

Understanding and Communicating CfA's Impacts: A Guide [WIP]

Amelia Hoover Green and Laurel Eckhouse
This draft: August 2023

Note to readers: While this guide was still in draft at the time of Code for America's September 2023 layoffs, I am nevertheless extremely proud of our work. I received permission to share this internal draft on the condition that it not be shared further, so please do not cite or distribute. You may also find that links to internal resources are broken.

TL;DR

First and most importantly, this document is a **work in progress!** We welcome feedback.

Purpose of the framework: to show how every part of Code for America's project portfolio contributes to its impacts, and ensure that we can demonstrate all our impacts.

Purpose of this document: to describe Code for America's [theory of change](#) in a way that encompasses all its portfolio areas, and outline a general framework for planning impact assessments across the organization.

Our process: We examined [many frameworks](#) for impact evaluation, and [interviewed CfA colleagues](#) about how they observe CfA's impact in their own work. Finally, we synthesized these findings, attempting to balance best-practices research and the practical realities of our work.

Our working assumptions:

- Across all its projects, CfA is working toward **equitable poverty alleviation**.
- All CfA projects have short- and medium-term impacts that are important in their own right, as well as laying groundwork for more expansive goals.
- Different kinds of impacts require different approaches to measurement and evaluation.

Our key findings:

- Abstracting from any specific project or portfolio, CfA's theory of change is:
Baseline ⇒ CfA's direct work ⇒ CfA's direct impacts ⇒ new circumstances/opportunities for impact ⇒ broader changes in policy and practice ⇒ equitable poverty alleviation.
- CfA colleagues describe their theories of change somewhat differently across disciplines and programs, but these visions very seldom conflict. This is great news.
- A major point of agreement: both the quality of our products and the quality of the relationships we build are vital to our impact.

How we'll use this framework: This document outlines the impact framework, but may not be detailed enough for some readers. Throughout the document, we link to detailed [Appendices](#) for readers wishing to dive a bit deeper. In the coming weeks we also expect to create a practical guide.

Last but not least: in several places, you'll see guiding questions for implementation in gray boxes like this one.

Background

Sometimes, understanding the impact of CfA's work is straightforward (conceptually, if not logistically). In other situations, estimating impact in a way that balances rigor, engagement, program priorities, and our institutional goals and values is much harder. For example: How much does GetCalFresh influence practices in other states? How do we talk about impact when non-partner states adopt CfA best practices? Does GetYourRefund improve the health or educational prospects of kids in participating families? What can we tell donors about our contributions to broad social changes?

Some of our most important impacts are less direct, occur over longer timescales, or are difficult to quantify. As CfA grows—both in size and in scope of work—we need a framework for impact evaluation that encompasses many different kinds of work.

During Q4 of 2022 and Q1 of 2023, we examined a wide variety of theories of change/innovation, as well as impact evaluation strategies from government, science, tech incubators, and non-governmental organizations. (Readers can find more detail in the [Resources Appendix](#).) Next, to get a clearer picture of “impact” across CfA's portfolio areas, we conducted interviews, asking respondents about how they observed impact and about “missing impacts”—important CfA contributions that were difficult to measure or describe. (Readers can find more detail in the [Contributors Appendix](#).)

Based on this background research, we articulated a multi-pathway theory of change describing how (we believe) all CfA projects contribute to equitable poverty alleviation. (See a visualization of the theory of change [here](#), including both the overarching theory of change and some specific examples.) **The theory of change, in brief:** Achieving our long-term goals requires both concrete successes and long-term relationships of trust. Together, these elements create new circumstances, like broader awareness, changing public expectations, adoption of best practices, and organizational buy-in, which open the door to more expansive direct work. In the long run, these changing circumstances lead to broader political and policy changes, like growing advocacy coalitions, increased state capacity, and legislative and regulatory changes. All these set the stage for structural impacts, like equitable poverty alleviation.

Impacts across the theory of change

In this section, we describe in broad strokes how CfA's work contributes to equitable poverty alleviation. We expect that the elements of the theory of change will be ordered and organized differently across CfA portfolios. For example, GetCalFresh's initial direct impact laid the groundwork for state partnership and broader influence, whereas shrinking the criminal legal system can require policy changes *prior* to having direct impacts. In Safety Net Innovation Lab projects, partnership development generally precedes both direct impact and policy influence.

Because we expect this framework to cross CfA portfolios, we have introduced some new terms. Most importantly, *participants* means anyone who interacts directly with our work, including agency workers, administrators, childcare providers, applicants, clients; *partner organizations* might be states,

agencies, or other non-governmental organizations. (The [Definitions Appendix](#) includes these and other, more detailed definitions.)

To begin understanding how a project “shows up” in this framework, ask: What is our partner organization? (A state? An advocacy organization?) Who are the participants? (Clients? Workers?) What is the most direct work we’re doing for partners and participants?

Baseline

Measuring baseline is a key part of assessing impact. “Baseline” might mean different things to different projects, but always includes measures of poverty, inequality, and social mobility in the relevant jurisdiction(s); participant and program experience, outcomes, and pain points; and areas of priority among participants, workers, and organizational decision-makers. (See the [Indicators Appendix](#) for potential indicators/measurements of both baseline status and all the outcomes discussed below.)

Direct work

All CfA work streams have internal quality standards, separate from any external impact of the work. Quality measures vary across projects and disciplines, but might include technical performance, accessibility features, evidentiary standards, research ethics, best practices in communication or marketing, and so on. High quality work is a prerequisite for positive impacts.

Our project teams are multidisciplinary. What are indicators of quality for each discipline involved in the project? How does high quality work set up our external impacts?

Direct impacts

These are the impacts we most directly control. Notably, some are simple to measure using traditional quantitative measures; some are not, and will require interviews, documentary evidence, or other qualitative information.

- Our work **improves individual participant (client, member, caseworker, advocate) experiences and outcomes**. We might measure completion rates, subjective experience of an application or service, overall approval, trust, wait times, or times to completion.
- Our work **directly improves equity** by mitigating participation gaps related to discrimination and exclusion, like differential participation among demographic or geographic groups.
- Our work **improves outcomes for partner organizations**, like organizational reputation, efficiency gains, or lower staff turnover.
- Our work directly improves our **relationships with key partners** and **CfA’s organizational credibility**. Stakeholders notice, engage with, and trust CfA’s work.
- CfA continually improves **knowledge and expertise** about human-centered government, both internally and among participants and partners.

Many of us are used to thinking about direct impacts like funnel metrics, approval rates, or benefits distributed. Begin assessing impacts on relationships, knowledge, or expertise for your project by asking: Within our partner organization, which relationships are most important, and why? What specific knowledge and expertise do we want to provide?

New circumstances and opportunities

We expect our direct impacts to lead to new circumstances, and new opportunities for direct work, amplifying CfA's impact.

- CfA capacity-building work “sticks.” People and organizations **apply, adapt, and expand** the skills, practices, and knowledge they learn with/from CfA beyond the organization's direct engagement with CfA.
- Our direct impacts lead to **new work opportunities**, like inquiries from new organizations, contacts with higher-level decision makers, broader scopes of work, increased funding, or new domains of work.
- Partner organizations' positive opinions of CfA lead to **influence in non-partner organizations** and contexts.
- Among organizations that adopt CfA best practices, we observe increases in **public trust and public expectations**.

Broader changes to politics, policy and practice

CfA's goal is capacity-building and agenda-setting. We are not interested in becoming a long-term vendor of government services, and we do not seek to influence public opinion directly. Instead, as our influence grows, we aim to build interest, knowledge, and expertise among partner organizations, individual participants, and the broader public.

- CfA's work contributes to **public understanding of, and engagement with**, policies and practices that support equitable poverty alleviation. We want to show what works.
- By demonstrating best practices, CfA contributes to changing public expectations about the quality and value of government services. Changing public expectations **strengthen incentives** for policy-makers and administrators to support and implement equitable poverty alleviation strategies.
- We observe long-tail policy feedback effects: advocacy, policy implementation, and legislation can be mutually reinforcing. As a result, **new legislation** and **administrative and regulatory changes** emerge to support equitable poverty alleviation.
- Following policy changes that expand or equalize access, we expect backlash and retrenchment efforts. However, CfA's work creates evidence and communities of practice that support **broad, resilient advocacy coalitions** that can weather these changes and reinforce policies that address inequality.
- State **redistributive capacity increases**, improving access to public resources and reducing the administrative burdens that drive inequitable outcomes. In particular, redistributive capacity increases *relative to other forms of state capacity* such as coercion and incarceration.

Structural changes

Over the longer term, we expect that mutually reinforcing changes in politics, policy, and policy implementation will contribute to **equitable poverty alleviation**. This is a structural change that requires and reinforces other structural changes, like improvements in health equity, equal access to high-quality education, and equal access to justice.

Connecting the dots

Identifying and measuring outcomes is a key first step to demonstrating our impact. However, we also need to “connect the dots” between our work and the impact we’d like to claim, a task that becomes more difficult as our impacts become broader or less direct. (For detailed descriptions of issues in causal inference and research methods, see the [Methods Appendix](#).) This section outlines steps for demonstrating causal connections in real-world conditions. Importantly, while the section is written as *if* all our impact assessments were fully planned ahead of implementation, we know CfA projects sometimes pivot unexpectedly or have unanticipated outcomes.

Know your story

First, identify the outcome(s) of interest. What changes from the baseline do we expect to create or contribute to, in what order? What actions or decisions are necessary to move from one step to another? Then, diagram a concrete, specific causal story. For example, we might write:

CfA product in state X ⇒ Direct impact ⇒ Discussion in professional networks ⇒ State Y opportunity + CfA outreach ⇒ State Y partnership ⇒ Positive partnership experience ⇒ Advisory relationship ⇒ Influence with key decision-maker(s) ⇒ Decision ⇒ Policy implementation ⇒ Impact in state Y.

Clarifying exactly how an outcome is expected to occur (or did occur, or whether it might occur via multiple pathways) guides us toward more concrete understandings about what information can demonstrate the importance of our work to the outcome. (This is also a useful moment to think about potential stumbling blocks, allies, or requirements from outside the organization that might be relevant.)

Consider other stories

There are always other things going on: coincidences, competing influences, you name it. Laying out alternative explanations shows which parts of a causal theory are most important to demonstrating CfA’s causal role. For example, one plausible alternative to our *CfA product in state X ⇒ ... ⇒ Impact in state Y* pathway from above is something like:

Candidate whose viewpoint happens to align with CfA ⇒ change of administration in state Y ⇒ Change in agency leadership ⇒ Decision ⇒ Implementation ⇒ Impact.

In this story, CfA might still develop an advisory relationship with the decision-maker, but CfA did not *cause* the policy change in question; the new administration was always going to implement the change favored by CfA. If we didn’t contribute to a given change, we shouldn’t claim it.

Test the story

With a concrete understanding of our impact story (and plausible competing theories) in hand, the next step is figuring out which story is the right one. In the example above, the final links in the chain are similar. Whether we conclude that CfA caused the impact depends on upstream information about the decision-maker and their decision process. The key question for understanding the causal relationship (if any) between CfA's work and the outcome is: "how did this official come to this decision?" Knowing this, we can build an assessment plan that includes public statements, political platforms, stakeholder interviews, and other qualitative information about decision processes.

Tell the story

Many of our impact stories are less straightforward than this example; some causal stories include many opportunities for random chance, new circumstances, or political opposition (for example) to intervene. Our goal is to make strong, rigorous claims and communicate them engagingly, and this often means combining multiple chains of impact, forms of evidence, or ideas of causation. Our goal is a statement like "X million people across Y states are served by benefits systems built on CfA's model" that is supported by rigorous evidence.

Conclusion: Implementing the framework

CfA faces some unique impact-assessment challenges because of the breadth of its portfolio—and its ambitions. Some key impacts are difficult to identify and measure; some are slow, incremental, or indirect. This document, its appendices, the (forthcoming) Impact Assessment Planning Guide, and our overall research practice are designed to help teams build actionable, rigorous, compelling impact evaluation plans for any CfA project, and to keep long-run change on our radar even as we push for moment-to-moment excellence. Here are some elements of the implementation strategy:

- Gathering and integrating feedback. This document is—and is likely to remain—a work in progress. Comments, questions, and suggestions are always welcome, either in this document or via email or Slack. After an initial feedback period (roughly April 15–May 15, 2023), we hope to develop a regular cycle for incorporating feedback.
- Aligning with senior leadership and the Board about systems change measurement and reporting, and building these into our standard Board metrics.
- Building a quick, simple way for anyone in the organization to report an impact or outcome, whether or not that outcome was an expected one. This is especially important when the outcomes we see are "our words in a benefits administrator's mouth" or "tech vendor X successfully copied our work," rather than "the dashboard shows this many users."
- Training. This is a high-level document. We don't expect teams to dive in without assistance. Instead, we expect to roll out workshops to guide teams through the process of developing theories of change and creating impact assessment plans.

Appendices

(The appendices are arranged in alphabetical order by title.)

Contributors

Many CfA colleagues contributed, directly and indirectly, to this project. An initial meeting with members of the data science, qualitative research and product management teams led to a broader set of interviews with colleagues across the organization. Throughout, we've received invaluable help and feedback, both formally and informally. If you'd like to have a conversation about this draft, or about impact more generally – or if you think there's someone important missing from the lists below – please contact [Laurel Eckhouse](#) or [Amelia Hoover Green](#). We're also happy to receive written feedback or reflections via Slack or email.

Initial meeting attendees:

- Laurel Eckhouse
- Amelia Hoover Green
- Eric Giannella
- Blair Read
- Fahad Quraishi
- Leah Bross
- Nicole Rappin

Interviews

- Fahad Quraishi ✓
- Leah Bross ✓
- Kerry Rodden ✓
- Francesca Costa ✓
- Ben Treviño ✓
- Mari Miranda ✓
- Sara Soka ✓
- Dustin Palmer ✓
- Mike Cowden ✓
- Andrew Xie ✓
- Lauren Spiegel ✓
- Kelly McBride ✓
- Hashim Mteuzi ✓
- Danny Mintz ✓

Definitions

- *Participants*: individuals who interact directly with our work. Participants include beneficiaries/clients, agency workers, agency leaders, service providers, and so on.
- *Partner organizations* are the larger groups we work with, like government agencies, non-government organizations, or academic institutions.
- *Experiences*: How workers, clients, and other participants feel about their interactions with programs or organizations. This includes feelings of satisfaction, dignity, self-efficacy, perceived burdens, and overall approval.
- *Outcomes*: Whether workers, clients, and other participants achieve specific goals/milestones in their interactions with programs or organizations. Outcomes include program awareness, knowledge, acceptance, participation, renewal/retention, and timeliness.
- *Equity/equitable*: In this document we use “equitable” to mean policies and practices that address historical exclusion and discrimination. When work is equitable, it does not exacerbate—and should narrow—gaps associated with exclusion and discrimination. Our goal is to measure equity alongside broader impacts at every stage of our work.
- *Poverty*: Poverty is not having enough resources to access basic necessities of life, such as adequate food, clothing, housing, health care, education, access to justice, or autonomous decision-making.
- *Poverty alleviation*: Poverty alleviation happens when the effects of poverty are diminished, whether via money aid or via access to other key resources.
- *Equitable poverty alleviation*: Diminishing the effects of poverty in ways that do not exacerbate, and should narrow, gaps associated with historical exclusion and discrimination.
- *People helped*: The number of people who used our product, program, or service.
- *People with outcomes*: People who received a benefit via our product, program, or service.
- *Lift*: People who received benefits via our product, program, or service—who otherwise would NOT have received it.

Indicators

In the table below, we list a variety of indicators that might be used to assess one or more of the goals listed above in [Impacts across the theory of change](#). These are grouped by unit of analysis, i.e. by the kind of thing or group (individuals, organizations, geographies, etc.) whose characteristics we’re measuring. This list is far from exhaustive, but our hope is that the suggestions here inspire creative thinking about other potential sources of information. Notice, also, that **these are indicators of outcomes, not indicators of causes**. To show that CfA caused or was necessary to a given indirect outcome, consider the [Connecting the dots](#) section of the main document, and/or see a “worked example” for GCF, below in the Appendix [Using the framework](#).

Concept	Potential indicator(s)	Potential data source(s)
Advocacy: organizations center perspectives of marginalized people	% of Board of Directors, staff, etc. by race, gov benefits use, criminal legal system contact; marginalized-person perceptions.	Organizations’ websites/publications, interviews

Concept	Potential indicator(s)	Potential data source(s)
Advocacy: Growth of advocacy organizations around poverty & inequality (and related)	Number, size, and diversity of advocacy organizations around poverty & inequality; Total budget by year of advocacy orgs focused on poverty, equity, reparations	Organizations' internal data; qualitative interviews about organizational resources; academic and NGO research; media mentions; market research
Advocacy: Growing breadth, size, and power of political coalitions around poverty & inequality	Density of network connections among poverty advocates, other advocacy orgs (e.g. students, seniors) and political power holders: meetings taken, invitations to key gatherings, etc.	Academic and NGO research, media monitoring, interviews with other orgs
Advocacy: participant knowledge of advocacy opportunities	% of org/program participants who can name "a way to get involved"	Client surveys, qualitative interviews with CBO stakeholders
CfA access to partner data (trust indicator, scope indicator)	N (%) records, N (%) fields, data warehouse access, partner stakeholder confidence in CfA data management	Comparison of data sharing agreements over time; interviews
CfA employees highly skilled relative to peer institutions	% with terminal degree, % with top 5 skills in field	BLS data, internal data
CfA hands off working, sustainable systems to partners	Minimal downtime at cutover/handoff; partners maintain pace of updates/upgrades; partners report feeling confident; products/services continue seamlessly; budgets sufficient over time.	Comparison of initial plan/contract and system received; partner feedback
CfA employee and manager knowledge of program and discipline specifics/projects	Knowledge of key CfA goals, programs, processes at onboarding, 90 days, 6 months, yearly	CfA internal data
CfA media mentions/appearances increase	N mentions, type of mention (by stakeholder, quote for news story, as org...), reach of source	MarComms internal data
CfA mentions in stakeholder professional networks increase	N mentions in conference/meeting proceedings, % of stakeholders/leaders who report discussing CfA with others	Proceedings, qualitative interviews, referral data
CfA partnership inquiries increase in number and breadth	Number of inquiries, scope of inquiry	Inquiries data
CfA resources widely consulted	Clicks, downloads, time on relevant pages by geography, domain name of user	Website traffic
CfA advisory relationships expand and deepen.	Stakeholders request specific advice; stakeholders request out-of-scope advice; stakeholders broaden scope of advising agreements	Qualitative interviews, referral data, contracts
CfA partnership arrangements expand in size and scope	Budget, scope, length, and projected impact of partnerships over time; projects/partnership expansions; quality of relationship with key partner org personnel	Internal documents/contracts; qualitative interviews
CfA delivers products of increasing size and scope	Budget, staffing, people with outcomes (or systems touched)	Internal documents/ contracts; qualitative interviews
CfA social media engagements increase	Standard metrics: followers, links to CfA documents, profile views, clickthroughs, etc.	Social media engagement data
CfA workers keep up with advances in their fields	% of staff who spend 100% of professional development budget; Staff who obtain new certifications	Qualitative interviews; internal records
Individual/household: increasing trust in government	Standard survey questions re: trust in government, by program participation & demographics	Public opinion survey(s)
Individual/household positive relationships with org/prog workers	Potential applicant, applicant, participant, and worker answers to quality of relationship questions, over time	Qualitative interviews
Individual/household satisfaction with application process	% answering that application was "understandable," "simple," "made sense," was designed with their experience in mind; specific pain points & frustrations	Applicant surveys, UX interviews

Concept	Potential indicator(s)	Potential data source(s)
Individual/household access to state services/contact with state	Neighborhood access to quality schools; crime and violent crime rates; favorability and impressions of police, teachers, mail service, city services, etc.	Census, surveys, municipal budgets; qualitative interviewing
Individual/household application acceptance rate	% of completed applications approved, overall and by demographic group	Application and enrollment data
Individual/household community/political/state engagement	% answering that they went to a community meeting, volunteered, protested, voted, etc.; % stating they expect good service when they contact org/program; qualitative descriptions of (non-) engagement	Survey (both question answers + response rates), qualitative interviews
Individual/household perceived self-efficacy	% stating they understand what they need to do to get and stay on benefits, % who perceive approvals as "random," knowing whom to contact with problems	Surveys, qualitative interviews
Individual/household increase in application completion	% of applications submitted (of those receiving an application number) by program, demographics, etc.	Application data
Individual/household decreasing procedural denials	% of denials for procedural reasons, overall and by demographics	Application and enrollment data
Individual/household: administrative burden declines as reason for not participating	Detailed descriptions of decision-making processes around applying for, enrolling in, and renewing benefits/program membership	Qualitative interviews
Individual/household administrative burden of renewal decreases	N months of benefits prior to required report/renewal, by key groups; % clients reporting renewal was simple; average spell on benefits for eligible people	Agency data; client survey data
Individual/household: wait times decrease	Days between application submission and eligibility determination; days between eligibility determination and disbursement	Application, enrollment, disbursement data
Jurisdiction concentrated poverty	% of PUMAs, ZIPs, Census tracts with high poverty	Census
Jurisdiction Horizontal inequality (e.g. across race)	Household median income by race; household wealth by race	Census
Jurisdiction income inequality	% of yearly income held by top decile, quintile	Census
Jurisdiction overall and demography-specific poverty and food insecurity	% living <= 100% of FPG by relevant demographic groups; % of households reporting food insecurity by relevant demographics	Census
Jurisdiction political control	Partisan control of municipal, county, state, federal legislature and administration; perceptions of bureaucrats' political loyalties among public & clients	News media; interviews
Jurisdiction substandard housing	% of housing units with safety or other violations	County Health Rankings (underlying data)
Jurisdiction unemployment (+ by demographics)	% of working-age adults not employed, not in school, and/or under-employed	Census, Bureau of Labor Statistics
Jurisdiction wealth inequality	% of wealth held by top decile, quartile in jurisdiction	Census
Jurisdiction, Organization/program: Role of CBO's	Overall attitude of workers to CBOs and vice versa; perceptions of "good" CBOs, community perceptions	Qualitative interviews, media mentions
Jurisdiction: stakeholders/decision-makers/people in power over program	People others report you "need to convince," or holding positions of formal decision-making power; spoilers	Qualitative interviews with workers & other stakeholders
Non-partner organizations adopt CfA best practices	Staffing decisions in alignment; language adoption; shifts in budget; administrative or regulatory practices	Media monitoring, stakeholder interviews

Concept	Potential indicator(s)	Potential data source(s)
Non-partner organizations influenced by CfA or CfA partners	Leaders of non-partner orgs know of CfA, “heard good things about” CfA, consulted CfA resources in decision-making	Stakeholder interviews
Organization/program: eligibility by demographics	Estimated proportion of people & households eligible	Internal estimates (Census, FNS, etc.)
Organization/program tech capacity	Existing tech stack, personnel, budget, skills; capacity for growth	Internal documents, systems testing
Organization/program workforce	N workers, yearly turnover by position, satisfaction by position	Partner orgs’ internal data/surveys
Organization/program size, budget, prestige	N staff, N members/participants/clients, yearly budget; leaders’ perceptions of access/prestige among peer orgs (e.g. other state orgs, similar NGOs)	Budget, internal organizational data
Organization/program: decreasing participation gaps	Demographic composition of program participants vs. eligible population vs. general population, over time	Census, internal estimates, surveys, application and enrollment data
Organization/program: variations in service by location	Client satisfaction, approval rates vs. eligibility rates, by office location	Qualitative interviews, program data
Organization/program: implementation of CfA recommendations	Staffing decisions in alignment; language adoption; shifts in budget; administrative or regulatory practices; institutional isomorphism	Media monitoring; qualitative interviews
Organization/program: Actively seeking to lower admin burdens/barriers to entry	Applications for waivers, demonstration/pilot projects, use of flexibilities	Federal agency data, interviews
Organization/program: Knowledge-sharing valued within partner organization	% of workers who report sharing expertise within & across hierarchies, % who report learning something from a colleague lately	Internal surveys, qualitative interviews
Organization/program: Political, ecological, health, leadership crisis resilience	Interruptions in service related to crisis situations; client perceptions of org. adaptations	Qualitative interviews with workers, stakeholders, clients
Organization/program: responsive to participant feedback	% of clients who report org. Is responsive; worker descriptions of client feedback procedures	Surveys, qualitative interviews
Organization/program: Size and scope of outsourcing contracts	Number, budget, and scope of contracts with for-profit contractors	Internal documents/budgets/information
Policy: Adoption of equity metrics	N/% organizations (in jurisdiction, of X type, etc.) measuring and reporting equity data	Stakeholder interviews, publicly available agency data
Policy: Increased redistribution	Over-time change in tax revenue, progressiveness of taxation, proportion of budget dedicated to poverty alleviation	Budgets of relevant jurisdiction(s), policy/academic partners
Policy: Federal agency recommendations	Published recommendations/best practices/policies focus on lower administrative burdens, equity, or other CfA priorities	Media and web monitoring; agency documents
Policy: Reconsideration of means-testing thresholds	Policy-makers and administrators actively consider “living wage” (vs. “poverty”) guidelines	Stakeholder interviews
Political context: Content of campaign literature	Campaign literature discusses poverty, redistribution, equity as priorities. Declines in “welfare-baiting.”	Campaign literature, media monitoring, data from org. partners
Political context: Content of legislative floor speeches	N mentions of poverty, equitable poverty alleviation, benefits	Legislative records
Political context: Legislative progress of bills supporting equitable poverty alleviation	N bills introduced by jurisdiction; N co-sponsors; if bill is repeatedly introduced, new coalition members over time; roll call votes	Data from organizational partners

Concept	Potential indicator(s)	Potential data source(s)
Political context: Legislator, executive, and candidate public statements about poverty, poverty alleviation	N mentions/discussions of poverty and redistribution as policy priority	Legislative records, media monitoring
Public opinion: factual knowledge about poverty and poverty alleviation	% of survey respondents who correctly answer factual questions about poverty, by demographics, geography, and political affiliation	Public opinion surveys
Public opinion: Organization/program knowledge	% name recognition of key anti-poverty/equitable poverty alleviation orgs, by demographics, geography, and political affiliation	Public opinion survey(s), qualitative interviews
Public opinion: Organization/program favorability	% favorability of key anti-poverty/equitable poverty alleviation orgs, by demographics, geography, and political affiliation	Public opinion survey(s), qualitative interviews
Public opinion: Proportion of people stating poverty is a key political priority	% reporting poverty or inequality is major issue, by demographics, geography, and political affiliation	News media mentions, public opinion polling
Public opinion: stereotypes of programs/orgs & beneficiaries	% agreeing with positive vs. negative statements about public benefits & beneficiaries	Public opinion survey(s), qualitative interviews
Public opinions: perceptions and expectations of government services	% favorability of specific government services; % reporting expecting excellent customer service	Public opinion surveys
Stakeholder/leader: Federal agency awareness and consultation of CfA	N federal agency contacts, contracts; quality of federal official knowledge	Internal records, stakeholder interviews
Stakeholder/leader: CfA awareness among key professional groups	% of key groups (benefits administrators, state and federal cabinet officials, legislators) with knowledge of CfA or CfA's work	Surveys, qualitative interviews
Stakeholders/leader: Impressions of power/influence of direct CfA contacts	Respondents report that CfA's key organizational contacts hold sufficient power to steer key decisions/actions	Qualitative interviews
Stakeholders/leader: Partner recommendations of CfA	Partners and former partners report discussing CfA with professional colleagues or recommending CfA; new or potential partners report hearing of CfA in professional networks	Referral data, qualitative interviews
Worker frustrations/pain points	Pages/screens that frequently cause mistakes or "have to be done just right."	Qualitative interviews, observation, surveys
Worker autonomy (vs.org leaders)	Workers report that they feel empowered to solve problems on their own	Qualitative interviews, internal documents
Worker caseload and cases/period	N cases/worker/period; hours worked/worker/case	Internal data, interviews with workers, interviews with managers
Worker caseload management strategies, stress, burnout	Workers report feeling they can't do a good job, there's too much to get done, or they routinely take shortcuts; % workers who have considered leaving recently; % reporting work stress is a major factor in their life	Qualitative interviews about worker strategies and challenges; worker surveys
Workers and administrators knowledgeable about client perspectives	% saying they spend time reading client feedback and/or think about how to solve common client problems	Worker surveys, stakeholder interviews

Resources [WIP]

Theories of change/innovation [WIP]

Diffusion of innovation
Policy windows (Kingdon)
Policy feedback (Mitchell)
Theories of public opinion change
Laboratory of federalism

Frameworks for impact evaluation

We considered several impact frameworks in use across [government](#), [science](#), and [tech incubators/accelerators](#), particularly those focused on [social impacts](#) and [non-quantitative impacts](#). While the frameworks differ, at the most basic level each requires evaluators to specify a theory of change, understand what observations would confirm or disconfirm the theory at each stage, and identify issues, events, or actors that might affect the theory of change “from outside,” either accelerating or blocking progress.

Methods [WIP]

Our current strengths

CfA has done and is doing fabulous work [assessing and narrating our impact](#). This document builds on our existing practice, highlighting opportunities for synthesis across program areas and emphasizing the role of multi-method impact evaluation. For example, we drew inspiration from the Safety Net Innovation Lab [evaluation plan](#), which draws on both qualitative and quantitative evidence to assess tangible and less tangible impacts (e.g. relationship- and trust-building). Similarly, our [Field Guide to Benefits Enrollment](#) (see 2019 version [here](#)) combines quantitative evidence about web accessibility with qualitative information and visualizations about the experience of interacting with states’ online benefits portals.

For day-to-day impact assessment, we have a strong quantitative and experimental skillset across CfA. We are able to leverage randomized controlled trials to estimate the average effect of many of our digital interventions, like variations in text messages, form flows, or communications strategies. For non-experimental evaluations, we are often able to leverage quantitative data for retrospective understandings of, for example, approval dynamics or renewal/recertification completions.

Experiments (aka A/B testing, randomized controlled trials): We can conduct experiments when randomization is both possible and ethical. Experiments are great because they allow us to isolate the causal effect of a given change or intervention. For example, by randomly assigning some clients to (initially) receive text messages about renewals, we can understand how many clients will renew following a text *who otherwise wouldn't*. However, experiments cannot always tell us *why* an experimental effect occurs, or whether the effect is the same across key groups of clients. For this, we might turn to **qualitative interviews** or to **large-N quantitative observational analyses**.

Here, *observational* just means “not experimental.” In the text-message case, an observational study would involve allowing people to opt into text messages, and/or making a universal change at a particular point in time. The concern here can be summed up as “correlation versus causation:” when we let people opt in to text messages, it’s likely that those who opt in are systematically different from those who do not, so we can’t be sure that our text messages are “doing the work” of improving renewal rates (even when rates improve!). The usual toolbox for large-N quantitative analysis is based on **linear regression**. In effect, regression analysis allows us to look at how (much) different factors/variables contribute to a specific quantifiable outcome. While we don’t have the advantage of a clean experimental design, we *can* flexibly investigate the effects of many different factors. For example, we might be able to look at how demographic variables are associated with texting opt-in and renewals success.

CfA has less experience with multi-method impact evaluations, evaluating indirect impacts, and evaluating impacts requiring many coalition partners. Here, we add detail to our process for “connecting the dots” between CfA’s direct impacts and its broader, less direct impacts.

Process tracing

Process tracing is a set of techniques to consider causation in one or a few cases, where the theory of change may be complex or contingent on factors outside our control. Process tracing can be used to generate theories, test theories, or consider individual cases, but here we focus on process tracing to answer the general question “Did CfA influence this policy change?”

Process tracing relies on the elaboration of a causal mechanism between X and Y (as described in section 4.1 of [this explainer](#)) connecting our direct work (X) and the outcome of interest (Y), and including other directions/paths at each step in the causal chain, and other potential influences on Y. (Such as “It was a coincidence;” “They were responding to this other thing.”) Clarity about the causal mechanism allows us to rule out alternative explanations for Y, and to understand what evidence is required to confidently support an assertion that CfA deserves credit for Y.

Frequently, the evidence required for process tracing will be qualitative: interviews, documentary evidence, and so on. One goal of this technique is to target the right people for interviews. Another is to emphasize the role of random contingencies or structural features.

Small- and medium-N comparison

Another strategy for supporting a causal link is comparison. If we see changes in our project jurisdiction that don’t match changes elsewhere, we *may* be able to attribute those changes to our work. However, this requires (at least some) data from non-project jurisdictions. For example, if the negative impacts (confusion, benefits changes) of the end of the federal pandemic emergency are felt less strongly in CfA partner states, then it may be worth making closer comparisons. A gold-standard qualitative comparison is not seeking “proof” that CfA’s interventions were the decisive factor. Rather, it emphasizes that CfA’s interventions (such as outreach or communications with clients) were a necessary, contributing factor.

Analogies/evidence from elsewhere

CfA is a small organization relative to the size of its goals, so it can be difficult to identify “our” segment of large structural changes. Nevertheless, we believe it’s important to explain how our work contributes to the big, structural changes we’d like to see—including, but not necessarily limited to, equitable poverty alleviation.

For example, Social Security’s longevity in the face of multiple moments of crisis is partly explained by its creation of a constituency with a vested interest in its survival. But only partly; another part of the story is the democratization of retirement, which creates a class of individuals with leisure time for advocacy. Still another lies in the destigmatizing effects of more universal entitlements, and the relative ease of coordination around a national policy (as opposed to state benefits). To whatever extent our work creates a constituency, destigmatizes benefits, or supports citizen advocacy, it supports a strong safety net in general.

Using the framework: an example

Here, we draw out one example of the way teams might use this document to create rigorous impact assessments. Here, we consider GetCalFresh (GCF) as an example. (The GCF example is also shown in our [Theories of Change visualization](#).) A first task is to move from the organization’s broad theory of change to project specifics.

Baseline ⇒ CfA’s direct work ⇒ CfA’s direct impacts ⇒ new circumstances/opportunities for impact ⇒ broader policy, political, regulatory change ⇒ structural change.

We start by thinking about the baseline GCF was responding to; then we consider how GCF’s specific narrative of change – its partner organizations, its participants, its sequence of events – shows up in direct impacts, new circumstances and opportunities, and broader changes. It’s important here to recall that the stages of action and change might be reordered, repeated, or cyclical.

Stage	Description	Potential data sources	Methods for connecting the dots	Language
Baseline	<p>Benefits gaps by demographics, language and geography across CA jurisdictions</p> <p>High variation in client experience and outcome by county</p> <p>Difficult to use platform</p> <p>High churn</p> <p>Seemingly arbitrary eligibility decisions</p> <p>No clear opening for state partnership (?)</p>	<p>Publicly available data and documentation</p> <p>Academic research</p> <p>User research</p>	<p>We might want to do some analysis about <i>why</i> we are in the baseline situation, but most of the work is in connecting the dots.</p>	<p>“Participation gap”</p> <p>“Opportunities to serve more Californians”</p> <p>“Make sure all Californians get the benefits they deserve”</p> <p>“Benefits that could be unlocked”</p>

Stage	Description	Potential data sources	Methods for connecting the dots	Language
Direct work (1)	Accessibility and usability metrics Initial technical specifications	Standard tech metrics	NA	“An easier experience applying because X, Y, Z” “X% faster”
Initial direct impacts	Size of GCF user base, month-over-month and year-over-year increase Volume and content of user feedback Increasing volume of GCF apps to counties State/county awareness of GCF as "competitor"	In-house data Interviews Text analysis of county stakeholder communication	Experiments: how did [feature] affect [application completion]? Large-N observational re: correlates of app volume by county ⇒ inferences about potential causation Qualitative interviews, media monitoring	People helped, people with outcomes, equity statistics “An easier, more dignified experience” “People who started applications were X% more likely to finish them” “Helped X Californians apply faster, saving Y hours.”
Direct work (2)	Increasing contacts w state and county stakeholders Continued improvements to GCF features, accessibility and usability Growth and performance of client success team Increasing contacts with CBOs, advocates	Contact counts Interviews Tech metrics/specs CfA Hiring data Publicly available data; some county data	NA (measure quality of work)	“Building a relationship with X counties means we have the opportunity to help Y people more directly.” “Lowered our wait time for assistance to X.”
Additional direct impacts	Statewide awareness of GCF Increased state interest, cooperation, and trust (esp. In key counties) Improved client outcomes/ experiences: higher satisfaction and approval rates, lower churn, less variability across counties. Improved worker outcomes/ experiences: lower turnover, higher satisfaction, etc. Increased applications and improved outcomes/experiences among historically excluded groups	Surveys (public opinion, client) Contact counts Interviews Access to state data Application and enrollment data	Correlate public opinion changes with timing of GCF expansion, etc. Interview Q’s about reasons for contacting CfA, willingness & ability to work with Large-N analysis of impact of GCF improvements on client success by demographics	“Of the X% of Californians who know of GetCalFresh, Y% approve of the job GCF is doing.” “We met with X stakeholders from Y counties, opening the door to helping Z more Californians” “Clients who submit via GCF report X% higher satisfaction than those who report via legacy systems” Lift, people with outcomes
New circumstances	Broader public and policy-maker awareness of GCF, CfA High usage, satisfaction and expectations of GCF	Interviews (state partners, advocates, clients, public) User research Media mentions	Large N analysis of correlation between public awareness and GCF outreach campaigns User interviews suggest GCF (not something else) has changed their expectations about government programs	“GCF now handles X% of all applications for SNAP in California.”

Stage	Description	Potential data sources	Methods for connecting the dots	Language
	<p>Statewide adoption of GCF with relatively high acceptance/approval @ county offices.</p> <p>State interest in integrated benefits portal</p> <p>Relationships of trust with county & state decisionmakers</p>	Client surveys	Referrals state they observed our successes in CA	<p>“We had productive conversations with county-level stakeholders in X of Y counties, serving Z% of all Californians.”</p> <p>“New relationships allowed us to expand our outreach to X% more clients.”</p> <p>“X% of benefits administrators say CfA is a trusted partner.”</p>
New work opportunities	<p>Requests for consultation from new state partners</p> <p>CfA as key partner in BenefitsCal development</p> <ul style="list-style-type: none"> - Training - Engineering/tech specs - Sustainability <p>Broader data access to study effects of outreach efforts & GCF access</p>	<p>Survey distributed to new inquiries</p> <p>Contract details</p> <p>Worker training participation & outcomes</p> <p>Interviews with tech team</p> <p>Assessment of BC at cutover</p> <p>DSA details</p>	<p>Referrals discuss CfA successes</p> <p>BenefitsCal stakeholders discuss GCF knowledgeably</p> <p>Workers’ knowledge improves and sustains improvement after CfA training</p> <p>In DSA negotiations, CfA track record with security is relevant factor</p> <p>...plus all the experimental and large-N work that shows direct impact!</p>	<p>“Policymakers serving X states, collectively providing benefits to over Y recipients, are looking to GCF as a model.”</p> <p>“Improved state workers’ knowledge and perceived self-efficacy by X%”</p> <p>“By providing X million data points for our analysis, California shows that CfA is a trusted partner whose research has national impact.”</p>

Stage	Description	Potential data sources	Methods for connecting the dots	Language
Broader political, policy, regulatory changes	<p>Citizen advocacy/activism</p> <ul style="list-style-type: none"> - Student advocate interest/take-up - Increased client expectations <p>Political, policy, regulatory changes</p> <ul style="list-style-type: none"> - Increasing support for human centered government principles among CA leaders - Other states view GCF/BC as model <p>State capacity</p> <ul style="list-style-type: none"> - Successful handoff of BenefitsCal - Sustainability of BenefitsCal - Privatization-proofing: BenefitsCal operated to high standards by the state. - State leaders express alignment on HCG 	<p>Client satisfaction and outcome data; public opinion data; number of waivers sought;</p> <p>FNS takeup of state innovations; legislator scorecards; institutional isomorphisms; data on outsourcing contracts; interviews</p>	<p>Compare advocacy orgs' size, knowledge, and agendas before and after GCF widely adopted.</p> <p>Interview student activists to test whether GCF inspired anti-poverty activism/alliances on campus</p> <p>In survey data people with greater awareness of public benefits also have higher approval</p> <p>Contact with GCF is associated with increased trust in government, all else equal</p> <p>Workers & admins say they feel more confident/competent to provide high-quality service without outsourcing than before</p>	<p>"Our work in California has increased the number of people involved in benefits activism by X"</p> <p>"Our partner states collectively applied for X waivers, which have the potential to impact Y people."</p> <p>CfA's performance and lobbying have moved X legislators to co-sponsor Y bills..."</p> <p>"Our interventions to improve state capacity resulted in savings of over X dollars in outsourcing contracts"</p> <p>"X% of state leaders expressed agreement with principles of human centered government, a Y% increase over 5 years."</p>
Structural changes	<p>Declining post-benefits poverty in California, esp early-adopter jurisdictions</p> <p>Declines in poverty targeted to historically excluded groups</p> <p>Longer spells on benefits for eligible households</p> <p>Improved health and educational outcomes for poor people</p> <p>Stronger public support for benefits</p>	<p>Publicly available data (Census, state data); data via DSA w/ California</p> <p>Public opinion data</p> <p>Interview data (in-house, with stakeholders, via academic partners)</p> <p>Secondary literature</p>	<p>Across many measures, compare change over time in early adopter counties/consortia vs. later adopters</p> <p>Large N: correlation between timing of GCF/BC rollout and public opinion changes, all else equal</p> <p>Legislators/leaders report they feel increasing pressure from advocacy coalitions</p>	<p>"We're a key part of the coalition that raised X Californians out of poverty over Y years."</p> <p>"In our partner states, declines in poverty are most strongly felt among historically excluded groups."</p> <p>"Our work directly raised X people out of poverty, and indirectly contributed to Y person-years of increased life expectancy."</p> <p>"Data show that effective public benefits programs drove an X% increase in favorability ratings for benefits programs."</p>

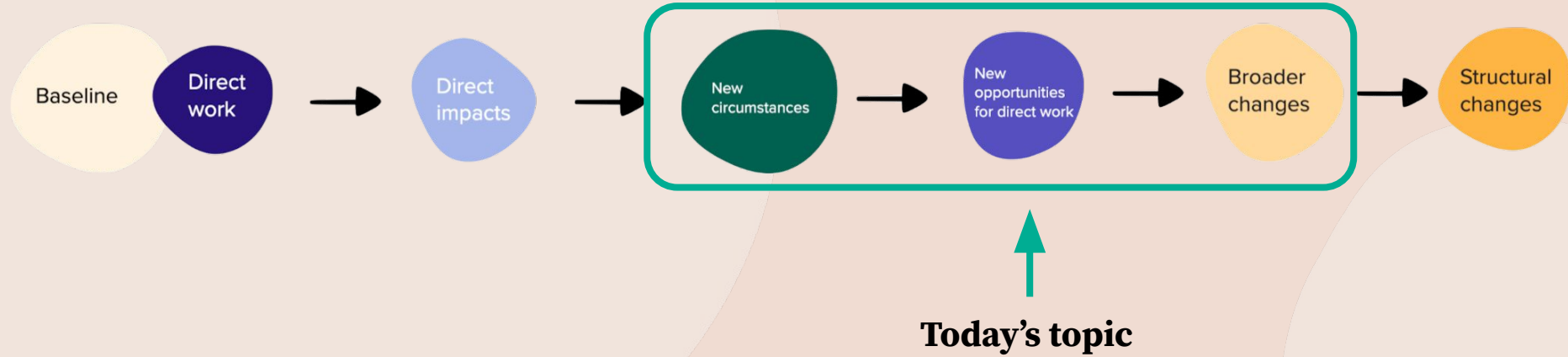
Measuring CfA's indirect impacts

Simple to complex cases

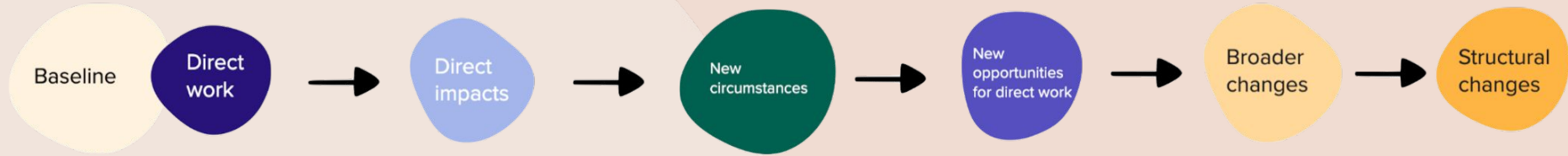
*Amelia Hoover Green & Laurel Eckhouse,
Code for America Data Science*

September 2023

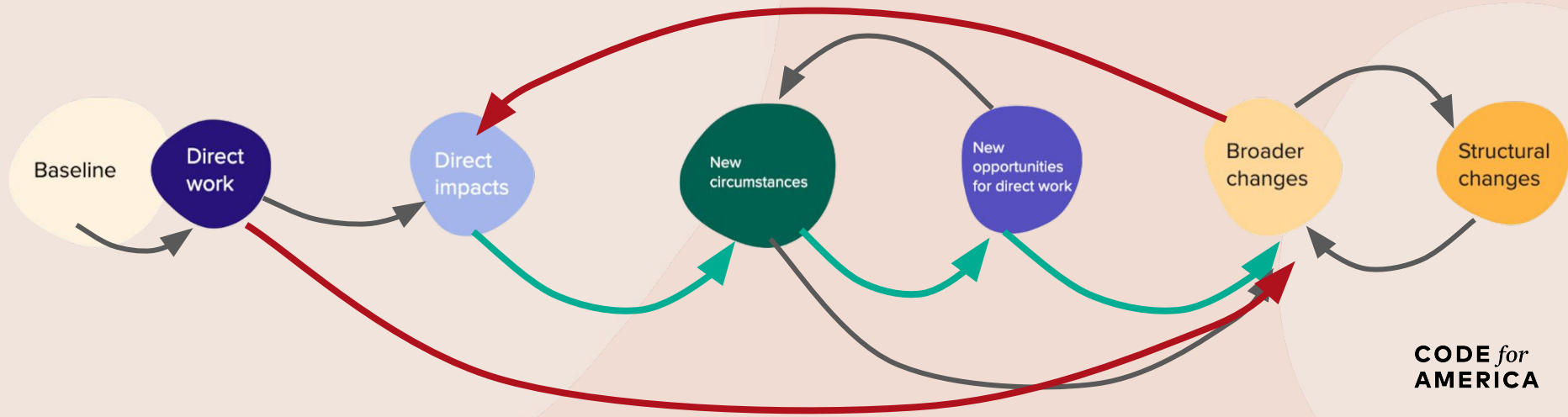
A cross-portfolio theory of change



A cross-portfolio theory of change



...with many pathways for impact



Impacts at each step

Our work

Internal measures of quality, separate from (but necessary to) impact

Direct impacts

Direct improvements in participant & partner org experience and outcomes, including relationships with CfA

New circumstances

Growing trust, credibility, and demonstrated impact creates new attention & opportunities

New direct work

Broader scope of work, opportunities to advise higher level decision-makers, greater access

Broader political/ policy changes

Changing incentives for policymakers and advocates, leading to changes in policy, public opinion, and public expectations.

Equitable poverty alleviation

Changing public opinion and a durable political coalition create sustainability for policy changes.

Connecting the dots

Our work → **Direct impacts**

Experimental inference for participant impact; observational analyses + process tracing to rule out alternative causal hypotheses.

Direct impacts → **New circumstances**

Process tracing (stakeholder and participant interviews, documentary evidence, administrative data) to rule out alternative causal hypotheses.

New circumstances → **New direct work**

Process tracing (stakeholder interviews, contract data, budget/SOW data) to rule out alternative causal hypotheses.

New direct work → **Broader political/policy changes**

Combination of structured comparison, process tracing and quantitative analysis (e.g. regression, quasi-experiment) of public opinion, policy, and regulatory data to rule out alternative hypotheses.

Broader political/policy changes → **Equitable poverty alleviation**

Framework goals

1. **Mission-driven *and* concrete:** specific milestones build to long-term goals.
2. **Practical *and* flexible:** teams find it simple and generative to apply and adapt the framework.
3. **Compelling *and* rigorous:** supports and transparently defends compelling claims.
4. **Cross-portfolio:** recognizes the importance of all CfA workstreams to our end goal.
5. **Equity-focused:** measures of equity included from the beginning.

Non-goals

- Quantitative data
- Fully developed theory of change

Framework goals \Rightarrow Indirect impact approach

1. **Mission-driven *and* concrete**

- \Rightarrow Clearly connected to equitable poverty alleviation.
- \Rightarrow Testable, tangible, and clear.

2. **Practical *and* flexible**

- \Rightarrow Choose 1-2 key indirect impact hypotheses per project to plan around.
- \Rightarrow Capture unexpected impacts via retro boards or 'serendipity tracker.'

3. **Compelling *and* rigorous**

- \Rightarrow When working in coalition or estimating indirect impact, we identify the full impact.
- \Rightarrow Then, use rigorous qualitative measures to describe exactly how we contributed.

Types of indirect impact

- **Post-handoff sustainability:** Following a direct engagement, a partner maintains products and practices we recommended as well as direct service improvements.
- **Working in coalition:** CfA is a key member of a broader coalition that creates impact via policy change or policy implementation.
- **Showing what's possible** / “demonstration effect”: Outside of a direct engagement, an agency adopts products or practices due to CfA influence.
- **Shaping the debate:** Outside of a direct engagement, agencies adopt CfA metrics, measures, language, or frameworks.
- **Public expectations:** The existence of CfA projects, products, and approaches changes public expectations about government performance.

Measuring indirect impact: options

- **People helped, benefits delivered:** Cumulative count of people receiving help and dollars delivered post-handoff or due to showing what's possible. Counting stops when systems no longer follow CfA best practices.
- **People affected, benefit dollars affected:** Number of people receiving benefits, and benefits delivered, via systems that adopted CfA language or metrics. Also: proportion of eligible people, proportion of benefits delivered.
- **Systems changed:** Number of X (agencies, states, policy guidance memos, laws, regulations, etc.) with clear evidence of CfA influence.
- **Overall sustainability:** Proportion of engagements that result in sustained changes.

Visualizing post-handoff impact

