

Experiment validity - Conclusion

Ivano Malavolta

Types of threat to validity

Theory



e.g. encoding algorithms

e.g. Energy efficiency



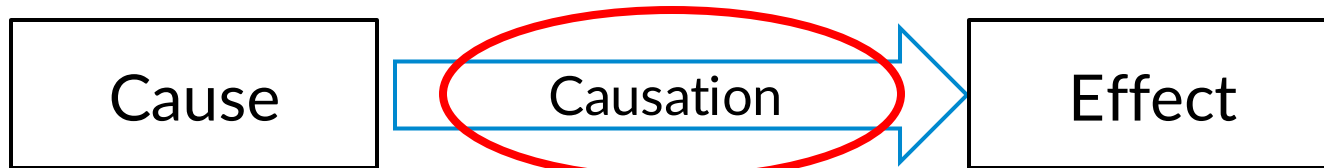
e.g. JPEG, PNG

e.g. energy per image

Observation

Types of threat to validity

Theory



e.g. encoding algorithms

External

e.g. Energy efficiency

Construct

Construct



e.g. JPEG

Internal

e.g. energy per image

Conclusion

Observation

Conclusion validity

Conclusion Validity: **statistical correctness and significance**

- Are my conclusions correct?
- Are my results significant enough?

Conclusion validity: types of threat

- Low statistical power
 - Results not statistically significant
 - There is a significant difference but the statistical test does not reveal it due to the low number of data points
- Violated assumptions of statistical tests
 - eg, many tests assume normally distributed samples
- Fishing and error rate
 - If you are combining multiple statistical tests, also their significance should be adapted (Bonferroni, etc.)

Conclusion validity: mitigation



Select appropriate tests



Aim for high levels of statistical power

Acknowledgements

Some contents of lecture extracted from:

- Giuseppe Procaccianti's lectures at VU