

External validity and the portability of findings from there to here

Clues on what to look for and how
in development

Vegard Iversen, University of Manchester
Heather Lanthorn, 3ie & Harvard School of Public Health

Disclaimers

- Work/thoughts-in-progress!
 - Feedback welcome and appreciated
 - Especially on defining the way forward, not just the problem

Disclaimers

- Work/thoughts-in-progress!
 - Feedback welcome and appreciated
 - Especially on defining the way forward, not just the problem
- Terms: ‘settings,’ ‘there’ and ‘here’
 - Geography/population, within and across countries
 - Points in time – generally, within implementation cycle
 - Implementing partners
 - Scale and affected stakeholders



Agenda

- Why external validity and transportability?
- What has been said?
 - What seem to be the gaps?
 - Example to illustrate some gaps
- What might a way forward look like?

Punchlines

- Disagree with Woolcock on onus of responsibility for external validity

Punchlines

- Disagree with Woolcock on onus of responsibility for external validity
- External validity is about more than external replications, multi-site studies

Punchlines

- Disagree with Woolcock on onus of responsibility for external validity
- External validity is about more than external replications, multi-site studies
- Theory (of Change) is what allows empirical findings grounded 'here' to be potentially useful to those making decisions 'there'

Punchlines

- Disagree with Woolcock on onus of responsibility for external validity
- External validity is about more than external replications, multi-site studies
- Theory (of Change) is what allows empirical findings grounded 'here' to be potentially useful to those making decisions 'there'
- Research designs and reporting guided by a Theory of Change and “pre-mortem” thinking
 - Allows others to better make assessments of whether what worked 'there' can work 'here' (e.g. White 2009)
 - Propose certain categories to cover

Agenda

- Why external validity and transportability?
 - (so what / why?)
- What has been said?
 - (what?)
- What might a way forward look like?
 - (now what?)

Why care about there if we're here?

- Paul Gertler said so.



◇ Why? ◇ What's been said? ◇ What's next? ◇

Why care about there if we're here?

- Paul Gertler said so.
- Increasing focus on using rigorous evidence (IE, SR), 'best practices' to inform decision-making 'here' about programs and policies
 - (continue, modify, expand, terminate... or rename)
 - Evidence-informed decision-making calls for evidence that is both internally and externally rigorous

Why care about there if we're here?

- Paul Gertler said so.
- Increasing focus on using rigorous evidence (IE, SR), 'best practices' to inform decision-making 'here' about programs and policies
 - (continue, modify, expand, terminate... or rename)
 - Evidence-informed decision-making calls for evidence that is both internally and externally rigorous
- But:
 - Often don't have rigorous evidence from 'here.'
 - What does it mean to be rigorous about external validity?

◇ Why? ◇ What's been said? ◇ What's next? ◇

Issues

- Discussions on internal validity, rigor often crowd out discussion of ‘setting,’ ‘external validity’
 - Relegated to discussion sections, critique of last-resort

Issues

- Discussions on internal validity, rigor often crowd out discussion of 'setting,' 'external validity'
 - Relegated to discussion sections, critique of last-resort
- Quest for internal validity may undermine ability to make claims about transportability

Issues

- Discussions on internal validity, rigor often crowd out discussion of ‘setting,’ ‘external validity’
 - Relegated to discussion sections, critique of last-resort
- Quest for internal validity may undermine ability to make claims about transportability
- We lack counterpart tools to make assessments about external validity and transportability
 - We lack a ‘a coherent, system-level imperative’ to take external validity and claim-making seriously (Woolcock 2013)

◇ Why? ◇ What’s been said? ◇ What’s next? ◇

Key commentators

- Cartwright and Hardie (2012) – need to think about:
- Woolcock (2013) – need to think about:

◇ Why? ◇ What's been said? ◇ What's next? ◇

Key commentators

- Cartwright and Hardie (2012) – need to think about:
 - ‘Causal principles’ (Program P), ‘support factors’
- Woolcock (2013) – need to think about:

◇ Why? ◇ What's been said? ◇ What's next? ◇

Key commentators

- Cartwright and Hardie (2012) – need to think about:
 - ‘Causal principles’ (Program P), ‘support factors’
- Woolcock (2013) – need to think about:
 - Causal density (from Manzi 2012) – property of a program (“P”)
 - Certainty about how Program P works, fixed protocol
 - Transaction intensity, spaces for misinterpretation and mistranslation
 - Extent of local discretion around implementation
 - Incentives and temptations for implementers to do otherwise

Key commentators

- Cartwright and Hardie (2012) – need to think about:
 - ‘Causal principles’ (Program P), ‘support factors’
- Woolcock (2013) – need to think about:
 - Causal density (from Manzi 2012) – property of a program (“P”)
 - Certainty about how Program P works, fixed protocol
 - Transaction intensity, spaces for misinterpretation and mistranslation
 - Extent of local discretion around implementation
 - Incentives and temptations for implementers to do otherwise
 - Implementation capacity

◇ Why? ◇ What’s been said? ◇ What’s next? ◇

Key commentators

- Cartwright and Hardie (2012) – need to think about:
 - ‘Causal principles’ (Program P), ‘support factors’
- Woolcock (2013) – need to think about:
 - Causal density (from Manzi 2012) – property of a program (“P”)
 - Certainty about how Program P works, fixed protocol
 - Transaction intensity, spaces for misinterpretation and mistranslation
 - Extent of local discretion around implementation
 - Incentives and temptations for implementers to do otherwise
 - Implementation capacity
 - Time trajectories and functional forms

◇ Why? ◇ What’s been said? ◇ What’s next? ◇

Example: Twaweza (Lieberman *et al*)

Implicit theory of change:

Programme staff provide information to households with children of targeted age group



Why?



What's been said?

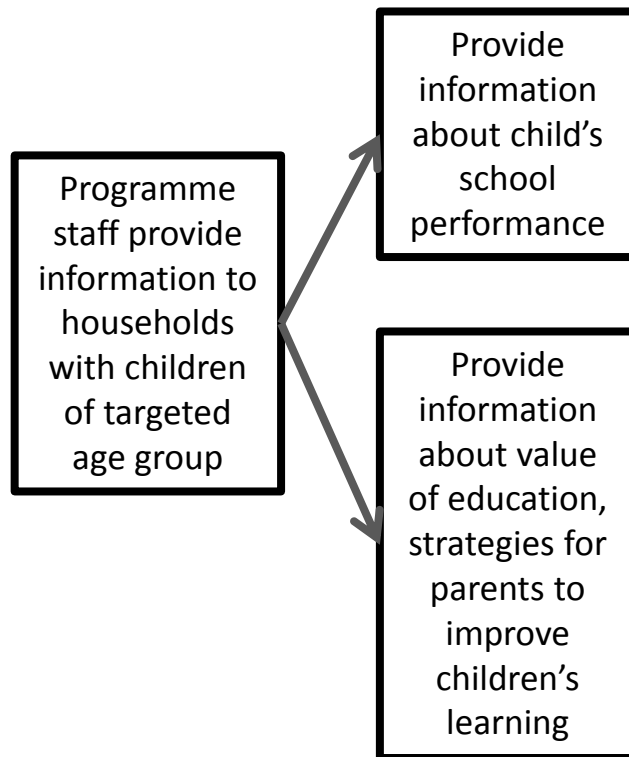


What's next?



Example: Twaweza (Lieberman *et al*)

Implicit theory of change:



Why?



What's been said?

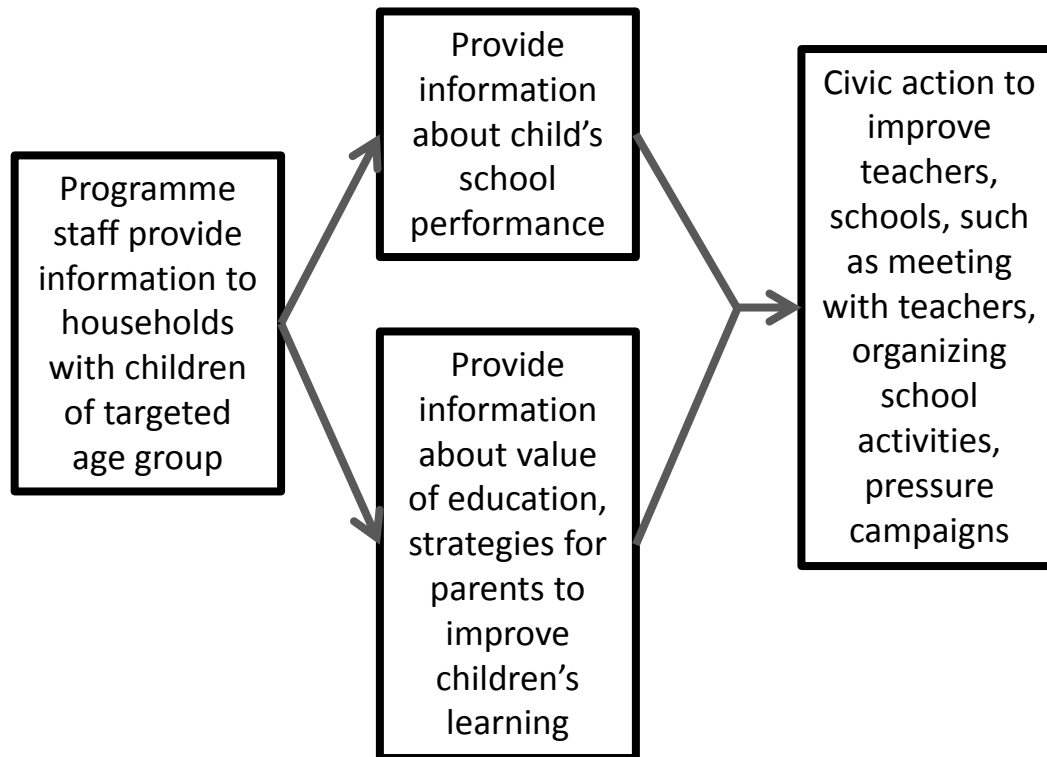


What's next?



Example: Twaweza (Lieberman *et al*)

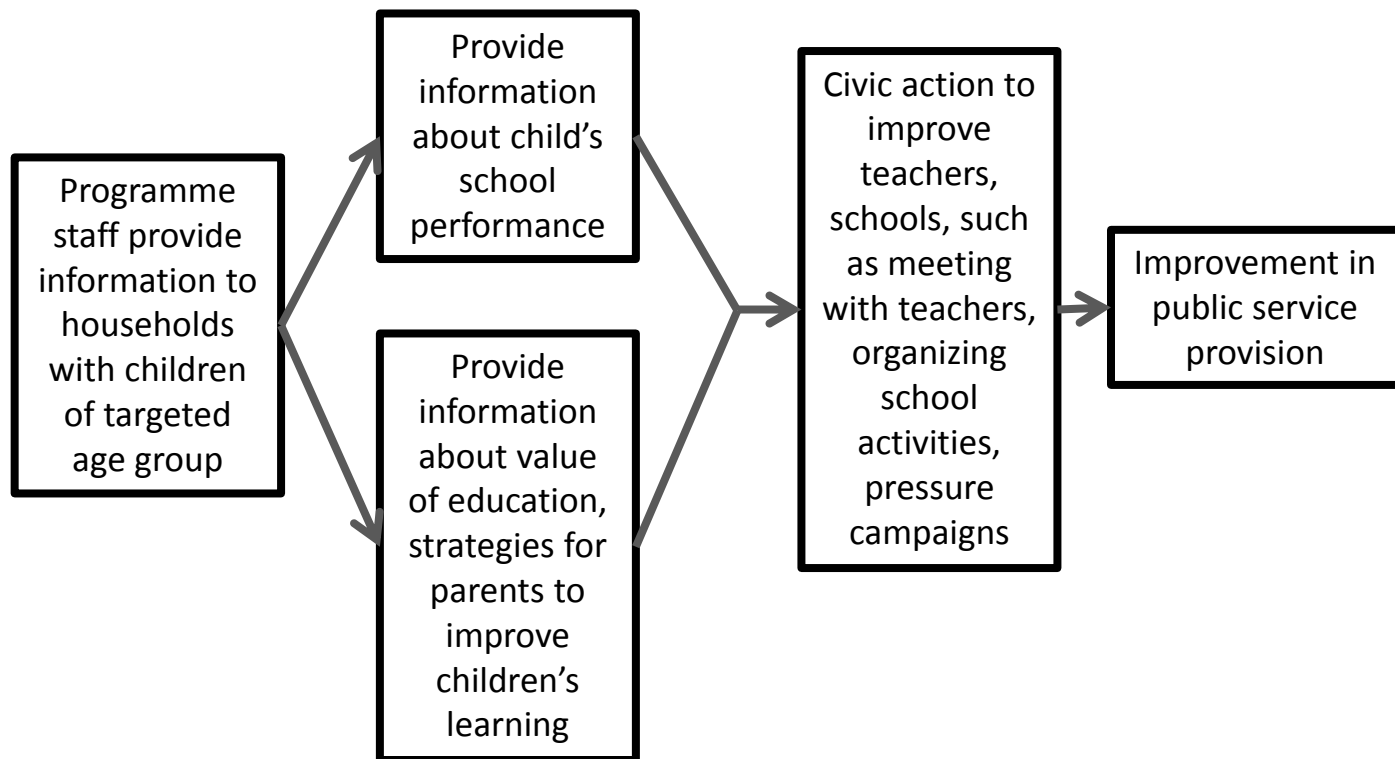
Implicit theory of change:



◇ Why? ◇ What's been said? ◇ What's next? ◇

Example: Twaweza (Lieberman *et al*)

Implicit theory of change:

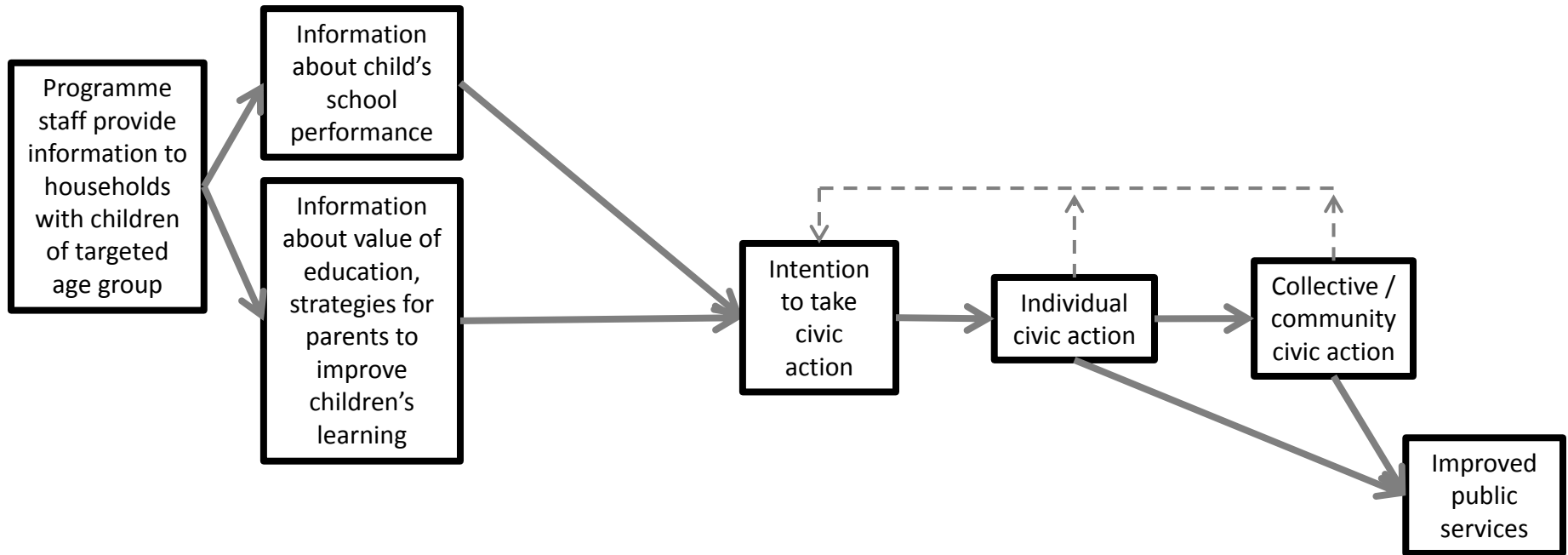


◇ Why? ◇ What's been said? ◇ What's next? ◇

Example: Twaweza (Lieberman *et al*)

- Find no treatment effect
- Engage in post-mortem thinking to create an ex post Theory of Change

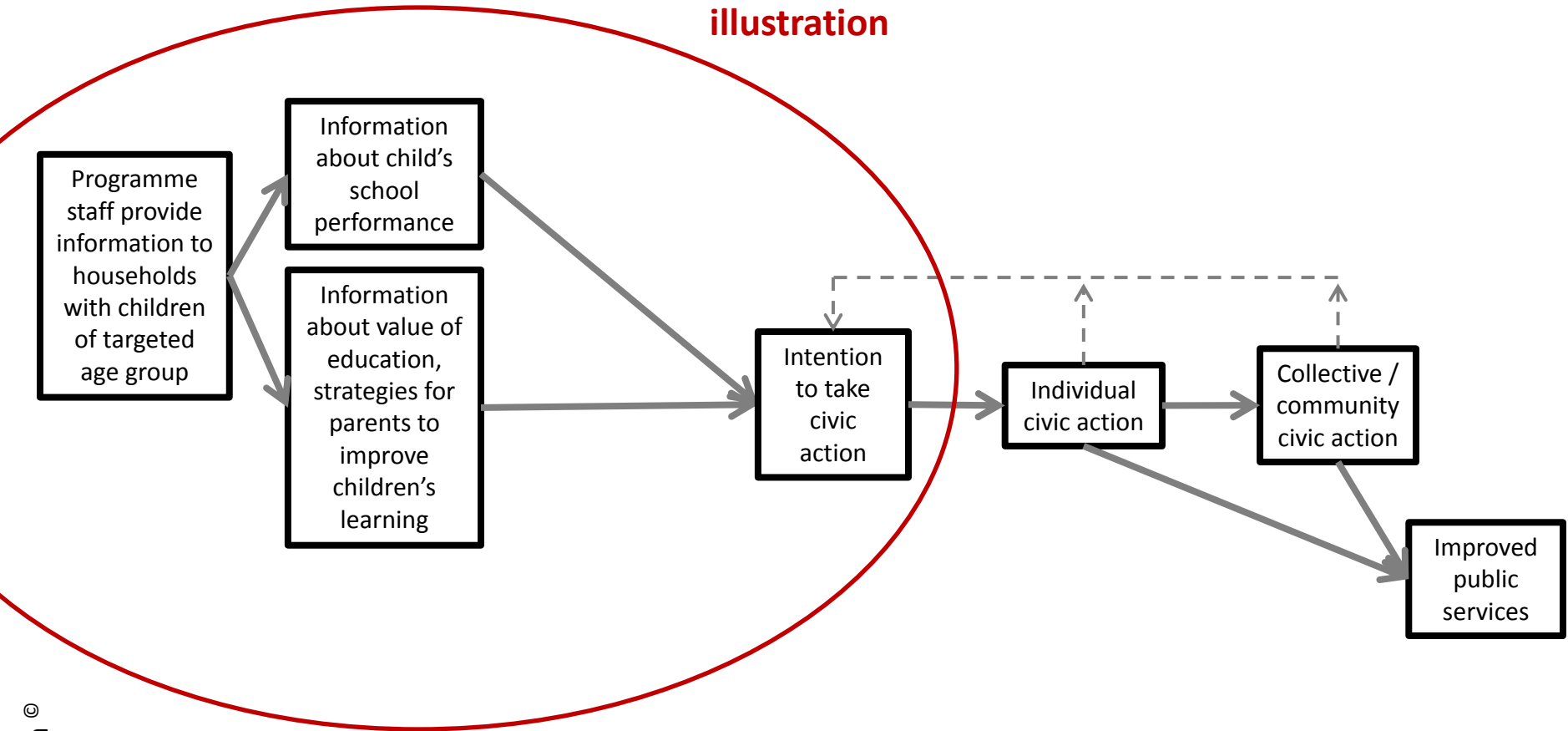
Twaweza (Lieberman *et al*): modified



◇ Why? ◇ What's been said? ◇ What's next? ◇

Twaweza (Lieberman *et al*): modified

illustration



Why?



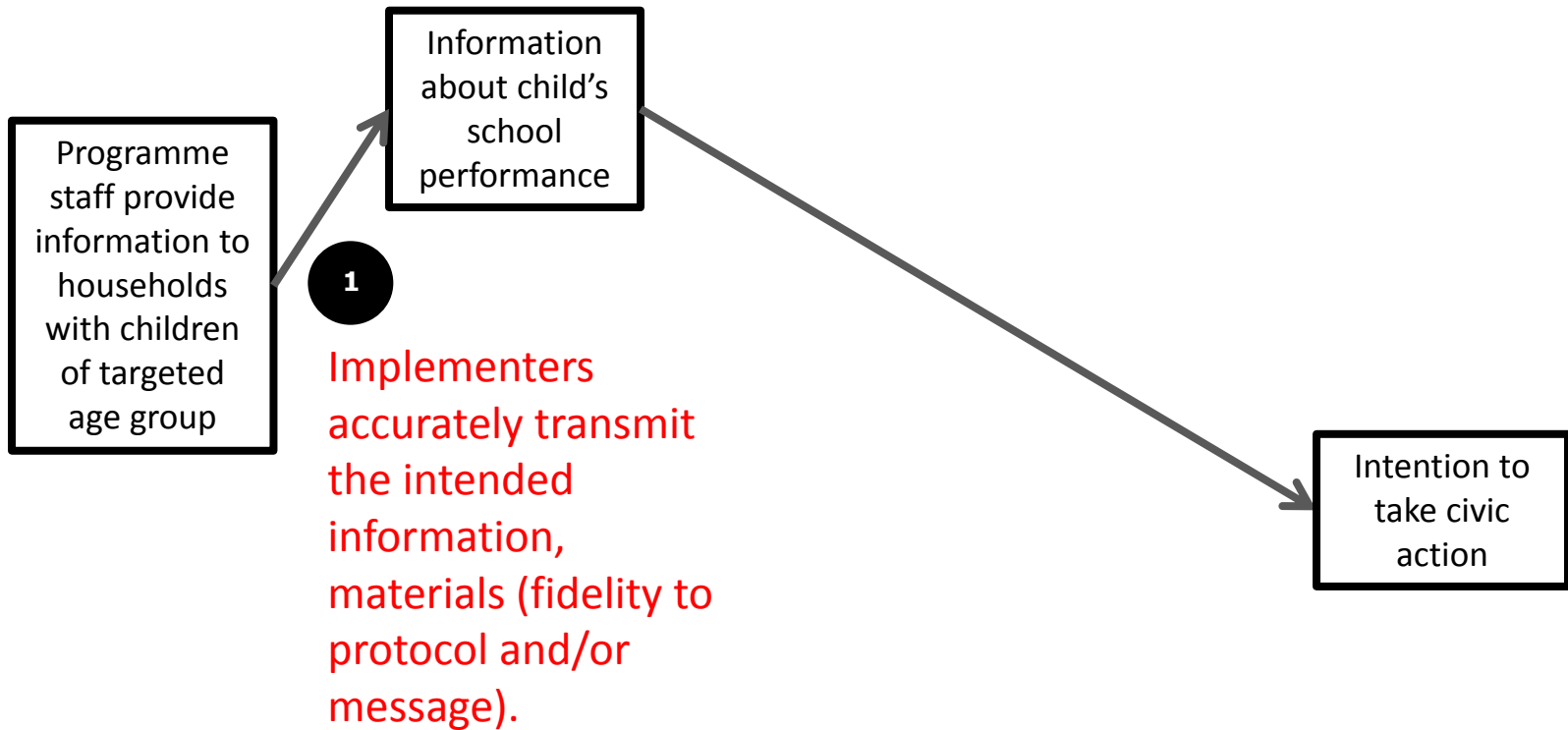
What's been said?



What's next?

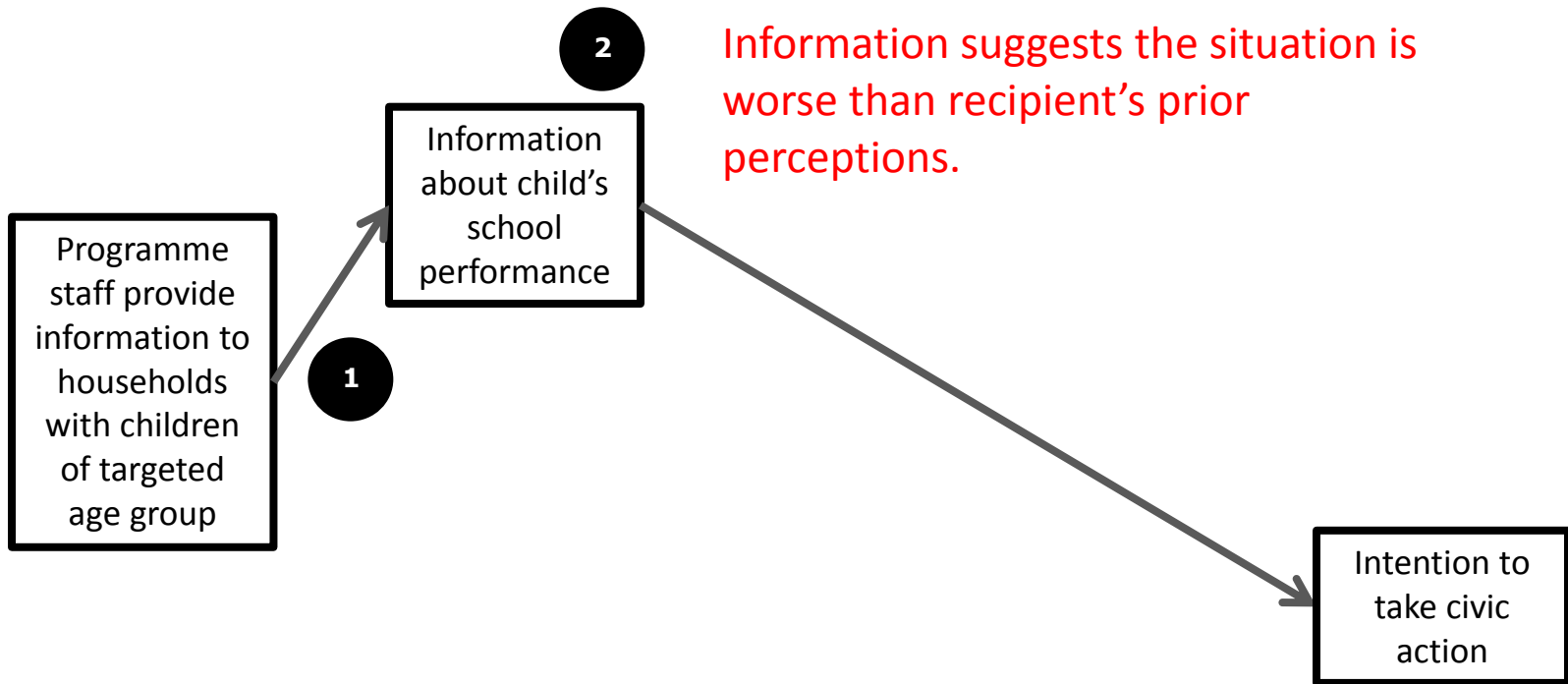


Twaweza (Lieberman *et al*): modified



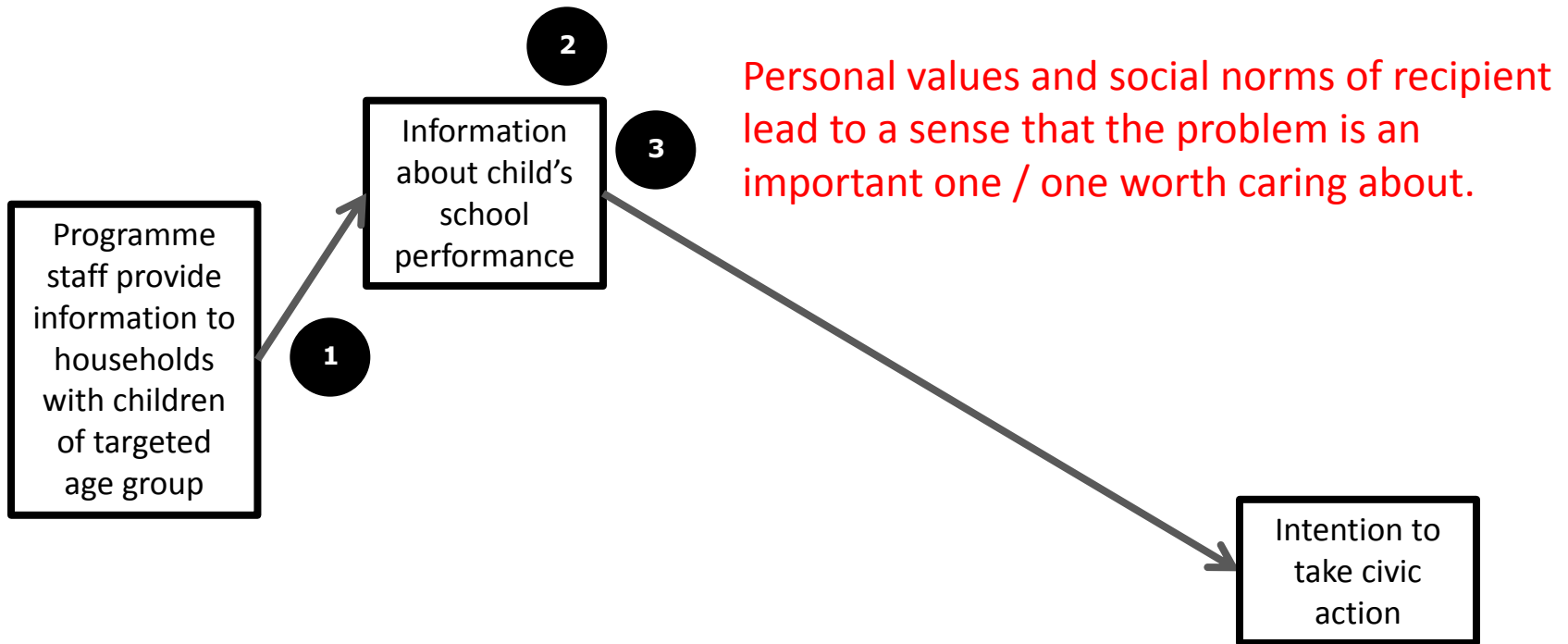
◇ Why? ◇ What's been said? ◇ What's next? ◇

Twaweza (Lieberman *et al*): modified



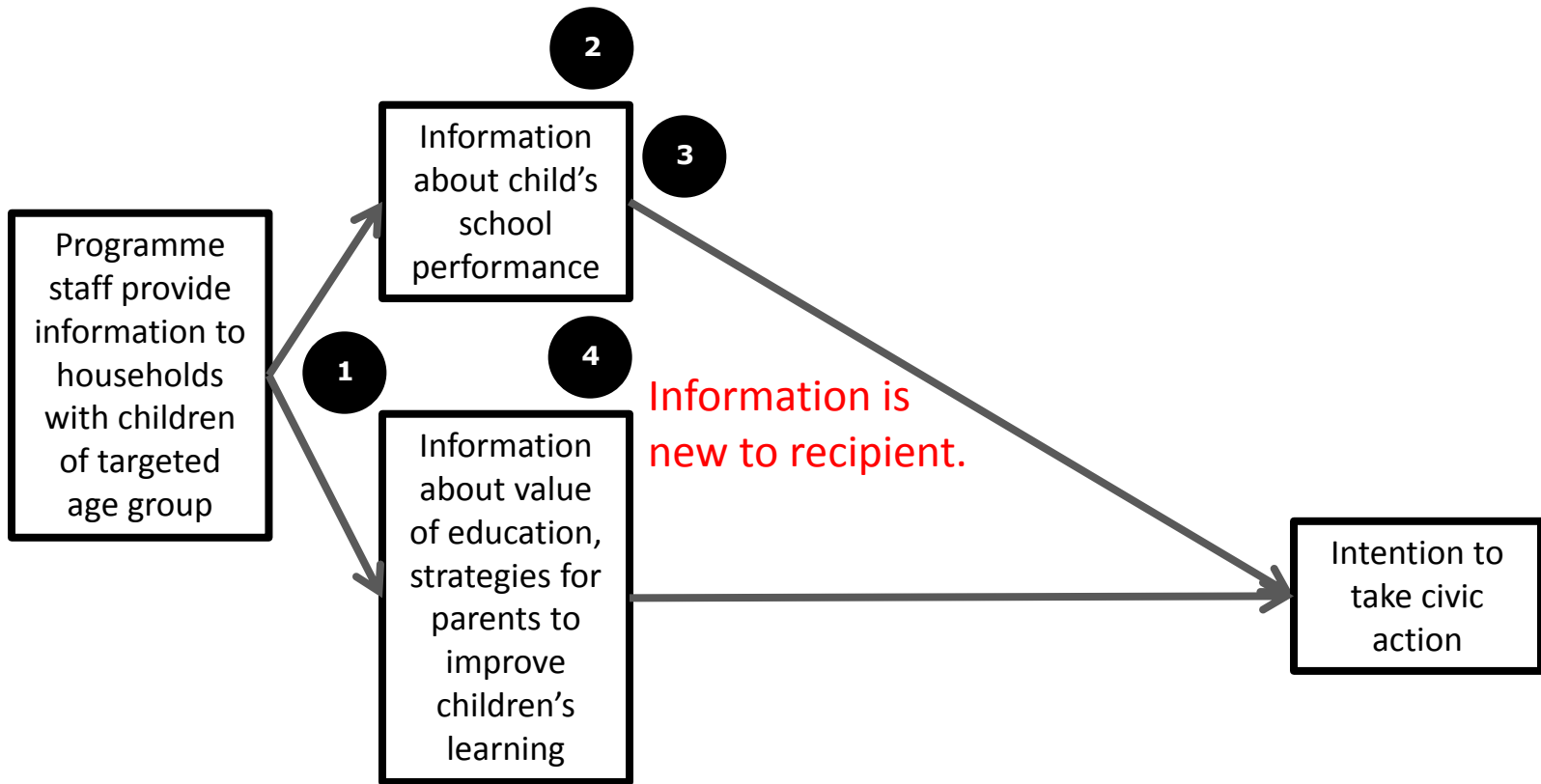
◇ Why? ◇ What's been said? ◇ What's next? ◇

Twaweza (Lieberman *et al*): modified



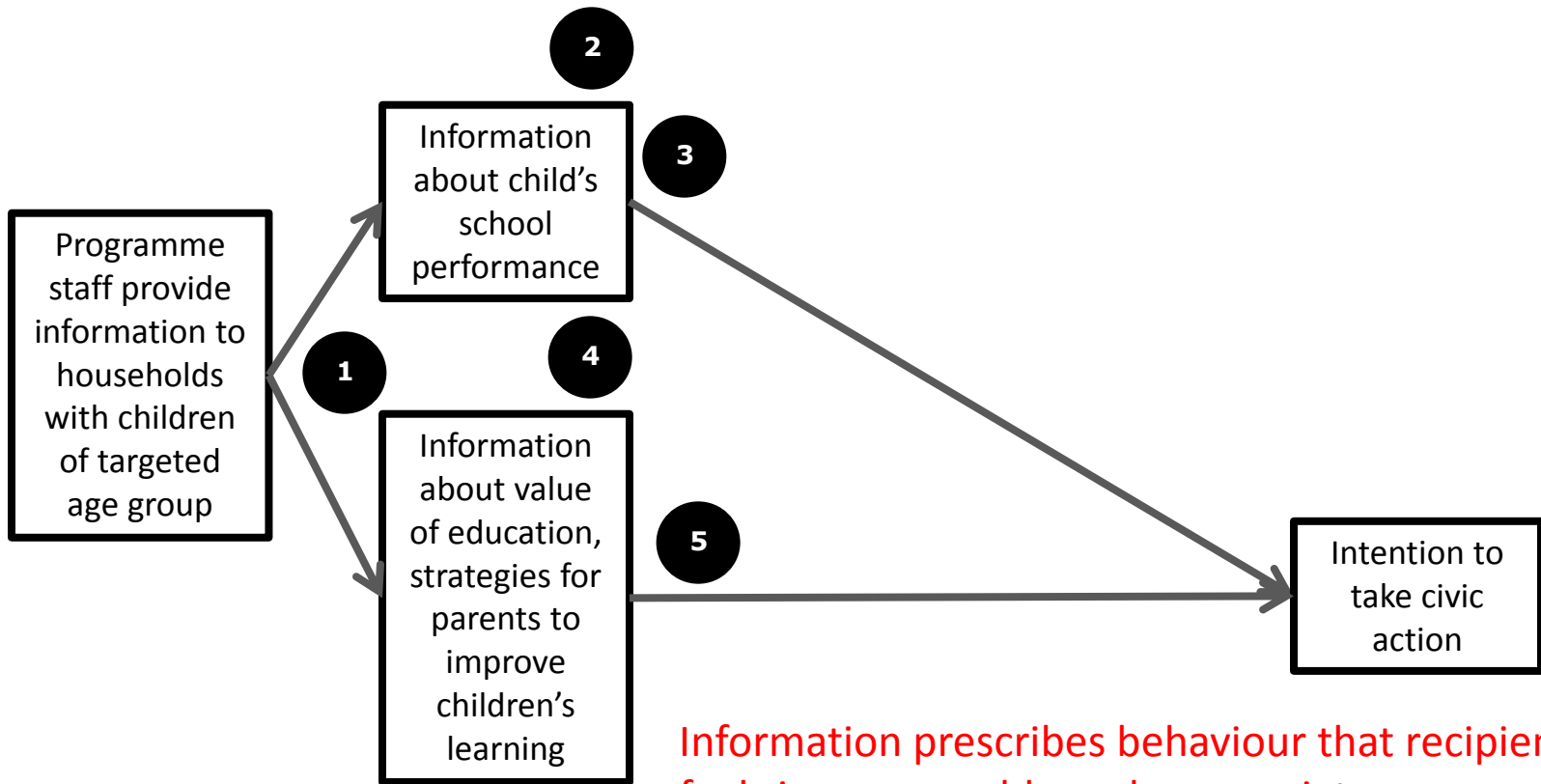
◇ Why? ◇ What's been said? ◇ What's next? ◇

Twaweza (Lieberman *et al*): modified



◇ Why? ◇ What's been said? ◇ What's next? ◇

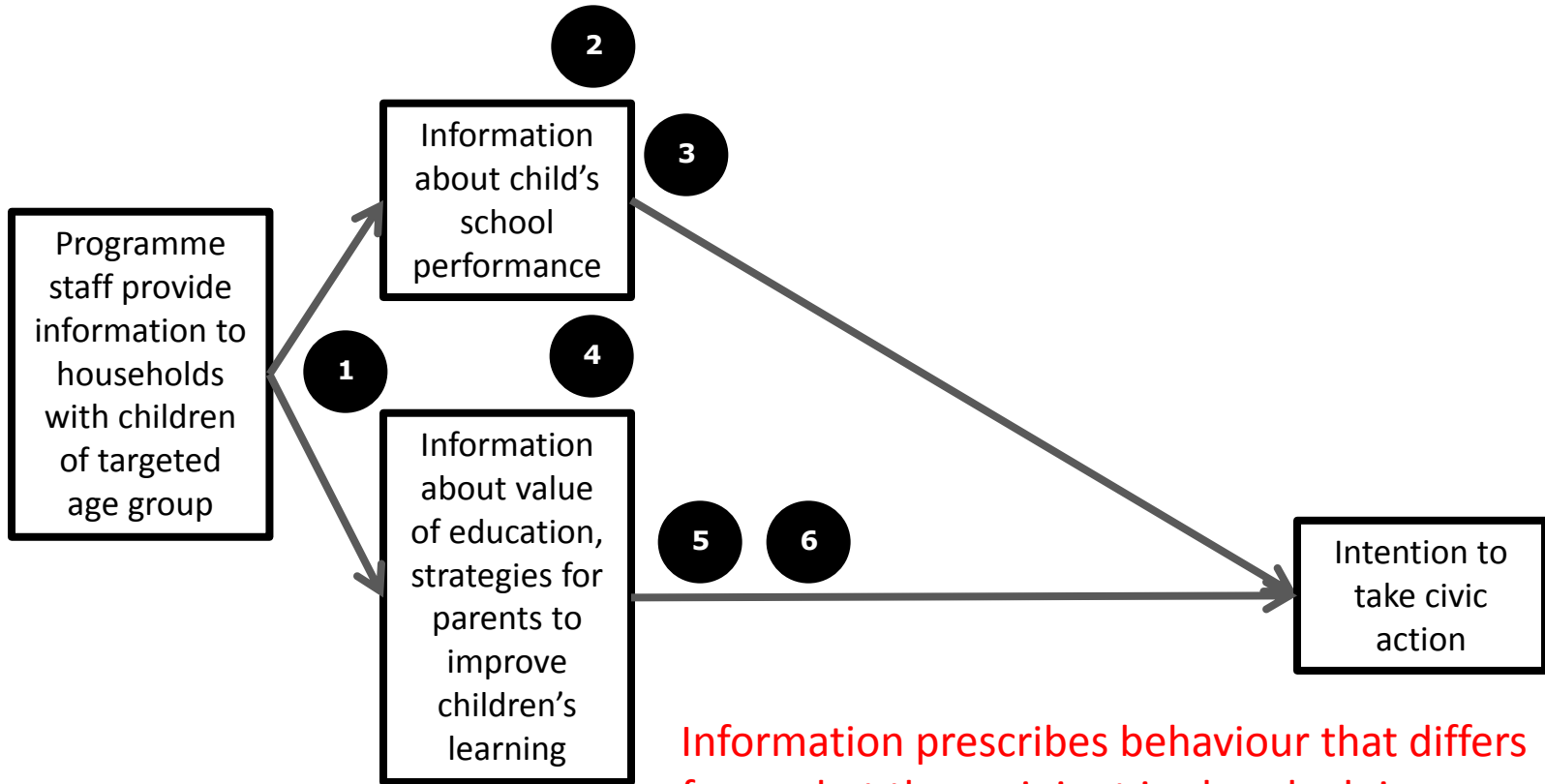
Twaweza (Lieberman *et al*): modified



Information prescribes behaviour that recipient feels is a reasonable and appropriate response to the identified problem.

◇ Why? ◇ What's been said? ◇ What's next? ◇

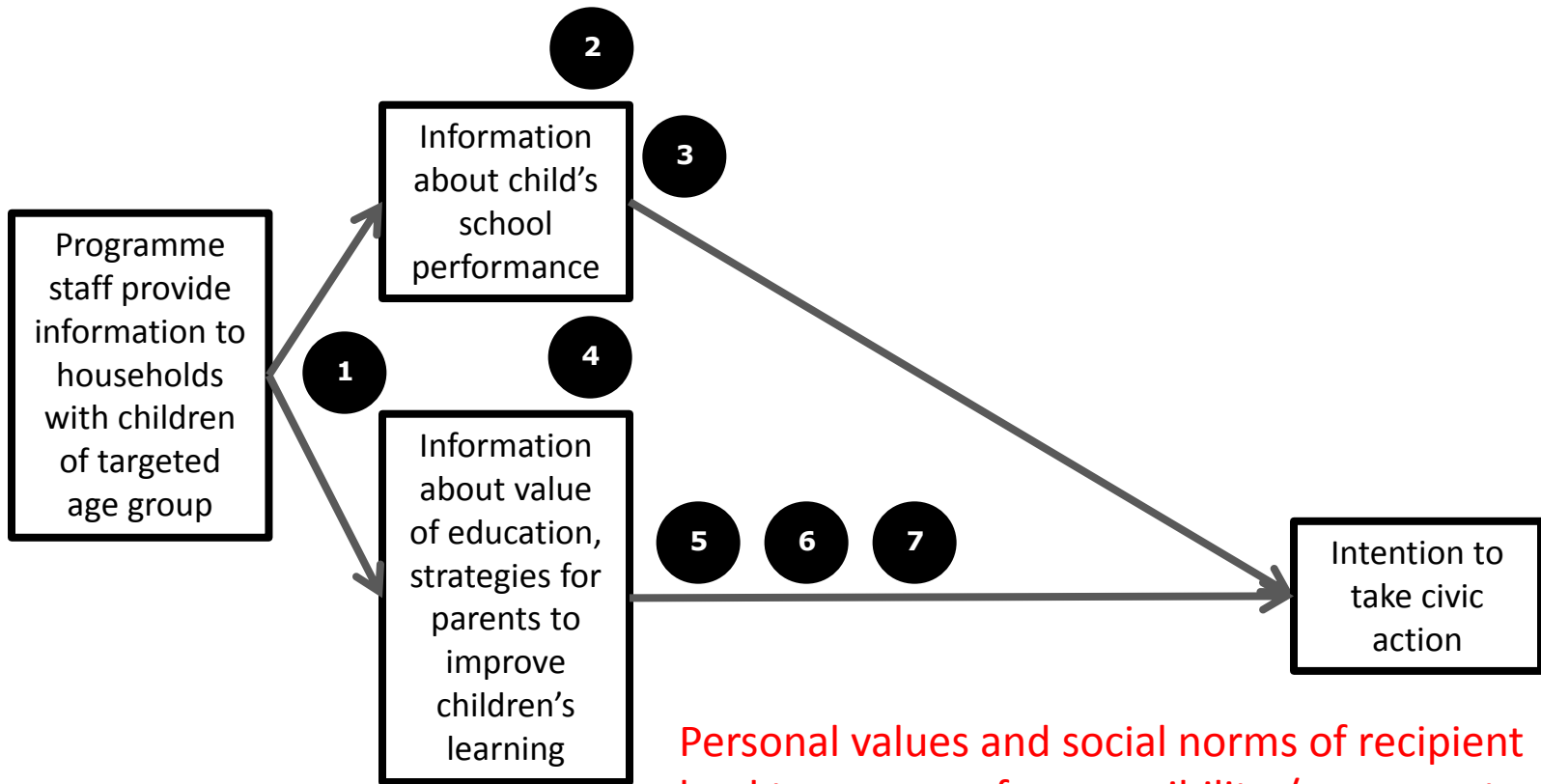
Twaweza (Lieberman *et al*): modified



Information prescribes behaviour that differs from what the recipient is already doing.

◇ Why? ◇ What's been said? ◇ What's next? ◇

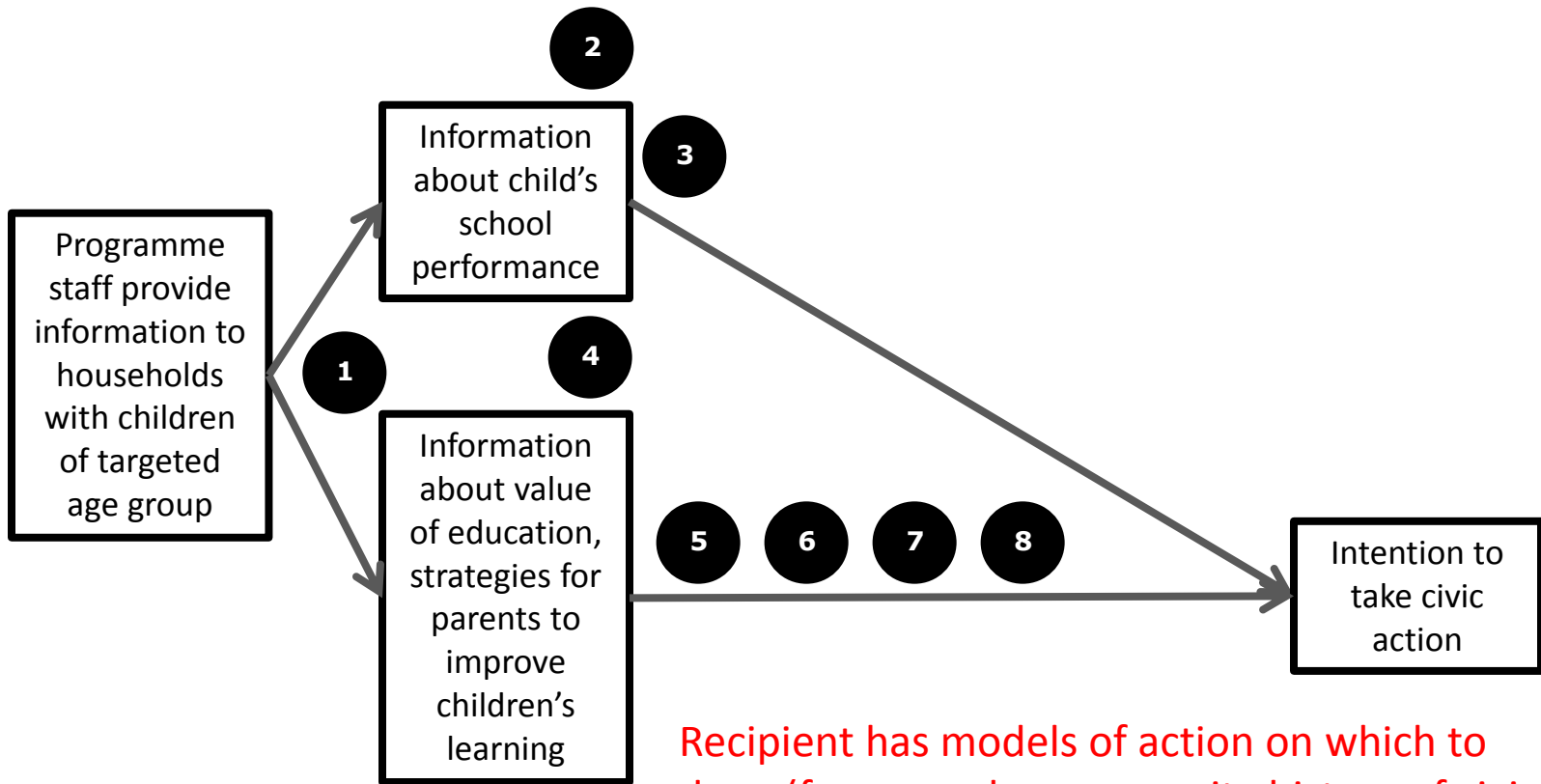
Twaweza (Lieberman *et al*): modified



Personal values and social norms of recipient lead to a sense of responsibility / pressure to do something about the problem.

◇ Why? ◇ What's been said? ◇ What's next? ◇

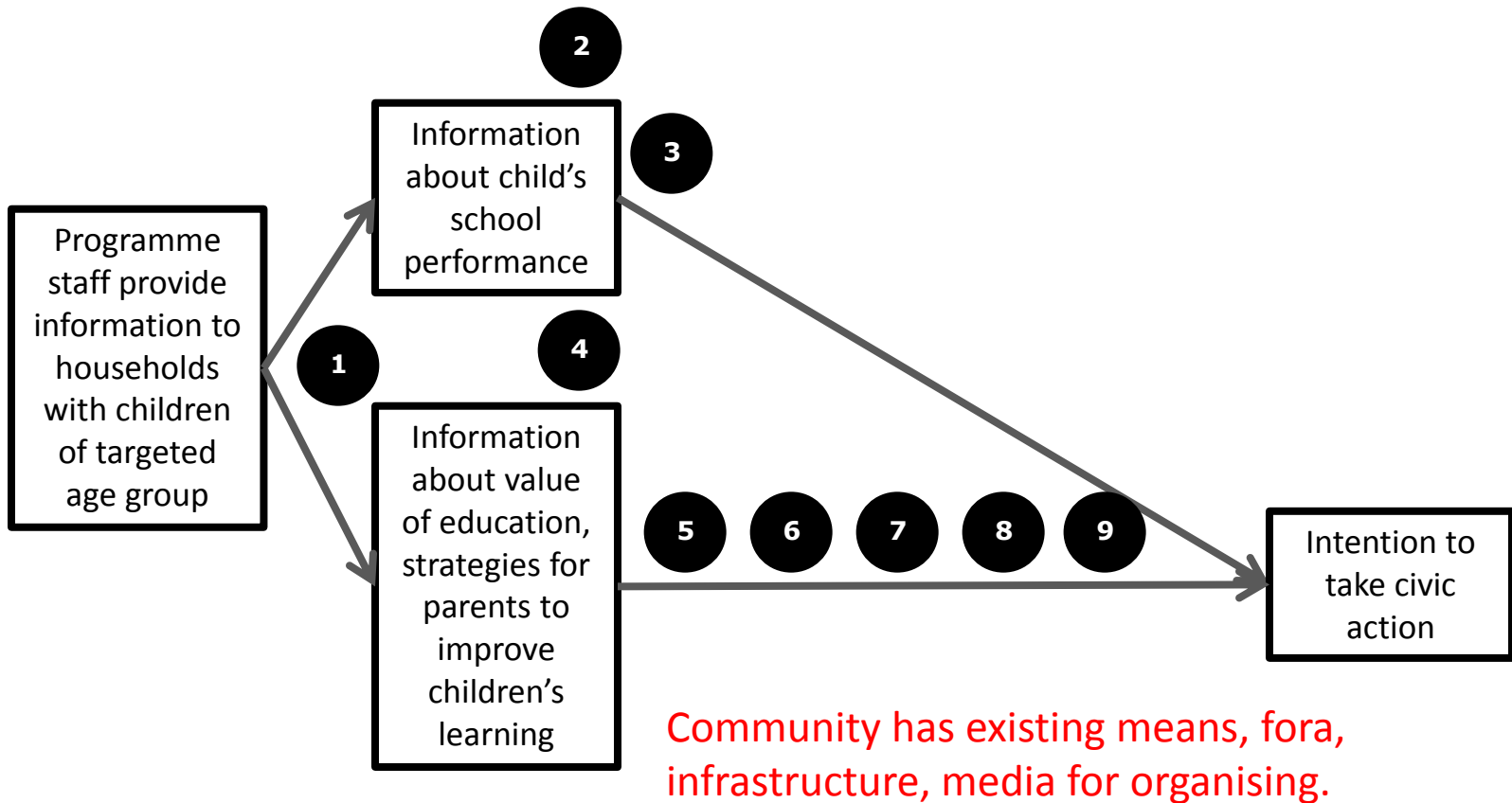
Twaweza (Lieberman *et al*): modified



Recipient has models of action on which to draw (for example, community history of civic (successful) civic action).

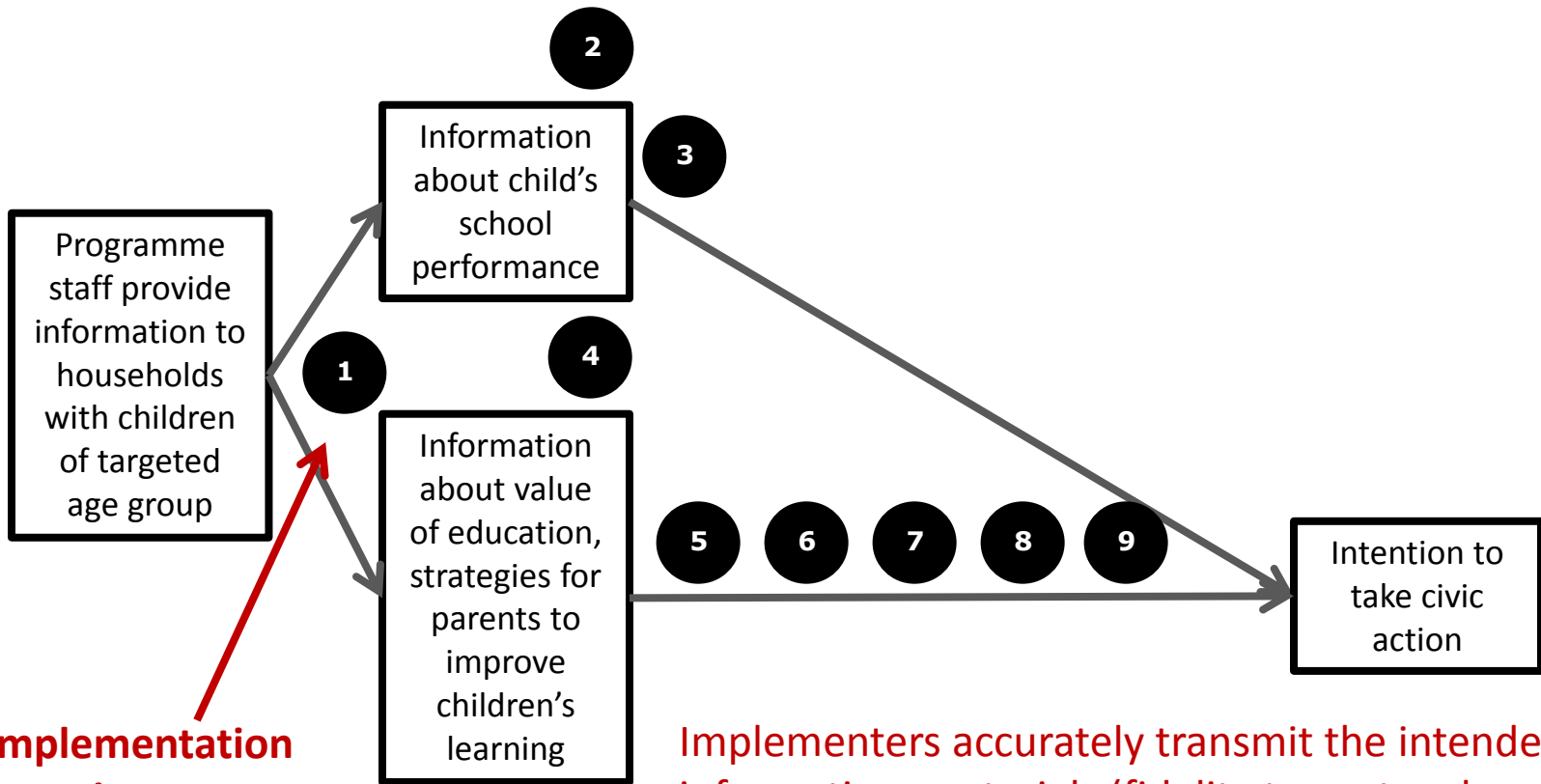
◇ Why? ◇ What's been said? ◇ What's next? ◇

Twaweza (Lieberman *et al*): modified



◇ Why? ◇ What's been said? ◇ What's next? ◇

Twaweza (Lieberman *et al*): modified

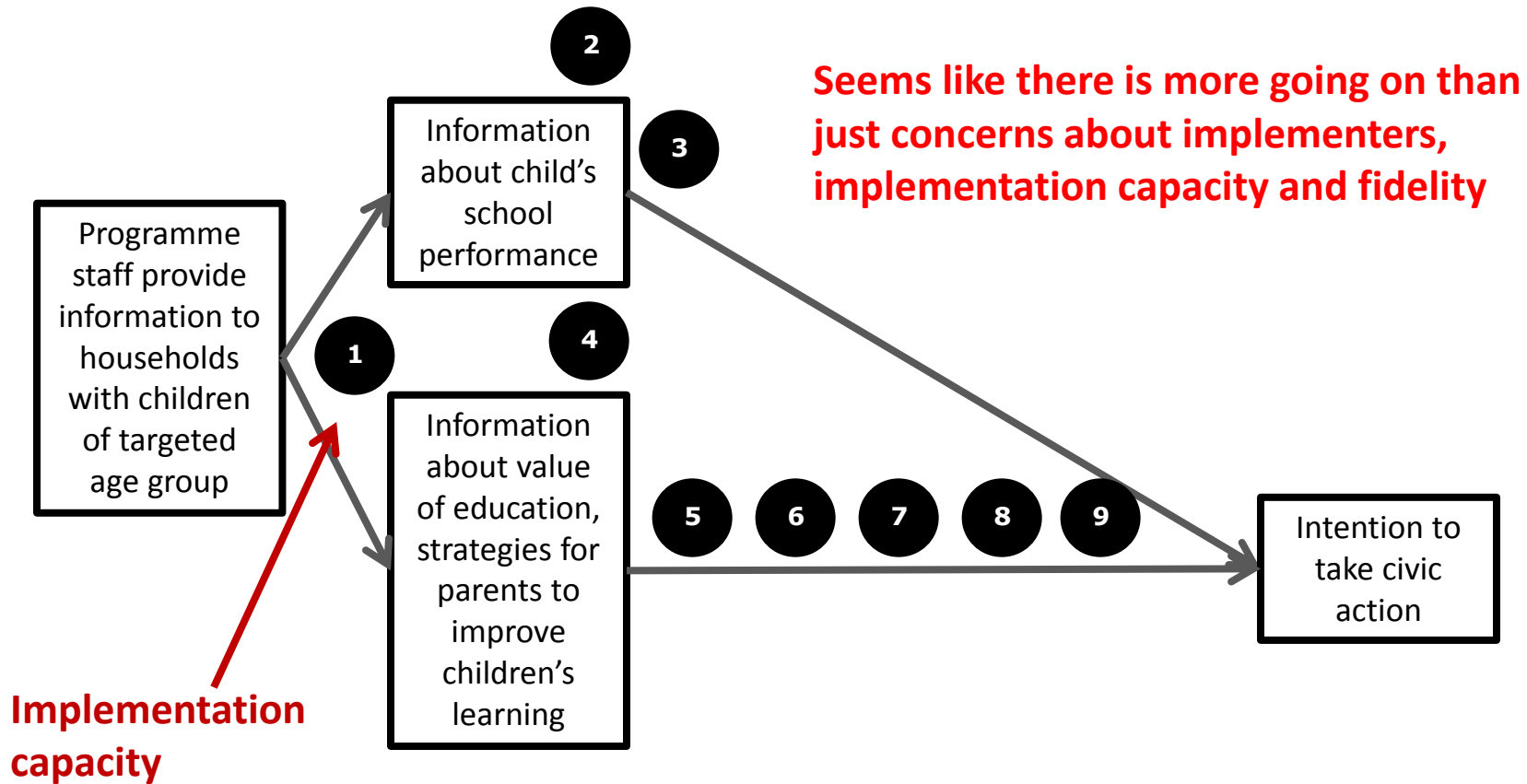


Implementation capacity

Implementers accurately transmit the intended information, materials (fidelity to protocol and/or message).

◇ Why? ◇ What's been said? ◇ What's next? ◇

Twaweza (Lieberman *et al*): modified



Wait, what's a ToC, external validity?

- Theory of change:
 - Both process and product
 - Combines program theory, logframe, & theories on social, behaviour change
 - Not full consensus on what it must include (e.g. Vogel, Valters)
 - Must make explicit assumptions about why inputs and activities will lead to short- and long-term outcomes in a given setting – i.e. hypotheses about how change happens
 - Should include a time component about how much change can be reasonably expected by when

◇ Why? ◇ What's been said? ◇ What's next? ◇

Wait, what's a ToC, external validity?

- External validity:
 - Technically, a property of a program
 - Does Program P (with x amount of input X) lead to the same magnitude change in outcome Y 'here' and 'there'
- Portability of findings (adapted from Cartwright):
 - More concerned about whether the mechanisms underlying Program P, which produced Y 'there,' can do the same 'here'
 - Do we understand enough about what was involved in producing Y 'there' to make informed decisions about whether it could work 'here'

◇ Why? ◇ What's been said? ◇ What's next? ◇

Punchline 1: onus for portability

- Woolcock (2013) says: “*Burden of proof should lie with those claiming that the result is, in fact, generalizable*”
- But, those making such claims require good understanding of the settings and processes that produced the evidence of interest
 - This puts some onus on researchers to systematically monitor and measure settings and processes – *and to report on it*

◇ Why? ◇ What’s been said? ◇ What’s next? ◇

Punchline 2: multi-site studies

- External replication (trying Program P ‘here’ and a multitude of ‘there’s) will achieve external validity
- Just replicating, without learning and adapting, is probably not sufficient for those trying to make decisions in yet-unstudied “there”s

Punchlines 3 & 4: theory rocks

- Theory is what guides generalization from the specific and grounded to the more abstract and portable
- “Pre-mortem” thinking and construction of ToC – which then guides research questions, monitoring plans, questionnaire items, etc – is requisite for producing evidence ‘here’ that those ‘there’ can reasonably use for decision-making

◇ Why? ◇ What’s been said? ◇ What’s next? ◇

Punchline 4: critical categories (speculative)

- Supply-side (adapted from Woolcock)
 - Implementation (bureaucratic, administrative) capacity
 - Scope for street-level improvisation/discretion in implementation and incentives to do this well / temptations to deviate

◇ Why? ◇ What's been said? ◇ What's next? ◇

Punchline 4: critical categories (speculative)

- Supply-side (adapted from Woolcock)
 - Implementation (bureaucratic, administrative) capacity
 - Scope for street-level improvisation/discretion in implementation and incentives to do this well / temptations to deviate
- Demand-side (adapted from Bongaarts, Lesthaeghe)
 - Readiness factors
 - Costs, opportunity costs, and economic advantage of addressing a problem with a new behavior
 - Willingness factors
 - Alignment with personal values, social norms with new behavior; belief in new behavior as appropriate and effective answer to the problem at hand
 - Ability factors
 - Awareness; community, intra-HH power hierarchies around decision-making; perceived access & efficacy

◇ Why? ◇ What's been said? ◇ What's next? ◇

Punchline 4: critical categories (speculative)

- Supply-side (adapted from Woolcock)
 - Implementation (bureaucratic, administrative) capacity
 - Scope for street-level improvisation/discretion in implementation and incentives to do this well / temptations to deviate
- Demand-side (adapted from Bongaarts, Lesthaeghe)
 - Readiness factors
 - Costs, opportunity costs, and economic advantage of addressing a problem with a new behavior
 - Willingness factors
 - Alignment with personal values, social norms with new behavior; belief in new behavior as appropriate and effective answer to the problem at hand
 - Ability factors
 - Awareness; community, intra-HH power hierarchies around decision-making; perceived access & efficacy
- Setting ‘landscape’
 - Topography, geography, climate, population density, demography
 - Built infrastructure
 - Institutional landscape – formal regulation, informal norms
 - Overlapping and nearby programs, policies (on-going, recent)

◇ Why? ◇ What’s been said? ◇ What’s next? ◇