

---

# Evolution; emergence; bricolage

*Dr. Tom Aston, Independent*

---

# The problem(s)

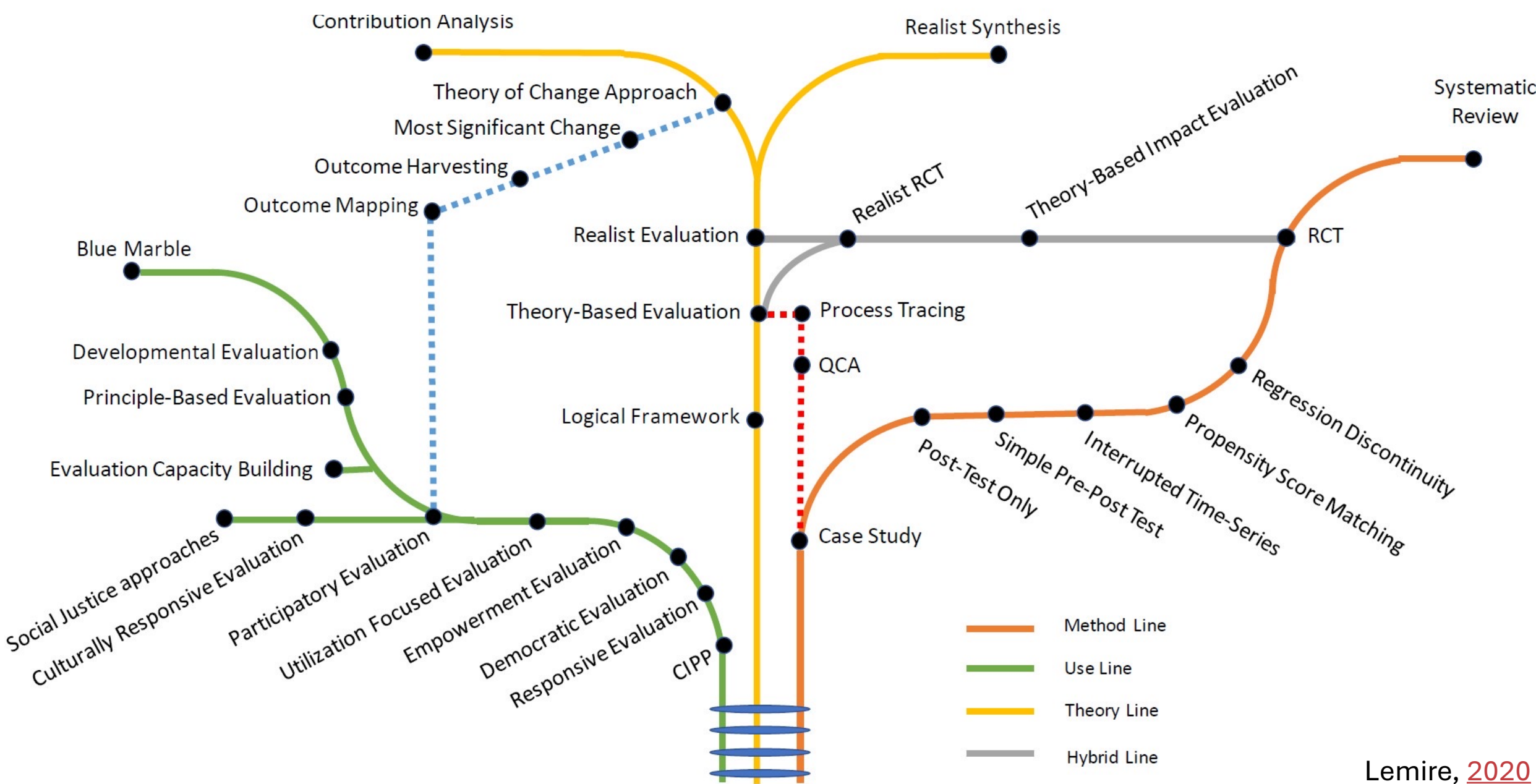
- Many evaluation practitioners are stuck with the **issues of complexity**
- There are a **growing multitude of methods**, and we don't always know which are best for our purposes
- How can we be both **pragmatic and rigorous**?

Table 1: Main characteristics of complex interventions

|                                     | Multicomponent interventions | Portfolio interventions | Long causal chain interventions | System-level interventions |
|-------------------------------------|------------------------------|-------------------------|---------------------------------|----------------------------|
| Number of components                | Red                          | Red                     | Green                           | Red                        |
| Number and difficulty of behaviours | Orange                       | Orange                  | Red                             | Orange                     |
| No. of groups and orgs. targeted    | Orange                       | Red                     | Green                           | Red                        |
| Number of sectors                   | Orange                       | Red                     | Green                           | Red                        |
| Number of stakeholders              | Orange                       | Red                     | Green                           | Red                        |
| Scale of intervention               | Green                        | Red                     | Green                           | Red                        |
| Level of standardisation            | Green                        | Red                     | Green                           | Red                        |
| Emergent outcomes                   | Orange                       | Orange                  | Green                           | Red                        |

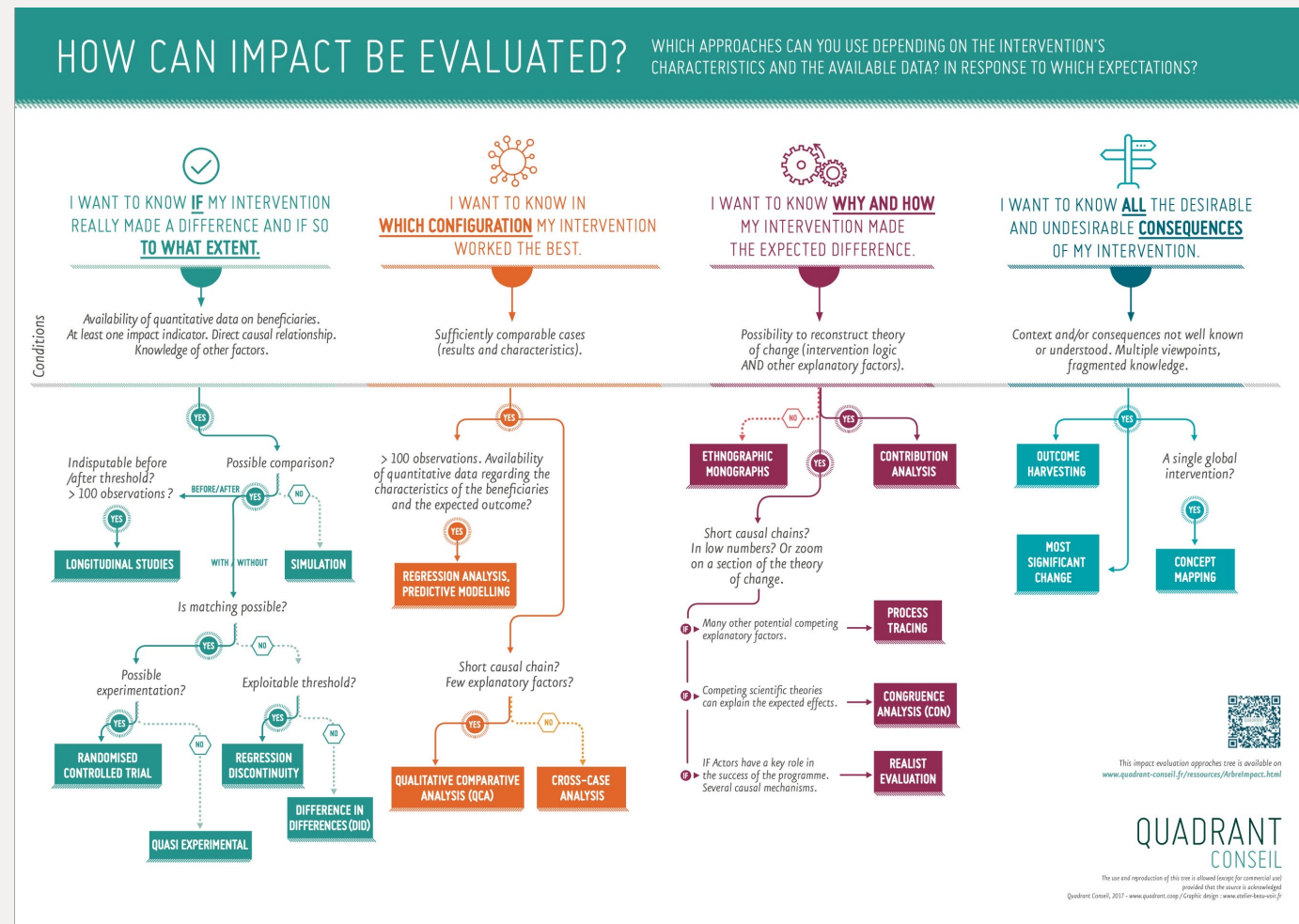
Note: Red cells represent a strong association between intervention type and complexity feature. Orange squares represent a moderate association. Green squares represent low association.

Masset et al. [2021](#)



# Evolution and emergence

- History of impact evaluation tells a story of evaluation more broadly (Delahais, [2022](#))
- A story of diversification, disruption, & hybridization



# Emerging lines

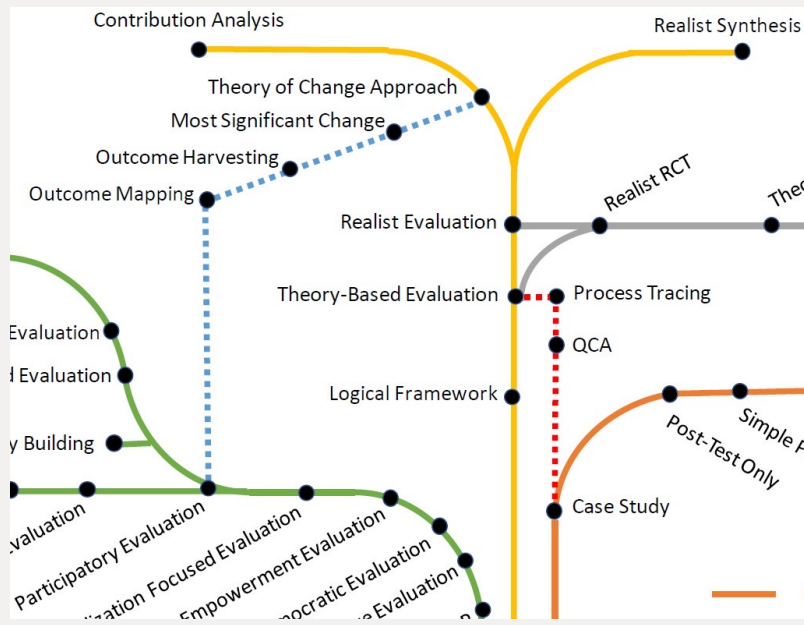
## Case Study

- Qualitative Comparative Analysis
- Process Tracing
- ... Theory-based evaluation

## Participatory Evaluation

- Outcome Harvesting
- Outcome Mapping
- Most Significant Change

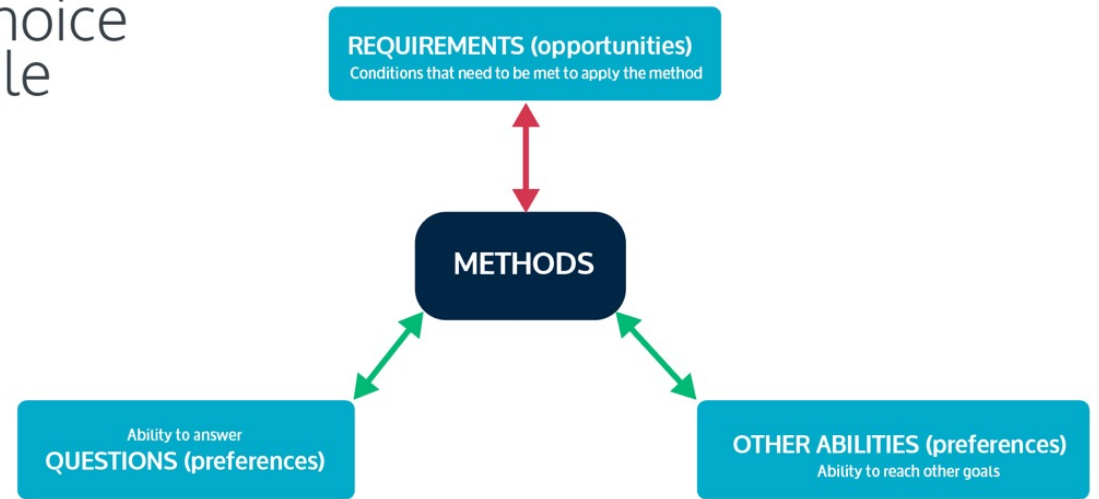
... Theory of change approach



# Pause to think

- **How** do you choose methods?
- What **criteria** do you use to choose those methods?

The Choice Triangle



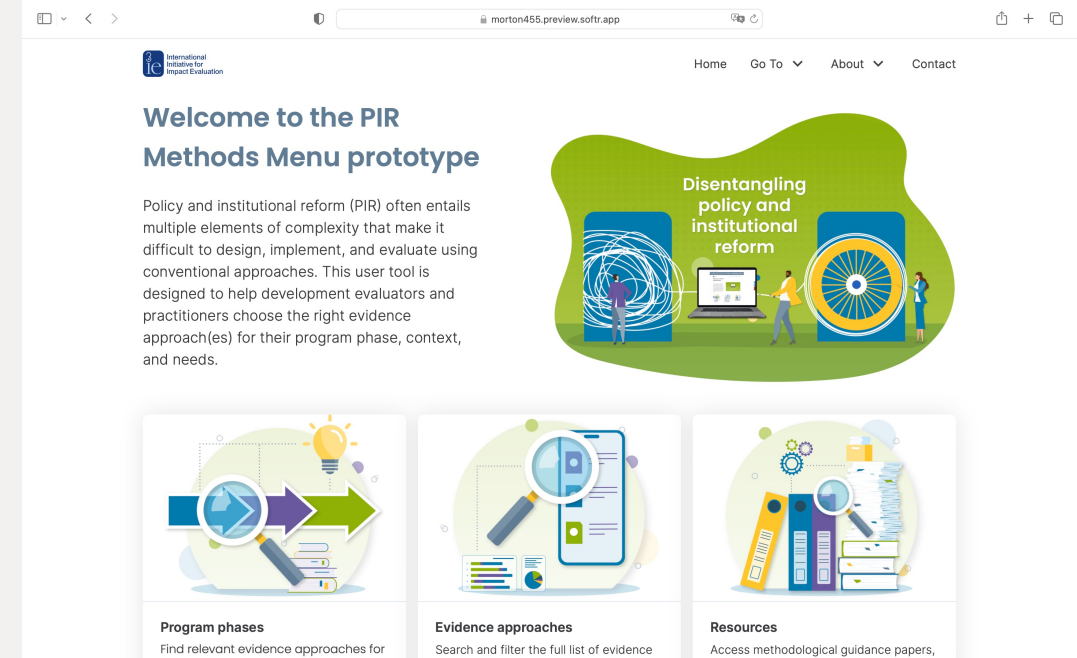
Befani, [2020](#)

# Choosing *between* methods

- **Either/or options** can be comforting, but misleading
- The common, but *insufficient*, answer proposed is “**mixed methods**”

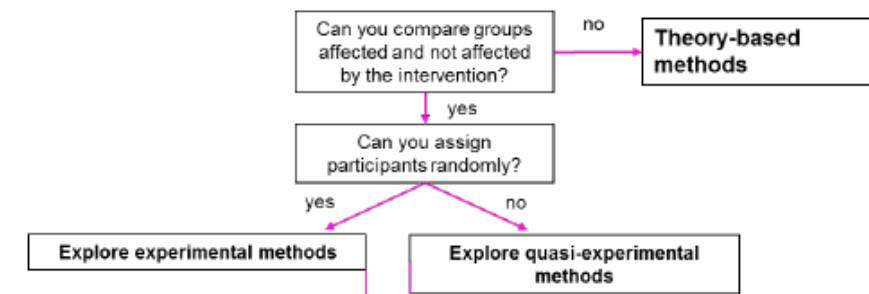
|   | A | B                              | C                        | D                    | E                            | F               | G                       | H                      | I                   | J                                   | K                              | L                           | M                  | N                                      | O                                 | P                     |
|---|---|--------------------------------|--------------------------|----------------------|------------------------------|-----------------|-------------------------|------------------------|---------------------|-------------------------------------|--------------------------------|-----------------------------|--------------------|--|-----------------------------------|-----------------------|
| <b>SUMMARY RESULTS - ALL STAGES</b>   |   | RCT (Randomised Control Trial) | Difference-in-Difference | Statistical Matching | Instrumental Variables (RDD) | Outcome Mapping | Most Significant Change | Soft Systems Modelling | Causal Loop Diagram | PSM (Participatory Systems Mapping) | BBN (Bayesian Belief Networks) | ABM (Agent Based Modelling) | Realist Evaluation | QCA (Qualitative Comparative Analysis) | Process Tracing/Bayesian Updating | Contribution Analysis |
| 1 Stage 1: Which Method is Best Suited to Answering My Key Evaluation Question(s)?  |   | 1                              | 80/100                   | 60/100               | 1                            | 20/100          | 40/100                  | 20/100                 | 20/100              | 20/100                              | 60/100                         | 80/100                      | 20/100             | 40/100                                 | 20/100                            | 20/100                |
| 2 Stage 2: Which method is most able to address my other interests?   |   | 81/100                         | 60/100                   | 69/100               | 74/100                       | 47/100          | 43/100                  | 57/100                 | 59/100              | 55/100                              | 59/100                         | 62/100                      | 62/100             | 50/100                                 | 59/100                            | 62/100                |
| 3 Stage 3: Which Method has the fewest essential methodological requirements that cannot be met by my intervention? (Which method is most feasible to use?) |   | 6                              | 6                        | 6                    | 4                            | 0               | 0                       | 0                      | 0                   | 0                                   | 0                              | 0                           | 0                  | 0                                      | 0                                 | 0                     |
| 4   |   |                                |                          |                      |                              |                 |                         |                        |                     |                                     |                                |                             |                    |  |                                   |                       |

Befani, [2020](#)



3ie, [2023](#)

Figure 3.1. Selecting experimental and quasi-experimental methods



HM Treasury, [2020](#): 47

# Increasing method hybridization

Experimental

Theory-based impact evaluation (White, [2009](#))

Theory-based systematic reviews (White, [2017](#))

Realist RCTs (Bonnell et al. [2012](#))

Configurational

QCA (Ragin, [1987](#))

Realist Evaluation & QCA (Befani et al. [2007](#))

Generative

Realist Evaluation (Pawson & Tilly, [1997](#))

Realist Synthesis (Pawson et al. [2004](#))

Contribution Analysis (Mayne, [2001](#))

Contribution Analysis & Process Tracing (Befani & Mayne, [2014](#))

Qualitative Impact Protocol (Copestake et al. [2019](#))

Outcome Evaluation Approach (Belcher et al. [2020](#))

Process Tracing (Van Evera, [1997](#))

Outcome Trajectory Evaluation (Douthwaite et al. [2023](#))

Participatory

Outcome Mapping (Carden et al. [2001](#))

Outcome Harvesting (Wilson Grau & Britt, [2012](#))

Outcome Evidencing, (Paz-Ybarnegaray & Douthwaite, [2016](#))

Contribution Rubrics (Aston, [2019](#))

Most Significant Change (Davies & Dart, [2005](#))

Collaborative Outcomes Reporting (Dart & Roberts, [2014](#))



# From comparison to integration

## *Qualitative Impact Protocol (QuIP)*

**Table 2.1** How the QuIP compares with other impact evaluation approaches: a summary

|   |  |
|---|--|
| Group 1. Approaches with some overlapping features with the QuIP                  | Appreciative enquiry; case studies; causal link monitoring; collaborative outcome reporting; critical systems heuristics; goal-free evaluation; outcome mapping; positive deviance; success case method; utilization focused evaluation.                             |
| Group 2. More quantitative approaches than the QuIP                               | Cost benefit analysis; difference-in-difference evaluation; qualitative comparative analysis; randomized control trials; social return on investment.  |
| Group 3. Broader approaches, with which the QuIP is congruent                     | Beneficiary assessment; contribution analysis; developmental evaluation; innovation history; institutional histories; outcome harvesting; process tracing; realist evaluation.   |
| Group 4. Approaches with stronger participatory and formative goals than the QuIP | Democratic evaluation; empowerment evaluation; horizontal evaluation; most significant change; participatory assessment of development; participatory impact assessment for learning and accountability; participatory evaluation and participatory rural appraisal. |

Copestake, [2019](#): 34

## *Outcome Trajectory Evaluation (OTE)*

**Table 4.** Similarities Between OTE and Other Approaches Used to Evaluate Policy Outcomes.

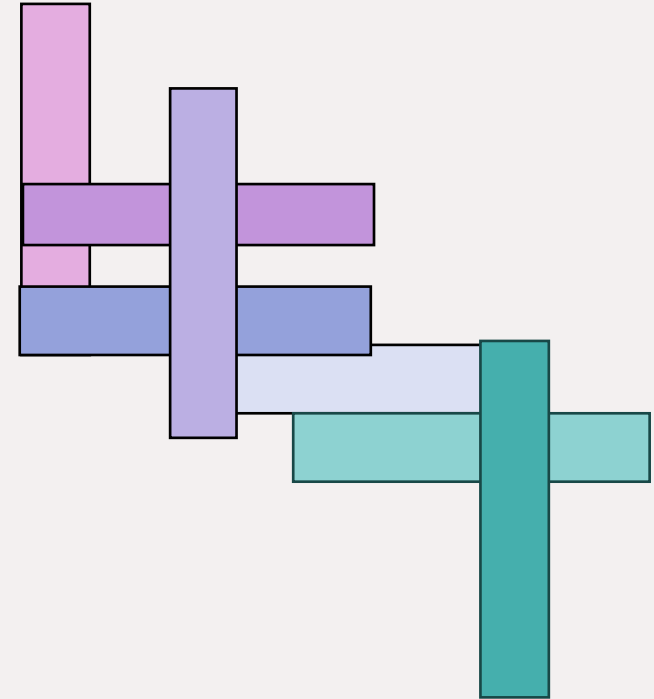
| Evaluation Approach                    | Similarities to OTE   |
|--|---|
| Process tracing (Collier, 2011)        | <ul style="list-style-type: none"> <li>- Focus on unfolding events or situations over time to make causal inferences;</li> <li>- The idea that causal inferences can be affirmed through building up a weight of evidence the robustness of which may be established through various tests (e.g., straw in the wind, smoking guns, etc.).</li> <li>- Use of criminal justice system analogies in explaining how the approach works</li> </ul> |
| Outcome harvesting (Wilson-Grau, 2018) | <ul style="list-style-type: none"> <li>- The practice of “back-casting” from an established outcome to understand what has contributed to it</li> <li>- Interviews with knowledgeable stakeholders to validate or repudiate causal claims</li> </ul>  |
| Contribution analysis (Mayne, 2012)    | <ul style="list-style-type: none"> <li>- Use of a contribution story, similar to the timeline used in OTE</li> <li>- The development and refinement of a ToC as a part of the analysis</li> </ul>   |
| Episode study (Carden, 2009)           | <ul style="list-style-type: none"> <li>- Back-casting from a well-defined policy change</li> <li>- Development of a historical narrative to explain the policy change along with important documents and events, and identifying key actors</li> </ul>  |

Douthwaite et al. [2023](#): 14

---

# Bricolage

- **Bricolage** is a way to combine (or triangulate) the best bits of methods
- Attempt to **reuse a heterogeneous repertoire** of available materials **to solve new problems** (Lévi-Strauss, [1968](#))
- **Not new**, but gaining prominence in recent years (Patton, [2011](#); Hargreaves, [2021](#); Aston *et al.* [2021](#); Aston & Apgar, [2022](#))



---

# Bricolage criteria

- **Reasoning:** Critical thinking, alternative explanations and interpretations, and search for outliers
- **Credibility:** Internal validity, contextually sensitive “probative value” of evidence
- **Responsiveness:** Reflect local stakeholders’ values and cultural context, sensitive to *their* experiences and definitions of success, and evaluation criteria
- **Utilization:** Actionable evidence, utilization of evaluation findings
- **Transferability:** External validity, focused on moderating factors and how the outcomes of an intervention are afforded by the context (Aston *et al.* [2021](#))

# Functional bricolage

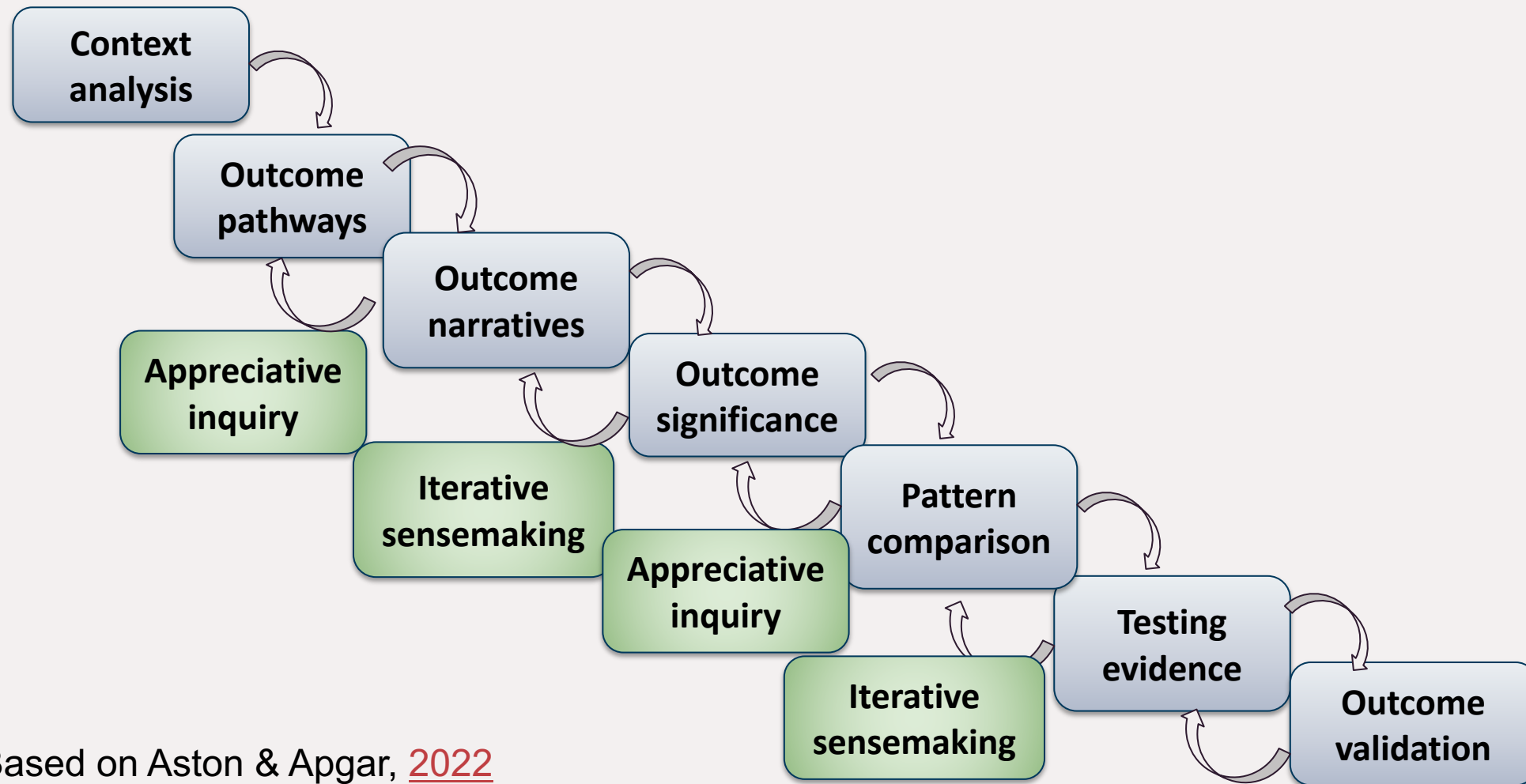
- “Threats to validity” – bolstering method weaknesses (Ton, [2012](#))
- Invert to focus on methodological strengths

Table 1 How functions support rigour criteria through methods

| Function  | Connection to rigour criteria  |
|---|--|
| <b>Context analysis</b> of potentially contributing factors and conditions  | Provides boundaries for causal mechanisms, enhances external validity of evaluative judgements, and supports transferability (e.g. context, mechanism, outcome configurations in Realist Evaluation).  |
| <b>Developing outcome pathways</b> for causal processes   | Helps to structure evaluative reasoning, enhances evaluability, and can bolster credibility of findings. If developed through a participatory process, this can contribute to responsiveness and utilisation (e.g. theory of change in Contribution Analysis, or developing causal chains in Process Tracing).                       |
| <b>Appreciative inquiry</b> of stakeholders to begin with an understanding of what matters to them, how they envision change, and what they want to learn about | Helps to ensure that evaluation is relevant and representative of stakeholders' experiences, and supports utilisation when orienting evaluation to learn from and about these experiences (e.g. collect significant change stories in Most Significant Change, or social inquiry in Collaborative Outcomes Reporting).               |
| <b>Articulating outcome narratives</b> to explain outcomes and contributions  | Enhances evaluative reasoning and the testability of effect patterns, which raises credibility, and can increase responsiveness if these are developed in a participatory way (e.g. draft outcome statements in Outcome Harvesting or progress markers in Outcome Mapping).  |
| <b>Appraising significance of outcomes</b> to explain why a change is important   | Enhances evaluative reasoning and improves responsiveness if appraisal is a participatory process (e.g. collecting significant change stories in Most Significant Change or collecting significance data for outcome statements in Outcome Harvesting).  |
| <b>Iterative sensemaking</b> of hypotheses and contribution claims  | Helps to strengthen reasoning and enhance credibility and can also be conducted in a participatory way to enhance responsiveness (e.g. assessing a contribution story and seeking additional evidence in Contribution Analysis or revising the conceptual model in Multiple Lines and Levels of Evidence).                           |
| <b>Testing strength of evidence</b> underpinning contribution claims  | Appraising evidence strength stress-tests evaluative reasoning. It strengthens internal validity and thus increases the credibility of findings and evaluative judgements (e.g. evidence tests and rival hypothesis assessment in Process Tracing; ruling out possible alternative explanations in General Elimination Methodology). |
| <b>Validation of outcomes</b> with communities, peers and/or experts  | Strengthens the credibility of evaluative judgements (internal validity). If conducted in a participatory way, it can contribute to responsiveness (e.g. outcome panel in Collaborative Outcomes Reporting; contribution trial in Process Tracing with Bayesian Updating).   |
| <b>Causal pattern comparison</b> between outcomes   | Can help identify trends and outliers, thereby contributing to evaluative reasoning and transferability of findings (e.g. data matrix and truth tables in Qualitative Comparative Analysis; SenseMaker's visualisation of patterns).   |
| <b>Supporting utilisation</b> of evaluation findings to inform future programming   | Enables use of evaluation findings and supports transferability through adapting programming and informing new programming (e.g. supports use in Outcome Harvesting).  |



# Finding the right combination: function, not form



Based on Aston & Apgar, [2022](#)

---

# Further reading

- Downes and Gullikson, [2020](#) and Schwandt and Gates, [2021](#) and Aston, [2022](#) on **valuing criteria**
- Stern *et al.* [2012](#); HM Treasury, [2020](#); Befani, [2020](#); and 3ie, [2023](#) on **selecting evaluation methods**
- Patton, [2011](#); Hargreaves, [2021](#) and Aston & Apgar, [2022](#) on **bricolage** with case-based methods
- Aston *et al.* [2021](#) and Raimondo, [2023](#) on **rigour in case-based methods**