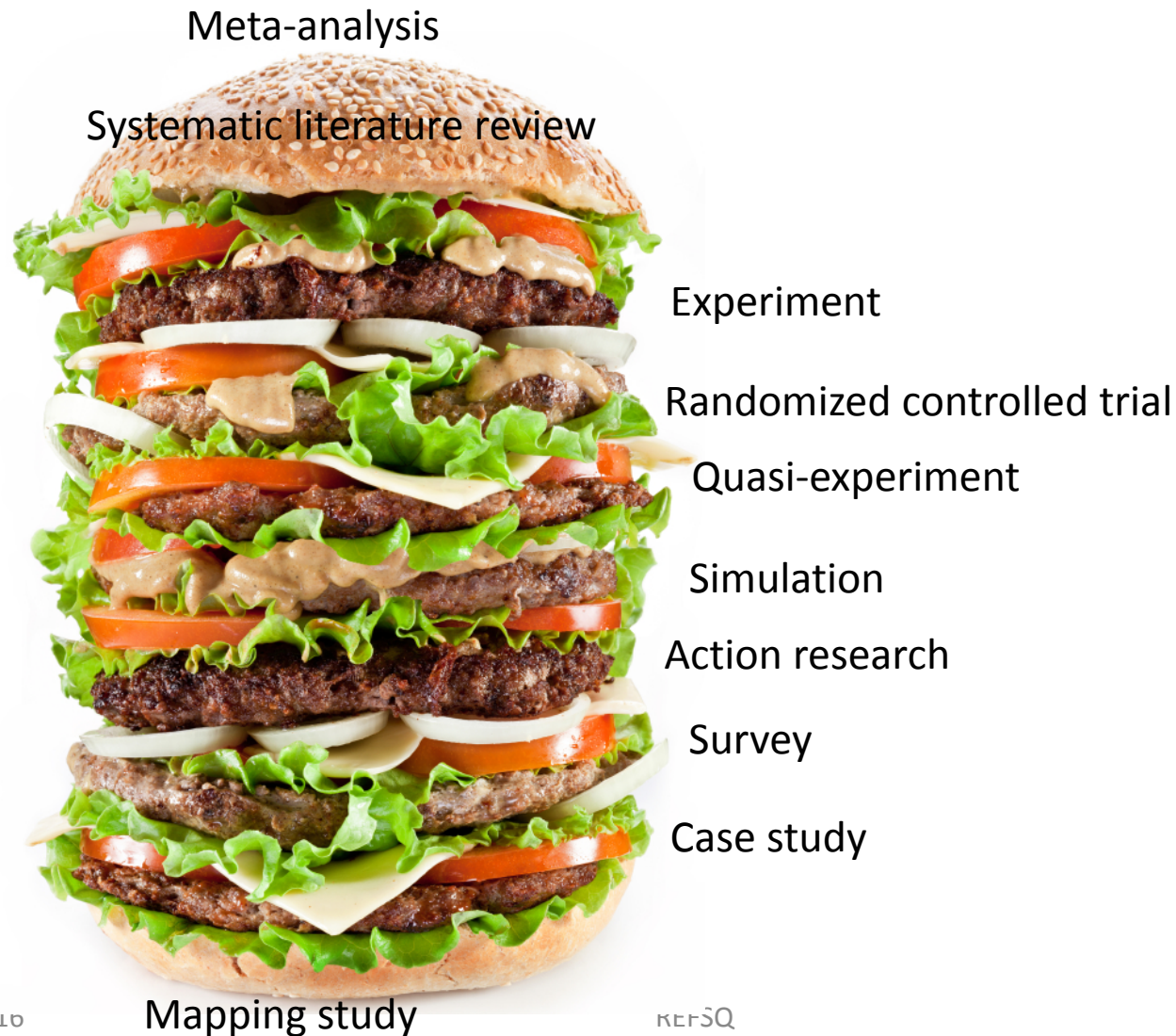
← Knowledge questions

First update to classification schema

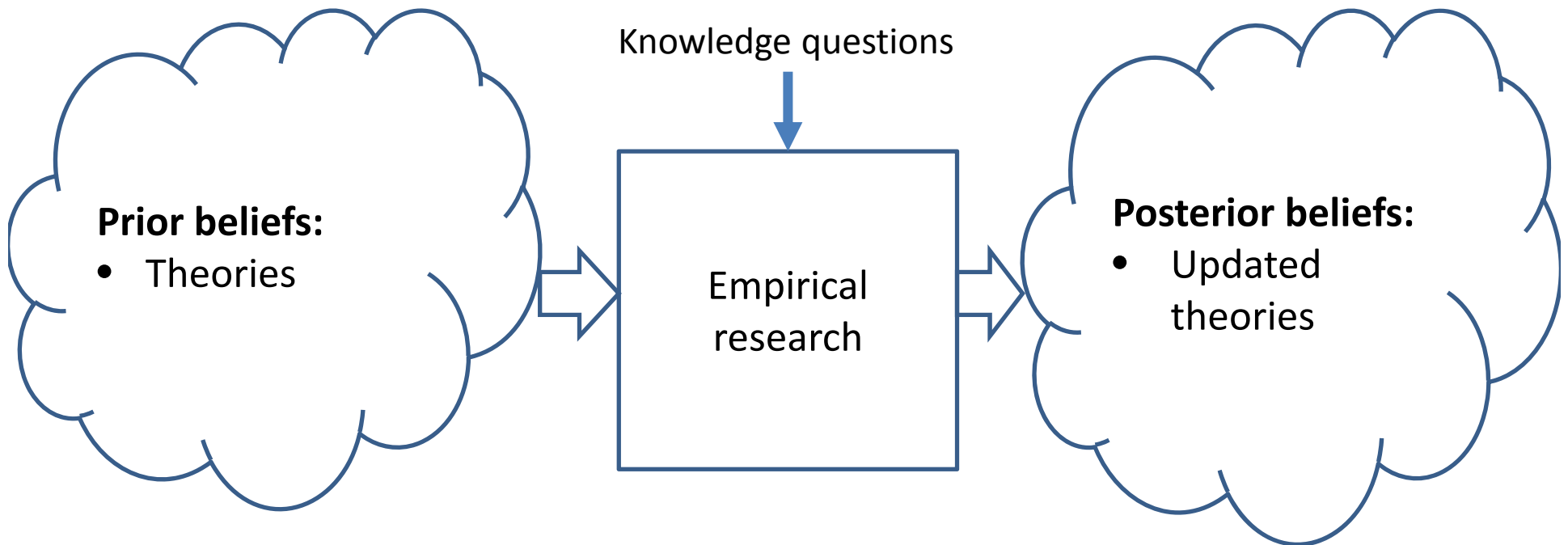
- **Problem research/implementation evaluation** → Investigate an artifact that is implemented in the real world
- **Treatment design** → Propose a new artifact design
- **Treatment validation** → Investigate an artifact that is still experimental



What kinds of empirical research methods?



Empirical research



- The goal of empirical research is to develop, test or refine theories
- Design science: theories about how an artifact works, what effects it has on a problem context, and why

What kinds of empirical research methods?

- Classification dimensions of empirical methods:
 - Field research or laboratory research
 - Observational or experimental
 - Case-based or sample-based

Research designs

	Observational study (no intervention)	Experimental study (intervention)
<p>Case-based: investigate single cases, look at architecture and mechanisms</p>	<p>Observational case study: study phenomena in the field</p>	<ul style="list-style-type: none"> • Action research: Work with stakeholders to analyze problem • Lab experiment: Test prototype in simulated context • Field experiment: test prototype in real problem context • Technical action research: Use prototype to treat real problem
<p>Sample-based: investigate samples drawn from a population, look at averages and variation</p>	<p>Survey of a study population.</p>	<ul style="list-style-type: none"> • Statistical difference-making experiment (treatment group – control group experiments in lab or field)

Other research designs

- **Mapping study:** classify what has been written about a topic
- **Systematic literature review:** summarize what has been written about a topic
- **Meta-analysis:** summarize and integrate what has been written about a topic
- **Mathematical analysis:** prove properties of an abstract conceptual structure

Empirical research papers should follow the **empirical cycle**

Description of research **Justification of choices made**

- Research questions Motivation and clarification of questions
 - Research design Validity of design for these questions
 - Research execution
 - Data analysis: Validity of analysis
 - description, (facts)
 - explanation, } (theory)
 - generalization }
- Validity = degree of support

Take a still deeper breath



← Design problems, design cycle

← Knowledge questions, empirical cycle

Now we can take a bite

Meta-analysis

- Problem investigation
- Treatment design
- Treatment validation
- Treatment implementation
- Implementation evaluation

Design problems, design cycle

Knowledge questions

Systematic

literature review

Experiment

Randomized
controlled trial

Quasi-
experiment

Simulation

Survey

Case study

Action research

Mapping study



Second update to classification schema

- **Problem research/implementation evaluation**
- **Treatment design**
- **Treatment validation**

Knowledge questions and design problems

- **Problem research/implementation evaluation Treatment validation**
Knowledge questions

- **Treatment design**
Design problems

Research goals

- **Knowledge-oriented papers**
 - Research goals: investigate existing real-world problem, or evaluate existing real-world implemented artifact, or validate newly designed artifact

- **Technical design papers**
 - Research goal: describe and explain new artifact design

Research methods & criteria

- **Knowledge-oriented papers**

- Research goals: investigate existing real-world problem, or evaluate existing real-world implemented artifact, or validate newly designed artifact
- Research method: analytical argument or empirical cycle (research questions, design, execution, data analysis)
- Criteria: interesting research questions, sound research design, clarity of presentation of measurements, validity (degree of support) of conclusions

- **Technical design papers**

- Research goal: describe and explain new artifact design
- Research method: design cycle (problem, treatment design, validity argument)
- Criteria: problem relevance, novelty of design, clarity of presentation of design, technical soundness of the design, analytical argument for validity.

Paper classification schema: REFSQ 16

- **Full research papers**, including literature reviews, evaluation research, solution proposals, and validation research.
 - **Experience reports**, describing positive and negative experiences.
 - **Vision papers**, stating where the research in the field should be heading towards.
 - **Problem statements**, describing open issues of practical or theoretical nature.
 - **Research previews**, reporting research results at an early stage.
 - **Knowledge-oriented papers**: various research goals, analytical or empirical methods (experiments, simulations, case studies, action research, mapping studies, syst. literature reviews, meta-analysis, ...)
 - **Technical design papers**
 - **Research programme?**
 - **Project proposal?**
 - **Research design, not yet executed?**
Artifact design, not elaborated?
-
- The diagram consists of three blue arrows originating from the first bullet point. One arrow points from the word 'evaluation' to the first bullet point of the second column. A second arrow points from the phrase 'solution proposals, and validation research' to the second bullet point of the second column. A third arrow points from the word 'validation' to the third bullet point of the second column.



The end