

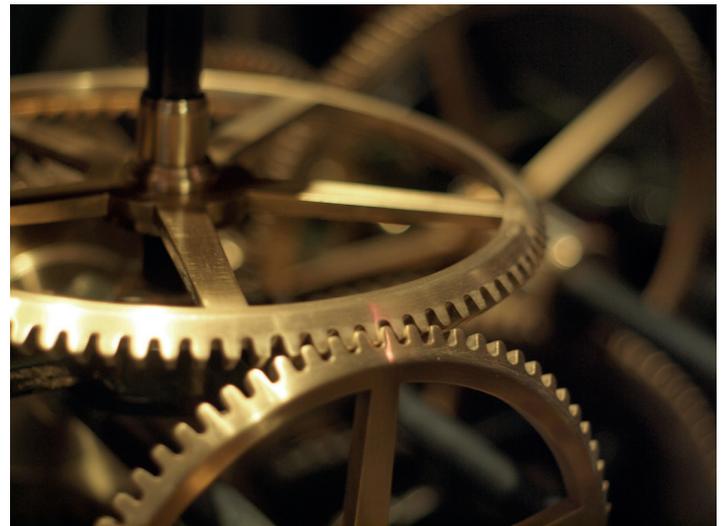
External validity and policy adaptation: a five-step guide to mechanism mapping

by Dr Martin J. Williams, Associate Professor at the Blavatnik School of Government

With the recent boom in impact evaluations around the world, policymakers in many sectors now have at their disposal an overwhelming amount of evidence about “what works” - or at least what worked in a particular context. Yet as impact evaluations have multiplied, it has become apparent that “the same” policy can have very different effects in different populations (Vivaldi 2016). Similarly, policies shown to be effective in small trials have not always been as effective when implemented at scale, even in the same country (Bold *et al* 2016).

This is the problem of the *external validity*. The limited external validity of impact evaluation evidence poses challenges for policymakers: how can one know if a policy will have the same effect in this implementation context as it did elsewhere? And to what extent should policymakers copy the design of policies that have worked elsewhere, rather than use local information to try to adapt them to fit the local context?¹

This memo proposes a simple and flexible framework for thinking about these questions, and about external validity more broadly. A policy can have a different impact in a new context than it had in a previous context if *part of a policy’s theory of change interacts with a difference in contexts*. A policy’s theory of change is a mapping of its intended mechanism spanning inputs to activities, outputs, intermediate outcomes, and final outcomes. Whether this mechanism works as intended depends at each step on the validity of a set of contextual assumptions. While these assumptions may have been true of the context in which a policy had previously been shown to work, whether the policy will have the same effects in a new context depends on whether these same contextual assumptions hold. Since context can include a wide range of factors - location, target group, implementing organization, scale, time period, the existence of related policy interventions, etc. - and the theory of change includes factors related to implementation as well as impact, this simple framework can be used to analyze a wide range of factors affecting policy impact.



Mechanism mapping in 5 steps:

1. Map out the policy’s intended theory of change (i.e. mechanism).
2. Below this, map the contextual assumptions that must hold for the theory of change to work.
3. Map the actual characteristics of the context and compare them to the assumptions.
4. Adapt the policy to address any mismatches between assumptions and characteristics.
5. Repeat steps 1-4 for the adapted policy, iterating in more detail until satisfied that all major policy design decisions fit the local context.

¹ For other useful perspectives on this subject, see Cartwright and Hardie (2014), Bates and Glennerster (2017), and Leviton (2017). In order to focus on issues of external validity and policy transportation arising from real differences in context, this paper abstracts from the issues of the statistical or methodological accuracy of published impact evaluations. While these issues can also lead to differences in estimated policy impacts across contexts, they have been discussed extensively elsewhere and are conceptually distinct.

Mechanism mapping

Comparing the policy's theory of change against its underlying contextual assumptions – *mechanism mapping* – focuses policymakers' attention on the validity of these contextual assumptions. If a necessary assumption does not hold in the new context in the same way it held in the old context, then the mechanism will be interrupted and the policy's final impact will differ. The mechanism mapping process can also be applied to questions of policy scale-up, since implementing a policy at scale involves different contextual assumptions (e.g. implementation quality, resource requirements, general equilibrium effects, political economy) than a small pilot, even if the pilot was undertaken in the same geographical location.

Mechanism mapping ideally consists of a systematic process of seeking empirical evidence to support contextual assumptions through descriptive statistics, qualitative data, and evidence from relevant impact evaluations. However, where time or resource constraints make this unfeasible, mechanism mapping can also be useful as a

quick and informal desk exercise undertaken by a single policymaker. This simple and intuitive diagnostic process gives policymakers a flexible framework for marshalling all available empirical evidence from different sources and of different levels of rigor in a structured way in support of policy decisions. Whereas the lack of data has often hindered evidence-based policymaking in data-poor contexts, mechanism mapping's ability to integrate less formal types of evidence makes it particularly well suited to such contexts.

The process of mechanism mapping also feeds directly into policy adaptation, by identifying specific aspects of the policy that are likely to work less well (or potentially better) than in the policy's original context. Policy adaptations thus flow directly from a diagnostic of the relationship between the policy context and the policy's theory of change, so that adaptations are based on a combination of local, context-specific information and evaluation evidence from other contexts.

The five-steps guide

How should an analyst actually go about creating a mechanism map? This section presents an example of the process and walks through the five steps of the process. I illustrate the approach using Cartwright and Hardie's (2014) example of the Bangladesh Integrated Nutrition Programme (BINP), a mainly World Bank-funded project in the 1990s.² The design of BINP was copied exactly from the World

Bank's highly successful Tamil Nadu Integrated Nutrition Programme (TINP) – a clear example of evidence-based policy – yet BINP had little impact on its key outcomes. The mechanism map below will make clear how the same policy could be so effective in Tamil Nadu but ineffective in Bangladesh.

STEP 1: Map out the policy's intended Theory of Change (i.e. mechanism)

The first step of mechanism mapping is to lay out a policy's theory of change, or mechanism. This can be thought of as a causal chain leading from a policy's initial inputs to its intended final outcomes, via activities, outputs, and intermediate outcomes. For clarity, this article explains mechanism mapping using a simple, linear theory of change, but the mechanism mapping process can be used with whatever style of theory of change the analyst prefers.

The intended final outcome of BINP was to improve mother and infant nutrition. To do so, government was to provide two main outputs: nutritional advice delivered to pregnant and nursing mothers, and the distribution of supplementary food to mothers to take home. These outputs would lead to

the final outcome via two sets of intermediate outcomes: first, mothers' nutritional awareness would improve, alongside their receipt of the supplemental food; and second, mothers would then decide to use the supplemental food for themselves and their infants (as opposed to giving it to other family members, i.e. program "leakage"). In order to produce these outputs, the government required inputs of adequate financial resources to purchase the food and pay personnel, as well as a logistical system and potential pool of extension workers to deliver the food and nutritional advice. Key activities for transforming inputs into outputs could include procuring the food, hiring and training workers, and conducting outreach to eligible mothers.

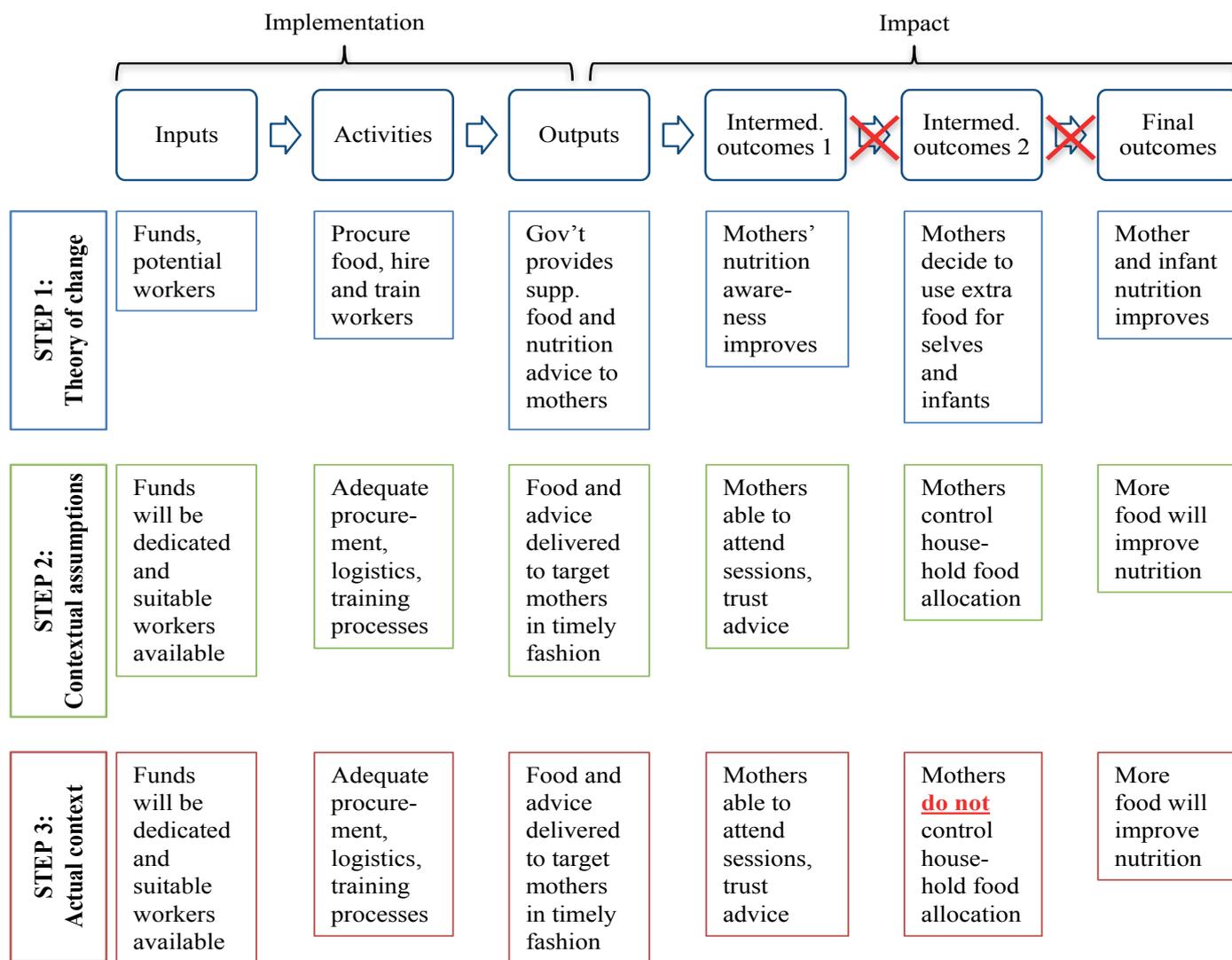
STEP 2: Map the contextual assumptions that must hold for the Theory of Change to work

The contextual assumptions required for this theory of change to work are listed in the second row of Figure 1. Sound implementation requires that government: dedicate adequate financial and human resources to the project;

procure and distribute food and hire workers effectively, including quality assurance as well as prevention of excessive corruption, and train workers adequately; and deliver these outputs to a pool of eligible mothers predictably and in a

² This example is based largely on Cartwright and Hardie's (2014, 80-84) excellent exposition, and draws also on Save the Children (2003), White (2005), World Bank (2005a), and World Bank (2005b). See Williams (2017) for more details. This memo's discussion of TINP and BINP, and their contexts, is of course simplified for clarity and brevity.

Figure 1: Mechanism Map for Bangladesh Integrated Nutrition Programme (BINP)



timely fashion. Impact then requires that mothers are able to attend the sessions and trust the advice they are being given; that mothers actually control household food allocation; and that the supplementary food, if consumed, will actually lead to the desired improvement in nutrition. In the Tamil Nadu context, these assumptions were presumably valid - hence the impact evaluation finding that TINP significantly improved mother and infant nutrition (World Bank 2005b).

One challenge is identifying which are the most salient contextual assumptions and characteristics to consider, since the high dimensionality of context makes it unfeasible to consider all aspects of context. Although this is ultimately a matter of judgment, two guidelines suggest themselves.

First, many dimensions of context are frequently salient and should be taken into consideration for almost any policy: demographic and socioeconomic characteristics of the target population; resource availability; political support and resistance; social and cultural norms; the effectiveness of implementing organizations; potential for corruption or resource diversion; geographic accessibility and other logistical issues; and so on. Second, important contextual factors specific to a particular policy are often suggested by the policy's theory of change. For instance, laying out BINP's theory of change makes it clear that decisionmaking over household food allocation is a key contextual assumption.

STEP 3: Map the actual characteristics of the context and compare them to the assumptions

The third row of Figure 1 contrasts these contextual assumptions to the actual contextual characteristics of the new context, in this case rural Bangladesh. The key contextual assumption that did not hold in Bangladesh was that mothers controlled household food allocation and would thus be able to act on their improved nutritional awareness:

whereas mothers were typically responsible for shopping and household food allocation decisions in rural Tamil Nadu, in rural Bangladesh men usually conducted the shopping and their mothers (the mothers-in-law of the pregnant or nursing women) controlled household food allocations (White 2005; Cartwright and Hardie 2014).

This broke the link between Intermediate Outcome 1 and Intermediate Outcome 2: while BINP succeeded in distributing food and nutritional advice to the mothers, and mothers' nutritional awareness did actually improve as a result, the program failed to improve mother and infant nutrition because most of the supplementary food went to other family members. Since Intermediate Outcome 2 was not achieved, neither was the Final Outcome. If the designers of BINP had carried out a mechanism mapping when transporting the successful TINP program to Bangladesh, perhaps they would have uncovered this crucial but implicit assumption.

Mechanism mapping can also be adapted to policies that are intended to lead to multiple final outcomes (e.g. a cash transfer that is intended to increase consumption and improve child school attendance) simply by creating multiple mechanism maps, one for each outcome. The theory of change may be the same for each outcome or may differ slightly in emphasizing the aspects of the mechanism that are more salient, but the key contextual assumptions and characteristics are likely to be different. The same procedure can also be used to assess the likelihood of negative outcomes or side effects of the policy, by placing these undesirable outcomes as the final outcome of the policy and assessing whether the policy mechanism and contextual characteristics and assumptions are likely to lead to them.

STEP 4: Adapt the policy to eliminate these mismatches between assumptions and characteristics

Since mechanism mapping as a diagnostic tool focuses on the interaction between a policy's theory of change and differences in context, the diagnosis of whether a policy is likely to be as effective in a new context as it was elsewhere inherently involves highlighting the aspects of the policy that should be targeted for adaptation. In the case of BINP, for example, Figure 1 makes it obvious that the key aspect where adaptation was necessary was the nutritional advice component, and specifically the individuals to whom this was targeted.

An obvious way to adapt the policy would be to extend the nutritional advice component to include the key decision makers about household food allocations besides mothers – their husbands and their mothers-in-law. This may also require changes to other parts of the program, since these new target populations may have to be reached in different ways, for instance through home visits. This in turn may imply other changes to the policy's theory of change, both in terms of effective implementation (greater resources needed, additional logistical and transportation issues) and also for these advice sessions to have the intended impact (e.g. home visits may raise different cultural or trust issues).

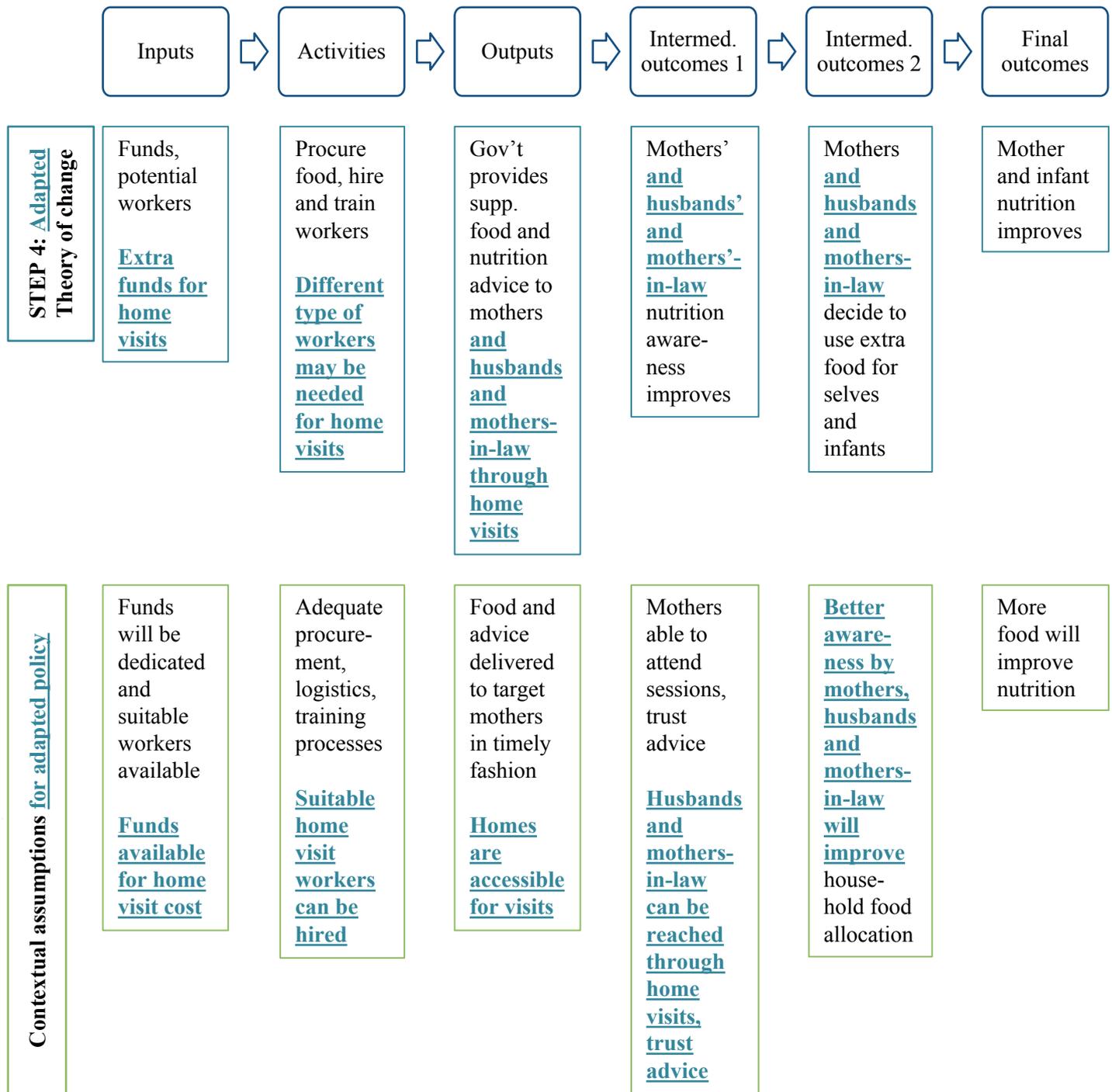
Empirical evidence has an important role to play in mechanism mapping. Most obviously, the contextual characteristics in the crucial bottom row are questions to which empirical answers - or at least suggestive evidence - may well exist. A mechanism mapper could, for example, examine budget data and political context to shed light on resource availability, investigate the performance of the implementing agency's procurement processes, conduct a survey of eligible mothers' trust of the state and baseline level of nutritional knowledge, undertake (or read existing) qualitative research on household food allocation decisions in rural Bangladesh, and discuss with public health experts the prevalence of diseases that might inhibit infants from absorbing nutrients properly present several excellent examples of using simple descriptive data to validate contextual assumptions. Bates and Glennerster (2017) gives some excellent examples of the use of evidence to identify contextual differences prior to transporting a program.

When the mechanism mapping is being conducted for a scale-up of a policy that has already been trialed on a small scale in the same location, the mechanism mapper may even have quite detailed evidence on these issues, and so the search for new empirical evidence can focus on the aspects of context that are changing with the larger-scale implementation: the effectiveness of the implementing agency, general equilibrium or spillover effects, political economy issues, etc.

Of course, many plausible adaptations could be made - the World Bank's own (2005a) review of BINP proposed eleven dramatically different "suggestions/options" for changing the program, from targeted social marketing strategies to discourage food sharing within households to giving women and communities incentives to meet pregnancy weight gain targets. As the following step discusses, mechanism mapping can be used to compare the feasibility and likely effectiveness of these options in a systematic fashion.

The other benefit of using mechanism mapping to suggest adaptations is that it makes it clear what aspects of the policy do not need to be adapted. For instance, Figure 1 makes clear that the other steps in BINP's theory of change fit well with the contextual assumptions and previous context in which the program had been evaluated, suggesting that there is little need for adaptation in these respects (except as necessitated by the adaptations in Figure 2). The design of the resulting adapted policy is thus informed both by evaluation evidence from other contexts - through the aspects of the original policy that were maintained in the new context - as well as by local, context-specific knowledge - through the aspects that were adapted.

Figure 2: Adapted Theory of Change and Assumptions for BINP



STEP 5: Repeat steps 1-4 for the adapted policy, iterating in more detail until satisfied that all major policy design decisions fit the local context

While the adaptation process can help mitigate the contextual differences – and possibly even improve on the original program – these adaptations also require new contextual assumptions to be effective, and these contextual assumptions must also be evaluated. Having proposed adaptations, the analyst should therefore repeat the mechanism mapping process in steps 1-4 for the adapted policy – or, in the case of multiple alternative adaptations, for the multiple versions of it. Although this policy memo does not illustrate this iteration process for the sake of brevity, iterating in this way is crucial to help the analyst to assess the feasibility and likely impact of each adaptation.

Where the contextual assumptions and characteristics do appear to match, the iteration process can be used to drill down into more detail. For instance, the recruitment and training of the field workers comprises numerous steps and assumptions, relating perhaps to the advertising of the positions, integrity of the hiring process, and quality of the available human resource pool. Thinking in detail about each of these could surface unexpected challenges and/or opportunities for improvement. In particular, the more precise that the analyst can be in diagnosing mismatches between the policy’s contextual assumptions and characteristics, the more precise she can be about exactly what adaptations are (or are not) needed.

Conclusion

Mechanism mapping is a simple and flexible tool to help policymakers identify external validity failures and design adaptations to address them. While evidence-based policymaking might use a successful impact evaluation from another context or a systematic review as a *starting point* for policy design, mechanism mapping can help policymakers make the adaptations necessary for the policy fit in their specific context.

The process of mechanism mapping is flexible enough that the same basic process can be undertaken either by an entire agency in systematic discussion with its stakeholders or by one analyst sitting at her desk for a short period, or any combination in between. Mechanism mapping seems especially well suited to participatory and collaborative policy design processes, since it is capable of integrating evidence of all varieties and the clients or beneficiaries of a policy may well be more aware of salient contextual assumptions or characteristics than policymakers. This contrasts with the top-down and largely technocratic approach to using quantitative evidence alone to inform policy choices.

Mechanism mapping complements another recent innovation: adaptive policymaking, which views policy design and evaluation as an iterative process of experimentation with tight feedback loops. Mechanism mapping could help connect experimentation to a more precise diagnosis of the barriers to the effectiveness of previous iterations of a policy, thus adding precision to the experimental search process. Similarly, since mechanism mapping lines up closely with the policy's theory of change, and monitoring and data collection strategies are typically based a policy's theory of change or logframe, the data that organizations generate during adaptive policymaking processes often closely align with the evidence required to make mechanism mapping more empirically rigorous.

While mechanism mapping is intended primarily as a tool for policymakers to use prospectively to predict the impact of a new policy, mechanism mapping is also of potential value to evaluators in two ways. First, it can be useful in the retrospective evaluation of policies by helping evaluators to show clearly the intended and actual mechanism(s) through which a policy had its impact (or non-impact). Second, prospective mechanism mapping can also help evaluators design trials to ensure that they collect the data necessary to assess each of the contextual assumptions *ex post*, along with potential undesirable outcomes and the alternative mechanisms that might bring them about.

Finally, mechanism mapping should be understood as a tool to help policymakers structure their judgment about policy transportation and adaptation, not a scientific procedure for determining whether or not a policy will work. It relies on policymakers' judgment in the identification of salient contextual assumptions, in deciding which potential interactions are significant enough to warrant adaptation, and in designing appropriate adaptations. By making causal links and contextual assumptions explicit, aim of mechanism mapping is to structure and improve policymakers' own judgment in the pursuit of better use of evidence in policymaking.

For a more in-depth discussion of the conceptual foundations of the approach and its relations to other approaches to external validity, please see Dr Williams's working paper <https://www.bsg.ox.ac.uk/research/working-paper-series/external-validity-and-policy-adaptation>

About the author

[Dr Martin Williams](#) is an Associate Professor in Public Management at the Blavatnik School of Government and Research Fellow at Green Templeton College. His research explores management, policy implementation, and political economy, mostly focused on Africa. His work involves close collaboration with governments to answer questions that are of interest to policymakers and academics alike, such as why some ministries work better than others, why many development projects never get finished, and what can be done to improve performance.

References

- Bates, Mary Ann, and Rachel Glennerster. 2017. "The Generalizability Puzzle." *Stanford Social Innovation Review*, Summer.
- Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, Alice Ng'ang'a, and Justin Sandefur. 2016. "Experimental Evidence on Scaling Up Education Reforms in Kenya." *Mimeo*, November.
- Cartwright, Nancy, and Jeremy Hardie. 2014. *Evidence-Based Policy: A Practical Guide to Doing It Better*. Oxford: Oxford University Press.
- Leviton, Laura C. 2017. "Generalizing about Public Health Interventions: A Mixed-Methods Approach to External Validity." *Annual Review of Public Health* 38: 371-91.
- Save the Children. 2003. "Thin on the Ground: Questioning the Evidence Behind World Bank-funded Community Nutrition Projects in Bangladesh, Ethiopia, and Uganda." *Policy Report*.
- Vivalt, Eva. 2016. "How Much Can We Generalize from Impact Evaluations?" *Mimeo*, May.
- White, Howard. 2005. "Comment on Contributions Regarding the Impact of the Bangladesh Integrated Nutrition Project." *Health Policy and Planning* 20(6): 408-11.
- Williams, Martin. 2017. "External Validity and Policy Adaptation: From Impact Evaluation to Policy Design." BSG Working Paper 2017-019, July. Available at: <http://www.bsg.ox.ac.uk/research/working-paper-series/external-validity-and-policy-adaptation>
- World Bank. 2005a. "The Bangladesh Integrated Nutrition Project: Effectiveness and Lessons." *Bangladesh Development Series Paper No. 8*, December.
- World Bank. 2005b. "Project Performance Assessment Report: Bangladesh Integrated Nutrition Project." *Report No. 32563*, 13 June.