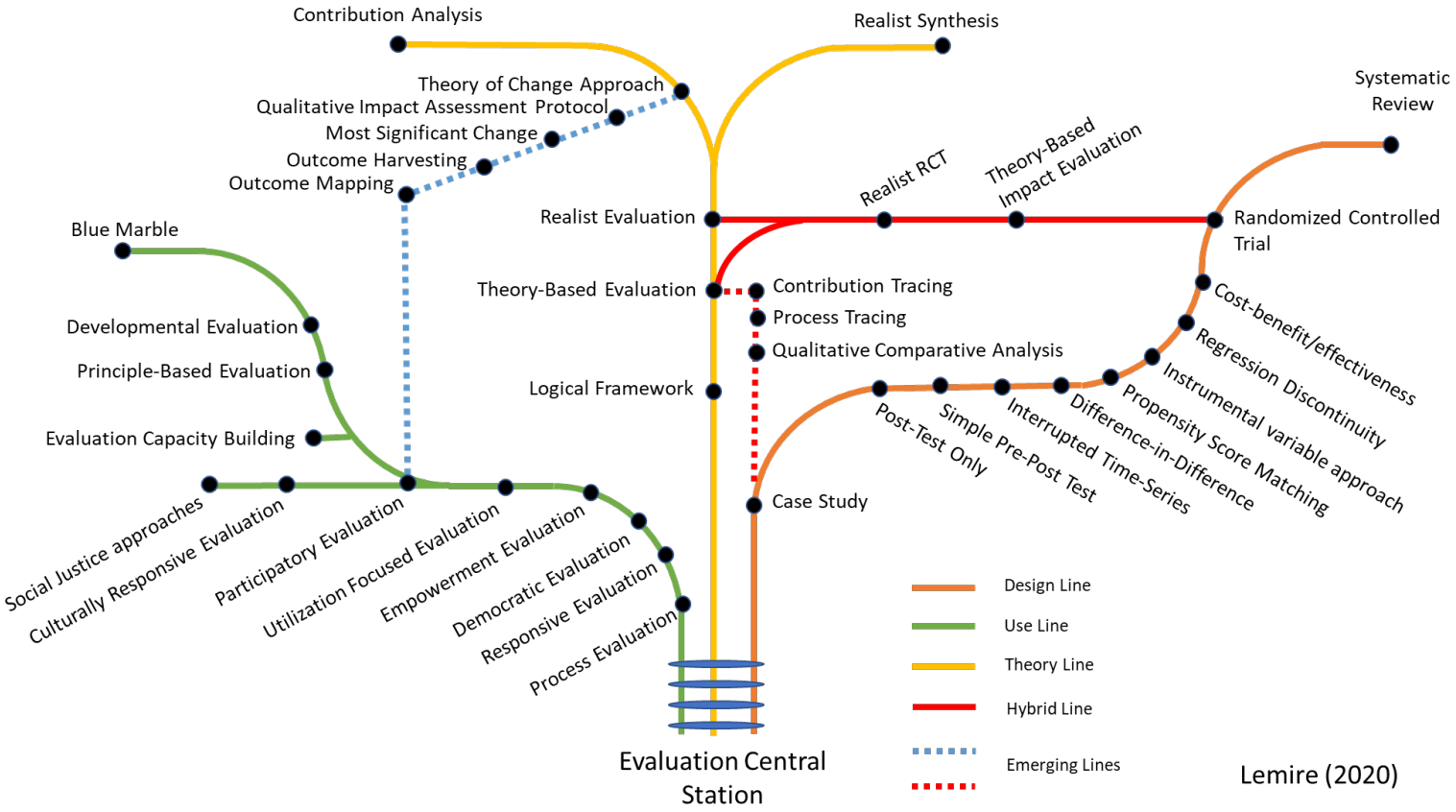
BOLD
THINKERS
DRIVING
REAL-WORLD
IMPACT

Practical Strategies For Testing and Refining Program Theories

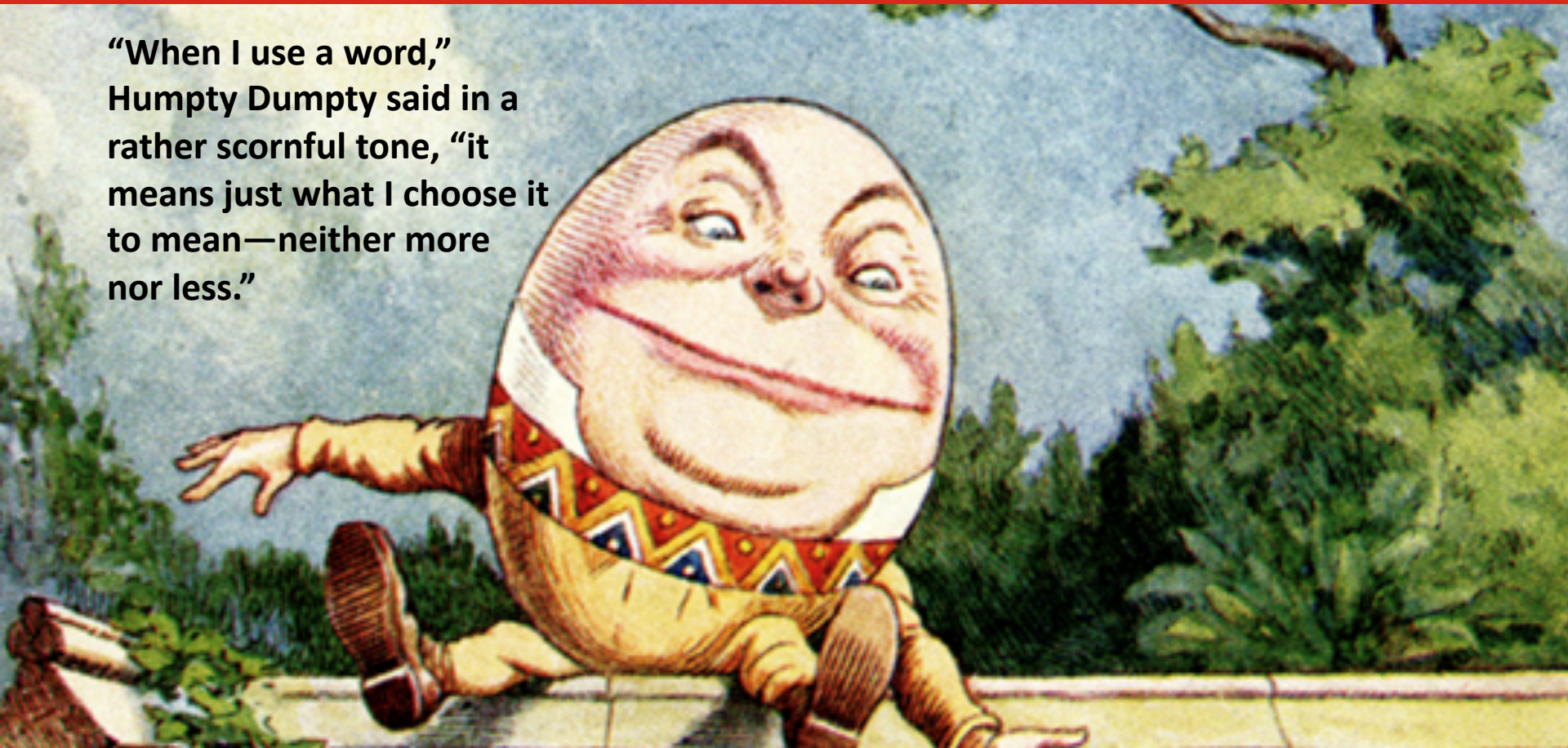
DES Learning Seminar
Sebastian Lemire

September 6
13:30-16:30

Evaluation Metro Map



What is Thing Called a Program Theory?



Today's Agenda

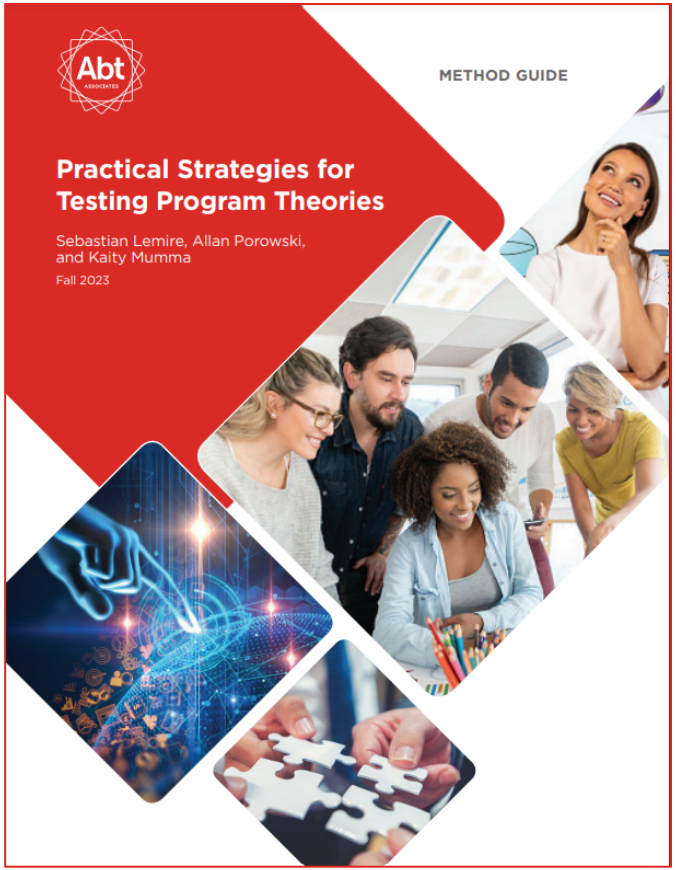
- Theoretical Assessment
  - Logic Analysis
- Data Collection and Coding
  - Realist Interview
  - Linked Coding

Break

- Analytical Strategies
  - Relevant Explanation Finder
  - Contribution Scores
  - Qualitative Comparative Analysis

Break

  - Structural Equation Modeling
- Reflective Practice

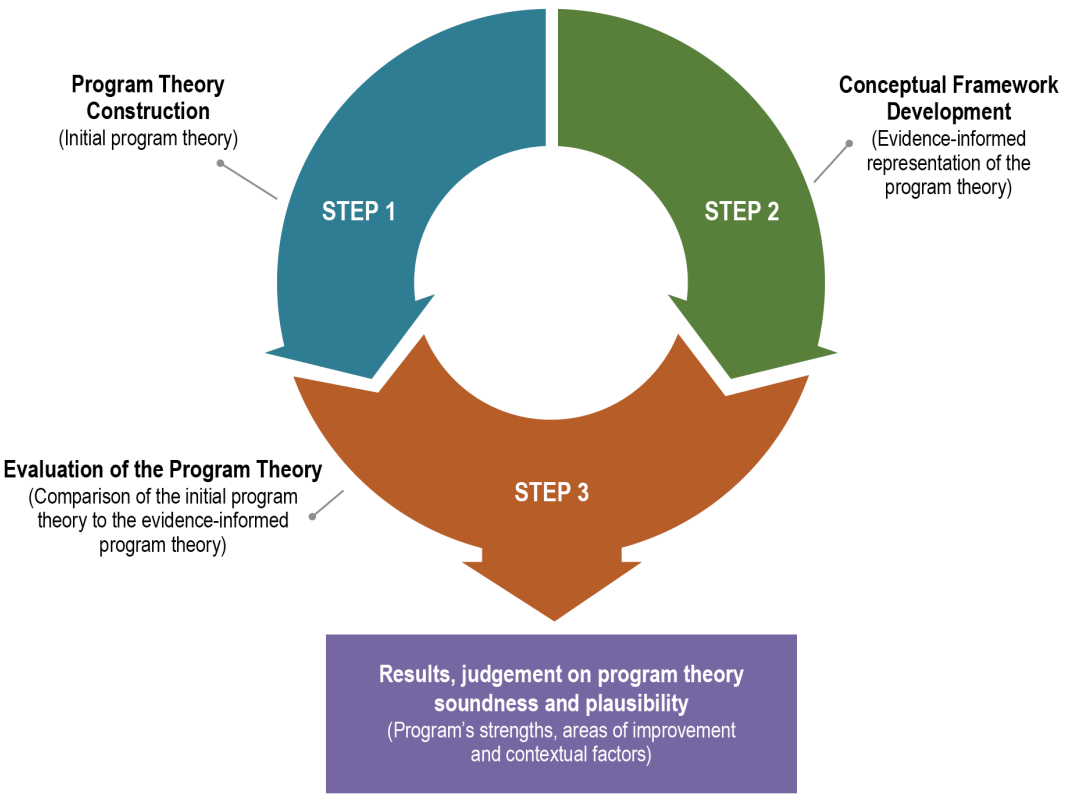


BOLD THINKERS DRIVING REAL-WORLD IMPACT

Theoretical Assessment

Logic Analysis

Logic Analysis—Main Steps



Source: Hurtubise et al. (2020)

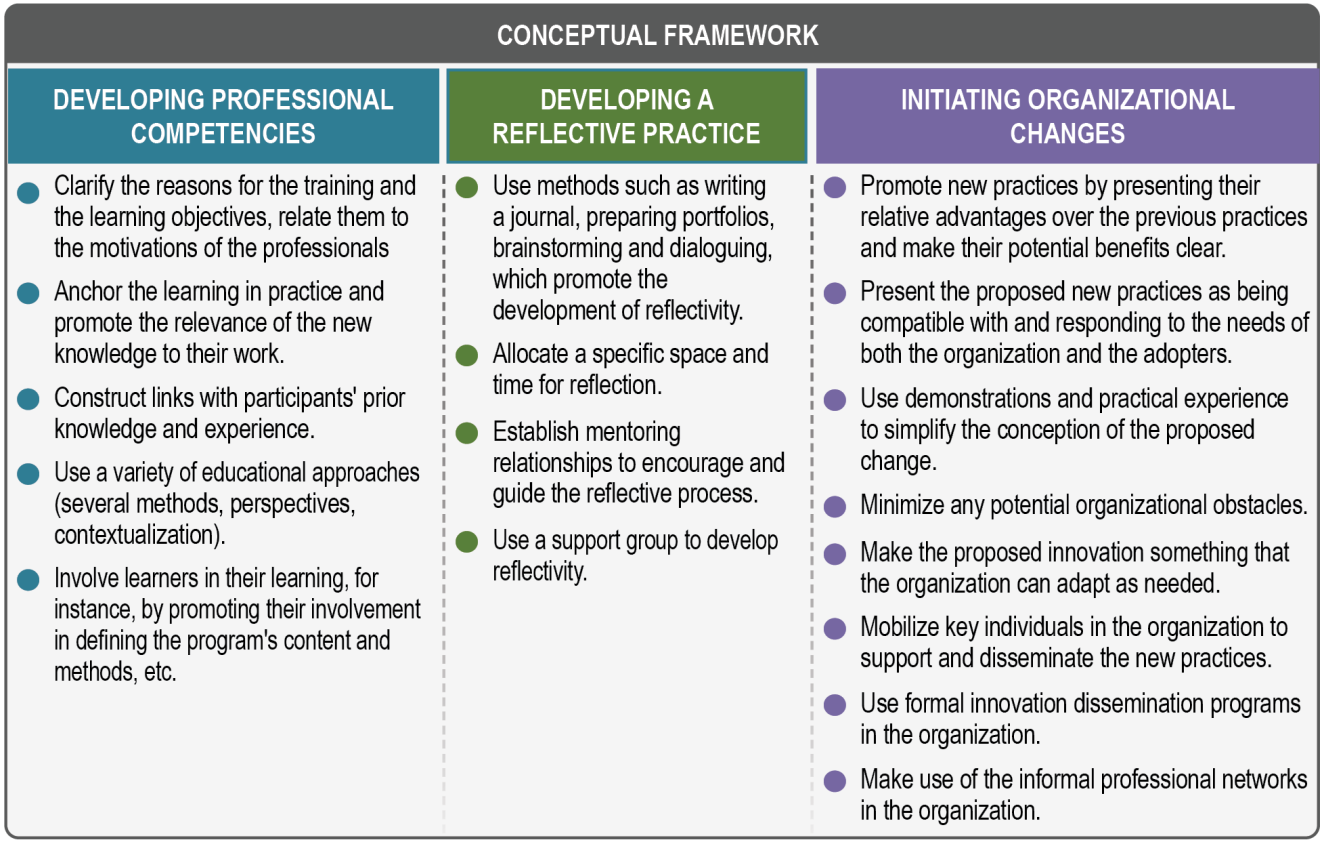
Direct and Reverse Logic Analysis

There are two types of Logic Analysis: direct and reverse.

Direct logic analysis examines whether a specific program design (as reflected in a program theory) is logically connected to a set of desired outcomes.

Reverse logic analysis identifies the program designs that can be logically connected to a set of desired outcomes.

Logic Analysis—Case Application



Source: Tremblay et al. (2013)

Logic Analysis—Benefits and Limitations

Benefits

- Theory knitting makes sense!
- Plausibility check on program design (and alternative designs!)
- Combined with a participatory approach on step 3 could be interesting!

Limitations

- Lack of procedural guidance on how to compare framework with program theory
- Few published case examples

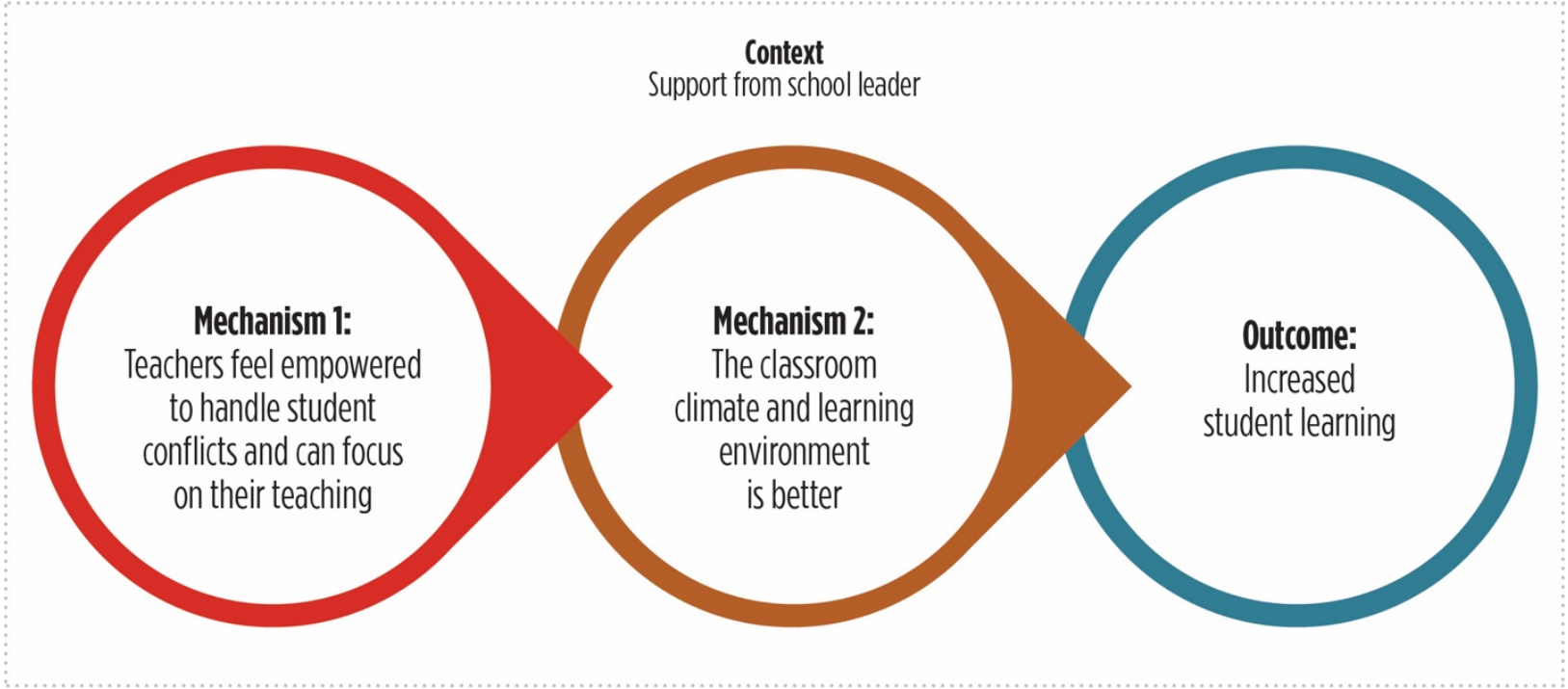
BOLD THINKERS DRIVING REAL-WORLD IMPACT

Data Collection and Coding

Realist Interviews

Logic Model Coding

Realist Interview



Source: Lemire, Porowski, Mumma (2023)

Realist Interview—Main Steps

1. Develop initial CMOs—based on “open” interviews with program staff.

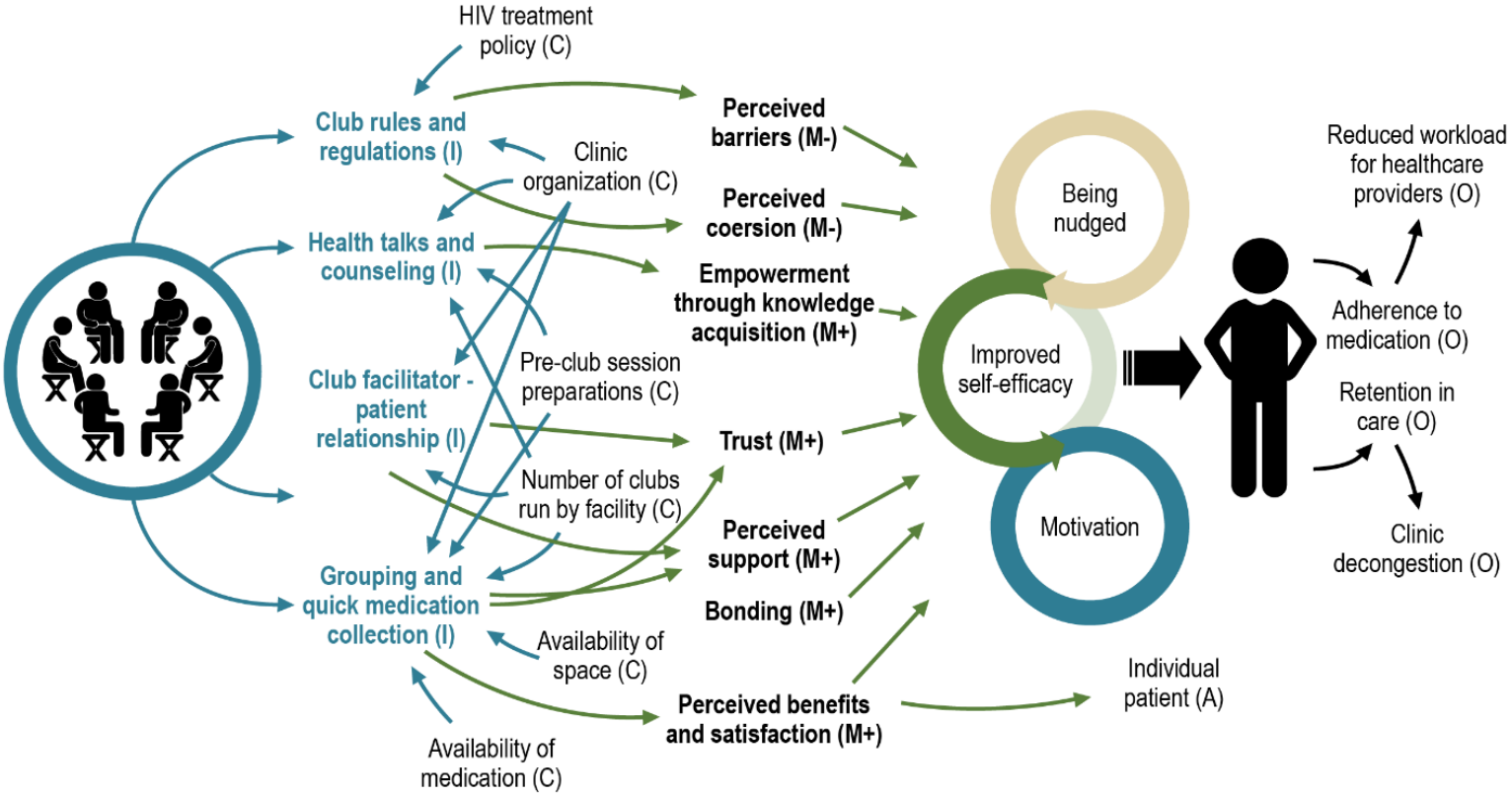
2. Refine CMOs based on CMO-centered interviews.

3. Confirm subset of CMOs based on CMO-centered interviews

Table 1. Topic guide for qualitative interviews with hospital staff in charge of patient 1 discharge.

	QUESTION	LOGIC
1	Could you explain your reasoning when organizing this patient’s discharge?	Questions 1–3 are introductory, to get them talking.
2	What ideally should happen to this patient in terms of discharge?	Questions 1–3 are introductory, to get them talking.
3	Could you talk to me about any difficulties that can alter your plans?	Questions 1–3 are introductory, to get them talking.
4	What characteristics in this patient would suggest that they may be likely to get delayed? I am thinking age, mental health condition, finances …	Exploring Context 1 (patient characteristics)
5	What characteristics in the way the staff works with this patient will help create a faster discharge?	Questions 5 & 6: Exploring Context 2 (Staff characteristics)
6	How do you think the level of experience in discharge planning in the members of the team has influenced the way this case has developed?	Testing Context 2 – staff experience. Questions 6–9. Looking also for mechanisms leading to people being transferred out of hospital faster. Asking about other staff first and then about themselves.
7	There seem to be external factors affecting the way this patient’s discharge plan progresses, I am not talking now about family but more about things like how the bed situation (lack of) in the ward may influence some of the decisions made …	Exploring Context 4 – Characteristics of the infrastructure. Questions 6–9. Looking also for mechanisms leading to people being transferred out of hospital to accelerate hospital discharges. Asking about others first and then about themselves.
8	How do you think the new fine system has impacted how social services staff dealt with this case? I am thinking that they may be doing things differently than they used to do before the new programme was implemented?	Questions 6–9. Looking for mechanisms. Asking about other staff first and then asking about themselves.

Realist Interview



Source: Mukumbang et al. (2020)

Realist Interview

Benefits

- Participant input is front and center!
- Group interaction

Limitations

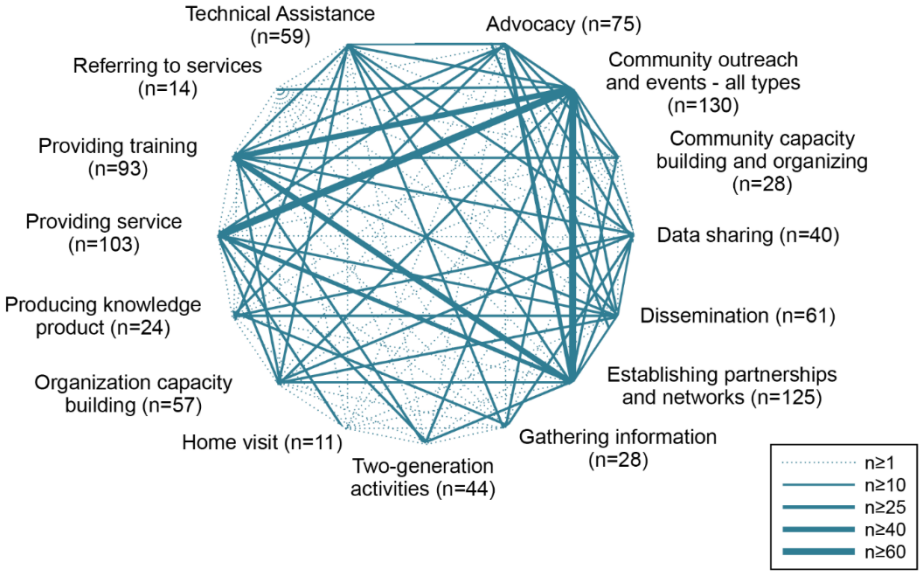
- Repeated interviews can be difficult
- Limited procedural guidance for step 3 (confirming CMOs)

Logic Model Coding

Source: Powell et al. (2023)

Logic Model Coding—Main Steps

1. Familiarization
2. Causal coding
3. Synthesis
4. Comparison



Source: Wu et al. (2019)

Logic Model Coding—Case Application

Coding Example:

A lot of our moms are first time moms, we do **have people coming back a second time,** but I think it benefits from having the community parent there**. It is that more comfortable person to talk to … they speak your language or come from the country you are from.**
[context—outcome—mechanism]

The mechanisms linked to frequent attendance that were specific to CPs were:
—CP is a person the participants can identify with
—CP makes personal connections with participants and ‘‘becomes like a friend’’
—CPs nonjudgmental attitude to parents
—CP refers to other parenting programs
—CP does follow-up call at home
—CPs do outreach in the community

Source: Jackson & Kolla (2012)

Logic Model Coding—Case Application

FACILITATORS

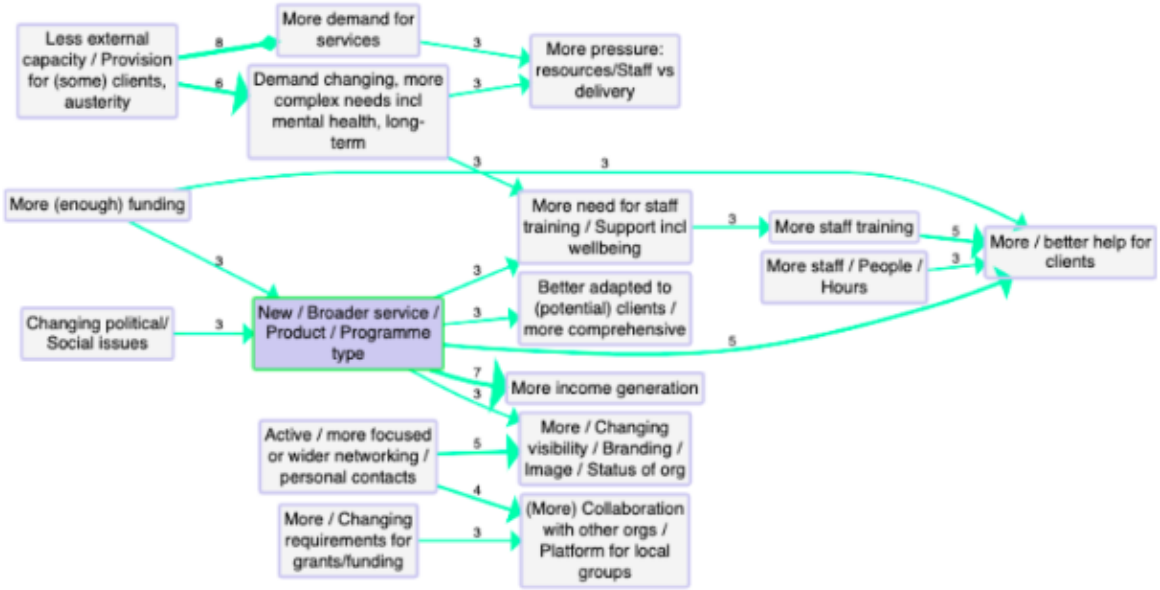
CONTEXT +	MECHANISM =	OUTCOME
• Strong family networks in rural areas • Variety of good-quality support providers • Professionals and parents provide supportive environment for the young person based on positive risk taking	• Flexibility with self-directed support to use individual budgets to employ family members • Accessible information (right information at right time for individual) • Self-advocacy supports individuals to articulate choice	• Some choice but limited to informal networks • Informed choice a reality • Informed choice a reality

BARRIERS

CONTEXT +	MECHANISM =	OUTCOME
• Traditional menu of services • Low expectations and risk-averse culture among professionals and parents • Health and social care integration that does not include child services/ child and adult services do not work collabratively	• Information limited to existing services • Information provided only on 'tried and tested' services • Information is service specific and not accessible to all young people with disability	• Choice limited to existing services • Choice limited to the statue quo • Young people and families lack both adequate information at the right time, and continuity of professional support to make informed choices

Source: Mitchell (2015)

Logic Model Coding—Benefits and Limitations



16 organisations said that they had introduced a new programme or broadened their service in some way

Source: Powell & Renmant (no date)

Let’s take a break!

Break

BOLD THINKERS DRIVING REAL-WORLD IMPACT

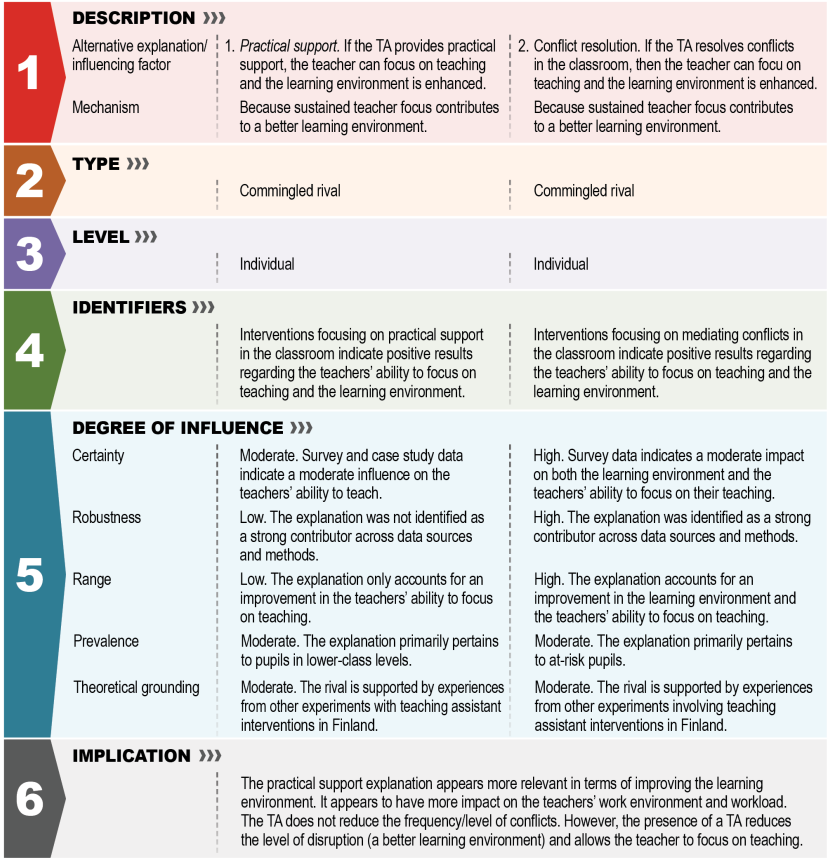
Analytical Strategies

Relevant Explanation Finder

Contribution Scores

Structural Equation Modeling

Relevant Explanation Finder (REF)—Main Steps



1	**DESCRIPTION** ›››		
	Alternative explanation/ influencing factor	1. *Practical support.* If the TA provides practical support, the teacher can focus on teaching and the learning environment is enhanced.	2. Conflict resolution. If the TA resolves conflicts in the classroom, then the teacher can focu on teaching and the learning environment is enhanced.
	Mechanism	Because sustained teacher focus contributes to a better learning environment.	Because sustained teacher focus contributes to a better learning environment.
2	**TYPE** ›››	Commingled rival	Commingled rival
3	**LEVEL** ›››	Individual	Individual
4	**IDENTIFIERS** ›››	Interventions focusing on practical support in the classroom indicate positive results regarding the teachers' ability to focus on teaching and the learning environment.	Interventions focusing on mediating conflicts in the classroom indicate positive results regarding the teachers' ability to focus on teaching and the learning environment.
5	**DEGREE OF INFLUENCE** ›››		
	Certainty	Moderate. Survey and case study data indicate a moderate influence on the teachers' ability to teach.	High. Survey data indicates a moderate impact on both the learning environment and the teachers' ability to focus on their teaching.
	Robustness	Low. The explanation was not identified as a strong contributor across data sources and methods.	High. The explanation was identified as a strong contributor across data sources and methods.
	Range	Low. The explanation only accounts for an improvement in the teachers' ability to focus on teaching.	High. The explanation accounts for an improvement in the learning environment and the teachers' ability to focus on teaching.
	Prevalence	Moderate. The explanation primarily pertains to pupils in lower-class levels.	Moderate. The explanation primarily pertains to at-risk pupils.
	Theoretical grounding	Moderate. The rival is supported by experiences from other experiments with teaching assistant interventions in Finland.	Moderate. The rival is supported by experiences from other experiments involving teaching assistant interventions in Finland.
6	**IMPLICATION** ›››	The practical support explanation appears more relevant in terms of improving the learning environment. It appears to have more impact on the teachers' work environment and workload. The TA does not reduce the frequency/level of conflicts. However, the presence of a TA reduces the level of disruption (a better learning environment) and allows the teacher to focus on teaching.	

- **Certainty**: The degree to which the observed outcome pattern matches the one predicted by the explanation/influencing factor.
- **Robustness**: The degree to which the explanation/influencing factor is identified as a significant contributor across a broad range of data sources and data collection methods.
- **Range**: The degree to which the explanation/influencing factor contributes to a broad range of the outcomes of interest.
- **Prevalence**: The degree to which the explanation/influencing factor contributes to the outcomes of interest across a wide range of implementation environments and target groups (e.g., different implementation sites and/or types of intervention).
- **Theoretical grounding**: The degree to which the explanation/influencing factor is informed by theory (i.e., existing theories linked to the explanation/influencing factors).

REF—Case Application

Degree of Influence	Explanation	Measure
Certainty	The degree to which the observed outcome matches the one predicted accounting for program assumptions and mechanisms	**Low** – little evidence (less than a quarter of the studies/evaluations) from the data sources confirm that the observed outcome matches the outcome described in the logic model **Medium** – approximately half (of the studies/ evaluations) the evidence from the data sources confirm that the observed outcome matches the one predicted **High** – it was noted across a range of different data sources that the observed outcome matches the one predicted
Robustness	The degree to which the assumption (or mechanism) is identified as a significant contributor to achieving program results	**Low** – across data sources there is limited evidence (less than a quarter of studies/evaluations) to show that the assumption (or mechanism) is a significant contributor to achieving results **Medium** – across data sources there is limited evidence (approximately half of the studies/ evaluations) that the assumption (or mechanism) is a significant contributor to achieving results **High** – It was noted across a range of different data sources that the assumption (or mechanism) is a significant contributor to achieving results
Prevalence	The degree to which the assumption (or mechanism) contributes to the outcomes of interest across a wide range of implementation sites	**Low** – assumption (or mechanism) affects limited implementation sites (less than a quarter studies/ evaluations) **Medium** – assumption (or mechanism) affects a range of different implementation sites (approximately half) **High** – assumption (or mechanism) affects a majority of implementation sites across NSW
Range	The degree to which the assumption (or mechanism) contributes to a broad range (e.g. impacts on one or more outcome) of outcomes.	**Low** – assumption (or mechanism) that affects one outcome of interest **Medium** – assumption (or mechanism) that affects half of the outcomes of interest **High** – assumption (or mechanism) that affect the majority of outcomes
Evidence based/ Evidence informed	The credibility and rigour of evidence which can verify or support findings from the analysis	**Low** – Untested local studies **Medium** – Case studies **High** – Experimental/quasi experimental studies

Alternative Explanations and Influencing Factors

Alternative explanation is an explanation for an outcome other than the program being evaluated. To illustrate, an alternative explanation can be a competing program that directly influence the observed outcome.

Influencing factor refers to aspects of the program setting and context that influences, positively or negatively, the ability of the program to generate the desired outcome(s).

REF—Benefits and Limitations

- Easy to use
- Promotes structure and transparency
- Evidence appraisal could be further formalized

Types of Explanations

Primary explanation – a program mechanism identified and purported to be the target intervention mechanism that accounts for and explains the observed outcomes.

Direct rival – a mechanism, different from the target program mechanism, that accounts for and explains the observed outcomes.

Commingled rival – other mechanisms, along with the program mechanism, that both contribute to and explain the observed outcomes.

Contribution Scores—Main Steps

1. Survey development.
2. Administer survey.
3. Convert ratings into contribution scores.
4. Refine program theory.

Key Terms

Contribution scores quantified ratings based on program participant perception of program influence on an outcome.

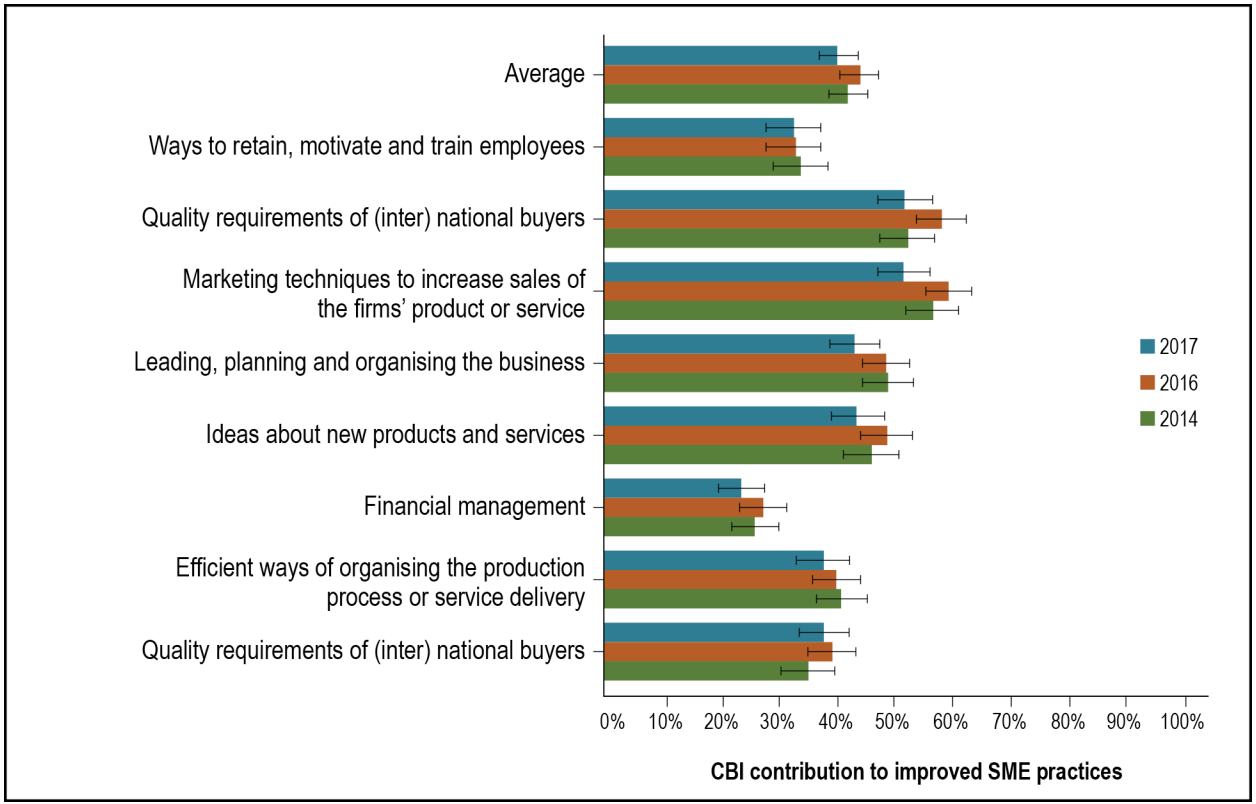
Causal pathway – a sequence of causal steps connecting specific program activities with one or more outcomes, depicted as a sequence of mechanisms in the program theory.

Contribution Scores—Case Application

CONVERSION TABLE			
Answer to question 2: How have your company's practices in this area changed over the past 12 months?	**Answer to question 3: Has [CBI/PUM] influenced this change?**	**Contribution rank (0-8)**	**Contribution score (%)**
Strong decrease	No effect	0	0
Decrease	No effect	0	0
No change	No effect	0	0
Increase	No effect	0	0
Strong increase	No effect	0	0
Increase	Very little	1	13
Strong increase	Very little	2	25
Increase	Some	2	38
Strong increase	Some	4	50
Increase	Quite a bit	5	63
Strong increase	Quite a bit	6	75
Increase	A lot	7	88
Strong increase	A lot	8	100

Source: Ton et al. (2023)

Contribution Scores—Case Application



Source: Ton et al. (2023)

Contribution Scores—Benefits and Limitations

- Direct way of gauging stakeholder assessment of perceived contribution
- More applications are called for!

Contribution Score Interpretation

Contribution score=0: There is no increase in knowledge or no program contribution.

Contribution score 1-50: There is an increase in knowledge (practice) and the program influenced this increase slightly or to some degree.

Contribution score 51-100: There is an increase in knowledge (practice) and PUM influenced the increase substantially or very substantially.

Qualitative Comparative Analysis (QCA)

1. Define the causal conditions and outcomes of interest;
2. Assemble relevant data on each case included in the analysis
3. Code each case according to the presence or absence (dichotomously or by degree) of each causal condition and outcome;
4. Use QCA software to summarize all the different causal configurations present among the cases;
5. Use QCA software to simplify the identified configurations into the essential set of causal recipes eliciting a positive or negative outcome;
6. Examine the consistency and empirical coverage of these recipes; and
7. Reexamine the individual cases represented by each of the identified causal recipes to better understand the nature of the latter.

QCA Software

QCA—Case Application

Study	Housing	Harm reduction	Supportive services	Client choice	Outcome
TSE(2000)	1	.67	1	1	1
GUL(2003)	1	1	1	1	.67
TSE(2003)	1	1	1	1	1
TSE(2004)	1	1	1	1	.67
GRE(2005)	1	1	1	.67	1
SIE(2006)	.67	.67	.67	.67	.67
STE(2007)	1	1	1	1	.67
TSA(2010)	1	.33	.33	.33	0
HAN(2011)	1	.33	.33	.33	1
APP(2012)	1	.33	1	.33	1
MON(2013)	1	.67	1	1	.67
PAL(2013)	1	.67	1	1	1
SOM(2015)	1	.67	1	1	1
STE(2015)	1	.67	1	1	.67
AUB(2016)	1	.67	1	1	.67
BRO(2016)	.67	.67	.67	.67	1.

Matched configuration of factors eliciting a positive outcome

Lemire & Christie (2019)

QCA—Case Application

	Coverage	Unique coverage	Consistency
~CHOICE*SERVICES*~HARM*HOUSING	0.13	0.03	0.83
CHOICE*SERVICES*HARM*HOUSING	0.76	0.66	0.88
Solution Coverage: 0.79			
Solution consistency: 0.88			

1. *~Choice*Services*~Harm*Housing*: Housing First programs with a strong fidelity to immediate housing and supportive services components combined with low fidelity to client choice and harm reduction promote housing tenure;
2. *Choice*Services*Harm*Housing*: Housing First programs with high fidelity to all four program components: provision of immediate housing, supported serviced, harm reduction, and client choice (i.e., the full Housing First model).

Lemire & Christie (2019)

QCA—Benefits and Limitations

Formation Stage	Consistency	Coverage
Coalition team leadership	1.0	1.0
Coordinator leadership	0.67	0.80
Team cohesion	0.50	0.75
Task focus	0.67	1.0
CPWI benefits	0.83	0.83
~CPWI costs	0.33	0.67
Sustainability planning	0.17	1.0
Maintenance stage		
Coalition team leadership	0.71	0.83
Coordinator leadership	0.57	1.0
Team cohesion	0.71	0.71
Task focus	0.71	0.83
CPWI benefits	0.57	0.57
~CPWI costs	0.57	0.67
Sustainability planning	0.57	0.80
Institutionalization stage		
Coalition team leadership	0.80	0.80
Coordinator leadership	0.40	0.57
Team cohesion	0.40	0.50
Task focus	0.60	0.67
CPWI benefits	0.40	0.57
~CPWI costs	0.30	0.43
Sustainability planning	0.50	0.63

Source: Jenkins et al. (2020)

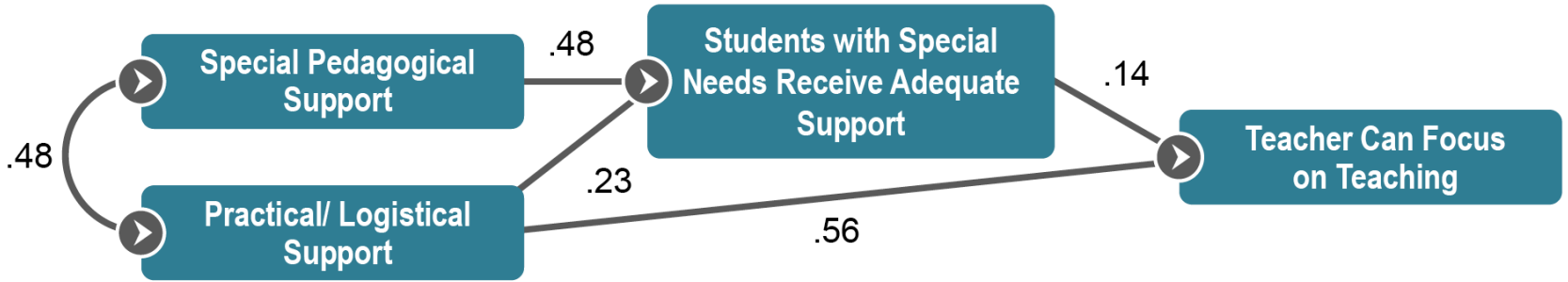
Consistency and Coverage

Consistency refers to the percentage of causal configurations of similar composition which result in the same outcome value. If the consistency of a configuration is low, it is not supported by empirical evidence. Therefore, it should be considered less relevant than other configurations with higher consistency

Coverage refers to the number of cases for which a configuration is valid. Unlike consistency, the fact that a configuration coverage is low does not imply less relevance.

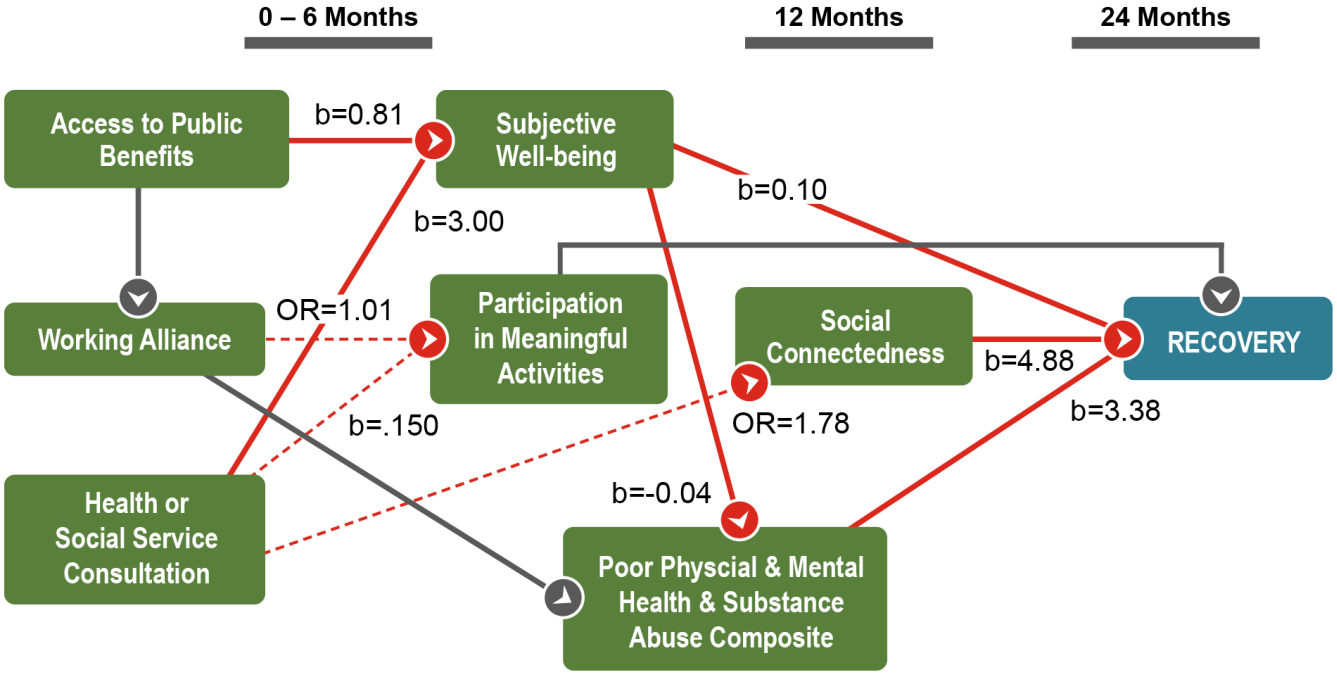
Let’s take a break!

Structural Equation Modeling



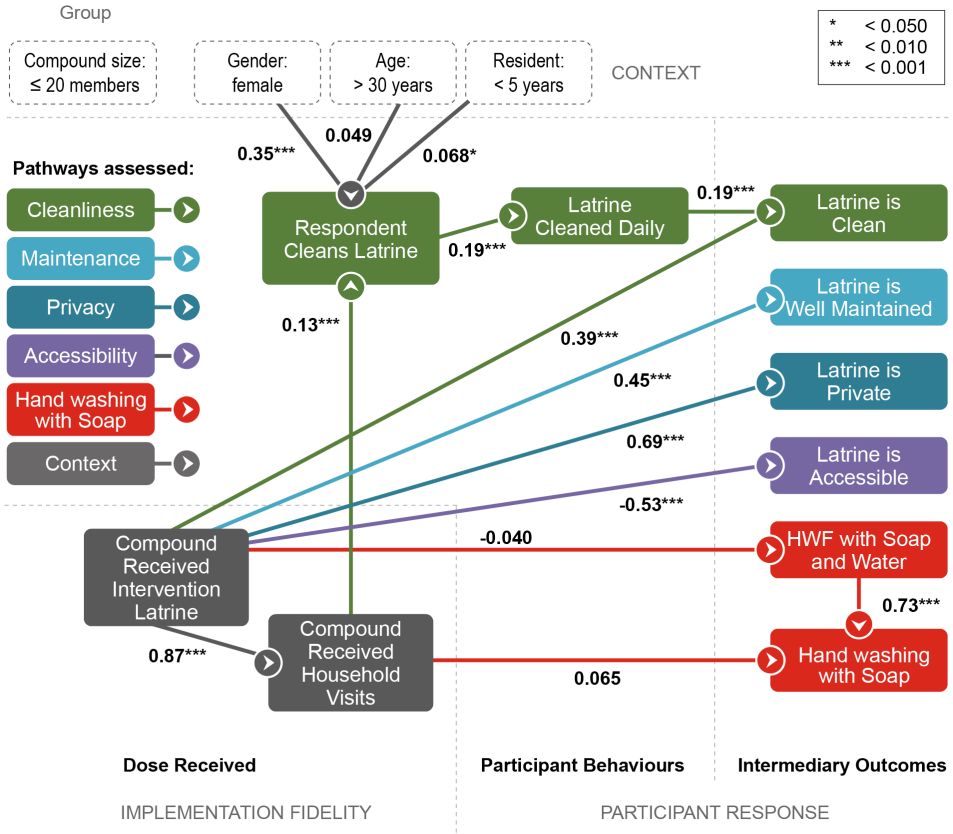
Source: Lemire, Porowski, Mumma (2023)

Structural Equation Modeling



Source: Adapted from O'Campo et al. (2022)

Structural Equation Modeling



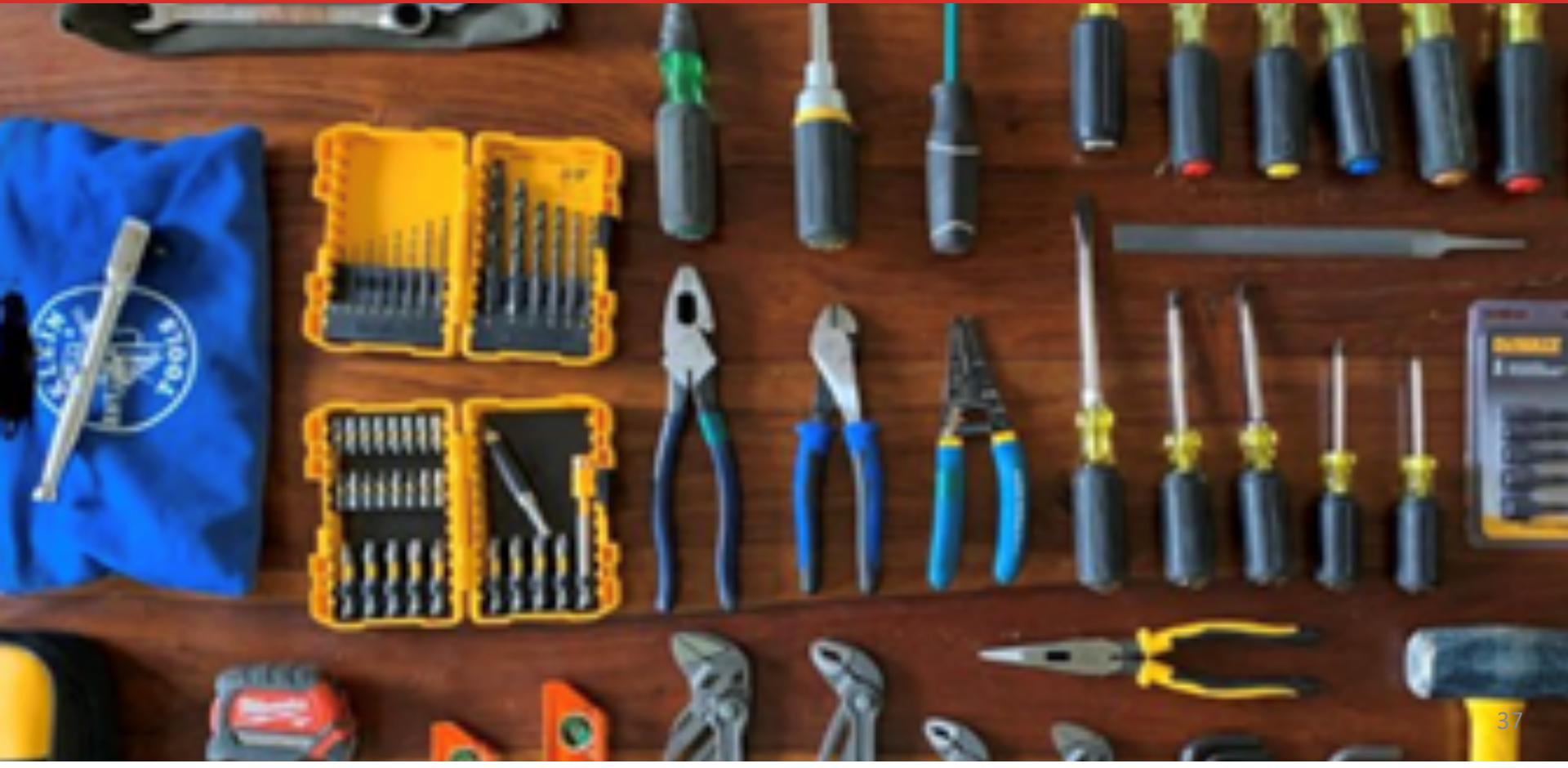
Source: Adapted from Bick et al. (2021)

BOLD THINKERS DRIVING REAL-WORLD IMPACT

Reflective Practice

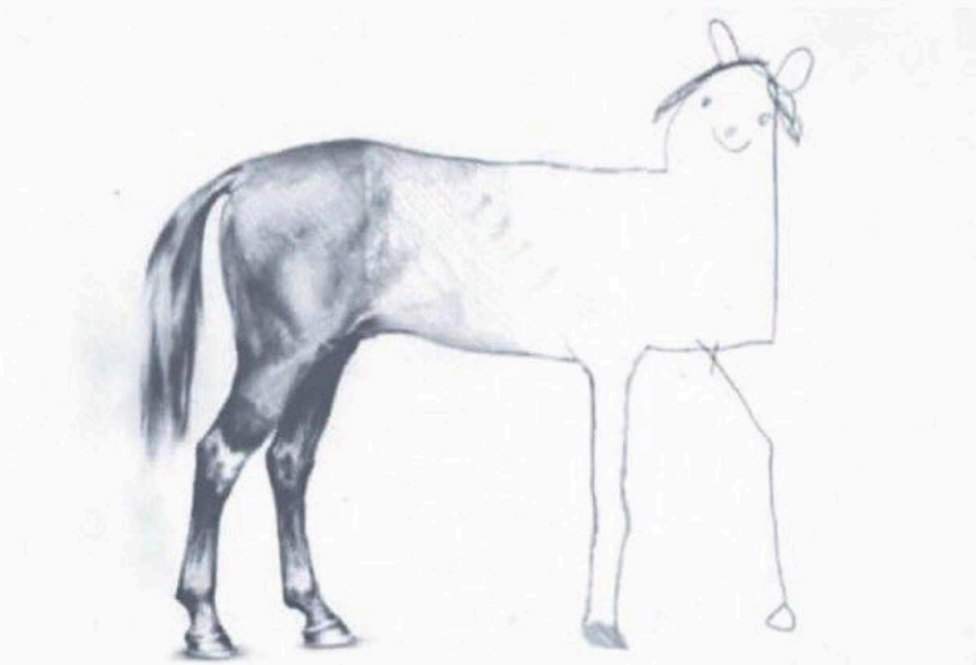
From Rigor by Design to Rigor in Thinking

Reflective Practice



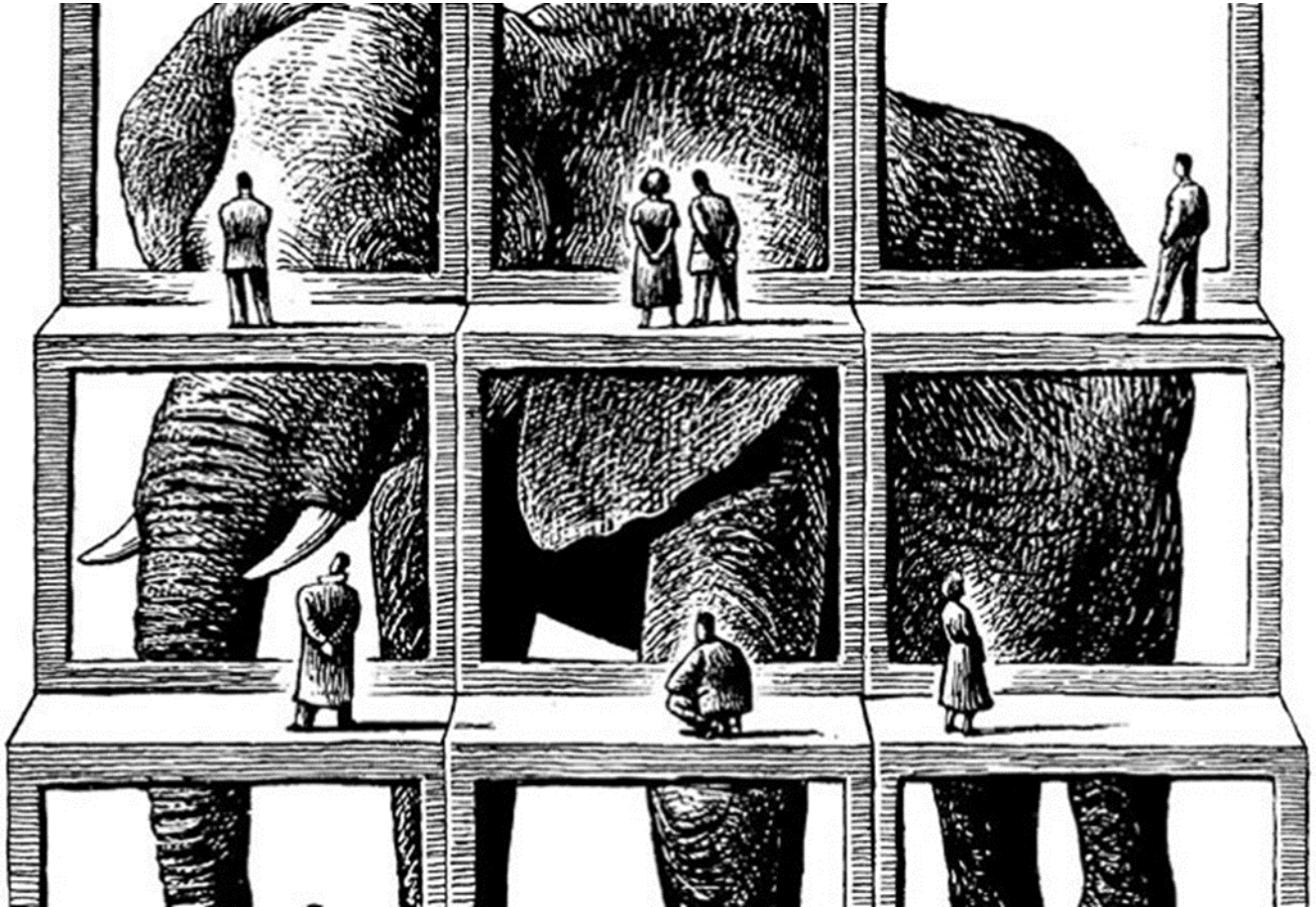
Real-world evaluation

Design of an evaluation → **Actual Implementation**



Ali Bati (2018)

Evidentiary Pluralism



Drill Down

1. **Core components**. Are there program components that define what the program is and/or constitute the primary driver for program impact?

2. **Critical assumptions**. Are there critical assumptions in the program theory, such as causal connections (arrows) and mechanisms that are necessary for the program to be successful and/or vulnerable to be influenced negatively by external factors? Actively examining assumptions around social norms and belief systems is critical for promoting equity.

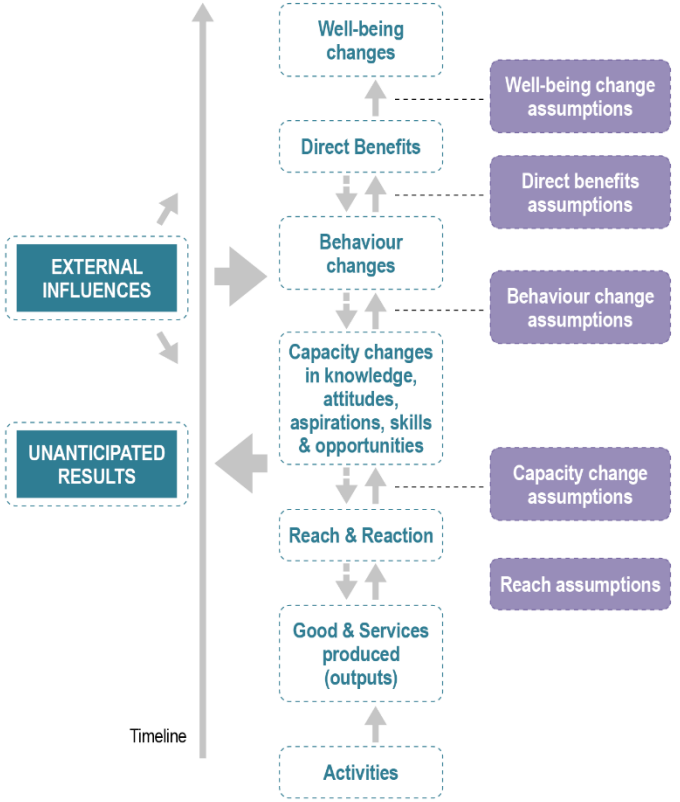
3. **Stakeholder/client relevance**. Which aspects of the program theory are particularly relevant for stakeholders or clients?

4. **Potential for new learning**. Which aspects of the program theory hold the greatest promise of new learning?

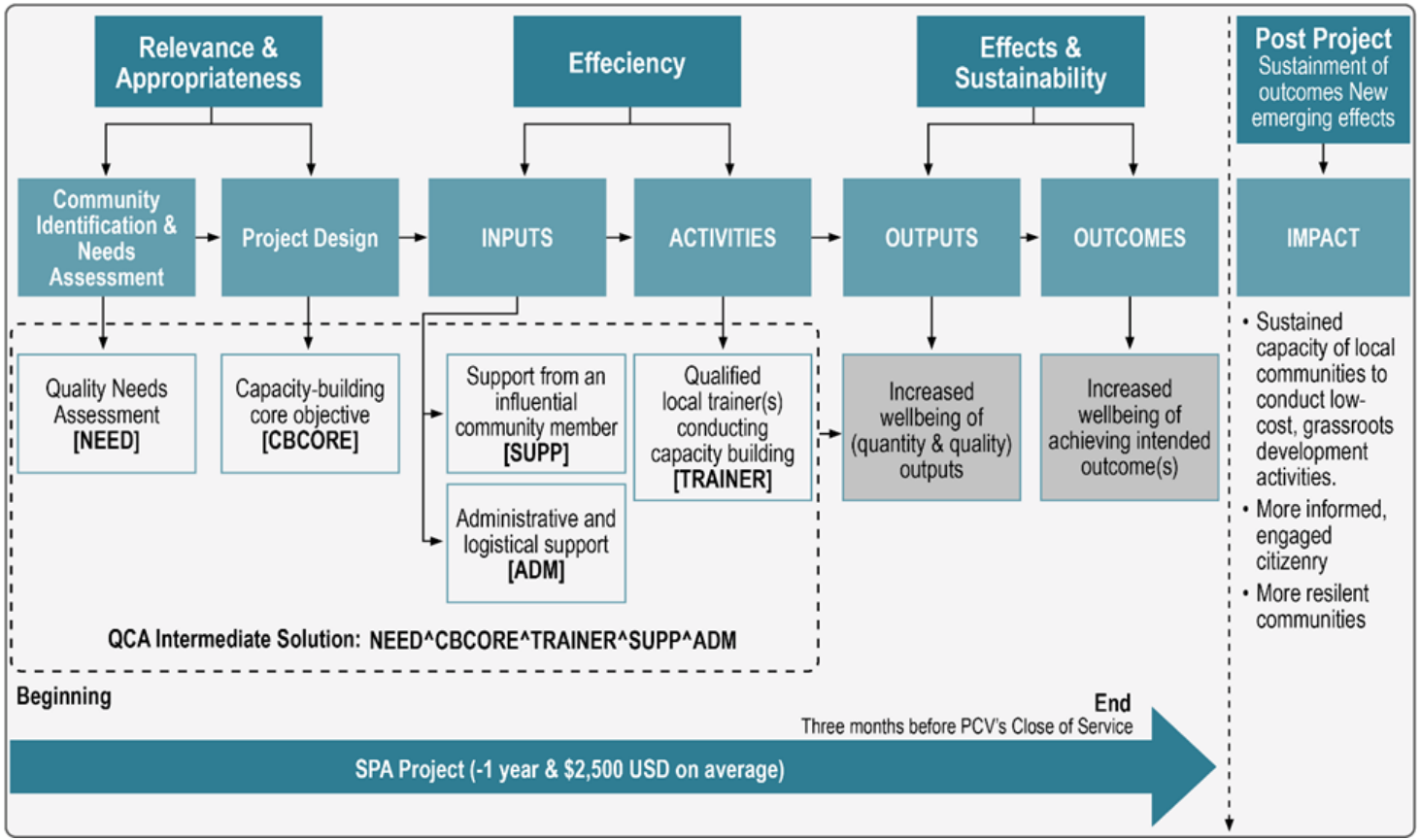
5. **Adverse consequences**. Are there aspects of the program theory that could potentially result in adverse or negative unintended consequences for anyone influenced by the program? This is particularly important from an equity perspective, to ensure the program is not doing harm.

Be Specific About Assumptions

- **Reach assumptions** are the events and conditions needed to occur if the outputs delivered are to reach and be positively received by the target group.
- **Capacity change assumptions** are the events and conditions needed to occur if the outputs that reach the target populations are to result in changes in their knowledge, attitudes, skills, aspirations, and opportunities; that is, their capacity to do things differently.
- **Behavior change assumptions** are the events and conditions needed to occur if the changes in the capacities of the target groups are to result in actual changes in their practices.
- **Direct benefits assumptions** are the events and conditions needed to occur if the behavior changes are to result in direct benefits for the target groups.
- **Well-being change assumptions** are the events and conditions needed to occur if the direct benefits are going to lead to changes in the well-being of the target group.



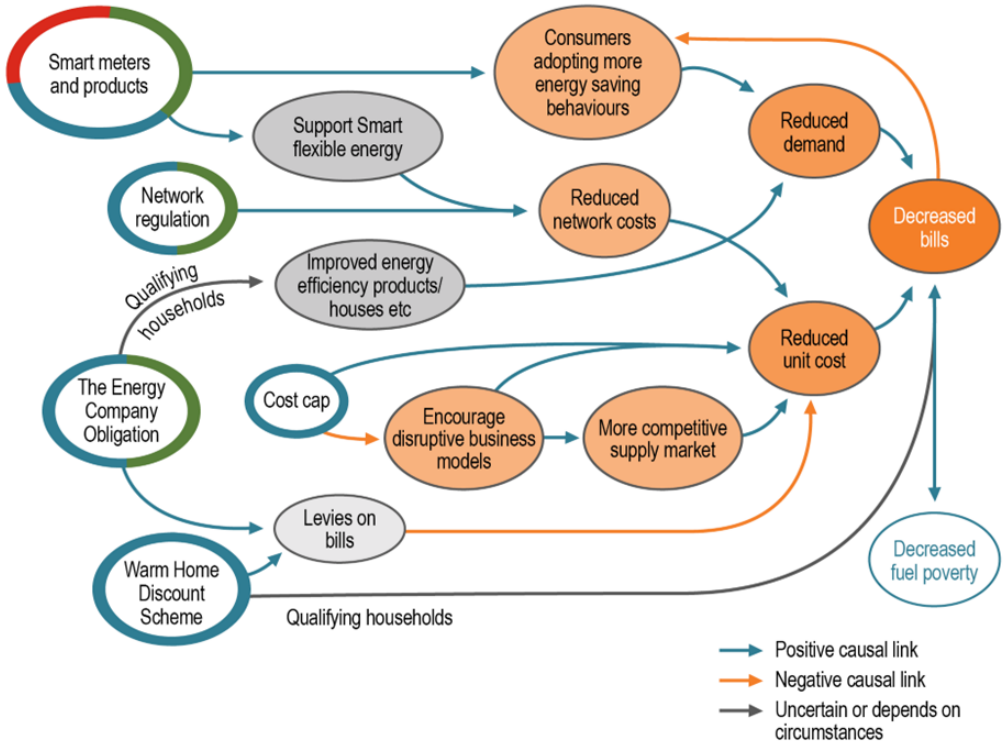
Visualize Evidence



Source: Dershem et al. (2023)

Mix and Match Visual Techniques

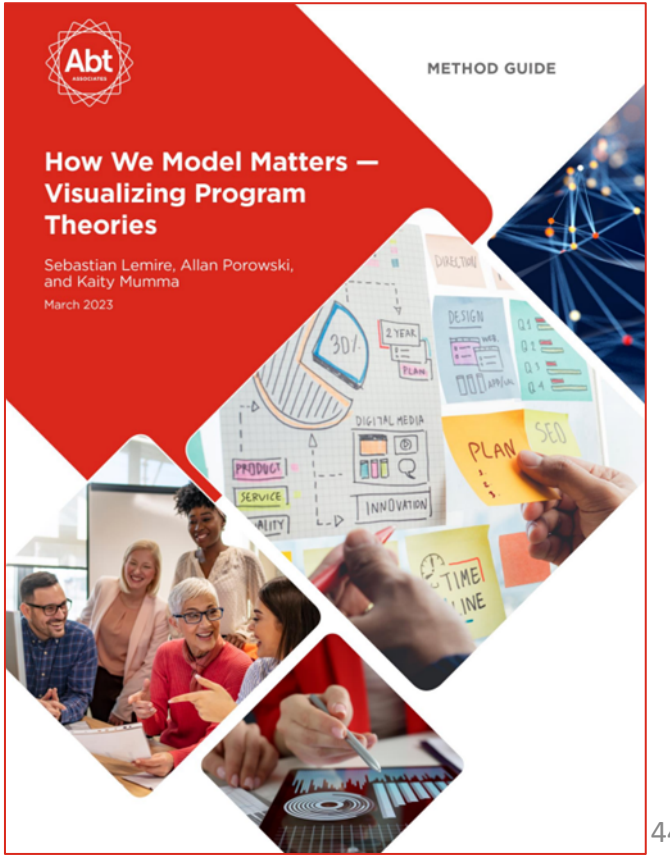
- Color-coded lines to indicate degrees of evidence or positive versus negative causal links
- Line thickness to distinguish between short- and medium-term outcomes
- Double bars "||" to indicate delayed outcomes
- Plus "+" and minus "–" signs or icons, such as smiley/frowny faces to reflect the direction and polarity of causal connections
- Differently shaped boxes or color-coding for program components, context, mechanisms, and outcomes



Source: Adapted from Wilkinson et al. (2021)

Resources

- Blogpost on six types of program theories
- Playbook with six types of program theories
- Abt's Method Guide on Visualizing Program Theories



BOLD
THINKERS
DRIVING
REAL-WORLD
IMPACT

abtassociates.com