

“How can recent thinking on methodologies and technologies be put in practice, so that robust and reliable impact measurement contributes to enhanced learning and improved quality of interventions?”

# Two sub-questions

- How to develop...
  1. *“robust and reliable impact measurement”*
- That will contribute to...
  2. *“enhanced learning and improved quality of interventions”*

# Lots of studies of evaluation use

- Raab, et al, 2014. *Review of evaluation approaches and methods for interventions related to violence against women and girls (VAWG)*. [39 evaluations]
- Herbert, 2014. *Researching Evaluation Influence A Review of the Literature*. [28 studies]
- Ledermann, 2012. *Exploring the Necessary Conditions for Evaluation Use in Program Change*. [11 evaluations]
- Johnson, et al 2009. *Research on Evaluation Use A Review of the Empirical Literature From 1986 to 2005*. [41 studies]
- Shulha, et al, 1997. *Evaluation Use: Theory, Research, and Practice Since 1986* [41 studies]
- Cousins & Leithwood, 1986. *Current Empirical Research on Evaluation Utilization*. [65 studies]

# Many factors affect M&E use

## 1. Quality of M&E process and product

- Design – purpose & approach
- Participation and ownership
- Planning (being timely)
- Evidence
- Follow up mechanisms
- Evaluator credibility

## 2. Relational factors

- Personal (e.g. relationship of trust with evaluator), interpersonal and collective level (e.g. within
- organisation or with stakeholders)
- Role and influence of an evaluation unit
- Networks, communities of practice

## 3. Organisational factors

- Culture
- Structure
- Knowledge management & organisational learning

## 4. External influences

- Indirectly involved stakeholders (not direct users) whose actions can affect the use (or non-use) of an evaluation.

Visser, I., Kusters, C., Roefs, M., Bulzer, N., 2014.

Conference: *Improving the use of monitoring & evaluation processes and findings.*

- **At best, quality of evaluation methods is only one of a number of factors affecting use**
  - “Findings point to the importance of stakeholder involvement in facilitating evaluation use and suggest that engagement, interaction, and communication between evaluation clients and evaluators is critical to the meaningful use of evaluations” (Johnson, 2009)
  - Method attributes = 2 of 5 conditions associated with evaluation use, in study of 39 evaluations (Raab, et al, 2014)

- **Robust, reliable, rigorous methods do not seem to be as important as methodological debates might suggest**
- “Utility as first and most important evaluation standard” - the Wageningen position
  - But what value is greater use of evaluation findings if it is based on weak data and poor analysis?
- In practice there will be trade-offs. E.g.
  - RCT by specialists may improve quality of evidence base but have weak stakeholder engagement,
  - Participatory evaluation may have good stakeholder engagement but weak data and analysis

# The way forward...?

## A convergent approach...

1. Enable more stakeholders to know how to do good data collection and analysis
  - Promote “evaluative thinking”\*\*
    - This has wider relevance: informed citizenship
2. Find methods that fit the circumstances
  - A. Complexity-consistent view causation
  - B. Data that is more readily available
  - C. Appropriate analytic tools

# 1. Evaluative thinking

- Assumption: With more understanding of the data at hand, and of ways of analysing it, stakeholders will be more likely to make use of findings.
- Three areas of possible stakeholder engagement to discuss later:
  - SoC collection and analysis
  - Survey planning, collection and analysis
  - Linking SoC and Survey data and analysis



## 2. Appropriate methods... (some of the) recent thinking

- Stern, et al, 2012. *Broadening the Range of Designs and Methods for Impact Evaluations*
- Goertz & Mahoney, 2012. *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences.*
- Kotu & Deshpande, 2014. *Predictive Analytics and Data Mining*

# A. Better ideas about causation...

- Identify *different types* of causal conditions
  - **Sufficient but Unnecessary conditions**
    - A bike will get me to the station, but there are other ways of getting there as well
  - **Necessary but Insufficient conditions**
    - I may have to take the car, but there are other things needed as well e.g. parking space
  - **Necessary and Sufficient conditions**
    - Sometimes a bike will be the only option, and that will be enough
  - **Configurations of conditions that are S, N, or N&S**
    - If I have a car and a parking space that will be sufficient to get me to the station

# All can be subject to *logical* tests

Is “high literacy” Necessary or Sufficient for a country to be a “parliamentary democracy”?

- A claim of Necessity can be disproved by one contrary case of a country without high literacy but with parliamentary democracy
- A claim of Sufficiency can be disproved by one contrary case of a country with high literacy but without parliamentary democracy

Cross-tabulation of survey data can identify S, N or S&N conditions associated with an outcome of interest

A single SoC or data point can disprove a claim

# A. Better ideas about causation...

Causal configurations...(combinations of conditions)

- Most outcomes are not the results of a single cause, but of a configuration of multiple causal conditions (= conjunctural causation)
- An outcome can result from more than one type of configuration (= equifinality)
- The absence of an outcome may be for reasons other than the absence of the causal configuration (= asymmetry)

## B. Tools to identify causal configurations

### 1. Qualitative Comparative Analysis (QCA)

- Where there is a plenty of relevant theory

### 2. Decision Tree modelling

- Where there is more data than theory

### 3. Ethnographic Decision Tree modelling

- Where specific stakeholder views are important

All can use the same kind of data and their results can be compared with each other

# C. Use of a wider range of types of data

WOMEN'S REPRESENTATION IN PARLIAMENT

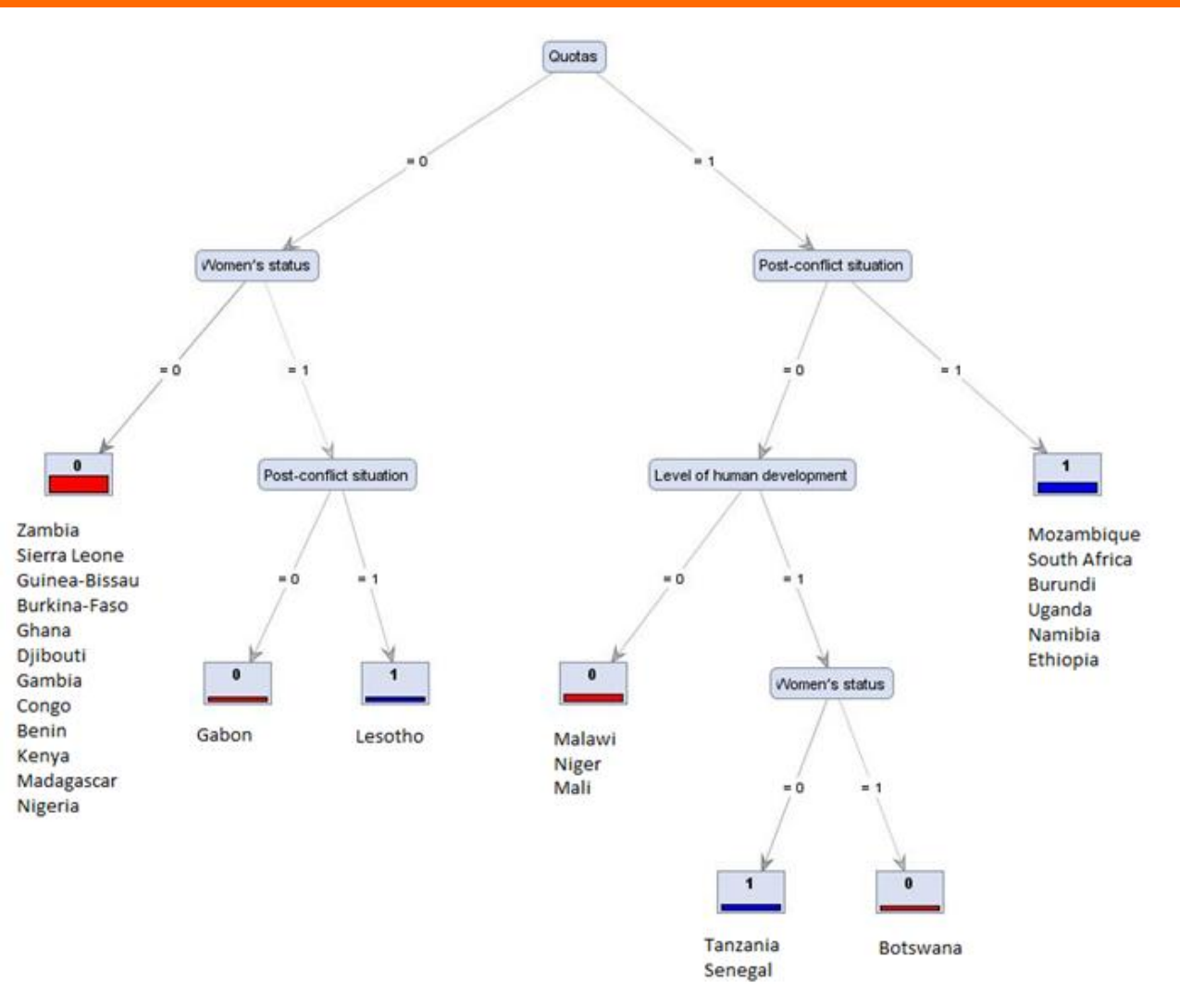
Country	Electoral system	Quotas	Women's status	Level of human development	Post-conflict situation	% Women in national parliament
Mozambique	1	1	0	0	1	1
South Africa	1	1	1	1	1	1
Burundi	1	1	0	0	1	1
Tanzania	0	1	0	1	0	1
Uganda	0	1	1	1	1	1
Namibia	1	1	1	1	1	1
Lesotho	0	0	1	1	1	1
Senegal	0	1	0	1	0	1
Ethiopia	0	1	0	0	1	1
Zambia	0	0	0	1	0	0
Sierra Leone	1	0	0	0	1	0
Guinea-Bissau	1	0	0	0	1	0
Malawi	0	1	1	0	0	0
Gabon	0	0	1	1	0	0
Niger	0	1	0	0	0	0
Burkina Faso	1	0	0	0	1	0
Botswana	0	1	1	1	0	0
Ghana	0	0	0	1	0	0
Djibouti	0	0	0	1	1	0
Mali	0	1	0	0	0	0
Gambia	0	0	0	1	0	0
Congo	0	0	0	1	1	0
Benin	1	0	0	1	0	0
Kenya	0	0	0	1	0	0
Madagascar	0	0	0	1	0	0
Nigeria	0	0	0	1	0	0

Categorical data, where measurement is not practical

Dichotomised measurements, where they are available

Krook, M. L. 2010. "Women's Representation in Parliament: A Qualitative Comparative Analysis." *Political Studies* 58 (5): 886–908.

# = Simple enough models: A Decision Tree



Nodes = conditions/attributes

Branches = configurations of condition

Leaves = outcomes associated with each configuration

Branches do NOT= causal pathways. Conditions are not in temporal order

# How good is good enough?

- Configurational models vary in
  - Proportion of cases they apply to
  - Accuracy of prediction of outcomes in those cases
- Simpler models tend to have higher coverage but lower accuracy
- Needs for accuracy vary
  - Surgeons will want very high levels
  - Gamblers can use anything above 50%
  - Simple-to-use models are also in demand



# Flexibility yes, but complacency, no!

- Reliability, as in replicability of findings, is a major problem in many fields of research
  - Ioannidis, 2005. *Why Most Published Research Findings Are False.*
- Ways forward? (relevant to WCP?)
  - Open data & wider public participation
  - Transparent process of analysis

*“With enough eyeballs (and data) all bugs are shallow”*

# The World Citizens Panel opportunity

- Use across-case analysis to find configurations of conditions associated with an outcome
  - Association is a Necessary but Insufficient basis for a claim of causation
    - Association + Mechanism = Plausible claim of causation
- Use within-cases analysis to uncover causal mechanisms at work

Can Stories of Change and Surveys provide useful information on Mechanisms and Associations respectively?