How to do Process Tracing

A Method for Testing “How Change Happened” in Complex and Dynamic Settings

J ewlya Lynn
Sarah Stachowiak
Jennifer Beyers
From the Authors

This piece was written based on our experiences implementing process tracing when our experience showed that existing materials on the method had a lot more conceptual than practical information. We’ve approached this as people with some successful (and some unsuccessful) experience with the method itself, alongside deep experience in evaluating initiatives and strategies in complex and dynamic settings.

We focus not on the Bayesian side of process tracing but rather on how this can be implemented in a way that’s more participatory and lifts up the experiences and wisdom of those closest to the work and the problems being tackled. We hope this contributes to and helps make more approachable the important work of political scientists and methodologists upon which this work sits.
Introduction

Evaluators work in the social innovation space, where inspiration is limitless and resources are limited. Those seeking the support needed to test or scale their efforts are often asked to provide evidence of their contribution to effect equitable positive change. Evidence of such a contribution, however, is hard to come by—particularly for systems-change efforts within complex, dynamic contexts. Beyond evidence of contribution, social innovators in complex spaces also want to understand how and when change happens to support adaptation and future success.
What is Process Tracing?

Process tracing is an evaluative methodology that helps people understand how a particular large-scale change actually happened within a complex, dynamic context. Relevant large-scale changes include policy changes, systems changes, and cultural shifts, which may have been either planned or unexpected. Importantly, the story of how the change happened includes both steps taken intentionally to achieve the change (e.g., a program) and other contributing events, forces, or factors in the larger context. By focusing on stories of change and the interrogation thereof, the process tracing method provides a structure for understanding contributions that are inclusive, rigorous, and open to emergence.

We, the authors of this brief, believe that greater use of this methodology by evaluators could help nonprofits, funders, and other social change agents achieve greater impact by better understanding how, under what conditions, and in what contexts positive changes have been achieved.

Understanding causal pathways that actually led to changes will lead to better assumptions in the future, stronger TOCs, and more meaningful ways to share lessons learned with others. Because of these benefits, we hope to make this methodology more accessible to more evaluators who work in a variety of settings and topic areas, and set evaluators up to center meaningful stakeholder engagement and equity within their implementation.

This brief explains a series of six practical steps for using the process tracing method in the context of complex systems-change efforts using participatory approaches, providing illustrative examples from our work and lifting up key considerations and potential pitfalls throughout. These steps are organized into three phases (Figure 1).
What is Process Tracing?

In the end, the steps described below need to result in a few key things:

- There should be a clear set of hypotheses about how change happened.
- Data from stakeholders that are refined and tested in partnership with stakeholders, should be used to test the hypotheses about how change happened.
- Information should be collected about what else could have led to the change, or alternate explanations, beyond the expected or hoped-for implementation of the initiative, program, or set of activities.

We strongly encourage reading about process tracing before applying these steps.

## The Phases and Steps Involved in Process Tracing

<table>
<thead>
<tr>
<th>Begin</th>
<th>Clarify</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laying a Strong Foundation</td>
<td>Developing Your Theory of Change</td>
<td>Getting Ready for Analysis</td>
</tr>
<tr>
<td>Understanding Your Case</td>
<td>Refining Your Theory of Change</td>
<td>Analysis, Synthesis, and Promoting Use</td>
</tr>
</tbody>
</table>
Getting Around this Guide

**Navigation Menu**

The left margin allows you to instantly navigate to any of the 6 Steps involved in Process Tracing. Once there, the title will appear in color as a reminder of which section you are in.

**Color Indicators**

Each of the 3 Phases has been assigned its own color: M V Y X | P R K L U | P M P H | P V U H U K. The segments that make up this vertical bar change color as you switch from one Phase to the next.

**Progress Bar**

As you learn about the 6 Steps of Process Tracing, a progress bar running the width of each spread (as demonstrated here in blue) is another feature revealing your position within each Phase.
As with all methods, process tracing is only as good as how it is implemented. We believe that evaluation must be inclusive and used to fight against systemic racism to be a tool for good. We therefore tried to avoid reinforcing white supremacist ways of thinking and acting by reflecting on ways the method could unintentionally do this and infusing all steps with practices that advance equity, such as the principles from the Equitable Evaluation Framework.™

Because multicase efforts require many of the same steps as a single case, in Appendix A we share, for those looking to apply process tracing within and across multiple sites, considerations and steps that are unique to multicase studies. Just look for, and click on, the icon below (arrow + dot) to be taken directly to Appendix A and the additional steps.

---

2 https://www.equitableeval.org/framework
Throughout this brief, we provide real-world examples of process tracing in action, using the study described below.

In 2011, John Kania and Mark Kramer published an article in the Stanford Social Innovation Review, laying out “collective impact” as an approach for solving entrenched social problems at scale. For some, the introduction of a defined framework for cross-sector collaboration provided a useful way to focus new and existing partnerships toward a common goal and, hopefully, greater impact. Based on its promise, many resources were directed toward use of the approach for creating population-level change in a variety of areas before it had been rigorously assessed.

In early 2017, the Collective Impact Forum (Forum), an initiative of FSG and the Aspen Institute Forum for Community Solutions (Aspen Institute), hired ORS Impact (ORS) and Spark Policy Institute (Spark) to conduct a field-wide study that could help answer a fundamental question: To what extent and under what conditions does the collective impact approach contribute to systems and population changes?

The study used process tracing to determine the degree to which collective impact contributed to significant and meaningful progress—at the population level—in solving the problems initiatives set out to address. Foci of the sites, based in the US and Canada, ranged from river health to childhood obesity to reducing opioid deaths.

Findings from the Collective Impact Study (CI Study) can be found in this executive summary or the full report. Readers of this brief may be particularly interested in Appendix A of the full report, where we detail the methodological approaches, and in the “New Directions for Evaluation” chapter, where we wrote about the methods and lessons learned.


1. BEGIN

Laying a Strong Foundation

Understanding Your Case

2. CLARIFY

Developing Your TOC

Refining Your TOC

3. TEST

Getting Ready for Analysis

Analysis, Synthesis, and Promoting Use

Phase 1: Begin

Begin
Purpose.

The process tracing methodology depends on specific data that are collected and organized to permit the analysis of a specific case. Rather than considering process tracing as a method you implement once you reach a project’s data collection phase, process tracing steps should be built in earlier, as part of what might typically be considered design and discovery phases. This is particularly important when you implement the method while applying the Equitable Evaluation Framework.

The initial work to define roles for your stakeholders and develop a clear causal statement and model of change will fundamentally guide and direct the data collection plan and methods.
Actions.

- Identify and recruit your advisors to the process
- Develop a clear statement about the causal relationship you are testing

CAUSAL RELATIONSHIP STATEMENT TEMPLATE

The deployment of ________________, which was implemented from ________________ in ________________ , has directly contributed to ________________ .
Collective Impact Callout: Causal Relationship Statements

The deployment of a collective impact approach, which was implemented from 2005 to 2010 in San Diego County, has directly contributed to decreased childhood obesity rates.

The deployment of a collective impact approach, which was implemented from 2011 to 2016 in the communities surrounding the Elizabeth River, has directly contributed to improved river health, including measurable decreases in polycyclic aromatic hydrocarbons in the water.
Actions.

Develop a model of the change pathway

In addition to hypothesizing interim outcomes, begin brainstorming types of external drivers that may help to explain why change happened. Think about both drivers that occur only within the geography or setting of the case and those that are the result of larger societal or political dynamics.
collective impact conditions| activities & strategies| early changes| systems changes| population-level change
---|---|---|---|---
Five conditions of collective impact (a common agenda, shared measurement systems, mutually reinforcing activities, continuous communication, and backbone support organizations) and the principles of collective impact, including the capacity to implement an L X[P[ H W W Y V H J O
Strategies, actions, and conditioning (resulting from the CI work or in the environment more broadly) including evidence of strategies designed to F L S K L J U _ W S P[^S
Preconditions such as strengthened partnerships and networks, new collaboration activities, new grant funding strengthened F L S K L J U _ W S P[^S not shifts in how institutions behave, but rather how individuals are working together and informally building shared capacity.
+P H L Y L U[ W L Z V M systems changes, including formal changes to institutions, formal changes across multiple institutions, and informal experiments that shift (but do not sustain shifts) in institutional practice, includes looking for evidence of systems changes intended to P U J Y L H Z L L X[^P[^S
Population-level J O U N L Z W L J[^P[^S to the goals of the initiative, including whether the change P U J Y L H Z L L X[^P[^S
Considerations for exploring different types of relevant external drivers:

- **Place-Specific Drivers**
  If your causal statement relates to a single city, consider external drivers such as local political leadership changes, new local ordinances or policies, or local economic shifts (e.g., the exit of a key employer).

- **National/State Context Drivers**
  At the same time, consider whether federal and state policies, larger economic context, narrative shifts occurring nationally, or other drivers from outside the city could have influenced the intervention context.

- **Issue-Specific Drivers**
  For example, if the ultimate impact you are testing is increased voter registration and turnout among youth, consider local drivers such as an increase in the number of polling stations and highly contentious, highly visible local races or ballot initiatives; state level drivers such as state policy changes that make it easier to register or statewide campaigns to increase turnout; and national drivers such as increased turnout across demographics due to the context of national elections.

The brainstorming at this stage is to help you think about the types of questions you may want to ask stakeholders and what to look for when reviewing documents (e.g., Did [interim outcome] lead to [ultimate impact]?). You can also use it to create an initial list of codes for your qualitative data coding process (e.g., possible indicators or types of an interim outcome). Keep in mind, though: There are very likely some interim outcomes and external drivers that are well outside your initial thinking. This work is not meant to define what is relevant but to prepare you to pay attention to many different things that may be relevant.

In the next section, we’ll go more in depth about what types of interview questions you might construct based on this early brainstorming.

For multi-case work, check out the specific additional steps in Appendix A.
Additional Considerations and Potential Pitfalls.

Take time

The biggest pitfall at this phase is rushing the time it takes to lay a strong foundation. Process tracing studies differ significantly from many other types of case research in their dependency on highly specific data—data that tells plausible stories about how change happened. If you complete your document review and later interviews and have a good sense of many different outcomes and strategies, but you cannot establish links between them—such as how a strategy, set of activities, or other interventions contributed to an outcome—it will be difficult to move into the later phases.
Understanding Your Case

Purpose.

Now that you have a conceptual framework for the types of things that may have influenced change, it's time to plan for and collect data that will help you build a case-specific TOC documenting how the change actually occurred. The decisions you make about what data, how much data, and whose stories to incorporate will influence the change process you document in the TOC. Throughout this data collection phase, actively solicit plausible explanations for how change happened apart from the initiative, efforts, or activities included in the hypothesis being tested. Ultimately, this will help you identify and collect data on key elements comprising two types of change processes:

1. Hypotheses about how change happened as a result of strategies/activities, outcomes, and/or external drivers that ultimately led to the policy or population-level change being investigated, and
2. Alternate explanations for how changes might have occurred.
Identify your data sources and methods

For example, instead of asking about the strategies that were used first, and then asking about the interim changes observed, and moving on to the ultimate impact, you might ask the informant to tell you the story of how that impact was achieved. As they tell the story, occasionally stop them to clarify a strategy they talked about by asking what resulted from it directly and indirectly. Check in about an interim outcome they mentioned as critical along the way and ask how that outcome was achieved. Check to see if a particular type of strategy or interim outcome you identified as important in your early brainstorming work was relevant to their work (thus prompting them to potentially think about something that wasn’t front of mind during the storytelling). This storytelling approach is more likely to help you learn the relationships between strategies and a chain of outcomes than treating each part of your TOC as a separate element.
In the CI Study, we reviewed documents prior to the first stakeholder interviews. This allowed us to both ask the story about what happened and prompt for specific details that we saw in the documents, using questions like:

**Were there changes...**

- in how partner organizations behaved, either internally or with other organizations or sectors?
- among individuals within organizations or the community that supported changes in [area of focus ]?

These questions about collaborations and partnerships also came from our early investigation of the types of outcomes predicted to be important when deploying collective impact models and often identified as relevant to achieving the types of major policy wins and/or systems changes that are precursors to population-level change.
Actions.

- **Identify your informants**
  Be bold about who you engage in interviews and small group discussions. Think about including people with culturally distinct perspectives—those with very different positions and levels of power, both antagonists and protagonists—and make sure that at least some of your informants don’t have a stake in proving a particular pathway or story true. Interview protocols that encourage informants to tell you the story of how change happened should be relevant even for those who weren’t part of implementing the specific strategies named in the causal relationship statement. Getting insights from those who don’t have a stake in proving those strategies successful can help in both learning the full story and generating credible findings.

- **Collect and code data**
  As you collect data from your informants, steadily code the data, making sure that, along with any other useful codes, your coding helps you pull out specifics about the strategies, interim outcomes, external drivers, and ultimate impacts. For multi-case work, check out the specific additional steps in Appendix A.
Additional Considerations and Potential Pitfalls.

Plan for data of varying quality

Often in complex systemic change strategies, there will be multiple, conflicting, or incomplete sources of insight to consider. So, don’t treat all data as equal. Creating a “quality” rubric for specific types of data can be helpful. For example, if your rubric signals that the data are weak for a given finding, it can trigger the decision to collect corroborating data, triangulate more thoroughly, and/or explore whether disconfirming evidence may suggest there is a different way to understand how change happened. Consider planning these rubrics early on (e.g., after developing the causal statement), but also recognize that until you know your sources of data, you may not be able to finalize your rubrics.

Probe to understand the story, but also leave room for the unexpected

When you create protocols for interviews or discussion groups, balance how much time you expect participants to spend answering specific, detailed questions with the time you give them to simply share the story of how change happened. It can be easy to want to use the limited time available to dig into the details, but the real challenge of process tracing is capturing each interviewee’s credible, plausible causal story of change. When you parse the outcomes and impacts too much in how you ask the questions, you lose that storyline, and it can be harder to discern differences and similarities in how stakeholders understand the story. Similarly, leave some budget and time to conduct a few unplanned interviews or dig deeper into secondary data. During your data collection, you may hear about people or sources who weren’t identified during earlier phases and who have unique perspectives and are critical for understanding the story of change.

Think through sequencing of and iterating between different types of data collection

You’re likely to engage in a mix of document review and new data collection. Document review can help you prepare for new data collection, helping you probe more deeply on issues raised but not covered sufficiently in the documents. On the other hand, document review risks grounding your thinking in existing or already written narratives of how change happened, and you may fail to capture nuances about the untold stories of what really happened. Beginning with new data collection could help with this, but it may be necessary to circle back around to one or more informants after document reviews are completed to clarify anything not fully explained in the documents.
In the CI Study, population- or ecosystem-level impact metrics were a critical part of the analysis. The rubric below was designed to help make judgments on how accurate and relevant these metrics were for the processes being traced.

<table>
<thead>
<tr>
<th>RUBRIC</th>
<th>Indicator</th>
</tr>
</thead>
</table>
  > There is baseline or comparison and follow-up data tracked.  
| Meaningfulness          | > The initiative has a comparison point and can describe why their change is meaningful and if they consider it to be on track.                                                                            |
  > The initiative has a comparison point and can describe why their change is meaningful and if they consider it to be on track. |
2. CLARIFY

Laying a Strong Foundation

Understanding Your Case

3. TEST

Getting Ready for Analysis

Analysis, Synthesis, and Promoting Use

Clarify
Purpose. With data collected (and coded) from stakeholders and documents, you're ready to document a theory of how change happened and refine it with stakeholders. The TOC you develop should include not only the strategies or interventions you originally named in your causal relationship statement, but also interim outcomes and external drivers.
Actions.

Write up the process by which change happened or important external factors. You'll need to be prepared to make some judgment calls about how to handle divergence and what the divergence ultimately means, including developing your plan for how to handle this before it happens. Ideally, these judgment calls are made in partnership with the stakeholders you engaged to help with the process tracing study. Additionally, when considering outcome achievement, it's important to reflect on who benefited, who was harmed, and the degree to which root causes have a bearing on understanding the relative strength or quality of outcome achievement. It's too easy in any evaluative effort to look only for positive change.

Your change memo should lay out the ways in which external drivers are thought to have impacted the change and start to lay out the alternate explanations for how change could have happened without the contribution of the hypothesized strategies or outcomes.

Based on your change memo, draft a TOC visual

The TOC visual is a powerful way to document the process articulated in the change memo. It's a tool for having conversations with stakeholders (detail in the next step) and making apparent each causal relationship. It shows, for example, how one strategy logically leads to an interim outcome, which then combines with another interim outcome and a new strategy to drive a new outcome.
It should be messy and full, with multiple lines indicating relationships. Like the change memo, it should include external drivers, not just the internal or targeted work of the initiative.

If you have multiple, divergent stories, you may want to visually include multiple possible pathways within one TOC (if the divergence is small) or construct multiple TOCs that clearly highlight where they diverge from each other (e.g., put in black-and-white all the common elements across the TOCs and put in color the elements that differ).

The TOC visual below shows the causal relationships evidenced in the data; for example, how strategies and external drivers led to—and possibly combined with—interim outcomes and led to the ultimate impact. It is an important tool for having conversations with stakeholders, which will help you refine the TOC prior to analysis.

For multi-case work, check out the specific additional steps in Appendix A.
Develop a theory of how change did happen based on evidence, not a theory of how change could have happened.

Although most evaluative work uses TOCs to predict how the future may unfold, in process tracing your TOC is how the past did unfold. This means, among other things, that everything on that TOC needs evidence to support it. It also means you want to avoid presuming anything about how change usually happens, but rather document the story you have been told.
Purpose.

Before you begin the process tracing analysis, review the TOC you developed with stakeholders, including having them assign weights that correspond to the strength of the contributions from different activities, outcomes, and external drivers or other elements of the TOC to the ultimate impact. This step is critical not only for centering the voices and experiences of those who were part of the change, but also for gathering evidence of what mattered the most in the context of the larger story about the change and further exploring alternate explanations.
Actions.

Engage stakeholders in reviewing and refining the TOC

- If meeting virtually, use sticky notes on a whiteboard app.
- Either way, set up the visual before the meeting.

- Ask participants to add and remove elements as needed to refine the TOC, and ask them about relationships between items.
- Ask participants to vote individually on which strategies and external drivers contributed the most to getting to the ultimate impact. This is when you begin to weigh the strength of different drivers of the ultimate impact and the interim outcomes along the way.

- Discuss why they voted as they did to understand how they came to their conclusions.
- Try to reach a consensus—or, at minimum, have good evidence from the discussion about competing points of view.

- Push on what else could explain the change, not just what may have helped accelerate or facilitate it (e.g. asking, “If all of these changes happened, but none of these efforts caused them, then what else could have led to the changes?” is one way to push for true alternate explanations, which are required in the analysis to come).

If you have one TOC to review, the meeting will likely take about two hours to implement, assuming all the work of assembling the visuals was done in advance. If you have multiple conflicting TOCs, plan to have all of them ready to be explored and expect to use considerably more time.

For multi-case work, check out the specific additional steps in Appendix A.
Additional Considerations and Potential Pitfalls.

**Facilitation skills are critical**

Effective group processes to do in-depth work like a TOC review meeting may work well with some process tracing requires strong facilitation skills. Groups, particularly those used to submitting the CI Study, the lead researcher on a site was not the lead facilitator for the group process. Rather, team members who were experienced facilitators and had the skills to navigate conflict and make complexity accessible were chosen to lead the TOC review meetings. Strong facilitation is critical to lift up all voices, explore areas of divergence, notice and address power dynamics, and allow for emergence of new ideas.

**Culturally responsive engagement is critical**

A TOC review meeting may work well with some groups, particularly those used to submitting logic models or TOCs to funders. It may work poorly in other settings, including settings where time isn't experienced as linear, oral storytelling deeply embedded in the culture, or participants have multiple primary languages. Facilitated approaches should be adapted to match the participants' culture to meaningfully, respectfully, and fully engage them in the process.

**Recognize the burden and provide value**

You're asking a lot of your participants in a process tracing TOC review process, including typically engaging for two hours or more and spending time with people who have very different—potentially uncomfortably different—points of view. Take time to work with your advisory group or others in the community to make alternate explanations, or causal relationships sure there is value for participants. In the CI Study, between specific outcomes that will help you we made sure each site received an individual site writeup they could share with funders and other partners to talk about their work, and we provided stipends for individual participation.

**Capture good notes and identify gaps**

The discussion during the TOC review meeting, including people's reasons for changing the visual, is critical data. Record or take notes during the meeting. Additionally, note where you may need to augment the data you've collected so far, especially to understand external drivers.
Test
Getting Ready for Analysis

**Purpose.** A crucial part of what brings quality and rigor to process tracing is the analytic process: systematically applying the process tracing tests to determine the certainty and uniqueness, or **inferential strength**, of the TOC toward achieved outcomes. With participatory processes, multiple data sources, and complex change scenarios, evaluators can struggle with organizing the data to efficiently and meaningfully do the analysis. These steps break down the process of organizing the data in ways that set you up for a successful analysis.
Getting Ready for Analysis

Generate your testable hypotheses

When developing hypotheses, be clear, specific, and focused around a particular causal link, such as from activities to outcome or across outcomes over time. It will be hard to assemble and analyze data for hypotheses that are too general; similarly, hypotheses that are too broad may need to be split apart so you can test the different linkages.

GETTING TO STRONGER HYPOTHESES

<table>
<thead>
<tr>
<th>Sample Hypothesis*</th>
<th>Issue/Resolution</th>
<th>Stronger Hypothesis/Hypotheses</th>
</tr>
</thead>
<tbody>
<tr>
<td>More funding led to greater capacity</td>
<td>Too general/Make it more specific</td>
<td>Because the program received additional funding in 2016-2020, they hired a trainer, had subs to cover their shifts, and built new skills to develop diverse youth leaders.</td>
</tr>
<tr>
<td>A trusting collaborative table, with funding and a diverse set of partners, collaborated to provide access to more pregnancy prevention programs.</td>
<td>Too many strategies grouped together with interim outcomes mixed in/Split up into multiple hypotheses</td>
<td>Funding helped diverse partners convene because they could provide stipends and childcare and could create a more consistent rhythm of meetings where they built greater trust. Partners around the table collaborated more deeply to provide more programming to the community because of the regular convenings.</td>
</tr>
<tr>
<td>The set of aligned activities of the school and early childhood changing wellness practices and policies in schools and daycares, capitalizing on local investments.</td>
<td>Strong—inclusion of local investment as external driver helps to further explain change</td>
<td></td>
</tr>
</tbody>
</table>

*Hypotheses can include any combination of strategies or external drivers to explain the relationship with specific interim outcomes or the ultimate impact within a causal chain.
In the CI Study, we worked with stakeholders to generate six to eight site-specific hypotheses based on the refined TOC and aligned with the generic CI change model. This process helped identify the most critical hypotheses behind the population-level change seen.

<table>
<thead>
<tr>
<th>Type of Hypotheses Based on Generic Change Model</th>
<th>Example Site-Specific Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>COLLECTIVE IMPACT</strong></td>
<td>Common agenda, mutually reinforcing activities (action plan and work groups) and backbone have (1) increased trust among partners, (2) created culture change among agencies (toward collaborative work), (3) facilitated development of local work, and (4) maintained high levels of political will.</td>
</tr>
<tr>
<td><strong>CONDITIONS TO EARLY CHANGES</strong></td>
<td>The early changes (legislative champions/political will, allies, deepened relationships/trust, and public engagement) together led to the adoption and implementation of legislation and legislatively mandated implementation/oversight bodies.</td>
</tr>
<tr>
<td><strong>EARLY CHANGES TO SYSTEMS CHANGE</strong></td>
<td>More aligned policies and practices within partner agencies/ to homeless clients and increased housing placements for homeless veterans and the chronically homeless.</td>
</tr>
<tr>
<td><strong>SYSTEMS CHANGES TO POPULATIONS CHANGES</strong></td>
<td>The extent of river cleanup and the changed relationship of the community to the water would not have occurred without the initiative.</td>
</tr>
<tr>
<td><strong>OVERALL THEORY OF CHANGE HYPOTHESIS</strong></td>
<td></td>
</tr>
</tbody>
</table>
Getting Ready for Analysis

1. BEGIN
   Laying a Strong Foundation

2. CLARIFY
   Developing Your TOC
   Refining Your TOC

3. TEST
   Getting Ready for Analysis
   Analysis, Synthesis, and Promoting Use

Actions.

Understand the different process tracing tests

Before entering into this discussion, note that the language in process tracing can be pretty difficult to wrap your head around. If you’d like to better understand the original framing of the four process tracing tests, we recommend Collier’s seminal piece, Understanding Process Tracing (2011) in that article, the author uses the example of a Sherlock Holmes murder mystery to explain the tests.

To help operationalize these tests and their corresponding levels of inferential strength, the table below describes the meaning of the tests relative to the change hypothesis under consideration (which includes the relationship between strategies/activities, other outcomes, and/or external drivers) and only alludes to the alternate explanation. We explain how to use the alternate explanation when testing the strength of the change hypotheses in the step that follows.

As a change hypothesis “passes” each test, it is stronger in its ability to fully explain the hypothesized relationship (i.e., has more inferential strength). In other words, a hypothesis cannot reach the “smoking gun” level of strength if it cannot also say that it met the standards of straw in the wind and hoop.
### The Process Tracing Tests and Their Associated Levels of Inferential Strength

<table>
<thead>
<tr>
<th>Test (process tracing term)</th>
<th>Level of Inferential Strength (explanatory language)</th>
<th>Re-stated level of inferential strength*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Straw in the Wind</td>
<td>The hypothesis is plausible but is neither certain nor unique.</td>
<td>The best we can say is <strong>THIS</strong> happened, then, plausibly, <strong>THAT</strong> would be the result. But we don't know if <strong>THIS</strong> or <strong>THAT</strong> really happened (e.g., face validity)</td>
</tr>
<tr>
<td>Hoop</td>
<td>The hypothesis is unique but not certain.</td>
<td>The best we can say is <strong>THIS</strong> happened and <strong>THAT</strong> happened, and we can't exclude the possibility that <strong>THIS</strong> drove <strong>THAT</strong> (e.g., correlation-level strength)</td>
</tr>
<tr>
<td>Smoking Gun</td>
<td>The hypothesis is certain but not unique.</td>
<td>The best we can say is <strong>THIS</strong> happened and we're fairly certain <strong>THAT</strong> resulted, but we can't exclude the possibility that something else may also have helped or led to <strong>THAT</strong> (e.g., strong contribution)</td>
</tr>
<tr>
<td>Double Decisive</td>
<td>The hypothesis is deemed to be certain and unique to explain the change.</td>
<td>The best we can say is <strong>THIS</strong> happened and <strong>THAT</strong> resulted from <strong>THIS</strong> specifically and only <strong>THAT</strong> (e.g., attribution)</td>
</tr>
</tbody>
</table>

*THIS = hypothesis set of drivers (some combination of intended strategies, external drivers, and prior outcomes)  
THAT = the outcome in the hypothesis
Actions.

Assemble relevant data
to test hypotheses

With at least five data elements for each hypothesis, you’ll be assembling a plethora of data points to test all your hypotheses. Creating a spreadsheet with the hypotheses as rows and the five elements as columns can provide a useful format for working through the process tracing tests. See Appendix C for a sample data assembly spreadsheet.

It’s not unusual to find gaps while assembling the data points for each hypothesis—gaps in data about outcomes, linkages, or alternate explanations. Take the time to fill gaps through additional data collection or document review, and build this into your timeline and process. You may also need to verify data points through secondary data sources or follow up with informants ahead of the process tracing analysis. The hypothesis chains should reflect the reality of change on the ground and just illustrate the hypothesized model of change being tested.

For multi-case work, check out the specific additional steps in Appendix A.

In addition to the summarized data, it can be helpful to have the analyst capture in a synthesis note their confidence in the data associated with each hypothesis.

Assemble similar data points for hypotheses representing later parts of the theory: short-term to medium-term outcomes, medium-term to long-term, long-term to ultimate impact, and any other relationships that are hypothesized (e.g., short-term to short-term outcomes; long-term outcomes that reinforce short-term outcomes). Additionally, beyond assessing the hypotheses for each linkage in the chain, it’s helpful to assess the strength of the overall theory for the case.
PREPARING FOR PROCESS TRACING ANALYSIS

Activities

+H[HVUX]HSP`
reach, amount, etc. of activities

Data on the link between activities and outcomes

Alternate explanation for why/how the outcome occurred

Outcomes

+H[HVUX]HSP`
of outcome(s) achieved, etc.

Data on likelihood, X\HSP` L\V M alternate explanation

How to Do Process Tracing
Additional Considerations and Potential Pitfalls.

Take care with alternate explanations

One of the greatest strengths—and gnarliest parts—of the process tracing data collection and analysis lies in the use of alternate explanations. They’re crucial for providing a point (or points) against which to determine the strength, certainty, and uniqueness of the proposed hypotheses of how change happened. In getting stories of change, it’s easy to collect an array of different factors that may or may not have influenced change: contextual factors that explain why change was accelerated or more feasible; other factors that needed to be present but do not provide a sufficient alternate explanation; or external factors that contribute to the strength or scale of the change seen. These pieces of information may be incorporated into your change hypotheses or alternate explanations, but on their own they’re insufficient answers to the crucial question: What else could explain how this observed change occurred, if not for the expected explanation (represented in the change hypotheses)? To answer that important question, you need information about alternate explanations—causes of change other than the expected one. As shown in Table 3, an external driver with a big explanatory role could sometimes be an alternate explanation. Sometimes there will be alternate explanations that are outside of the expected or hoped-for TOC that supplant the starting hypothesis.

Who had power and who lacked power?

When thinking through your hypotheses and analysis, consider how the process has unfolded and what your blind spots might be. Look critically at the data and think about whose story it tells and whose stories and experiences might be left out or hidden within the data if you’re considering only a set of dominant culture frames. Earlier work to engage stakeholders should help mitigate this, but as you consider the data in its totality, it’s important to reflect again on what is being lifted up and what is being left out of the analysis. This reflection is yet another point at which you may discover the need to pause and collect additional data or discuss your preliminary work with your advisors.

Be true to the story

While our causal chains might suggest a world in which change is predictable and linear, the real world doesn’t function that way. The rigor of the process won’t matter if the story you capture and assess doesn’t reflect the messiness of real-world change. You may hear and see stories of virtuous cycles, where later outcomes strengthen earlier outcomes that beget later outcomes. You may hear stories where change is not linear or is buoyed by a favorable change in the external environment. The value for learning from these projects is strengthened by embracing the uniqueness and messiness that might be at play, rather than flattening it out for the sake of simplifying analysis.
## Sample Change Hypotheses and Alternate Explanations

<table>
<thead>
<tr>
<th>Change Hypothesis*</th>
<th>Alternate Explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td>A strategy focused on hearts and minds (public will and science) led to a willingness of local jurisdictions to go above and beyond requirements in their cleanup efforts.</td>
<td>External driver explains change: Some practices improved due to regulation alone. Explanation that might supplant hypothesis: Practices improved due to shifts in elected and appointed leadership unrelated to the local activities.</td>
</tr>
<tr>
<td>The initiative worked through mutually reinforcing activities to advance strategies that led to systems changes (programming in afterschool; curriculum in school; increased availability, accessibility, and acceptability of contraction); federal funding allowed for new strategies that otherwise would not have been part of the mutually reinforcing activities.</td>
<td>External driver explains change: Federal funding allowing for new strategies solely explains the systems changes seen. Explanation that might supplant the hypothesis: Media that was not a part of the initiative but was released and widely viewed during this period.</td>
</tr>
</tbody>
</table>

*May include external drivers.
Now you're ready to engage in a core part of the process tracing approach: systematically applying the process tracing tests to determine the certainty and uniqueness, or inferential strength, of the change pathways toward achieved outcomes. After that, you'll be ready for synthesis, within and across hypotheses and the TOC overall, and then promoting use of these results.

Purpose.

---

Assess each change hypothesis against the tests and document results.

Focused efforts to groups explain the result? If there was no strong alternate explanations or strong external drivers also contributing to change, the hypothesis would be deemed unique. When developing hypotheses, be clear, specific, and focused around a particular causal link, such as from activities to outcome or across outcomes over time. It will be hard to assemble and analyze data for hypotheses that are too general; similarly, hypotheses that are too broad may need to be split apart so you can test the different linkages.

**Matrix for Assessing the Certainty and Uniqueness of Evidence**

- **Evidence necessary for h**
  - HIGH
- **Evidence insufficient for h**
  - LOW
- **Evidence sufficient for h**
  - HIGH
- **Evidence not necessary for h**
  - LOW

<table>
<thead>
<tr>
<th>Source: Centre for Development Impact</th>
</tr>
</thead>
</table>

(continued)
Practically, it can be helpful to use the four-square matrix on page 39 to physically map each hypothesis relative to certainty and uniqueness without getting caught up in the idea of passing “tests.” Where do the data suggest the hypothesis falls? Some hypotheses will achieve the same level of inferential strength even though your assessment of their uniqueness or sufficiency differs somewhat. For example, two hypotheses may ultimately be characterized as having “low” uniqueness, even if you assessed one as being lower in uniqueness than the other. Table 4 on the right shows example evidence relative to each test or level of inferential strength for one change hypothesis.

In addition, it’s important to assess the strength of the individual parts of the TOC, but also to step back and assess the TOC as a whole. To do this, you’re engaging in similar analysis, but your hypothesis is that the overall theory, in whole, explains the result, versus the individual parts. While the individual linkages are useful and important to understand, looking only at the individual pieces can obscure a consequential understanding of the overall, cumulative work.

For each hypothesis, note in your data assembly spreadsheet what “test” or level of inferential strength each hypothesis achieved (e.g., hoop, smoking gun).
## Applying the Four Tests to a Specific Hypothesis

<table>
<thead>
<tr>
<th>Test</th>
<th>Example of evidence that would support the hypothesis</th>
<th>What this evidence does and does not indicate vis a vis certainty and uniqueness</th>
<th>Data sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Straw in the Wind</td>
<td>Policymaker comments after the media campaign ran included key phrases and framing often used in the campaign.</td>
<td>This information is helpful but does not prove the media campaign was the reason policymakers used the language, nor that the campaign contributed to the policy change.</td>
<td>Secondary analysis of policymaker speeches/press releases; analysis of secondary data about campaign activities.</td>
</tr>
<tr>
<td>Hoop</td>
<td>Policymakers were aware of the campaign, and their own framing of the issue used elements of the campaign.</td>
<td>The campaign cannot be excluded as a potential way change was achieved; evidence to link the media campaign to the outcome.</td>
<td>Bellwether interview data.</td>
</tr>
<tr>
<td>Smoking Gun</td>
<td>7VSP'THRLYZPKLU[PLK[OLTLKPH campaign as a critical turning point in their understanding and support for the issue, though they also noted the national conversation around health care and the governor’s race as other factors.</td>
<td>;OLLPKLJLPZZU implicitly suggests that the media campaign contributed and rules out alternate explanations to explain the outcome.</td>
<td>Bellwether interview data, secondary data (e.g., timing of campaign and changes in statements), data about prevalence/visibility of alternate explanations.</td>
</tr>
<tr>
<td>Doubly Decisive</td>
<td>7VSP'THRLYZPKLU[PLK[OLTLKPH campaign as a critical turning point in their understanding and support for the issue, leading to their vote for the policy, noting that there was a lack of other discussion, visibility of the issue, or opportunities to learn about it, and without the media campaign, there would have been no action.</td>
<td>;OLLPKLJLPZZU implies that the media campaign contributed and rules out alternate explanations to explain the outcome.</td>
<td>Bellwether interview data.</td>
</tr>
</tbody>
</table>
Thus, while the final part of the chain was “doubly decisive” (i.e., it was clear that the adopted curriculum was contributing to better academic results), the overall inferential strength for the connection to collective impact passed the “hoop” test (i.e., the initiative was a player in how the curriculum was adopted and implemented, but many other factors also contributed).

Alternately, other initiatives had less strong inferential strength as you moved through the TOC (i.e., hypotheses about connections between activities to early changes are stronger than those related to systems changes leading to population-level changes). However, the strength of the impact of the initiative in earlier stages strengthened the overall inference around collective impact being a key part of the causal mechanism.

In the CI Study, at one site we found that an adopted curriculum was highly likely to have led to improvements in academic achievement, but the earlier linkage of how the initiative contributed to it was less strong compared to alternate explanations.
Within your spreadsheet, you'll now have the results of the analysis for each hypothesis and for the overall change story. Now is the time to describe what contributed to the change and with what level of certainty. Here you can augment the descriptive story of change you have gathered to date to include where the hypotheses about change held up, where and in what ways external drivers played a role, and the degree to which alternate explanations played a role.

In more complex systems-change settings, the likely story of change includes multiple contributing factors, necessary preconditions, and a confluence of activities and outcomes that were necessary for change. Where the hypothesized relationship is less strong, it will be helpful to understand what else contributed. Where the hypothesized relationship is strong, it may be helpful to describe or explore important aspects of the context or capacities required. Overall, the goal here is to deepen learning about what was effective in this case and help those involved and interested in related efforts to apply the lessons in future work.
Collective Impact Callout: Excerpt from Elizabeth River Site Report

Strength of the Causal Pathway

Overall, there was compelling evidence the collective impact approach had a strong connection to the population change, with low plausibility of an alternate explanation. The initiative’s strong backbone and leadership, combined with an inclusive, participatory, shared-credit leadership style that placed no blame but rather focused on joint ownership of solutions, was necessary for engaging diverse partners and attracting funding. The initiative’s strategy, which focused on changing hearts and minds, led to improved practices, new work groups and programs, improved ongoing measuring and monitoring practices, and a willingness of local jurisdictions to go above and beyond.

The initiative’s strong backbone and leadership, combined with an inclusive, participatory, shared-credit leadership style that placed no blame but rather focused on joint ownership of solutions, was necessary for engaging diverse partners and attracting funding. The initiative’s strategy, which focused on changing hearts and minds, led to improved practices, new work groups and programs, improved ongoing measuring and monitoring practices, and a willingness of local jurisdictions to go above and beyond.

Overall, there was compelling evidence the collective impact approach had a strong connection to the population change, with low plausibility of an alternate explanation. The initiative’s strong backbone and leadership, combined with an inclusive, participatory, shared-credit leadership style that placed no blame but rather focused on joint ownership of solutions, was necessary for engaging diverse partners and attracting funding. The initiative’s strategy, which focused on changing hearts and minds, led to improved practices, new work groups and programs, improved ongoing measuring and monitoring practices, and a willingness of local jurisdictions to go above and beyond.

Media and public engagement campaigns were critical for building commitment, ownership, and trust. Overall funding, combined with committed cross-sector partnerships, has become cleaner; it has reinforced commitment among partners and a positive relationship and sense of ownership about the river.
Foster learning about results among stakeholders

While the language around the process tracing tests can be alienating or distracting, the benefit of the results is deeper learning about complex change initiatives among those who want to learn, improve, scale, or tackle new related issues. Actively finding ways to make meaning and use what is learned is important. Consider the following:

> Share preliminary results with your advisors. Prior to formal documentation of results, discuss findings with your advisors. This is an opportunity to explain how their contributions were used, correct inaccuracies, and address doubts. Like the evaluation itself, this discussion should be as inclusive as possible and create adequate space for reflection and active, free, and meaningful participation (Bamberger & Segone, 2011).

Additional Considerations and Potential Pitfalls.

**Consider a team approach**

In the Collective Impact Study, many people played different roles. In the end, we found it helpful for the people conducting the process tracing tests to be strong analysts who hadn’t been directly involved in the data collection and TOC review process. This provided some strength of triangulation across analysts; it also allowed the focus to be on the data rather than the personalities or charisma of different individuals within site data collection processes. At the same time, each case had a site lead who reviewed the analysis and could push back and add nuance as necessary. It may be helpful to think about how to differentiate roles and connections to sites if you have a multiperson team. Also think about how your advisors and stakeholders may be involved and engaged throughout these processes.

**Leave good “tracks”**

Conduct analyses and syntheses that can be understood and replicated by others. Be sure to document criteria, data sources, and gaps filled along the way.
Consider other types of contribution, too

Process tracing seeks to help understand how change happened, but there are additional nuanced ways to think about contribution on a broader spectrum. If you find that a particular effort, initiative, or program doesn’t seem to have caused the change seen, consider whether it may have contributed in other, less visible ways through this process. In No Royal Road, evaluators Jim Coe and Rhonda Schlangen offer a spectrum of types of contribution beyond “lead contributor” that may be useful for explaining the different types of roles of an effort or initiative in complex change efforts, like policy change or other community and ecosystem-level initiatives seeking population-level changes.

These include recognition of roles such as “seed sower,” an actor that initiates a campaign and sows the seeds that grow into something that gathers momentum and takes on a life of its own, or “over-the-line getter,” an actor that makes a discrete or niche contribution that’s undeniably vital but may not show up in a process tracing hypothesis.12

Closing Considerations

When you’re ready to operationalize and use the process tracing methodology, consider these final additional overarching thoughts.

When should I use process tracing?

Know some key limitations

» Process tracing isn’t fast or low budget. The process requires meaningful engagement with stakeholders, iteration over time, and multiple data sources.

» Process tracing is robust, but fairly narrow in scope. It will provide a strong answer to the degree and ways in which expected activities contributed to a proven result, or not. While you will also learn other things related to traditional process and outcome evaluation questions, it’s important to go into it with clarity about what the process will generate. In the CI Study, for example, we could state with greater certainty that collective impact meaningfully contributed toward population-level change in eight sites where an ultimate impact had occurred and fidelity to the collective impact model was strong. The study did not generate findings generalizable to all collective impact efforts (i.e., collective impact will always result in population-level change). We also did not answer every question that could be of interest about how collective impact was working. If hypothesized causation isn’t found, you’ll come out of it understanding what did lead to change in a case (or across a set of cases).

When should I use process tracing?

» Use it when untangling how change happened in more complex change scenarios.

» Use it when you’re certain that the ultimate impact you’re seeking to understand has occurred (e.g., policy change, population-level outcome).

» Do not use it when looking only at shorter-term outcomes for which there is a relatively straightforward causal relationship.
Expect to iterate

- Be prepared to enter into process tracing knowing that the work will be iterative. As you learn more, you may identify new informants, new relationships to consider, and different ways of understanding other contributing factors/external drivers or alternate explanations.
- Create a workplan, timeline, and budget that accounts for this reality.

Have a team that can cover various skills and roles

- The method overall requires a range of skills, especially if you want to meaningfully engage an array of stakeholders.
- The work is well-served by staff who are strong interviewers. Focused but still open-ended inquiry is important when initially understanding the story and creating the change memo and TOC.
- Strong facilitation skills are needed to effectively use the advisory group and engage in good group processes for testing the TOC.
Conclusion
Too often, evaluations of complex, messy systems-change efforts do little more than describe what happened. Deeper inquiry to understand and interrogate how change really happened leads to evaluations that truly help to sharpen social innovators’ assumptions, deepen their learning, and enhance their ability to adapt and respond to future contexts and scenarios. We also think this methodology can be consistent with Equitable Evaluation Framework principles and make evaluation and learning a tool for advancing equity—one that is conducted in collaboration with and not on communities. Implemented well, this methodology is participatory and can help put more power in the hands of those closest to and impacted by the work, centering the stories, knowledge, and experience that exist outside the boundaries of funders’ strategies or TOCs.
We realized it would be most helpful to lay out the key steps of process tracing for a single case or result, because even in multi-case study, you’ll do many of these same steps for each case. However, we also wanted to share additional guidance and lessons learned for the steps when conducting a multi-case study.

In this table, we provide additional notes and considerations for each stage. The table is intended to be used in tandem with the main content for that step.
Step 2
In studies with multiple cases, chances are you won't have the resources or time for a deep investigation in every site. You can still successfully capture change stories, but you'll need to be very careful and strategic about who you interview and the documents you collect. In the CI Study, we had two key informants in this first stage, alongside a wide range of documents (dozens of documents from some of the sites). To save time, we did a cursory review of each document to decide whether and how it would be useful and which parts to code.
How to Do Process Tracing

Appendix A

Clarify

Testing Your Theory of Change

Steps 1 and 2
In studies with multiple cases, you conduct this step no differently than you would with a single case. You'll do the testing unique to each case. Make sure you test in similar ways across sites and with similar breadth and diversity of stakeholders. A common protocol and list of stakeholder types can be helpful to develop prior to recruiting participants into the testing process.

MULTI-CASE ADDITIONAL CONSIDERATIONS
Steps 1 and 2
In studies with multiple cases, you'll have a change memo for each case, followed by a TOC for each case. At this point, you don't need to develop any cross-case memos or visuals, nor will you do anything differently while you're testing the TOCs in the next phase. You'll return to the cross-case level work during the analysis later in the process, during Phases 5 and 6.

Developing Your Theory of Change

1. Engage stakeholders in testing and refining the TOC.
2. Develop a detailed protocol and prepare your visual presentation of the TOC in advance.

ACTIONS
1. Write up the process by which change happened (change memo).
2. Using your change memo, draft a TOC.
Step 2

For multi-case studies, you'll likely want to analyze the strength of each case AND the strength of the individual steps in the TOC. For example, in the CI Study, we made determinations about the inferential strength of collective impact and the population-level impact. We also analyzed across sites how strong the linkages were from the collective impact conditions to the short-term outcomes, from short-term outcomes to systems change, and from systems changes to population-level changes. Therefore, you'll want to organize and conduct the process tracing analysis by case/site and then reorganize the data by grouping together by the type of hypothesis—for example, Excel worksheets where you have “like” hypotheses (e.g., activities to short-term outcome, short-term outcome to mid-term outcome).

Additionally, having multiple cases will let you do additional kinds of quantitative analysis to explore cross-site themes and patterns through frequencies and cross-tabs.

Step 2

In the multi-case context, you may want to share site-specific reports with stakeholders for each case. Indeed, the promise of such a report may be offered as an incentive for participating in the study.

1. Assess each hypothesis against the tests and document the results.
2. Synthesize the results.
3. Foster use among stakeholders.

ACTIONS

1. Identify and recruit your advisors to the study.
2. Develop a clear statement about the causal relationship you are testing.
3. Develop a model of the change pathway.

MULTI-CASE ADDITIONAL CONSIDERATIONS

Step 1

If you're working across multiple cases, you'll need to complete this step for each case. It's important to agree across team members and sites on the level or altitude of the hypotheses being developed at this stage so they are comparable throughout the TOC and across cases.

Step 3

Develop the same sets of data for each site so that a robust site-level analysis can be completed before you turn to cross-site analysis. Additionally, just like you agreed on the level of hypotheses being developed, it's important to agree across team members and sites on the level or altitude of the data being captured. You don't want determinations of certainty or uniqueness in the process tracing tests to be an artifact of how data were captured versus the virtue of the change itself. For example, say that one analyst captures only cursory data about evaluation data showing that participants in a program achieved outcomes, and another analyst provides more detailed data on the degree of achievement and the degree to which results were seen across different segments, and shares confidence in the evaluation. The analysis of the hypothesis could be seen as more “certain” in the latter case, even if the data were equivalent in the former case.
Appendix B

A facilitation guide for a process tracing theory of change review with stakeholders

### The purpose of the meeting, introductions, and definitions

1. Introduce all participants and research team.
2. Introduce the purpose/frame up the meeting:
   - Provide some framing around being critical: “We want this exercise to be a pressure test of how changes have unfolded over the past [number] years. I encourage you to put on your skeptic/critical hat to share your thoughts on how the initiative might have impacted change as well as how other factors—outside of the initiative—have impacted change.
   - Provide some framing around this being a “safe space”
3. Share informed consent language.
4. Define key terms, including population outcome, systems change, early outcome, strategies, collective impact (CI) conditions, other drivers (have handout or put on flipchart paper):
   - Changes in what you’re ultimately trying to influence (health, educational, or social outcomes; environmental improvement; etc.)
   - Formal or informal changes within or across organizations (changes to policy or organizational practices, new programs, etc.)
   - Changes in practices or how people are working together
   - How the initiative is being implemented
   - Things happening outside of the initiative—for example, other initiatives, policies, funding—that could have contributed to changes in the population outcome

15 minutes
1. Walk them at a high level through the storyline of the systems changes that contributed to population-level change.

2. Physically put the pieces up on the wall as you're talking them through it (all pre-populated)

3. Provide some framing around the population change: what data support it, including the most recent year in which data were reported (so you have a time frame to anchor in)

4. Ask people to react to all of the systems changes—what resonates, what really describes how change happened, and what might they state differently?

5. Continue to refine until there's a general calming of energy in the room and people are feeling good about the flow on the wall. Hopefully, some things have been added, moved, and removed.

6. Test the Narrative. Ask the room, “Looking at what's on the wall, are there any systems changes that could be removed, and we likely would have seen the population-level change anyway?”

   Review each systems change that doesn't have a star on it and check on the importance of each one. If there's strong agreement in the room, pull the sticky note off the wall. If there are mixed feelings, leave it where it is and move on.

   What was the impact of that change? How did it drive outcomes? (to make sure this is a driver of the population change you're talking about) “When did it happen? (to make sure that the change has been accomplished)

• Unpack, try to understand if you're dealing with a difference of definition/description or a true difference in belief about what drove the change. If the former, try to clarify to get everyone on the same page. If the latter, get people to process whether this is about level of contribution or about whether it was a contribution at all (“Are we rejecting part of the causal chain, or are we suggesting it isn't as significant because of some other driver?)

• Depending on the response, you may be adding a new systems change or reframing the one already up.
1. Focusing on the systems changes that they identified as directly contributing to the population-level changes, walk them at a high level through the storyline of how strategies contributed to those systems changes, including the “early wins” along the way (interim outcomes). Let them know where the data came from, why it's important, and that you'll be unpacking and challenging it.

2. Physically put the pieces up on the wall as you're talking them through it (all pre-populated):

   - ORANGE. Early wins (interim outcomes)
   - GREEN. Strategies

3. Ask people to react now to the beginning of the map—what resonates, what really describes how change happened, and what might they state differently?

   - When something resonates, acknowledge it, point to where it is on the visual, and put a star in the upper corner of the sticky note using the marker.
   - When something doesn’t resonate, check in with the room—is there a sense of agreement that a change is needed?
     - IF YES. Make the change using the appropriately colored sticky notes (note-taker: note this shift in the causal relationship has happened).
     - IF NO. Unpack, trying to understand if you’re dealing with a difference of definition/description or a true difference in belief about what drove the change. If the former, try to clarify to get everyone on the same page. If the latter, get people to process whether this is about level of contribution or about whether it was a contribution at all (“Are we rejecting part of the causal chain, or are we suggesting it isn’t as significant because of some other driver?”).

4. Depending on response, you may be adding a new driver or reframing the one already up.

5. Continue to refine until there's a general calming of energy in the room and people are feeling good about the flow on the wall. Hopefully, some things have been added, moved, and removed.

6. Test the Narrative. Looking at what’s on the wall, are there any strategies or early changes that could be removed and we likely would have seen the relevant systems changes anyway?

   - Review each sticky that doesn't have a star on it and check the importance of each one. A potentially good probe: “Would you have gotten to systems changes without X?”
   - If there's strong agreement in the room, pull the sticky note off the wall.
   - If there are mixed feelings, leave it where it is and move on.
4. Explore the “story” of other drivers

1. Let’s imagine that even though all of this good work happened, none of it was the reason the population-level outcome changed. Instead, other factors were major drivers of the change. We know some of them from our pre-interviews:

2. Go through each driver, one by one: “Was this something that directly caused the population- or systems-level change, enabled the change, or in some way mediated it—for example, sped it up or slowed it down, but didn’t cause it directly?”

3. What other drivers, outside the control of the initiative, were critical to achieving the systems- and population-level changes? A potentially good probe: “What happened that you did not influence but were very glad it happened?”
## Establish the “weight” of influence of early changes, systems changes, etc.

### 1. Introduce the idea: “We have a story with many different drivers, and we want to understand the weight of the influence of the key actions they took compared to other drivers. We're going to engage in a process of weighting the early changes, systems changes, and alternate drivers according to how influential they were in driving the population change.”

### 2. Introduce the individual weighting activity (to provide data for discussion)

### 3. Lead people in group processing of individual weights:

- Many different colors of dots are on one sticky note, showing that many people voted on this sticky note: “It looks like there is a lot of consensus this was a critical influence.” (Then move onto the next, unless someone responds otherwise.)
- Many different colors of dots are on one sticky note, with some colors having multiple dots, some having only one dot: “It looks like a lot of people agree this was very influential, but there isn’t necessarily agreement on how critical. Can someone who put more than one dot share your thinking?” (You want to get to a general sense of agreement on the “weight.”)
- About half the colors are present on a sticky note and about half are missing: “It looks like there’s some mixed opinion in the room about whether this was a critical influence. Does anyone want to share your thoughts about why it’s critical? What would happen if X went away?” (Your goal is to get to a general sense of agreement on whether it was or was not—capture both sides.)

45 minutes
4. Group Weighting. After the individual ratings and discussion, say: “Given our individual votes, I’m hearing ... sure you have general agreement that people can “live with” the story—no thumbs down, and not many sideways thumbs.

• Only a few colors are present on the sticky: “Only a couple people feel strongly about this being an influencer. Does anyone who put a dot up want to make an argument to the room about why it was critical?” (Absent a lot of head nodding and murmured agreement, thank the person and move on after the pitch is made.)

• No colors are present on a sticky: Ignore.

• When relevant, probe for proof: “How do you know, or what evidence do you have, that [you reached a lot of people with the marketing campaign, doctors changed their practices, the backbone was a key driver of the work, etc.].”
Divide people into two groups—one to play the “champions” and one to play the “skeptics.”

1. Say: “I want [indicate group/people] to take on the role of the initiative champion, and provide some concrete examples to us of how we’ve gotten from these strategies to the early and systems changes you’ve prioritized [name them].”

2. Say: “I want [indicate group/people] to take on the role of the skeptics, who believe the prioritized alternate drivers [name them] created the population change.”

3. Say: “Each side needs to provide their best ‘proof’ using concrete examples.”

4. Say: “I encourage you to channel folks who are not in the room—groups who have either championed or been critical of the work of the initiative over the years. Also, it may be helpful to consider what those most directly impacted by [type of issues] might think.”

5. Give each group seven minutes to make their case.

6. Listen and acknowledge. If you feel there’s an area where little attention has been paid for example, data suggests an external driver is important) take the roles of the skeptic/champion and ask people for data for/against. When relevant, probe for proof: “How do you know, and what evidence do you have?”

Test the whole “story” for 25 minutes.
### Wrap-up

1. **Ask:** “One last question as we’re wrapping up: To what extent do you feel like the perspective we have around the table represents the full range of stakeholders who can speak to the initiative? Can everyone give a thumbs up (full range) sideways thumb (some key perspectives are missing) or thumbs down (many perspectives are missing)?”

2. **Ask those who were thumbs down or sideways:** “Who could share alternate perspectives that would help us gain a better picture of what’s happening in the initiative? Please provide names and contact info” (for additional verification if needed)

3. **Thank them for the dialogue.** Let them know how the results will be used and what they’ll receive back and when.
Appendix C

Single-Case Spreadsheet

<table>
<thead>
<tr>
<th>Change Hypothesis</th>
<th>[Hypothesis 1]</th>
<th>[Hypothesis 2]</th>
<th>[Last Hypothesis]</th>
<th>[Overall TOC]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change Hypothesis Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alternate Explanation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alternate Explanation Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analyst Confidence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Passed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Multi-Case Example: In a multi-case effort, after conducting tests at the case level, as illustrated above, you'll likely want to look at patterns across sites for hypotheses that are at similar parts of the theory of change. In the Collective Impact Study (CI Study), for example, we were able to look for patterns across sites about how collective impact conditions led to early changes, how early changes led to systems changes, and how systems changes led to population-level impacts. On the next page is a sample of what the spreadsheet looked like when looking for patterns across similar types of hypotheses.
### Collective Impact Conditions to Early Changes Hypotheses

<table>
<thead>
<tr>
<th>Site</th>
<th>Site 1</th>
<th>Site 2</th>
<th>Site 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Passed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strength of Collective Impact Implementation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Types of Early Changes Achieved</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Types of Alternate Explanations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analyst Confidence</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Glossary

Alternate Explanations

External Drivers

Case

Inferential Strength

Rubric

Change Memo
How to Do Process Tracing

References


About the Authors

This piece was collaboratively developed by Jewlya Lynn (PolicySolve), Sarah Stachowiak and Jennifer Beyers (ORS Impact).

Jewlya Lynn is a facilitator, strategist, advisor, and researcher who works with leaders dedicated to making a difference in the world by solving complex, systemic problems. Her recent work has focused on nuclear threat reduction, forced labor and human trafficking, education systems reform, and climate change.

Sarah Stachowiak loves to help people seeking social impact wrestle with the gnarly, hard to assess ideas and how to measure and learn more along the way. Her consulting work spans a variety of topics from arctic fisheries to K-12 education reform, and types of strategies, such as building fields, changing narratives and shifting ecosystems.

Jennifer Beyers takes deep satisfaction in applying her long-honed design and analysis to support clients’ social impact work. Her work has focused on youth development and capacity-building in education, public libraries, and Jewish philanthropy.
This piece was written and published with a lot of support along the way. We particularly grateful to Houa Lee of ORS Impact and Julia Coffman of the Center for Evaluation Innovation for reviewing an early draft and providing feedback to help us make this most useful for other readers. We also very appreciative of Dr. Derek Beach of Aarhus University, political scientist and expert in process tracing methodology, for providing a review from a deeper methodological point of view. We were grateful for all the feedback, we, the authors, take full responsibility for the content within.

We also grateful to the Center for Evaluation Innovation for supporting editing and design, graphic designer Kurt Niedermeier for his beautiful and visually accessible layout, and Dragonfly Editorial for making the content clear and consistent.
Thank you.