

Causal Pathways Evaluation: Values-Driven Rigor and Quality

Causal Pathways Initiative Training

Today's trainers: Florencia Guerzovich and Drew Koleros

Content developed as a collaborative effort by initiative network members and staff including: Jewlya Lynn, Carlisle Levine, Marina Apgar, and Carolina De La Rosa Mateo with support and content from Tom Aston, Julia Coffman, Heather Britt, Yulianto Dewata, Abdoul Karim Coulibaly, Steve Powell, and Fiona Remnant.

The Causal Pathways Initiative

Making visible the "black box" of philanthropic and systems change strategies, helping us collectively see how systems are (or are not) changing

An international network of evaluators, methodologists, philanthropic leaders, and more.

Focused on supporting philanthropy, other funders and their evaluation partners by **building awareness, will, and skills to use evaluation approaches that can make sense of causal relationships** without depending on more traditional experimental and quasi-experimental approaches.

Our Trainers



Drew Kleros
Mathematica
Cambridge, MA
USA



Florencia Guerzovich
Independent Consultant
Florianopolis, Brazil

Who is in the room?

Poll

1. Your organization type
2. Your role
3. Your level of experience with causal pathways evaluations

Causal pathways evaluations make visible the "black box" of philanthropic and systems change strategies, helping us collectively see how systems are (or are not) changing.

Agenda

- Welcome
- Rethinking rigor
- Evaluator practices
- Values
- Inclusive and participatory approaches
- Using rubrics
- Conclusion

Learning Objectives

- Understand how to strengthen quality and rigor in a causal pathways evaluation through:
 - Reflecting on evaluator practices
 - Integrating values in assessments of quality and rigor
 - Employing inclusive and participatory approaches
 - Applying rubrics

Rethinking rigor



Carlo Spoldi

Why traditional notions of rigor fall short for causal pathways evaluations

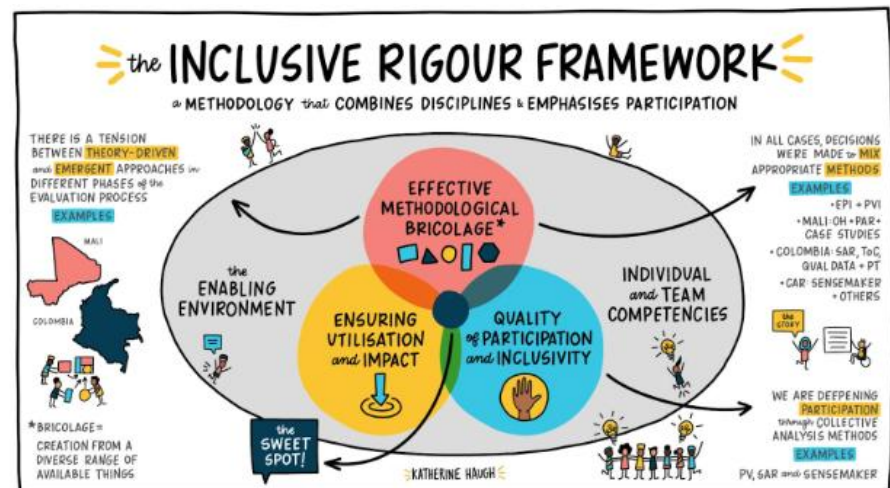
- More conventional evaluation approaches equate rigor as a methodological concern.
 - These methods are often underpinned by assumptions about control, stability, fidelity, replicability, etc.
- These fall short when the context of the evaluation is high in complexity
 - Too rigid, linear, and context-blind
- Rigor in systems change efforts need to reflect a broader set of dimensions, such as:
 - Complexity
 - Inclusion
 - Emergent and plural causal pathways

Causal Pathways Rigor Guidance

Causal Pathways Initiative guidance: How do we define and support quality and rigor in Causal Pathways evaluation?

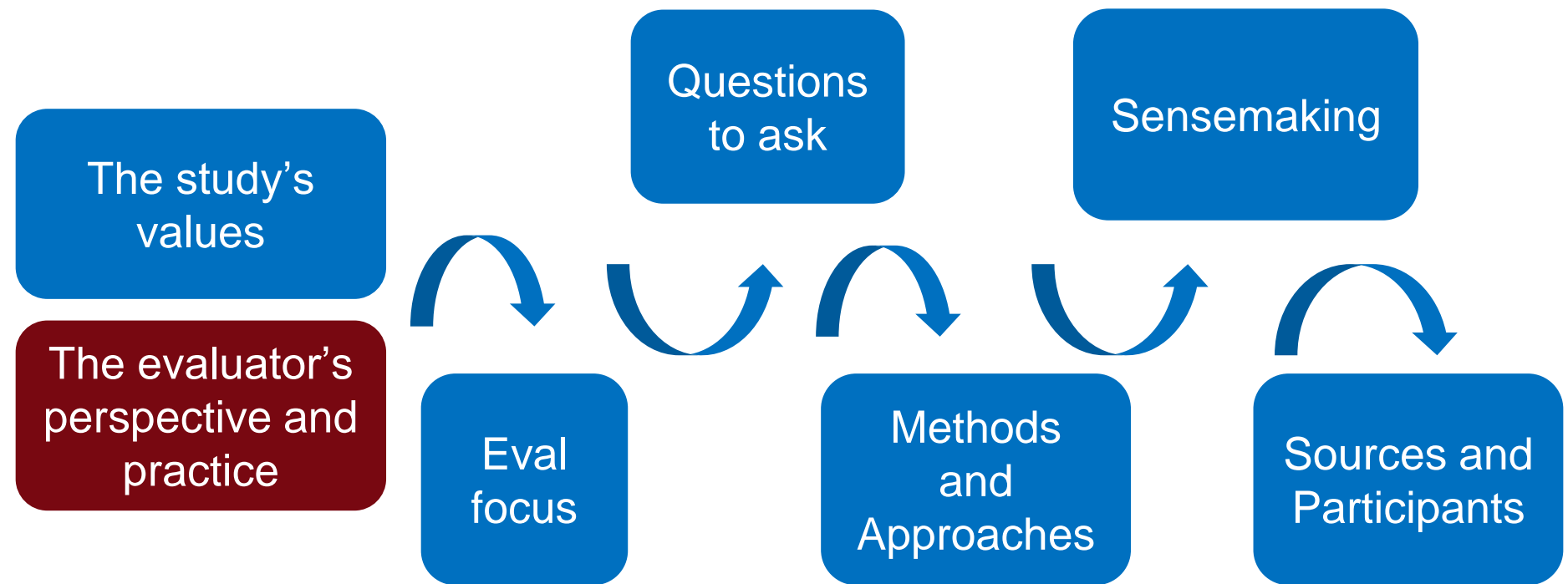
(2025, Marina Apgar and Tom Aston)

Grounded in
the concepts
from....



<https://tinyurl.com/inclusiverigorcolab>

Our values inform all stages of our evaluations and our practices as evaluators



This diagram is adapted from Apgar, M., Aston, T., Snijder, M., & Zwollo, T. (2024). Raising the Bar: Improving How to Assess Evidence Quality in Evaluating Systems-Change Efforts. *The Foundation Review*, 16(2). <https://doi.org/10.9707/1944-5660.1712>

Evaluator Practices to Strengthen Rigor



Alex Williamson

Two key practices we bring as causal pathways evaluators

Reflexivity:

Recognizing and
interrogating
evaluator
positionality and
bias

Credibility:

Building
trustworthiness
through methods
transparency
and diverse
evidence

Reflexivity as a practice

Reflexivity

Credibility

- Map your own position and potential biases
- Reflect collectively with participants and co-evaluators
- Build reflexivity into analysis—not just process

Credibility as a practice

Reflexivity

Credibility

1. Articulating explicit causal pathways
2. Paying attention to a range of outcomes and impacts
3. Understanding contextual variation
4. Using an iterative, bricolage approach
5. Drawing on a range of causal inference strategies
6. Dialogue with stakeholders

Grounding in Values to Strengthen Rigor



An example of values from the Causal Pathways Initiative rigor framework

Equity

- Inclusion, representation, and power-sharing
- Attending to how interventions affect different populations in different ways

Learning & Utilization

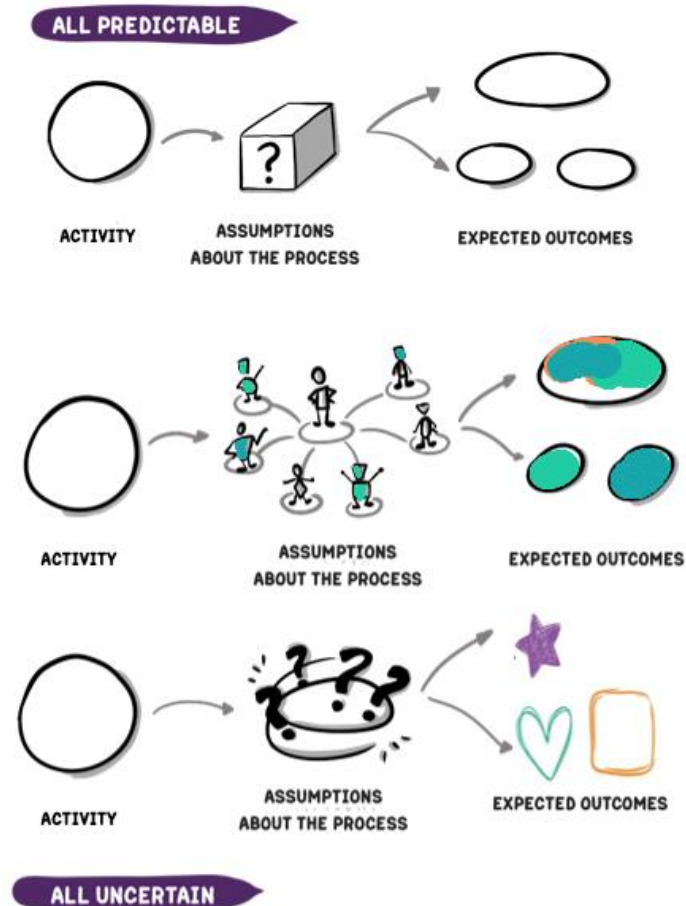
- Iterative - leveraging learnings along the way
- Producing evidence that can inform strategy

Complexity

- Context aware and non-linear
- Attending to both intended and unintended effects

Why complexity as a value in causal pathways evaluations?

- Multiple causes, actors, and factors interact to shape outcomes
- Change is non-linear, dynamic, and often emergent
- Context deeply influences how and why change occurs
- Feedback loops, time lags, and adaptation shape results
- There may be no single “right” pathway—multiple truths can coexist



Source: Guerzovich and Aston (2024) adapted from [Guerzovich and Wadeson, 2024](#).

Operationalizing the value of complexity

- Use multiple sources, perspectives methods, and data types
- Be transparent about limitations and judgement calls
- Ensure findings make sense to those closest to the work (and ideally have them generate the findings and sensemake them with you)
- Ensure the findings honor and make visible the complexity instead of summarizing and explaining it away

Mini case study: An example of values from a causal pathways evaluation of virtual and hybrid school models



Background on causal pathways evaluation

- National narrative in the U.S. post-COVID that virtual and hybrid learning “does not work” to accelerate learning outcomes
- Causal pathways evaluation of four virtual and/or hybrid school models in the country that have previously demonstrated evidence of achieving learning outcomes
- Evaluation focuses on understanding how and why these models work differently for different types of students (and under what conditions)
- Intended to generate rigorous evidence to inform improvements to current school leaders from the four sites and school leaders implementing similar approaches

Surfacing values on the evaluation

- REFLECT
 - Think about a specific experience where high quality evidence was used to understand how a program works. What made it a positive experience? What values mattered?
- INDIVIDUAL EXERCISE
 - Each evaluation team member developed their own narrative and surfaced their individual values
- COLLECTIVE PRIORITIZATION
 - Group facilitated session to cluster similar values and prioritize “top values”

Example of values identified

Collaboration with program users

- Collaboration
 - Participation: engaging users in the process
-

Critical reflection across multiple sources of evidence

- Critical reflection – Findings can be verified, triangulated in a variety of ways
 - Multiple data sources and attention to how those data relate, support, refute each other
-

Elevating lived experience and multiple voices

- Connection to program's understanding and experience
 - Ensuring multiple voices and perspectives are represented
 - Elevate participant experience / lived experience
 - Multiple voices are included
-

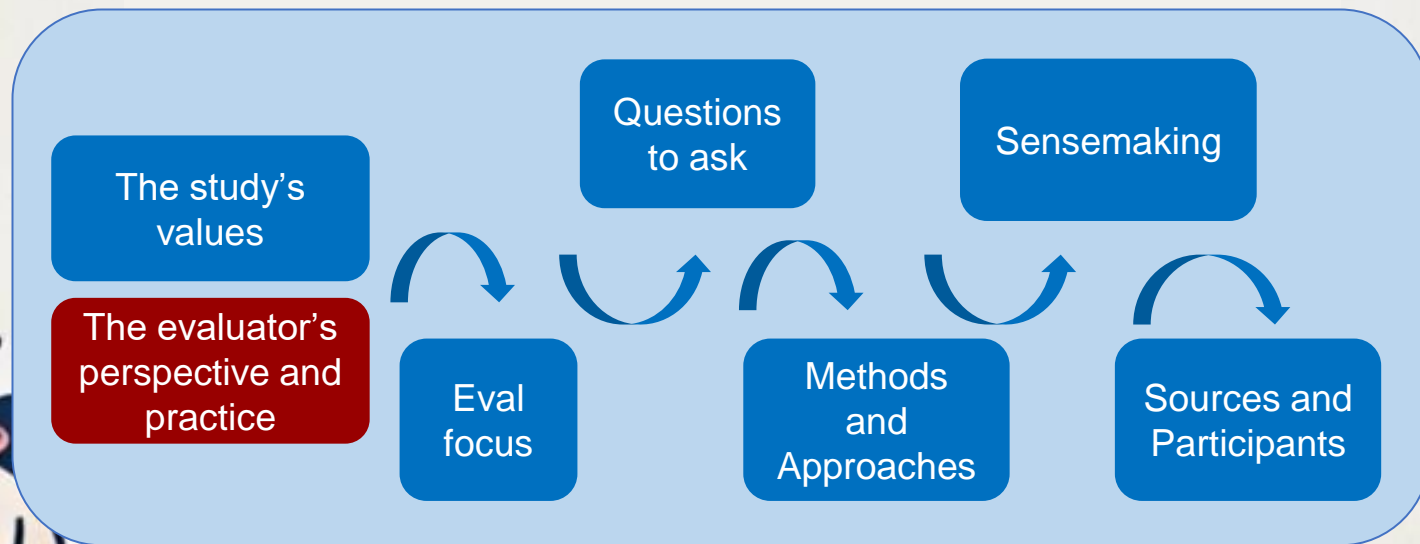
Usefulness for school model improvement

- Generating actionable findings that can support adaptation/improvement
- Usefulness – Is the evidence usable by / useful for the program?
- Actionable – Produce something that can be used by sites / by others

SMALL GROUP DISCUSSION

Think about a specific evaluation you've been part of:

- If you centered “honoring complexity” as a value in the work, how might it have changed your approach in different stages of the evaluation?

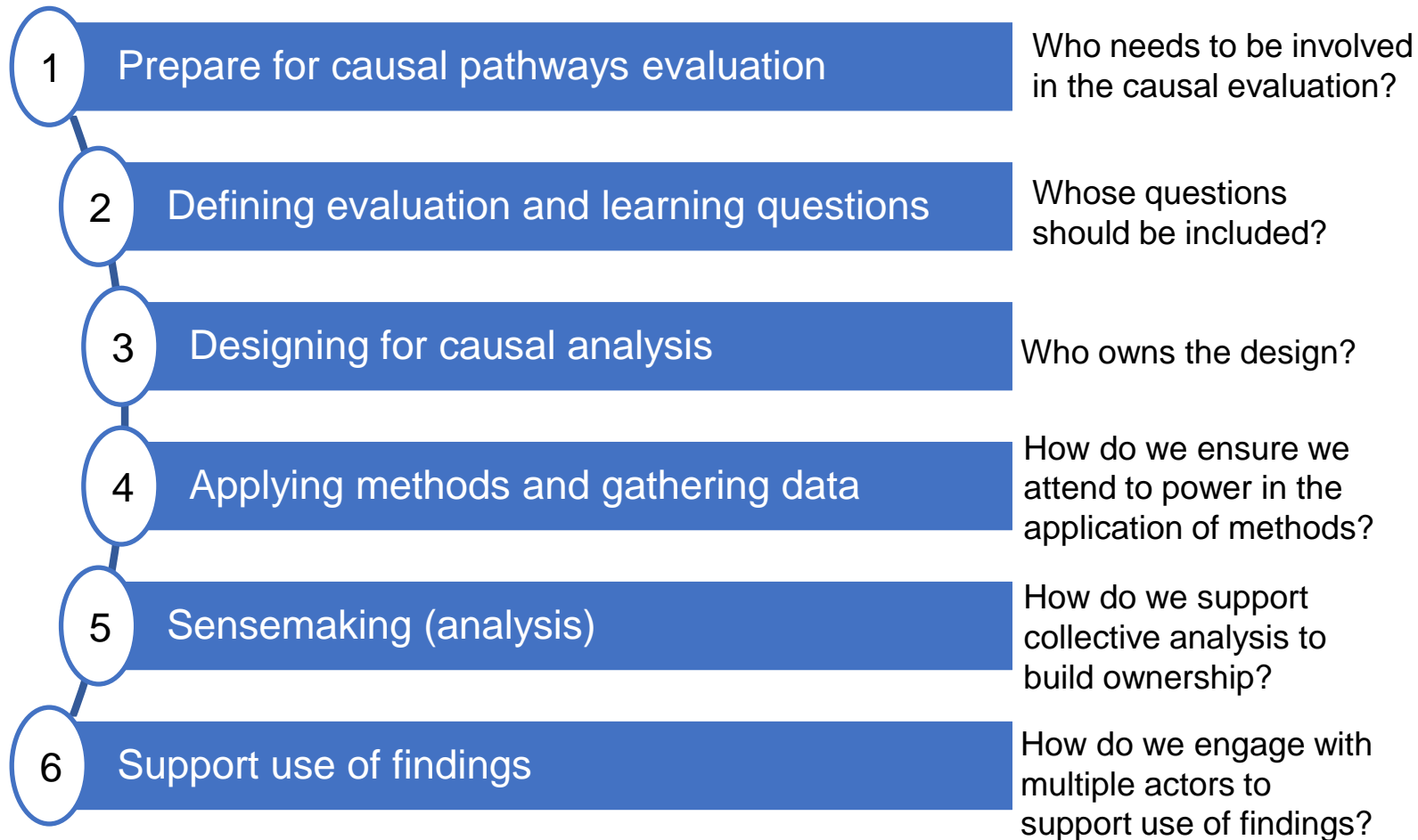


Inclusivity and Participation to Strengthen Rigor

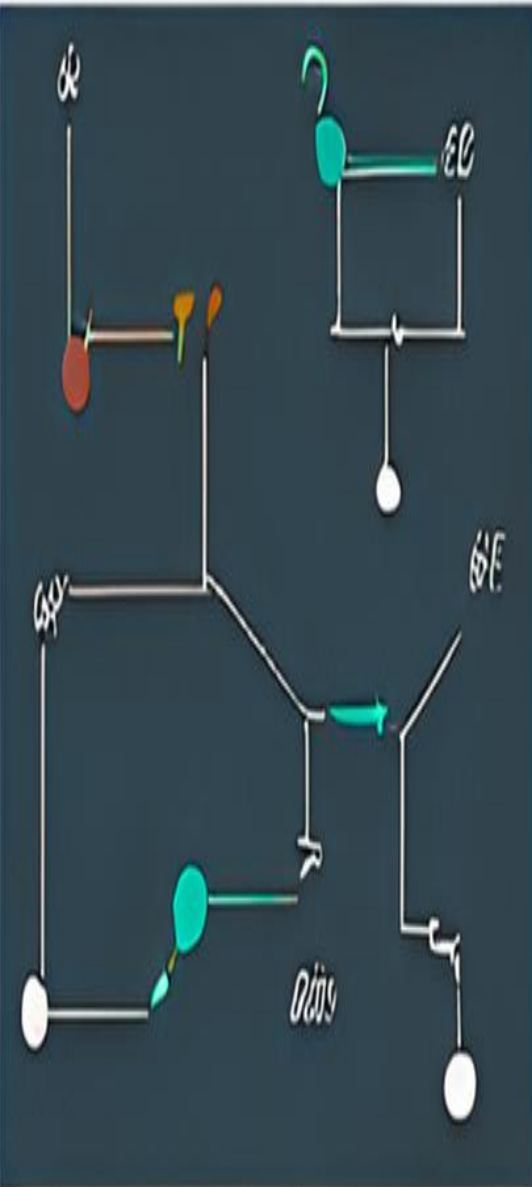


Jemal Demeke, 2025
[Art and Equity in the Age of AI](#)

Participation is possible and helpful in all aspects of a causal pathways evaluation



A critical place for inclusion and participation: Sensemaking the causal pathways



Start with ***causal questions***, not findings: “What caused what, for whom, and in what context?”

Visualize possible causal chains (e.g., flow diagrams, causal maps, story timelines)

Facilitate testing of causal claims: What else might explain this? What had to be present?

Use “causal probes”: “What made this outcome more likely?” “What was necessary but not sufficient?”

Incorporate rival explanations and counterfactual thinking

Engage multiple perspectives to strengthen or challenge hypotheses

Using Rubrics to Strengthen Rigor

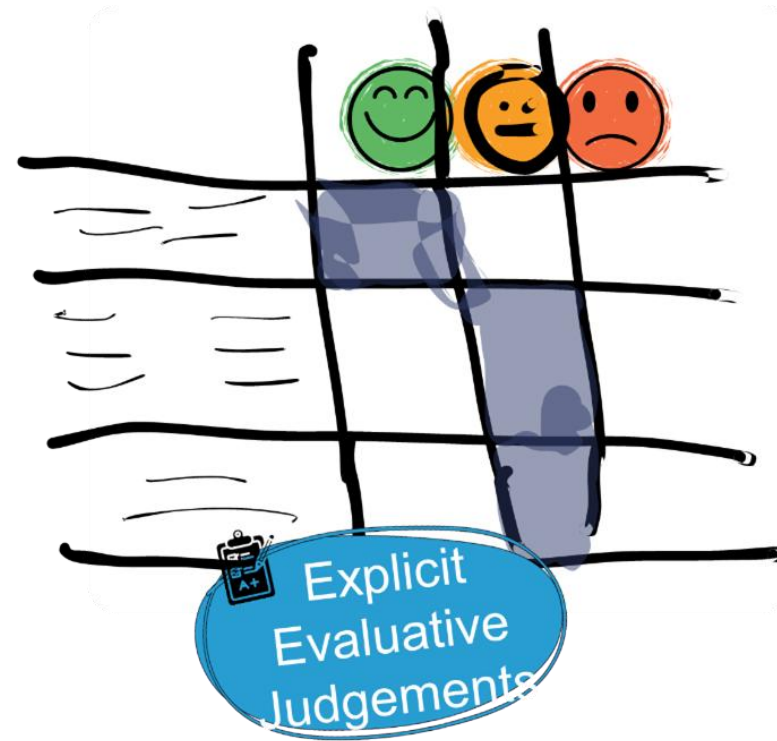


Alberto Boncompagni, "Inheritance of the Twentieth Century"

Rubrics: A tool to operationalize values and strengthen rigor

A structured framework for assessing change that makes clear:

- The aspects of performance of focus and interest
- Sequential levels of performance related to the outcomes
- What performance looks like for each level (criteria)



Rubrics: Evaluating with purpose



1. Rubrics can help us judge overall methodological approach.
2. Rubrics can help us make judgements tied to specific study evidence and content.
3. Rubrics aid causal claim strength decisions.
4. Rubrics help us operationalize our values.

Sustainability Relational Rubric

SCORE 01	<p>No evidence of any use/application/adaptation of element(s) of or insights from a collaborative social accountability process by any priority stakeholders and/or public sector institutions. No evidence of stakeholder interest, dialogue of alignment.</p> <p>The unit of measurement for this indicator in the GPSA's Results Framework is 0%. Therefore, a score of 0% would be provided for the indicator in the Results Framework and considered as 'no uptake'.</p>	0% UPTAKE
SCORE 02	<p>Evidence of interest by priority stakeholders and/or public sector institutions expressed publicly or privately about learning from a collaborative social accountability process in the project.</p> <p>In this instance, a score of 25% would be provided for the indicator in the GPSA's Results Framework.</p>	25% UPTAKE
SCORE 03	<p>Evidence that priority stakeholders and/or public sector institutions have expressed where to adopt, adapt and/or sustain elements or insights from a collaborative social accountability process and how this could be incorporated in some way into other operations, programs, policies (i.e., concrete entry points have been identified).</p> <p>In this instance, a score of 50% would be provided for the indicator in the GPSA's Results Framework.</p>	50% UPTAKE
SCORE 04	<p>Evidence of dialogue² with priority stakeholders and/or public sector institutions on how to adopt, adapt and/or sustain elements of the collaborative social accountability process in future operations, policies, or programs.</p> <p>In this instance, a score of 75% would be provided for the indicator in the GPSA's Results Framework.</p>	75% UPTAKE
SCORE 05	<p>Evidence of actions taken by priority stakeholders and/or public sector institutions on adoption, adaption and/or sustaining elements of a collaborative social accountability process in other operations, policies, or programs. Triangulation of data with at least 2 sources of evidence to confirm is required.</p> <p>In this instance, a score of 100% would be provided for the indicator in the GPSA's Results Framework.</p>	100% UPTAKE

No evidence

Evidence

Triangulation of at least 2 sources

Rubrics: A tool to operationalize values and strengthen quality amid complexity

1. Rubrics are completely new to me
2. I've used rubrics a bit in my work but not within a causal pathways evaluation
3. I've used rubrics in some of the ways listed here

Quick Poll: Which of these three uses of rubrics have you tried in your own work (or evaluations you've overseen/commissioned)?

Examples of dimensions of rubrics that help us explore the strength of the causal claims

- Transparency – Is it clear how data were collected and analyzed?
- Triangulation – Are diverse sources and methods used?
- Plausibility – Are causal claims supported and coherent?
- Uniqueness – Are rival explanations considered?
- Representativeness – Whose perspectives are included?

Examples of a rubric related to the strength of a causal claim

Table 1.1 Plausibility rubric (by performance level)

1	2	3	4	5
Unclear, illogical, or contradictory explanation connecting intervention to outcome.	Explanation indicates a possible connection between intervention and outcome.	Explanation is clear, logical and temporally consistent, and suggests a likely association between intervention and outcome.	Convincing explanation of how evidence connects intervention and outcome. Conclusions drawn tend to follow the data.	Highly convincing account, clearly and logically signposting key steps and specific data connecting intervention to outcome. Conclusions drawn unambiguously follow the data.

Mini case study: An example of values from a causal pathways evaluation of virtual and hybrid school models



Mini case study: An example of values from a causal pathways evaluation of virtual and hybrid school models

- Transparency
- Triangulation
- Plausibility
- Uniqueness
- Representativeness

Mini case study: An example of values from a causal pathways evaluation of virtual and hybrid school models

- Transparency
- Triangulation
- Plausibility
- Uniqueness
- Representativeness
- Transferability
- Ethics
- Responsiveness
- Independence
- Utilization

Mini case study: An example of values from a causal pathways evaluation of virtual and hybrid school models



Collaboration with program users



Critical reflection across multiple sources of evidence

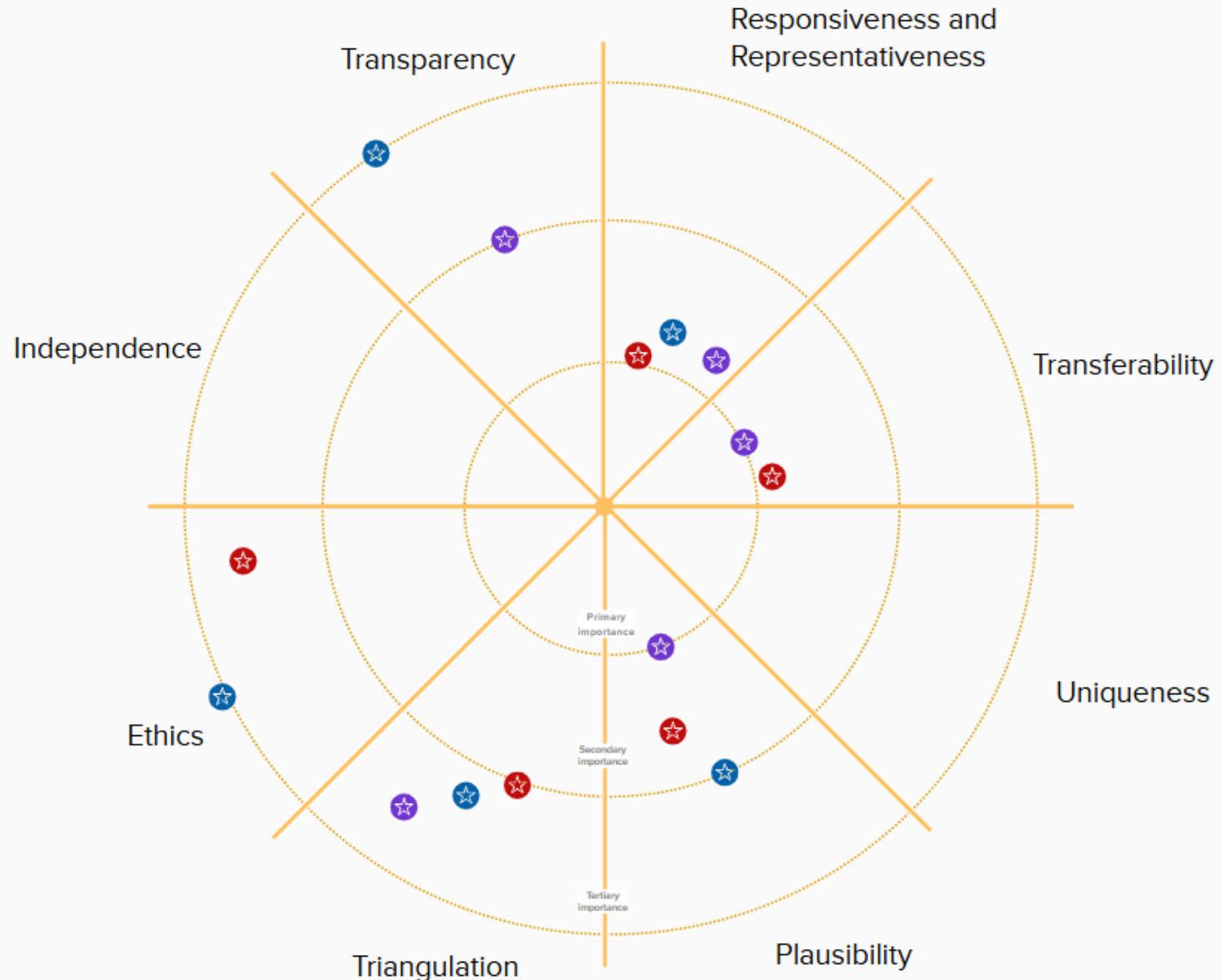


Elevating lived experience and multiple voices



Usefulness for school model improvement

Mini case study: An example of values from a causal pathways evaluation of virtual and hybrid school models



Mini case study: An example of values from a causal pathways evaluation of virtual and hybrid school models

Criteria	Description	Example approaches to support criteria
TRIANGULATION	Use of multiple lines and sources of evidence to inform the causal claims made, and consistency of findings obtained through different means	<ul style="list-style-type: none"> • Multiple lines of evidence (types of data) • Multiple data sources (individuals), sites, and time periods • Multiple analysts • Multiple theories • Multiple methods
PLAUSIBILITY	Coherence and strength of narrative about how intervention leads to outcomes, including how it follows the data (i.e., clear, logical explanation of observed causal steps between intervention and observed outcomes)	<ul style="list-style-type: none"> • Prolonged engagement • Persistent observation (focus on key elements) • Peer debriefing • Member checking
TRANSFERABILITY	Applicability of change narrative to other contexts and settings, and provision of information that allows others to judge the applicability of the evidence to their own situations	<ul style="list-style-type: none"> • Thick description of research context and participants
RESPONSIVENESS and REPRESENTATIVENESS	<p>Use of methods and processes to engage multiple voices and respond to local values and needs</p> <p>AND</p> <p>Engagement of those targeted by and involved in the intervention in different parts of the process that generated the evidence, and centrality of their perspectives in the evidence that is presented</p>	<ul style="list-style-type: none"> • Stakeholder mapping • Community engagement / participatory approaches, especially in design • Participatory approaches in data gathering, analysis, interpretation, and reporting • Sampling strategies • Gathering data directly or indirectly from all participant types • Separate analysis by participant type

Mini case study: An example of values from a causal pathways evaluation of virtual and hybrid school models

Table 1.3. Triangulation rubric (by performance level)

1	2	3	4	5
No evidence corroborates the connection between intervention and outcome. Other evidence contradicts the proposed connection.	A single line of evidence from a single source* supports the claim.	Multiple lines* of evidence from a single source corroborate the connection between intervention and outcome.	Multiple lines of evidence from multiple sources corroborate the connection between intervention and outcome.	Multiple lines of high-quality** evidence from multiple sources corroborate the connection between intervention and outcome.

Applying it to your work

SMALL GROUP DISCUSSION

Think about a specific evaluation you know well, ideally one that sought to make causal claims (or would have been strengthened by doing so):

- What values were (or were not) made explicit?
- One is one concrete thing I could have done to increase the quality and rigor of the study, drawing on the insights I learned today?

Share your example with your small group and reflect on the examples together.



Questions and Commitments

Your questions...

Taking Learning into Action

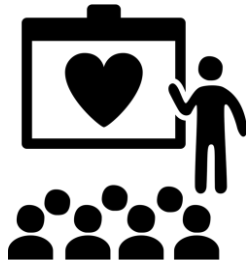
As you reflect on today's content, what is **one thing you can act on in the next three months?**

You might consider:

- Continuing to learn, leveraging Causal Pathways Initiative case studies, Better Evaluation, and other resources
- Sharing something you learned with a colleague
- Applying something you learned today in a current or new project

Where to find more about exploring causal pathways

Resources from the Causal Pathways Initiative

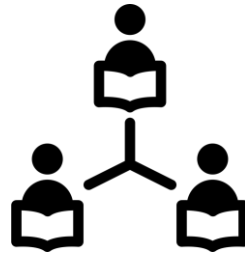


Presentations & trainings to build understanding and will

American Evaluation Association

Available to attend other events by request

Virtual 101 level training available on request



Resources to support understanding and action

BetterEvaluation.com resource hub on causal pathways evaluation

Case studies to provide stories and more detailed examples

Book chapter with step-by-step guidance on how to plan a causal pathways evaluation



Learning and acting together with support

Brain Trust to help funders work through tough questions with field experts

ABOUT THIS THEME**Theme type**[Cross-cutting themes](#)**Tags**[Causal Pathways](#)

Causal pathways

Contributing partner:



Causal Pathways Initiative

A causal pathways perspective on evaluation focuses on understanding how, why, and under what conditions change happens or has happened.

It is used to understand the interconnected chains of causal links that lead to a range of outcomes and impacts. These causal pathways are likely to involve multiple actors, contributing factors, events, and actions, not only the activities associated with the program, project, or policy being evaluated or its stated objectives.

Overview

Rather than being a specific approach, causal pathways evaluation might be best understood as a perspective on evaluation, which can draw on a combination of existing evaluation approaches, processes and methods. It uses a range of types of evidence, especially participant voices and narratives, and emphasises the use of participatory processes.

Causal pathways evaluation, as understood by the [Causal Pathways Initiative](#), can be distinguished by the following features:

THANK YOU

