



Counterfactual impact evaluation and policy—challenges revisited

Rahel Kahlert, Ph.D.

Rigorous Impact Evaluation in Europe

A Research Conference in Honor of Alberto Martini

May 20-22, 2018

Collegio Carlo Alberto

Torino, Italy

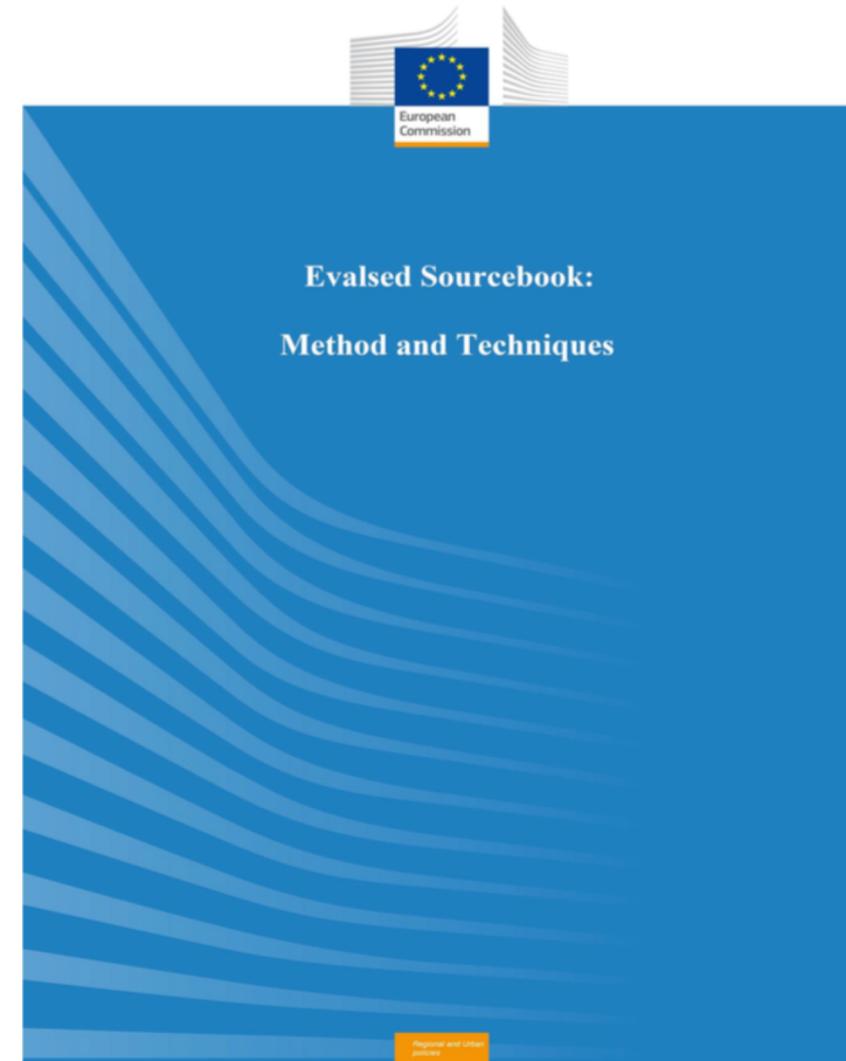
Context: EVALSED chapter

Directorate General for Regional Policy: EVALSED Sourcebook, Methods and Techniques (2013)

EVALSED: The resource for the EVALuation of Socio-Economic Development

New chapter on impact evaluation (p.47-119):

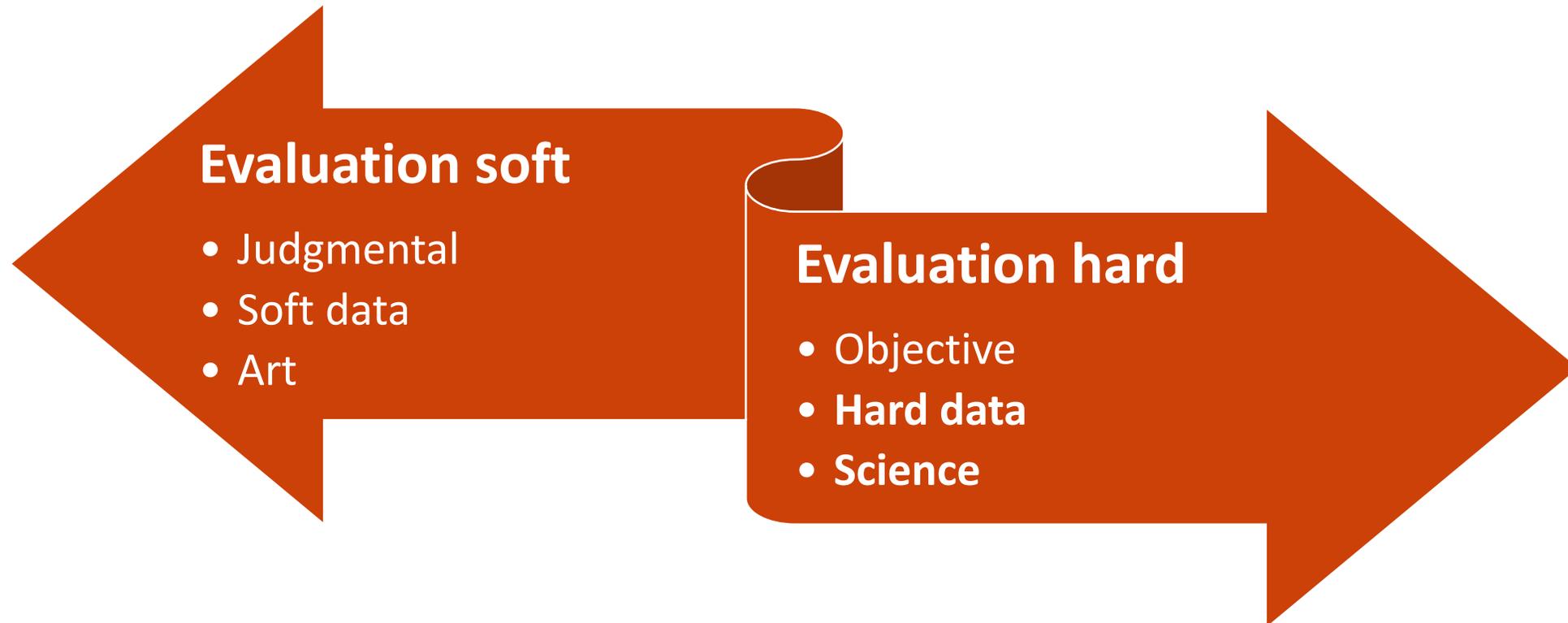
- Theory-based impact evaluation (Frans Leeuw)
- Counterfactual Impact Evaluation (Alberto Martini)



Context: Professionalization of Evaluation in Europe

- European Evaluation Society since 1996
- Network of Evaluation Societies in Europe (26 members)
- National standards (e.g., Swiss since 2001; German since 2001, French since 2003)
- University-based education programmes
- Guidebooks for decision-makers (Evalsed)

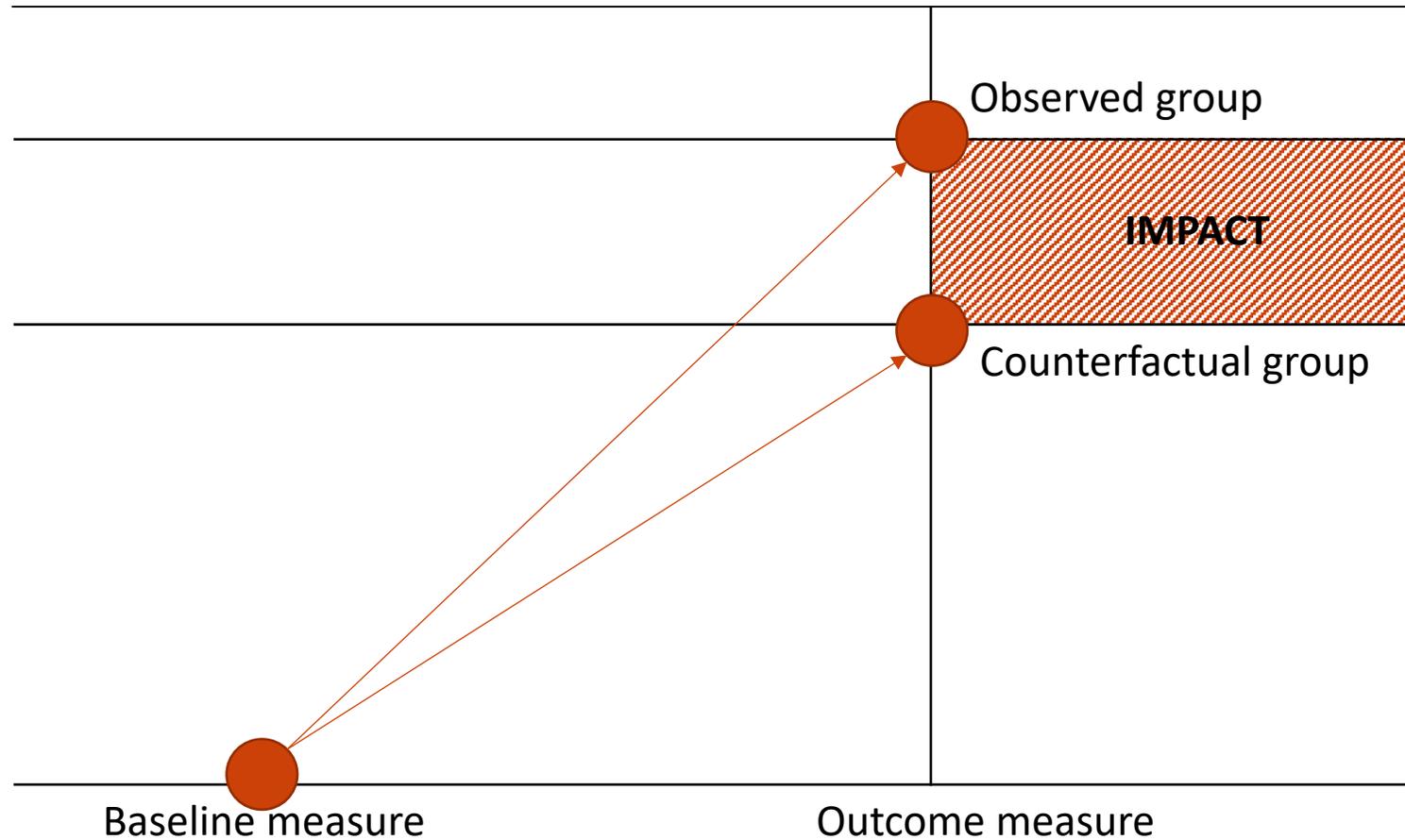
Context: Myths and misconceptions



Martini's definition of CIE

- Quantification of effect:
 - “quantification whether a given intervention produces desired effects on some pre-established dimension of interest” (77)
 - Difference between the outcome occurred *after* an intervention has taken place and the outcome that would have occurred *in the absence* of the intervention”

= Impact as net effect



Two definitions of impact

separate concepts, different implications

Impact as net effect

- World Bank
- JPAL
- Drug agencies

Impact as long-term effect

- OECD DAC definition
- NONIE

Impact as long-term effects

Long-term effects produced can be multifold:

- Positive and negative
- Primary and secondary
- Direct and indirect
- Intended and unintended
- Economic, socio-cultural, institutional, environmental, technological

Adapted from OECD DAC Glossary of Key Terms in Evaluation and Results Based Management

Objective: More nuanced understanding of CIE

For more informed decisions and better policy solutions

- By discussing the eight challenges that Martini lists in the EVALSED Sourcebook (p 81)
- By putting these in a wider policy and practice context

8 Challenges

1	Impact estimates as difficult to generalize
2	Being costly
3	Requiring close monitoring
4	Posing ethical challenges that are politically sensitive
5	Requiring a fairly stable environment
6	Requiring an intervention being fairly simple
7	Implementation altering the intervention (randomization bias)
8	Implementation introducing non-compliance bias

Challenge 1: Impact estimates as difficult to generalize

Champbell 1957

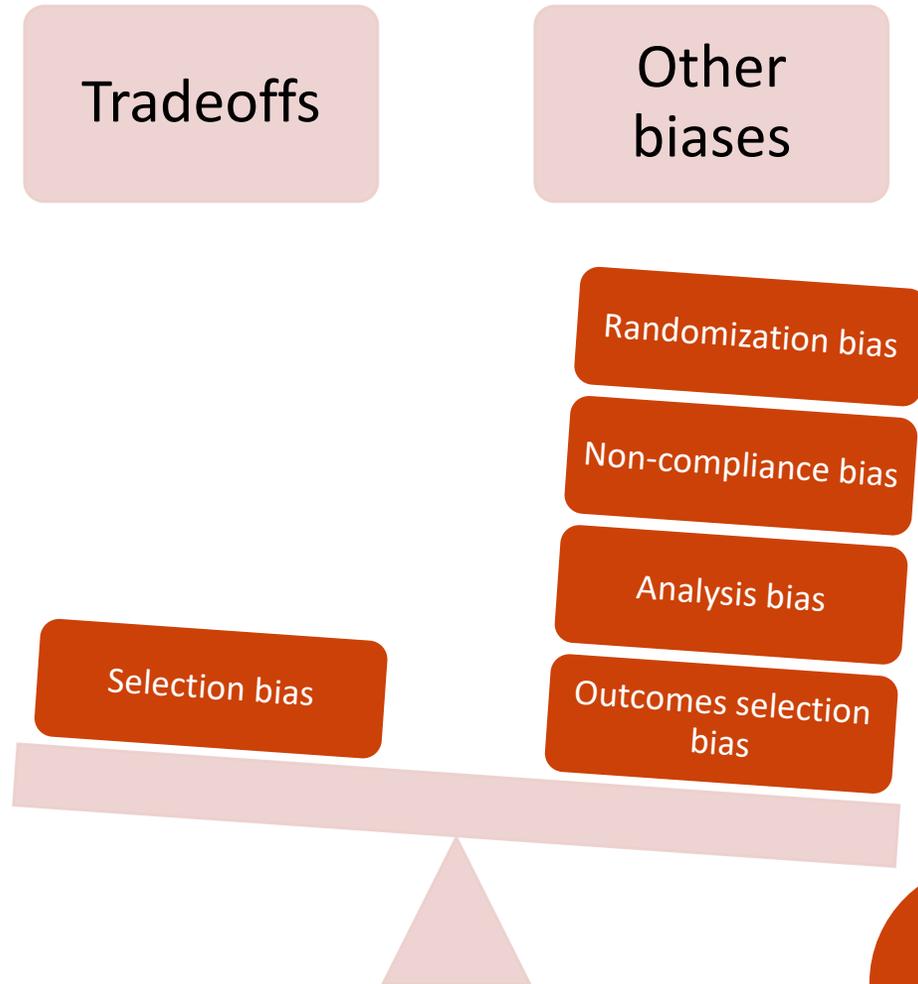
Internal validity

- basic minimum
- primacy

External validity

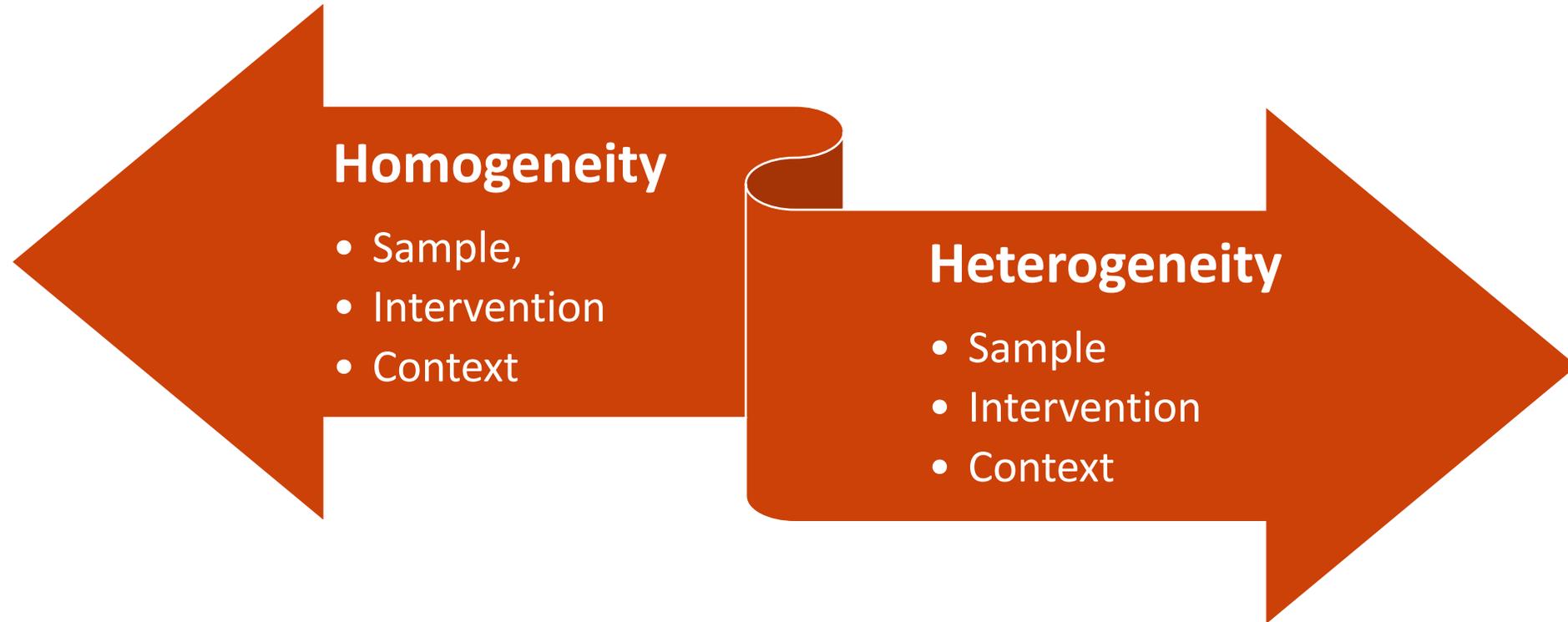
- importance for decision making

Biases: randomization bias (7), non-compliance bias (8)



Use of qualitative-interpretive approaches to detect various biases

Simple intervention (6), stable environment (5)



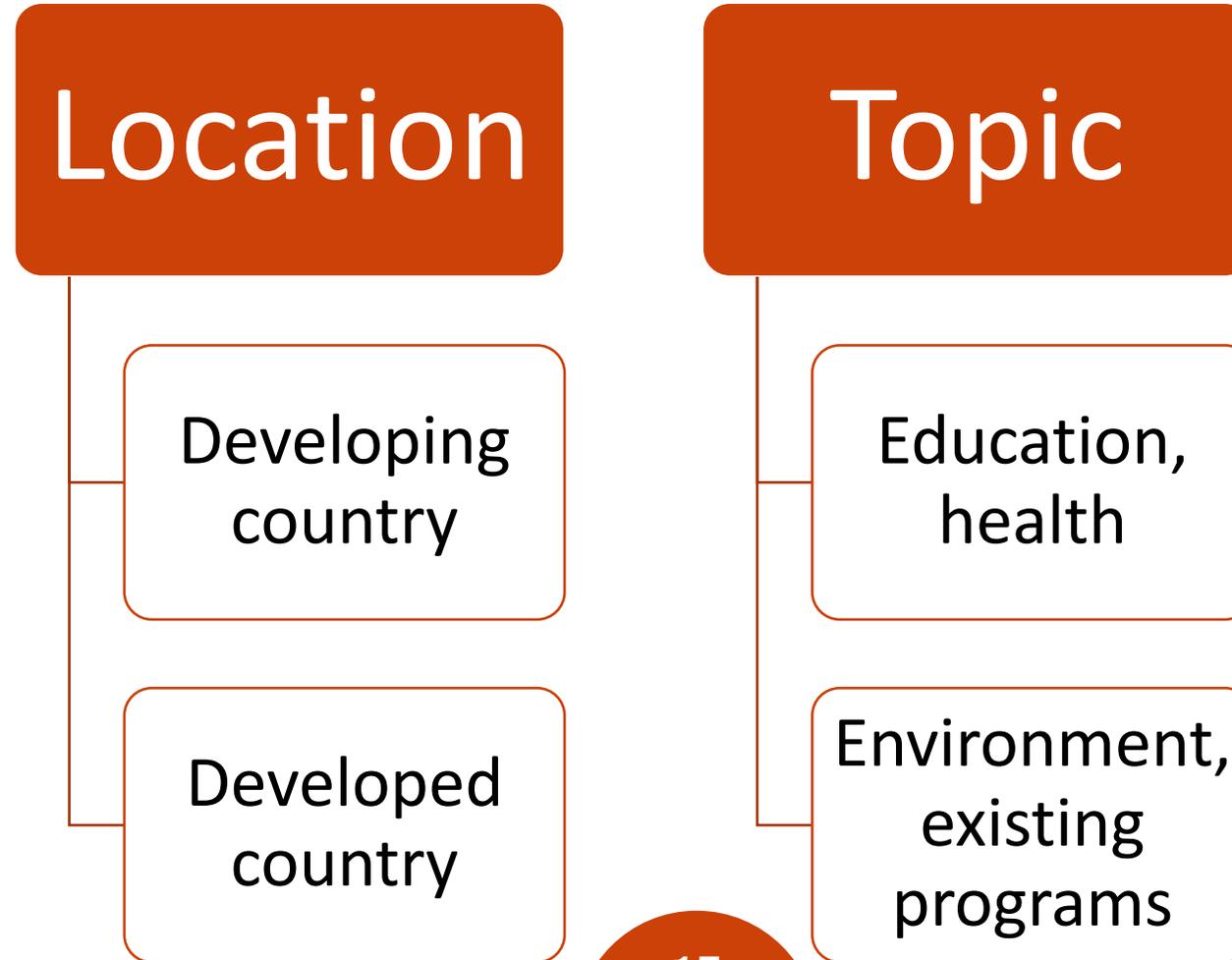
- How much out of control can experiments be? (Hawe et al 2004)
- Heterogeneity in subjects, intervention, context may lead to insignificant findings.
- Even if findings are significant, what elements of intervention worked.

Challenge 4: Posing ethical challenges that are politically sensitive

- As soon as intervention proves to be effective, it is not ethical to perform an experiment.
- Cf. Tuskegee experiments
- Challenge: Many interventions do already exist, are not brand new. There is usually some preliminary evidence whether they are effective or not.

Cost factor (2) can become a cost bias

Plus possible intervention selection bias



Addressing challenges 1: Clarify qualitative-interpretive components of experiments

Prior: Analysis of prior research

Select the hypothesis to be tested

Decide on heterogeneity of sample, required effect size

Predict potential effects

Interpret findings and determine how representative these are

Post: Generate policy conclusions

Conclusion: Addressing challenges 2

- Include qualitative-interpretive approaches into impact evaluation to interpret biases and open black box
- Clarify representativeness of sample
- Create interdisciplinary evaluation teams with quantitative, qualitative and mixed-methods skills
- Encourage systematic reviews to compare impact findings across intervention contexts

Selected references

Evaluated Sourcebook: Methods and Techniques, incl. Martini's Counterfactual IE (2013)

http://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/evaluation_sourcebook.pdf

Jump@School. Testing a model to contrast early school leaving (2016)

https://www.zsi.at/object/project/3257/attach/0_Final_Publication_-_EN.pdf

Kahlert (2012) Randomized controlled trials to evaluate impact. Their challenges and policy implications in medicine, education and international development, Texas Digital Library,

<https://repositories.lib.utexas.edu/handle/2152/19463>

Piso and Tüchler (2016). Antwort auf die Stellungnahme der AGAKAR zur prospektiven Kardiorehabilitationsstudie.

http://eprints.hta.lbg.ac.at/1101/3/Stellungnahme_der_AGAKAR_und_AW_des_LBI-HTA.pdf

Scriven (2008) A Summative Evaluation of RCT Methodology: & An Alternative Approach to Causal Research. Journal of MultiDisciplinary Evaluation, 9(5), 11–24.

Alejandro R. Jadad, Murray W. Enkin (2008) Bias in randomized controlled trials. Chapter 3 in: Randomized Controlled Trials: Questions, Answers, and Musings, 29–47.

Comments and questions are welcome!

Rahel Kahlert

kahlert@euro.centre.org

European Centre for Social Welfare Policy and Research