

Methodology Deep Dive

Applying Process Tracing to Policy Advocacy Evaluation

JANUARY 31, 2024

EVALUATION OVERVIEW

In 2023, the Robert Wood Johnson Foundation (RWJF) worked with ORS Impact (ORS) to conduct an evaluation to understand progress during the 117th Congress on three federal policies: the Advance Premium Tax Credit (APTC), the Child Tax Credit (CTC), and closing the Medicaid coverage gap. The evaluation seeks to understand how change happened, answer cause-and-effect questions, and better understand contributions to change. This evaluation is summarized in four briefs: three individual policy briefs and this methods brief.

This evaluation looks at the advocacy ecosystem as a whole during the 117th Congress. This means that the activities detailed in these briefs include not only activities undertaken and/or supported by RWJF, but also activities conducted by other individuals and organizations in the larger advocacy ecosystem and/or by RWJF grantees with other sources of funding. While some of the grantees that RWJF supports might engage in lobbying through their broader strategies, the Foundation does not earmark any funds for lobbying. RWJF is a 501(c)(3) private foundation that does not lobby or support or oppose candidates for public office or political parties; political party information is provided in these briefs only for context.

OVERVIEW OF THIS BRIEF

This brief summarizes what we did and what we learned about: (1) collecting and analyzing data, (2) developing and testing hypotheses, and (3) reporting to conduct process tracing on the three policies. The primary audience for this brief is the evaluation field, particularly evaluators considering applying process tracing to a policy advocacy evaluation. Our hope is that sharing our approach and lessons learned can make it easier for others to implement this method with high quality for other advocacy and policy change evaluations.



METHODOLOGY

The evaluation took an explicitly causal approach to understanding progress for policy change. Too often policy change evaluations only provide descriptive stories about which changes occurred without helping to uncover how change happened. Rather than describing which changes occurred or assessing a hypothesized theory of change, this evaluation focuses instead on understanding *how* and *why* progress was made (or not). To do this, the evaluation used a method called *process tracing* to understand how a particular large-scale change actually happens within a complex, dynamic context. Importantly, the story of how the change happened includes both steps taken intentionally to achieve the change (i.e., activities and strategies) and other contributing events, forces, or factors in the larger context.

The data for this evaluation came from **51 interviews and two participatory workshops**, including not only RWJF staff (n=9) and grantees (n=19), but also external partners (n=14) and administration/Hill staff (n=9).¹ Because process tracing seeks to understand the causal linkages of an established outcome, we **focused on the 117th Congress** to establish the evaluation boundaries for exploration.

About Process Tracing

A January 2022 [article in *The Foundation Review*](#) called for more use of causal analysis in philanthropic evaluation, arguing that “*philanthropy more frequently needs to examine causal relationships, using a growing suite of methodological approaches that make this possible in complex systems. Causal methodologies can challenge and strengthen the often uncontested beliefs that underlie philanthropic interventions, while offering evidence about enabling contexts and system drivers. Strong causal analysis considers not only the funder’s model and assumptions, but also the beliefs others hold about how and why change occurs, opening the door to more equitable and less biased ways of understanding change.*”

Building on this momentum around and interest in causal evaluation, ORS used process tracing in this work to encourage a participatory and iterative approach to telling the causal story. Process tracing explores competing hypotheses about plausible explanations of the causes of a given outcome (in this case, inclusion or exclusion from federal policy vehicles). Using rigorous analysis, process tracing **assesses the degree of contribution that can be connected to each hypothesis or cause**. In addition to evaluating activities and outcomes, ORS undertook *process induction* to identify salient, plausible explanations for the outcomes, and used *process verification* to assess the extent to which each explanation is supported or not supported by the available evidence.

For more information about how to do process tracing, see [this](#) process tracing guide by Jewlya Lynn, Sarah Stachowiak, and Jenn Beyers.

¹ RWJF staff are people who work at the Foundation; these interviewees were not confined to those within the Policy Department. RWJF grantees are organizations that have received a grant from RWJF’s Policy Department since 2021, although not necessarily around this work. Partners are people and organizations who worked closely with the Foundation to lead and design strategy, but did not receive grants from RWJF’s Policy Department. Administration/Hill staff are people who worked on the Hill or in the Biden administration during the 117th Congress.



KEY FINDINGS

Notably, the three policies had different results in the 117th Congress. CTC and APTC were both expanded in the American Rescue Plan Act (ARPA), but only APTC’s expansion was extended in the Inflation Reduction Act (IRA) passed the following year. CTC’s expansion was not extended after the ARPA provisions expired. Neither the ARPA nor IRA included any provisions to address the Medicaid coverage gap. Currently, the Medicaid coverage gap remains unaddressed at the federal level.

Table 1 summarizes ORS’ key learnings and challenges from implementing process tracing in this policy advocacy evaluation. The following sections provide more detail about what we did and what we learned at each phase of the work.

Table 1 | Summary of RWJF’s Non-Grantmaking Activities During the 117th Congress²

LEARNINGS	<ul style="list-style-type: none"> • Process tracing was a good methodology for the two policies that had clear policy outcomes: one win and one that made it far along but ultimately fell out during final negotiations. We were not able to do process tracing for one policy area where there was no clear consensus on what was achieved across our range of participants. • Creating visual causal stories from a wide range of perspectives that were vetted in a participatory workshop was a good way to clarify and understand what contributed to important steps in the process. • Iterative rounds of data collection (i.e., document review, interviews, participatory workshops, follow-on interviews) helped us develop confidence in the causal stories and conduct the process tracing tests. • Hypotheses were often multifaceted, including both advocacy efforts and contextual factors contributing to other outcomes. • Using the discipline of the process tracing method was helpful in making evaluative judgements, but sharing findings required translation and finesse to make it more accessible to our audience.
------------------	---

² Note that this evaluation only considered federal-level solutions for closing the Medicaid coverage gap; it did not include individual state-level solutions.



CHALLENGES

- While we got a range of perspectives from those working in these advocacy areas and some of the advocacy targets, it was difficult to access individuals and organizations that do not support the policies. The strength of the causal stories also relied on being able to include the perspectives of a range of people with deep and different knowledge.
- There is an opportunity to consider different ways of structuring interviews to gather data that more easily lend themselves to creating a causal story. ORS is considering leaning more into John Kingdon's policy window theory in future protocol development, for example. By creating a bespoke causal story for each policy, we were true to the data we heard, but it was also harder to learn across the policies.
- Interviewees often could not identify a clear alternate explanation for the outcome. If anything, they were more likely to pose what-if scenarios (e.g., if they had done x, y would have happened), which we cannot compile data and evidence on. In the absence of a strong alternate hypothesis, we often found ourselves using a null hypothesis.
- The actual work of implementing process tracing tests is iterative and esoteric. It helped to think about the causal strength (i.e., plausible, can't be disproved, contributing but with other strong contributing factors, clearly the contribution) as a way of working through the tests.

1. DATA: WHAT WE DID AND WHAT WE LEARNED

What We Did

As stated earlier, **the primary data for process tracing came from 51 interviews and two participatory workshops.** Other data sources include strategic materials from the Foundation and articles published during/about the 117th Congress.

After collecting, coding, and synthesizing the interview and secondary data, ORS developed causal stories for each policy area. These causal stories are visual depictions of the high-level stories the ORS team synthesized for each policy, including how both intentional strategies and contextual factors contributed to the ultimate policy results around CTC, APTC, and closing the Medicaid coverage gap during the 117th Congress.



ORS then vetted and refined the causal stories in participatory workshops with RWJF staff and grantees. The workshop participants came to consensus for each causal story, considered alternative explanations for what happened, and engaged in an activity to help establish the weight and influence of different contextual factors, activities, and outcomes. They also identified missing voices from the conversation, with whom ORS followed up to conduct additional interviews before finalizing the dataset and causal stories.

What We Learned

Different policy outcomes lead to different data. For the policies that weren't included in the IRA (CTC and closing the Medicaid coverage gap), it was difficult for interviewees to share data about what worked. Interviewees often minimized the positive impacts of advocacy activities, other strategies, and/or contextual factors because the policy didn't pass in the end. They were often more focused and reflective on what didn't work, missteps along the way, what-ifs, and new/refined ideas for future opportunities. This led to data that felt quite different for each policy. Especially for the policies that weren't included in the IRA, we had relatively limited data that connected activities and outcomes. This wasn't necessarily because those connections weren't there; rather, interviewees often did not explicitly make those connections.

Since process tracing relies on identifying and gathering sufficient data about those connections, we specifically probed on those connections in the participatory workshops and later-stage interviews. For example, we asked, "What did [X activity] make possible?" or "How did [Y strategy] impact the work that came after it?"

There is an opportunity to consider different ways of structuring interviews. ORS is considering leaning more into John Kingdon's policy window theory in future policy advocacy causal evaluations, especially in data collection.³ Kingdon posits that policy change happens when multiple things come together: an issue seen as a policy problem, a solution, and a policy window. With hindsight, ORS wonders whether structuring our interview protocol around these three components might have led to data that more easily lent itself to creating a causal story. Our protocol was more linear—"What contextual factors were at play during the 117th Congress? What advocacy activities happened? What outcomes did they lead to?"—and additional questions like "To what extent did a policy window exist around CTC in the 117th Congress? To what do you attribute that?" might offer slightly different data.

By creating a bespoke causal story for each policy, we were true to the data we heard, but it was also harder to learn across the policies. Others have used mid-range theories as an organizing framework, and that might be useful to consider in future evaluations as well.⁴

³ Kingdon, J. W. (1995). *Agendas, alternatives and public policy*. Addison-Wesley.

⁴ For example, see: Douthwaite, B., Proietti, C., Polar, V., & Thiele, G. (2022). Using theory to understand how policy change happens: Insights from Agricultural Research for Development. *Research Evaluation*, 32(2), 200–212. <https://doi.org/10.1093/reseval/rvac038>



Work benefits from many voices and iterations. Our approach of doing many interviews, facilitating a group process in the participatory workshops, and conducting additional interviews to fill gaps in perspectives or the overall story worked well. The iterative approach allowed us to gut check early thinking about the causal story, identify missing voices, and probe on missing or incomplete parts of each policy's story. For example, since each interviewee spoke about different aspects of a policy's story, the data was widespread, and it was not necessarily clear how all the pieces came together. After ORS took a first pass at bringing the full story together, having RWJF staff and grantees who serve in more coordinating roles take a step back and help us clarify the full story—especially the sequencing and the relationships between different components—was very helpful.

The participatory workshops also benefited from many voices, especially grantee partners who are close to the work. Since RWJF staff who participated are often more distant from the day-to-day research and advocacy work, they often had less to contribute to the specifics of the causal stories than grantee participants. It was also helpful when grantee participants could be in conversation with one another to remember specifics of the work, add nuance to each other's perspectives, and/or offer alternative explanations or interpretations from one another.

It was very difficult, however, to schedule interviews with individuals and organizations who do not support the policies. Data about the opposition to each policy often came from RWJF staff and grantees who have a perspective on the full field.

2. HYPOTHESES & TESTS: WHAT WE DID AND WHAT WE LEARNED

What We Did

Following the finalization of the causal story for each policy, ORS analyzed our data for relationships between the specific activities and strategies (including, but not limited to, those supported by RWJF and its awardees) and contextual factors that workshop participants weighted as most important. Using these relationships, **we developed hypotheses describing how activities or contextual factors led to outcomes** and how earlier changes led to later changes. **For CTC and APTC, ORS then considered these relationships using process tracing.** This means that ORS assessed the prioritized hypotheses according to the strength of data plus the uniqueness of the contribution.⁵

⁵ See Lynn et al.'s [process tracing brief](#) to learn more about how to run process tracing tests.



What We Learned

Process tracing requires a widely agreed-upon outcome. Unlike the other two policy areas, we did not use process tracing on the findings around closing the Medicaid coverage gap because it did not offer sufficient coalescence around a single policy proposal or a clear enough longer-term outcome. For process tracing to provide meaningful insights, the evaluation needs an agreed-upon outcome that the evaluator can work backwards from, but our data did not include a clear anchoring point from which to do the backward testing. That said, interviewees did share interesting and substantive data about both what happened and the work going forward that ORS summarized to inform ongoing work at RWJF and among advocates.

Finding the right level and specificity of hypotheses—including alternate hypotheses—is iterative work. It took more time and iteration to land our hypotheses than we had anticipated. A few lessons learned included:

- While the Lynn et al.'s [process tracing guide](#) says that “*hypotheses that are too broad may need to be split apart so you can test the different linkages,*” it is quite common in policy advocacy work to use many different strategies and tactics to achieve a single outcome, so it’s difficult to pull those pieces apart. We found ourselves *frequently stacking activities or contextual factors into a single hypothesis* (e.g., “Past advocacy efforts around healthcare affordability and the Democratic trifecta in the 117th Congress directly contributed to the inclusion of APTC in the ARPA.”).
- **We had to be careful not to create what-if alternate hypotheses (e.g., if they had done x, y would have happened).** When asked to share alternate explanations for an outcome, interviewees often spoke to what-if scenarios and ideas about what to do differently next time, especially about policies that didn’t pass. Initial drafts included these types of alternate hypotheses, but it is not possible to collate data and evidence supporting a what-if scenario, so we ended up shifting most of those to null hypotheses (see below).
- **In the absence of a strong alternate hypothesis, we often found ourselves using a null hypothesis** (e.g., if the hypothesis was “research led to increased political will,” the alternate might be “while political will increased, research was not the reason why”). This was because interviewees often could not identify a clear alternate explanation for the outcome.

In running the process tracing tests, **it helped to think about the causal strength (i.e., plausible, can’t be disproved, contributing but with other strong contributing factors, clearly the contribution) as a way of working through the tests.** While the tests continue to be quite esoteric and require iteration to finalize, this approach grounded us in the question of “*What can we actually say with confidence?*” when we found ourselves caught up in the intricacies of the methodology.



3. REPORTING: WHAT WE DID AND WHAT WE LEARNED

What We Did

ORS summarized the evaluation findings into an individual policy brief for each of the three policies, which included an overview of the policy's history, the overall story of what happened, and reflections for the future. For APTC and CTC, we also included findings from the process tracing methodology about which advocacy activities and contextual factors made a difference to achieving progress in the 117th Congress.

What We Learned

Process tracing allowed us to answer the evaluation questions directly. Using process tracing, ORS was able to not only say which activities and contextual factors led to which outcomes for APTC and CTC, but also the strength of the data supporting those conclusions. These findings are different than for closing the Medicaid coverage gap, where we couldn't do process tracing, because the findings were more descriptive. For example, for CTC, we were able to say "*Research and stories demonstrating CTC's benefits from ARPA helped build support in Congress*" because that hypothesis passed a level 2 process tracing test.⁶ In comparison, for Medicaid, we could report the number of interviewees who cited research and data as an important advocacy activity (n=21), the number who cited credibility from data/science as an interim outcome (n=4), and say "*Some interviewees connected the increased credibility from data and science with Democratic policymakers increasing their support for the policy.*" While summarizing interviewee perspectives for Medicaid is helpful, there is no external validation of the accuracy of their perspectives—a key distinction from our analysis for CTC.

It is a balancing act to share process tracing test results without getting too in the weeds. The intended users of our evaluation findings were RWJF program staff, not evaluation people. These users are mostly interested in how the findings can inform their strategy and future policy advocacy work and less interested in the specific methodology that we used to get to those findings. They trust ORS and their internal evaluation and learning team's expertise to select and rigorously apply the appropriate methodology for the questions they're asking. Yet process tracing is a relatively sophisticated and jargon-filled methodology, so accurately reporting findings of the process tracing tests in a way that is helpful meant we needed to find ways to describe the causal strength in more colloquial ways. Here are examples of how we signified the causal strength of different findings in the briefs:

⁶ The level 2 test result means that we have strong evidence that both the change and alternate hypotheses contributed to the outcome. In other words, research/impact stories helped build support in Congress, but those activities are not the only explanations for that outcome occurring; additionally, we do not have data disproving the relationship.

Note that we use the language of levels 1-4 for the process tracing tests, but they are colloquially referred to as straw-in-the-wind (level 1), hoop (level 2), smoking gun (level 3), and doubly decisive (level 4).



- **Level 3:** “Pending premium increase notices **strongly contributed** to Congress’ willingness to prioritize an APTC extension in IRA negotiations.”
- **Level 2:** “Past advocacy efforts around healthcare affordability, the Democratic trifecta in the 117th Congress, and known inadequacies of the Affordable Care Act Marketplace **contributed** to the inclusion of APTC in ARPA.”

We did not have any level 4 (i.e., doubly decisive) test results and imagine they are quite unlikely in policy advocacy evaluation given the nonlinear and amorphous nature of the work. We did not report on level 1 test results in the “What Made a Difference” section of the three individual policy briefs because the process tracing test result did not indicate they had made a difference. The one exception was that participants in the APTC workshop had weighted polling as influential to the final APTC outcome, but no interviewees even mentioned polling as an advocacy activity relevant to APTC. To clear up that misconception, we included the following finding: “Polling was considered impactful, but that didn’t hold up when looking at the full dataset.”

We also shared the exact hypotheses, alternate hypotheses, test results, and interpretations in an appendix of each individual policy brief for readers interested in more detail.

CLOSING

From this experience, ORS recommends process tracing as a good method for evaluating a specific policy win or a clear, widely agreed-upon outcome in the policy process. It is not well suited if participants cannot agree on the outcome you are evaluating.

As interest continues to grow in applying causal evaluation methods to policy advocacy work, we hope this brief will help other evaluators build from our experiences and continue experimenting in this space. We see opportunities to: (1) continue applying process tracing to policy advocacy work to learn more about what works, when, and how; and (2) experiment with applying other causal methods beyond process tracing to policy advocacy work.

Authors: Melissa Howlett, Sarah Stachowiak, and Muniba Ahmad, with special thanks to Cassidy Webb and Chris Beauboeuf.

Recommended citation: Howlett, Melissa, Sarah Stachowiak, and Muniba Ahmad. 2024. “Methodology Deep Dive: Applying Process Tracing to Policy Advocacy Evaluation.” ORS Impact, January 31, 2024. <https://orsimpact.com/directory/Deep-Dive-Applying-Process-Tracing-to-Policy-Advocacy-Eval.htm>