Using Innovative Methods in Evaluation – What's Needed?

Patricia Rogers, PhD p.rogers@anzsog.edu.au

BetterEvaluation ANZSOG (Australia and New Zealand School of Government)

Australasian Evaluation Society conference

Canberra, Australia 4 September 2017

What do we mean by innovation in evaluation?

A change – at least in that setting-

That adds value – such as solving a problem, meeting a need, requiring fewer resources

Some types of innovation

Invention

New technology or new process

Transfer or translation

Bringing in an idea from another setting or another purpose, and possibly adapting it

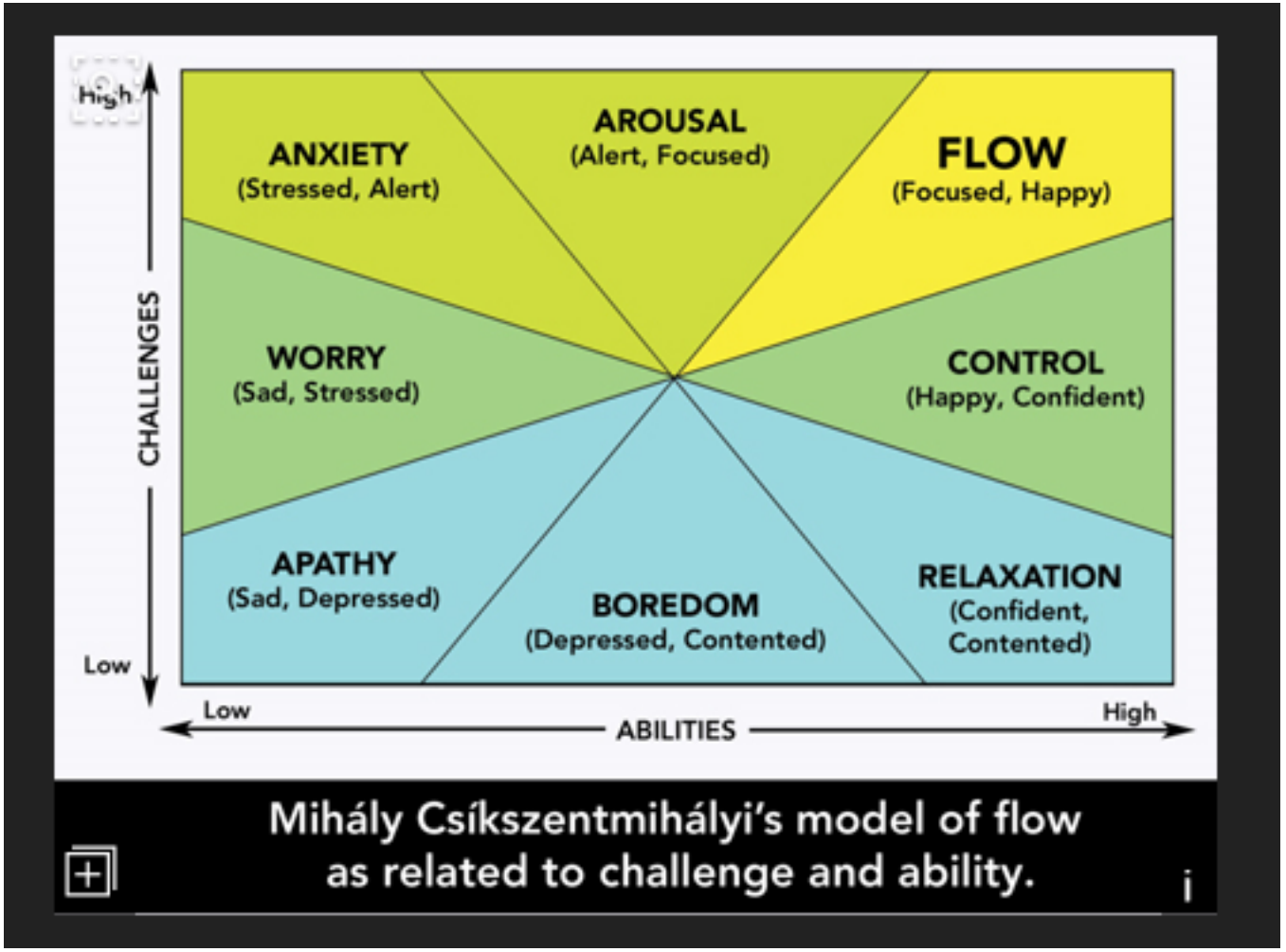
Bricolage

Gathering together existing elements in a new way

Systematisation

Documenting and making explicit and systematic some existing practices

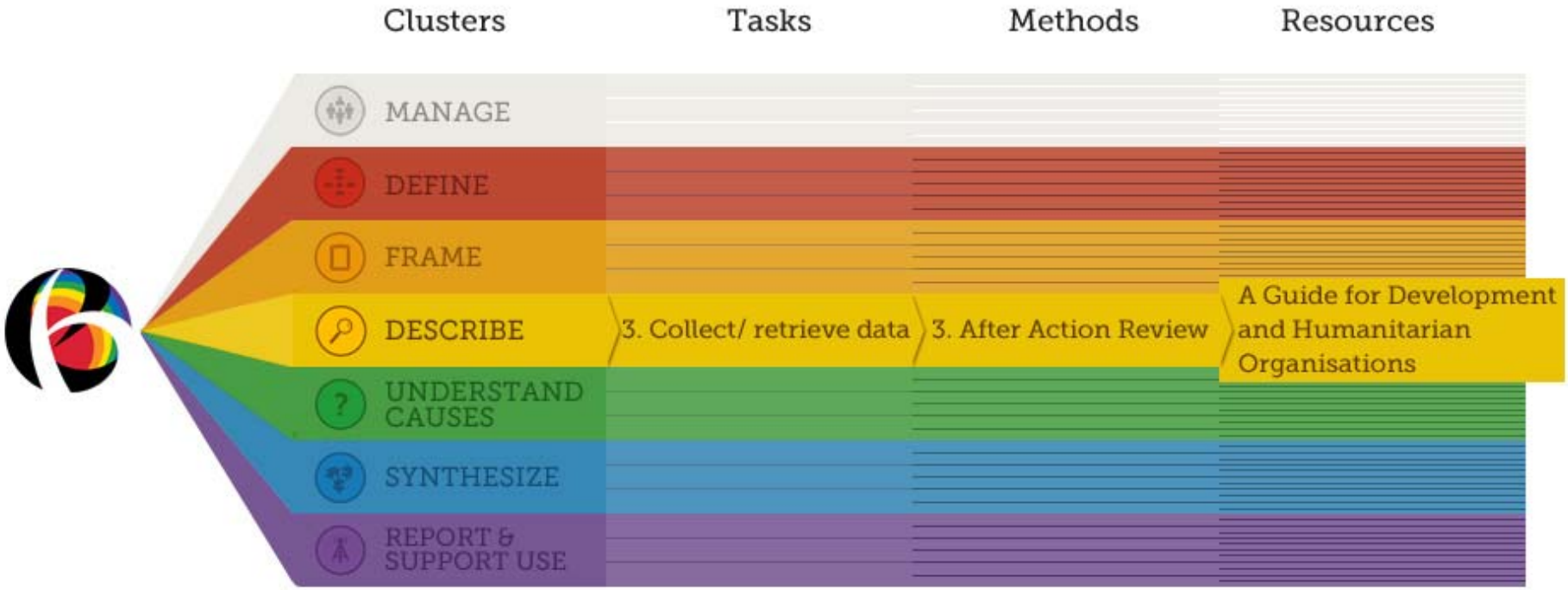
When does innovation arise?

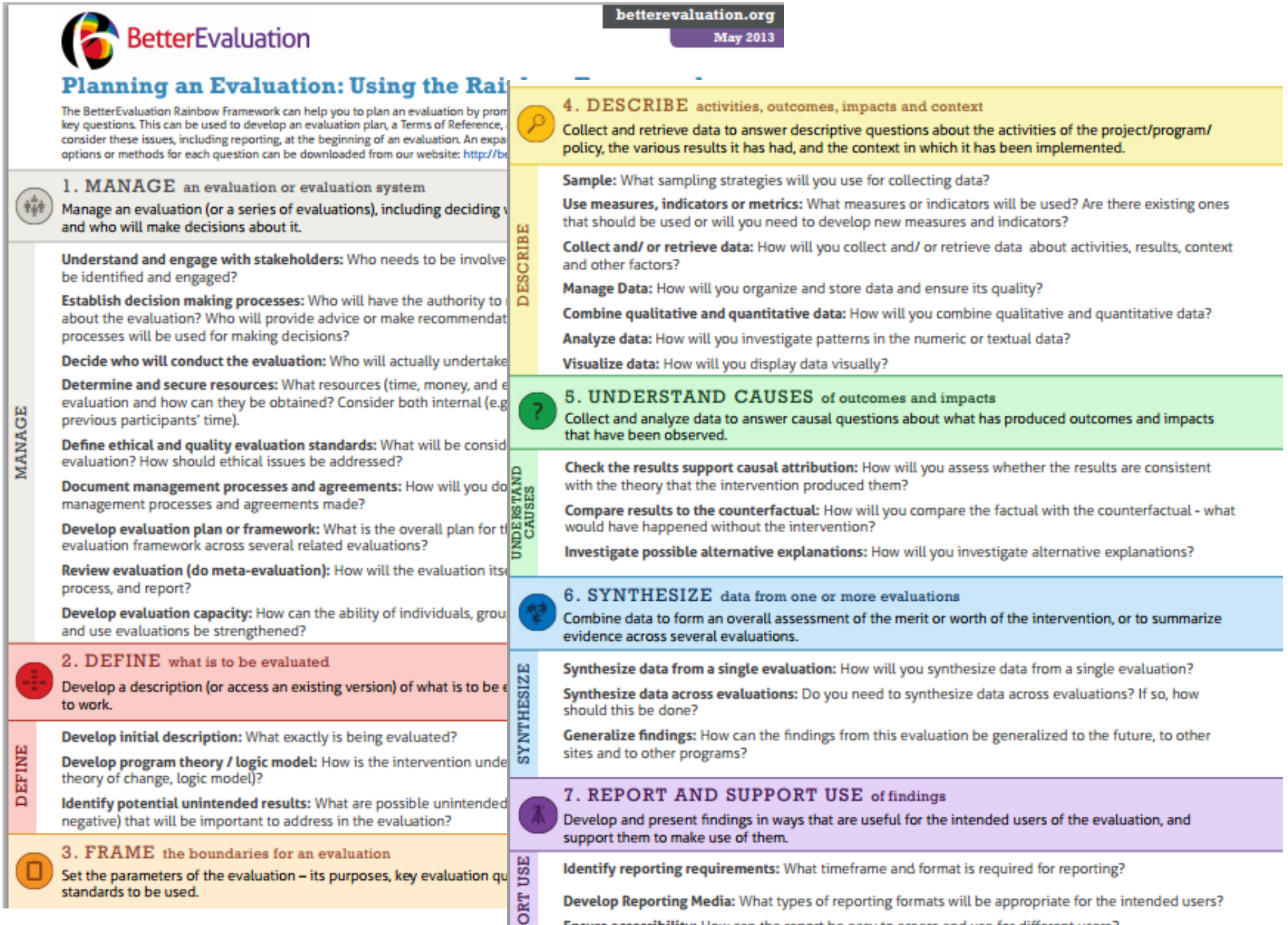


http://www.globoforce.com/gfblog/2013/happiness-flow-and-how-to-be-a-better-leader/

Where are some areas of evaluation practice where innovation might be needed?

Innovation across the range of evaluation tasks





BetterEvaluation

betterevaluation.org
May 2013

Planning an Evaluation: Using the Rai

The BetterEvaluation Rainbow Framework can help you to plan an evaluation by prom
key questions. This can be used to develop an evaluation plan, a Terms of Reference,
consider these issues, including reporting, at the beginning of an evaluation. An expa
options or methods for each question can be downloaded from our website: http://be

1. MANAGE an evaluation or evaluation system

Manage an evaluation (or a series of evaluations), including deciding
and who will make decisions about it.

MANAGE

Understand and engage with stakeholders: Who needs to be involve
be identified and engaged?

Establish decision making processes: Who will have the authority to
about the evaluation? Who will provide advice or make recommendat
processes will be used for making decisions?

Decide who will conduct the evaluation: Who will actually undertake

Determine and secure resources: What resources (time, money, and e
evaluation and how can they be obtained? Consider both internal (e.g
previous participants' time).

Define ethical and quality evaluation standards: What will be consid
evaluation? How should ethical issues be addressed?

Document management processes and agreements: How will you do
management processes and agreements made?

Develop evaluation plan or framework: What is the overall plan for t
evaluation framework across several related evaluations?

Review evaluation (do meta-evaluation): How will the evaluation itse
process, and report?

Develop evaluation capacity: How can the ability of individuals, grou
and use evaluations be strengthened?

2. DEFINE what is to be evaluated

Develop a description (or access an existing version) of what is to be e
to work.

DEFINE

Develop initial description: What exactly is being evaluated?

Develop program theory / logic model: How is the intervention unde
theory of change, logic model)?

Identify potential unintended results: What are possible unintended
negative) that will be important to address in the evaluation?

3. FRAME the boundaries for an evaluation

Set the parameters of the evaluation – its purposes, key evaluation qu
standards to be used.

4. DESCRIBE activities, outcomes, impacts and context

Collect and retrieve data to answer descriptive questions about the activities of the project/program/ policy, the various results it has had, and the context in which it has been implemented.

DESCRIBE

Sample: What sampling strategies will you use for collecting data?

Use measures, indicators or metrics: What measures or indicators will be used? Are there existing ones that should be used or will you need to develop new measures and indicators?

Collect and/ or retrieve data: How will you collect and/ or retrieve data about activities, results, context and other factors?

Manage Data: How will you organize and store data and ensure its quality?

Combine qualitative and quantitative data: How will you combine qualitative and quantitative data?

Analyze data: How will you investigate patterns in the numeric or textual data?

Visualize data: How will you display data visually?

5. UNDERSTAND CAUSES of outcomes and impacts

Collect and analyze data to answer causal questions about what has produced outcomes and impacts that have been observed.

UNDERSTAND CAUSES

Check the results support causal attribution: How will you assess whether the results are consistent with the theory that the intervention produced them?

Compare results to the counterfactual: How will you compare the factual with the counterfactual - what would have happened without the intervention?

Investigate possible alternative explanations: How will you investigate alternative explanations?

6. SYNTHESIZE data from one or more evaluations

Combine data to form an overall assessment of the merit or worth of the intervention, or to summarize evidence across several evaluations.

SYNTHESIZE

Synthesize data from a single evaluation: How will you synthesize data from a single evaluation?

Synthesize data across evaluations: Do you need to synthesize data across evaluations? If so, how should this be done?

Generalize findings: How can the findings from this evaluation be generalized to the future, to other sites and to other programs?

7. REPORT AND SUPPORT USE of findings

Develop and present findings in ways that are useful for the intended users of the evaluation, and support them to make use of them.

SUPPORT USE

Identify reporting requirements: What timeframe and format is required for reporting?

Develop Reporting Media: What types of reporting formats will be appropriate for the intended users?

Ensure accessibility: How can the report be easy to access and use for different users?

Develop recommendations: Will the evaluation include recommendations? How will these be

Some evaluation challenges where innovation might help

1. How can innovation help us to meet the challenges of the 4 competing imperatives of utility, validity, feasibility and propriety?
2. How can evaluations gather ‘good enough’ data in time to inform decisions?
3. How can evaluations ensure the values of intended beneficiaries are taken into account?
4. How can evaluations synthesise evidence and values systematically to draw evaluative conclusions?
5. How can evaluations make reasonable claims about the impacts of programs, given the influence of other factors?
6. How can evaluations draw causal conclusions when it is not possible to construct a credible counterfactual?
7. How can evaluation designs include an emergent element to accommodate emerging issues and understandings?

Some strategies for supporting appropriate choice of innovations

Rogers' theory of innovation - the importance of:

- Relative advantage – clear, unambiguous advantage in terms of effectiveness or cost-effectiveness
- Compatibility with existing systems
- Ease of use
- Trialability

What are other factors affecting the likely suitability of an innovation?

Some strategies for supporting appropriate use of innovations

1. Gather existing knowledge about the innovation, including possible technical support
2. Do a trial in one or more sites and visibly evaluate it
3. Share learnings
4. Provide support for further uptake

USAID Complexity-Aware Monitoring ‘Shepherding new methods

What are other strategies to support implementation of innovations?

For each of the following 4 innovations…

CHOOSING

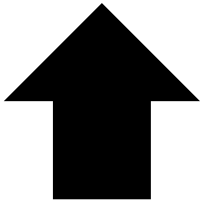
1. Do you think this might be useful for you to use? Why or why not?
2. What other information and assistance would be useful to have in order to make this choice?
3. How might you get this?

USING

1. What other information and assistance would be useful to have in order to use this innovation well?
2. How might you get this?

Innovation example 1: Negative program theory

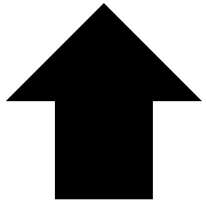
IMPROVED STUDENT LEARNING



HIGH STAKES TESTING OF STUDENTS

Innovation example 1: Negative program theory

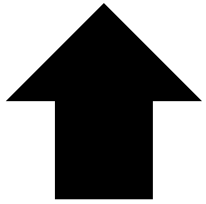
WORSE STUDENT MENTAL HEALTH



HIGH STAKES TESTING OF STUDENTS

Innovation example 1: Negative program theory

WORSE STUDENT LEARNING



HIGH STAKES TESTING OF STUDENTS

Innovation example 2: Big Data

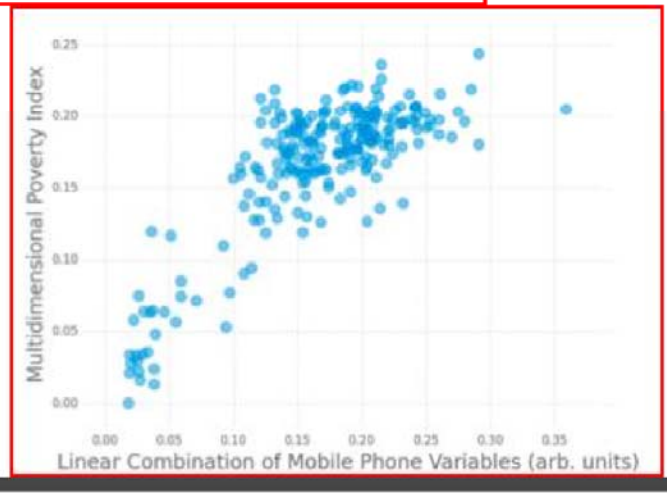
USING MOBILE PHONE DATA AND AIRTIME CREDIT PURCHASES TO ESTIMATE FOOD SECURITY

PARTNERS: UN WORLD FOOD PROGRAMME, UNIVERSITÉ CATHOLIQUE DE LOUVAIN, REAL IMPACT ANALYTICS

PROGRAMME AREA: FOOD SECURITY & AGRICULTURE

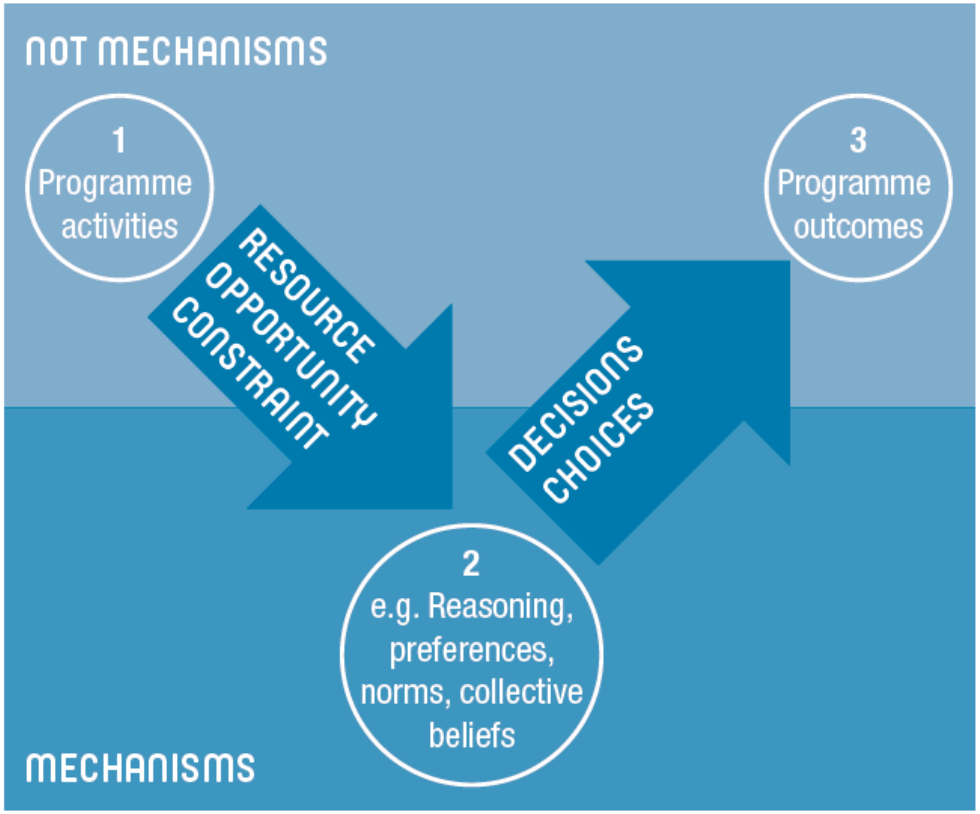
SUMMARY

This study assessed the potential use of mobile phone data as a proxy for food security and poverty indicators. It was conducted jointly with the UN World Food Programme (WFP), Université Catholique de Louvain in Belgium and Real Impact Analytics (a Belgian data analytics company). Data extracted from airtime credit purchases (or "top-ups") and mobile phone activity in an East African country was compared to a nationwide household survey conducted by WFP at the same time. Results showed high correlations between airtime credit purchases and survey results referring to consumption of several food items, such as vitamin-rich vegetables, meat or cereals. These findings demonstrated that airtime credit purchases could serve as a proxy indicator for food spending in market-dependent households. In addition, models based on anonymised mobile phone calling patterns and airtime credit purchases were shown to accurately estimate multidimensional poverty indicators. This preliminary research suggested that proxies derived from mobile phone data could provide valuable real-time information on the levels of several indicators related to food security, which could be integrated with early warning and monitoring systems, filling data gaps between survey intervals, and in situations where timely data is not possible or accessible.



Innovation example 3: Realist Evaluation

Generative explanation in realist programme evaluation



Wong, G., Westhorp, G., Pawson R, and Greenhalgh, T. (2012) Realist Synthesis RAMESES Training Materials. Reproduced with permission. Note that 'generative explanation' means 'explaining how causation works'.

Innovation example 4: EvalC3 – evaluating complex causal configurations

It enables users to:

a. identify sets of attributes that describe a project intervention & its context, and which are good predictors of the achievement of an outcome of interest,

b. compare and evaluate the performance of these predictive models, and

c. identify relevant cases for follow-up within-case investigations to uncover any causal mechanisms at work.

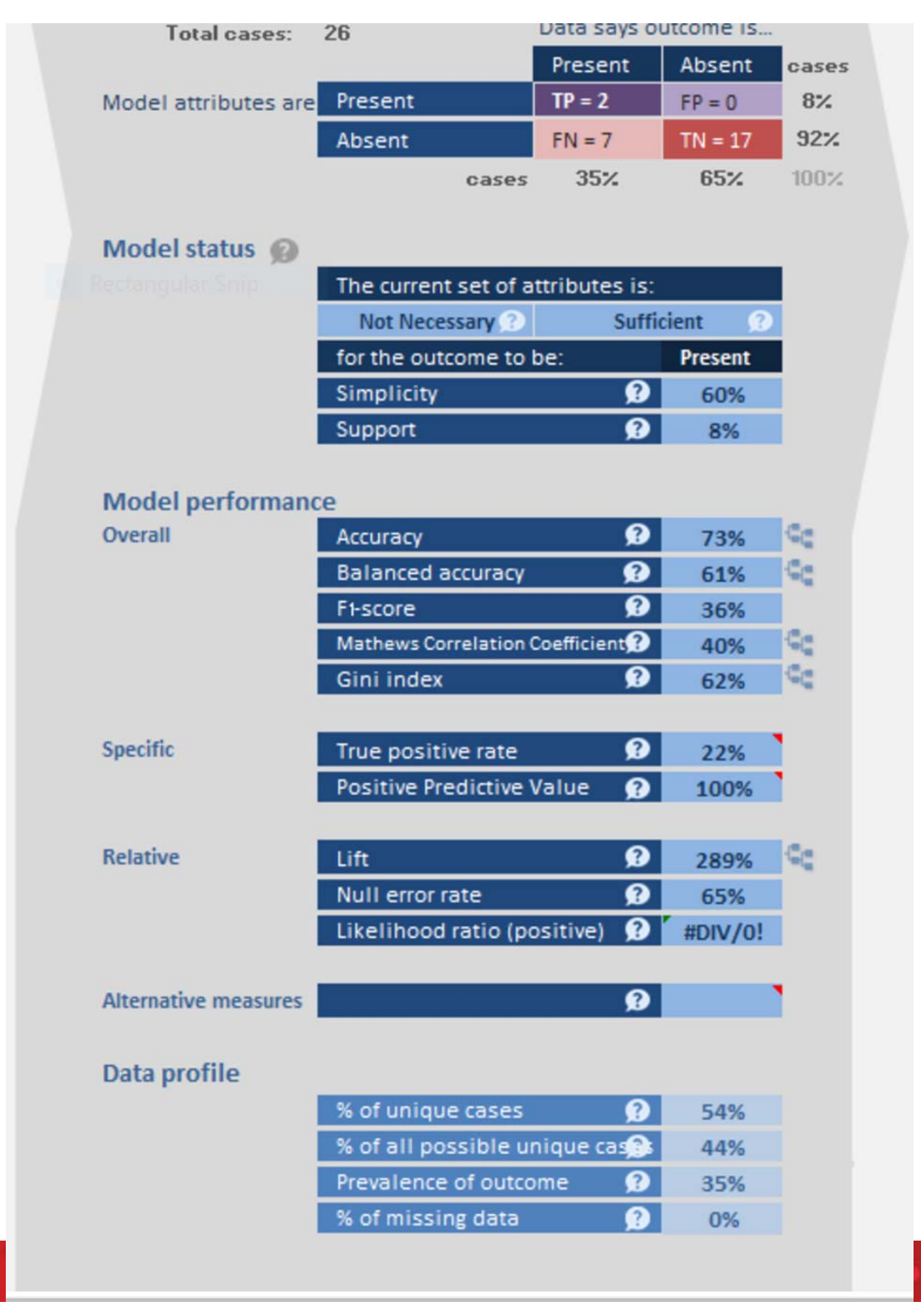
- Can be used during a project evaluation (also has other uses)
- **Causes of effects" analysis:** To identify what combination(s) of project activities (and their contexts) were associated with a significant improvement in beneficiaries lives.
- **"Effects of causes" analysis**: To identify what combinations of improvements in beneficiaries lives were associated with a specific project activity (or combination of)
- **To identify "positive deviants"** – cases where success is being achieved despite the fact that failure is the most common outcome
- https://evalc3.net/

Select Data

Get Started | Input Data | **Select Data** | Design & Evaluate | View Models | View Cases

Configurations: 14 Consistency: 100% Diversity: 44% Find optimal attributes

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



Total cases: 26

Data says outcome is...

Model attributes are	Present	Absent	cases
Present	TP = 2	FP = 0	8%
Absent	FN = 7	TN = 17	92%
cases	35%	65%	100%

Model status

The current set of attributes is:

Not Necessary	Sufficient

for the outcome to be:	Present
Simplicity	60%
Support	8%

Model performance

Overall

Accuracy	73%
Balanced accuracy	61%
F1-score	36%
Mathews Correlation Coefficient	40%
Gini index	62%

Specific

True positive rate	22%
Positive Predictive Value	100%

Relative

Lift	289%
Null error rate	65%
Likelihood ratio (positive)	#DIV/0!

Alternative measures

Data profile

% of unique cases	54%
% of all possible unique cases	44%
Prevalence of outcome	35%
% of missing data	0%

What other innovations might you consider?

8 Global Innovations in Measurement and Evaluation

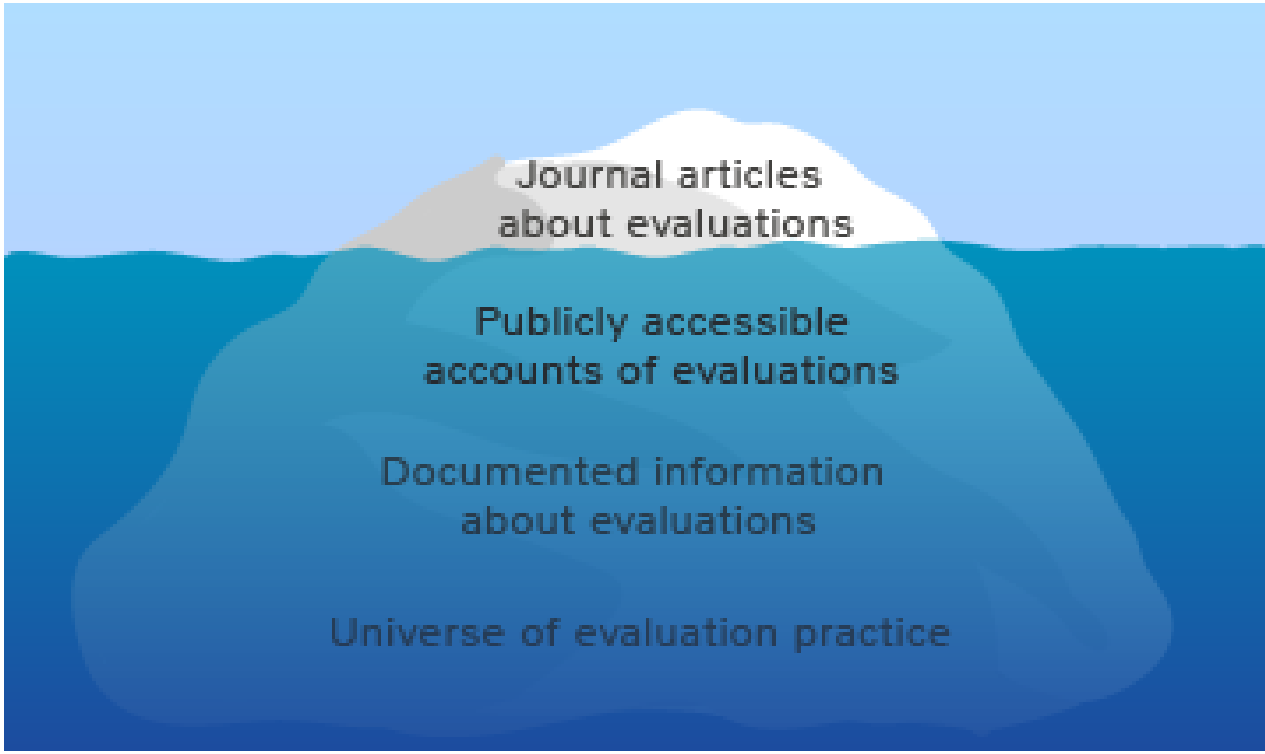
1. User-Centric Evaluation
2. Shared Measurement and Evaluation
3. Theory-Based Evaluation
4. Impact Management – so evaluation feeds back into programme delivery
5. Data Linkage
6. Big Data
7. Remote Sensing
8. Data visualisation

Some other innovations

1. Children as evaluators
2. Community scorecards
3. Positive deviance
4. Mobile device data collection
5. Community mapping
6. Citizen juries
7. Realist synthesis
8. Infographics
9. Rapid impact assessment
10. Co-design

Supporting innovation

- As individuals
- As organisations
- Through the AES and other organisations
- Through BetterEvaluation as a platform for co-creating and sharing knowledge about evaluation methods and processes



BetterEval: *World* | Select Language

BetterEvaluation

Sharing information to improve evaluation

Join us | Login | Contribute content

Start Here | Evaluation Options | Approaches | Themes | Resources | Blog | Search

BetterEvaluation

An international collaboration to improve evaluation practice and theory by sharing and generating information about options (methods or processes) and approaches.

Start here
to learn more about using BetterEvaluation

- Global Innovations in Measurement and Evaluation
- Illustrating models and theories of change
- BetterEvaluation is going to AES17 - Come say hello !
- Lessons from a trial of the Success Case Method

Find options

The Rainbow Framework organizes 300+ evaluation options into 7 clusters of tasks (shown to the right as coloured tabs).

MANAGE | DEFINE | FRAME | DESCRIBE | UNDERSTAND CAUSES | SYNTHESISE | REPORT & SUPPORT USE

MANAGE an evaluation or evaluation system

Manage an evaluation (or a series of evaluations), including deciding who will conduct the evaluation and who will make decisions about it. Read more.

1. Understand and engage stakeholders
2. Establish decision making processes
3. Decide who will conduct the evaluation
4. Determine and secure resources
5. Define ethical and quality evaluation standards
6. Document management processes and agreements
7. Develop planning documents for the evaluation
8. Review evaluation (do meta-evaluation)
9. Develop evaluation capacity

What's new on BetterEvaluation

Find new content and the latest events recently added to BetterEvaluation.

NEW EVENT:
Evaluating and managing for sustainable development impact

NEW RESOURCES:
Global Innovations in Measurement and Evaluation

NEW BLOG ENTRY:
2017 International Realist Conference

NEW BLOG ENTRY:
Illustrating models and theories of change

NEW BLOG ENTRY: