Experiment validity - Construct

Ivano Malavolta

VRIJE UNIVERSITEIT AMSTERDAM

LOOKING FURTHER

Experiment validity

Validity is the extent to which our results are **SOUND** and

APPLICABLE TO THE REAL WORLD

- We aim for **adequate** validity, not **universal** validity

  ○ What matters is our population of interest

- Validity is in *trade-off* with experiment scope

Ivano Malavolta / S2 group / Green Lab

VU

Threats Identification

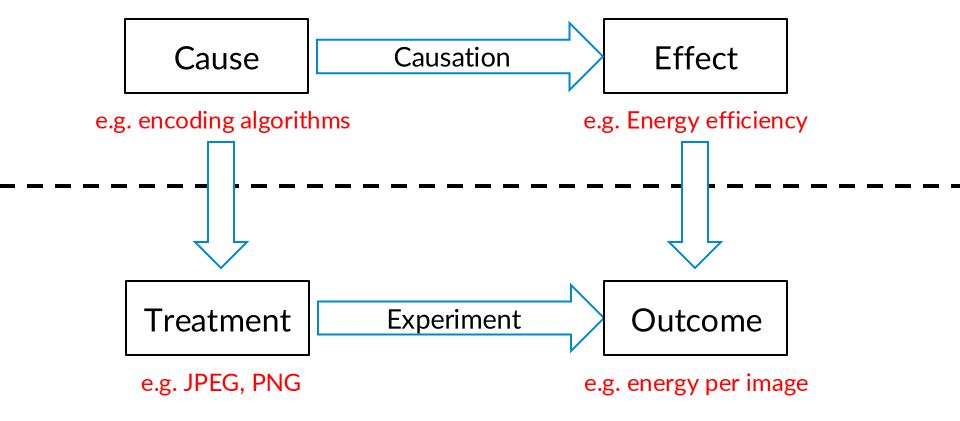- Identifying **threats** helps to plan for adequate validity

- Each threat needs appropriate **mitigation**

- Several classifications of validity threats:

  - Campbell and Stanley [1]

  - **Cook and Campbell [2]**

Ivano Malavolta / S2 group / Green Lab

VU

Types of threat to validity

Theory

Observation

Ivano Malavolta / S2 group / Green Lab

VU

Types of threat to validity

Theory

```
┌─────────────┐                    ┌─────────────┐
│    Cause    │ ====Causation====> │   Effect    │
└─────────────┘                    └─────────────┘
e.g. encoding algorithms           e.g. Energy efficiency

Construct  ↓                          ↓  Construct

┌─────────────┐                    ┌─────────────┐
│  Treatment  │ ===Experiment====> │   Outcome   │
└─────────────┘                    └─────────────┘
e.g. JPEG, PNG                     e.g. energy per image
```

Observation

Ivano Malavolta / S2 group / Green Lab

VU

Construct validity

Construct Validity: relation between theory and observation

- Have I defined my constructs properly?

  where constructs = dependent and independent
  variables, treatments, factors, co-factors, etc.

- Am I analyzing the correct variables for the effects?

Ivano Malavolta / S2 group / Green Lab

VU

Construct validity: types of threat

- **Inadequate preoperational explication of constructs**
  - construct not well defined before being translated into measures
  - Theory unclear
  - Comparing two methods, but not clear what does it mean that a method is better than another

- **Mono-operation bias**
  - I have only one independent variable, one single object or treatment
    $\rightarrow$ the experiment could not represent the theory

- **Mono-method bias**
  - When you use a single type of measures or observations
  - The experimenter may bias the measures

Ivano Malavolta / S2 group / Green Lab

VU

Construct validity: mitigation

✓ Early definition of constructs (GQM)

✓ Use appropriate experiment design
we will have a dedicated lecture about this

✓ Justify your choices for factors and treatments

✓ Introduce redundancy for cross-checks

VU

Readings

Claes Wohlin · Per Runeson
Martin Höst · Magnus C. Ohlsson
Björn Regnell · Anders Wesslén

Experimentation in Software Engineering

Springer

Chapters 7, 8.7, and 8.8.3

+ [FGCS paper](#) (Sections 5 and 6.3.2) on Canvas

+ All papers in the "Articles on performed experiments" folder on Canvas

(only the part related to Goal and research questions)

Ivano Malavolta / S2 group / Empirical software engineering

VU

Acknowledgements

Some contents of lecture extracted from:

● Giuseppe Procaccianti's lectures at VU

VU