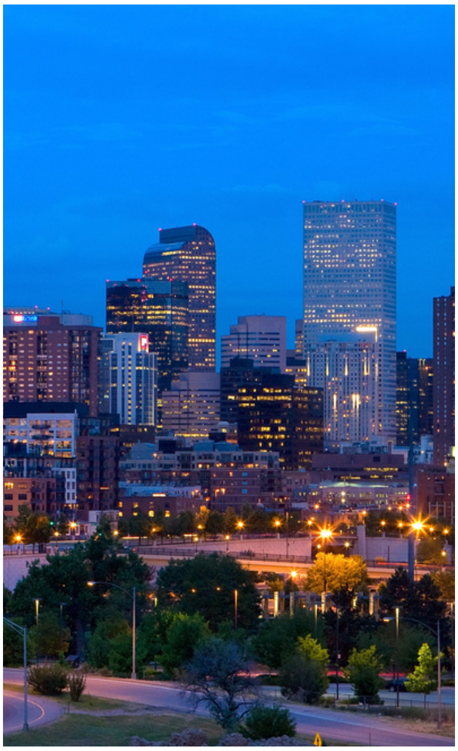
CEEMI COLORADO EQUITABLE ECONOMIC MOBILITY INITIATIVE

Building Evidence Questionnaire and Workbook

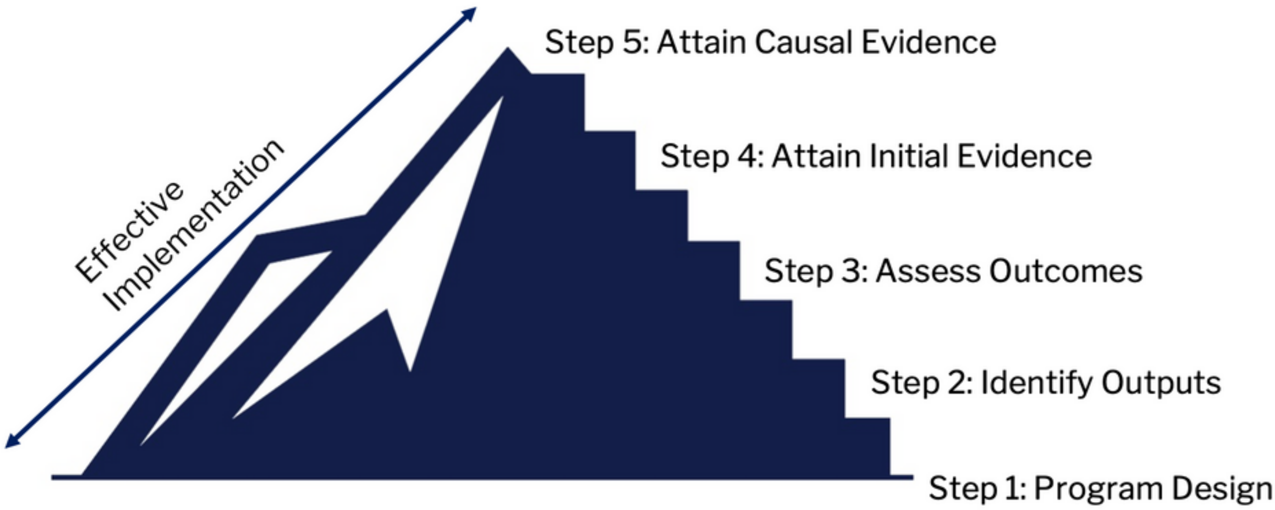
Background & Introduction



The Colorado Equitable Economic Initiative (CEEMI) is a nonprofit focused on scaling effective, proven workforce and postsecondary programs for Coloradans in or near poverty. CEEMI prioritizes working with programs, services, and pathways that have been (or want to be) rigorously evaluated. The organization focuses on leveraging public funds, including through advocacy, and providing technical assistance and support to state and local government agencies and nonprofits/providers in Colorado seeking to measure impact and build evidence of effectiveness.

Colorado is an exemplary state for prioritizing and advancing the use of data, research, and evidence through multiple avenues. One example of the prioritization of evidence through state policy is the development of the Colorado Evidence Continuum. The Colorado Workforce Development Council (CWDC) contracted with CEEMI to develop a tool for other CWDC grantees to better understand evidence and evaluation, and CEEMI identified the Colorado Evidence Continuum as a logical starting place.

The Evidence Continuum



The Evidence Continuum identifies five steps to build evidence; CEEMI has developed this tool to serve as a guide to help your organization identify where your program falls on the continuum and how you can advance from one step to the other. It is important to keep in mind that not all programs can, will, or need to advance to step five, but regardless of whether a program advances toward more rigorous evaluations in steps four or five, steps one through three still need to be completed first.

This tool aims to help your organization build evidence about whether your program is being implemented as designed and is improving important outcomes for the clients you serve. Having this kind of information is not only helpful for understanding your program's impact but can also make your program eligible for funding, both public and philanthropic, that prioritizes supporting programs shown to be effective through rigorous evaluation(s).

Definitions

Program: A program can be any practice, treatment, service, tool, etc., with replicable elements that is hypothesized to improve one or more specified outcomes for a clearly defined target population. *For example, Per Scholas is a rigorous skills training program for individuals typically left out of tech careers and connects skilled talent to leading businesses.*

Outputs: The activities produced/completed as part of your program's implementation with the goal of improving participants' outcomes (defined below). Outputs typically measure reach and enrollment related to your program activities and whether key components of your program are being implemented as intended. *For example, Per Scholas has over 17,000 graduates of their program.*

Outcomes: Measures of what your program is trying to improve for its target population. For example, the percentage of people in a new job within 6 months of training or annual earnings after receiving the training. *For example, typical Per Scholas graduates have a pre-training income of $10,000 and a post-training income of $42,000 or more.*

Impact: A measure of the change in outcomes that can be attributed to your program and not to other factors. Importantly, to measure your program's impact, you need to evaluate it in a study that compares their outcomes to those of a comparison group of highly similar individuals who did not receive the program. *For example, Per Scholas, participants have been shown in randomized controlled trials to earn $4,000-$6,000 more per year compared to a randomized control group of similar individuals who did not receive the program.*

Diversity, Equity, Inclusion, and Justice (DEIJ): Policies and programs that promote the representation and participation of populations that have been historically underrepresented or subject to discrimination because of their background or identity. Your program or organization may have its own DEIJ framework, which should be incorporated throughout your evidence building process.

Comparison/Control Group: These terms may be used interchangeably. A group of individuals who are highly similar in observable characteristics to program participants but who do/have not receive(d) the program being evaluated and instead receive(d) "services as usual" (i.e., any other available services besides the program being evaluated). As noted above, comparison/control groups are necessary for measuring a program's impact; this is because they provide an estimate of what would have happened to program participants if they had not received the program.

Quasi-Experimental Design (QED): A type of comparison group study whose comparison group is identified/formed through methods other than random assignment.

Randomized Control Trial (RCT): A type of control group study that randomly assigns some individuals eligible for a program to either a group that is offered the program (i.e., a program/treatment group) or a control group that is not. Random assignment provides confidence that there are no differences between the program/treatment and control groups. As a result, any differences in outcomes between the groups can confidently be attributed to the program as opposed to other factors. RCTs' ability to create groups that are highly similar not just on observable characteristics (e.g., employment history, demographics), but also unobservable ones (e.g., motivation) is why they are considered, across many fields, to be the most rigorous study design for measuring a program's impact.

Your Program's Information

Organization Name:

Program Name:

Program Description:

Step 1: Program Design

STEP 1: CLEARLY DEFINE WHAT YOU ARE DOING AND WHY

Create Theory of Change/
Gather Evidence
- Create logic model
- Create replication materials

Step 1: Program Design

Does your organization have a Theory of Change?

☐ Yes ☐ No

If yes...

What is your long-term outcome goal?

What problem are you trying to solve? How will you know you are successful?

What is your intended population? (e.g. justice-involved youth, women of color, men ages 18-25, etc.)

Have you included your intended population in your program design? If not, what can you change to include their voices?

If yes...

What are your essential program elements?

Why are those elements expected to lead to change in the desired outcomes?

If yes...

Have you codified program materials so that the program can be consistently replicated?

Have you developed Specific, Measurable, Attainable, Relevant, and Time-bound (SMART) goals? If so, What are they?

If no theory of change.....

What do you need to start the process?

Which stakeholders do you think should be involved?

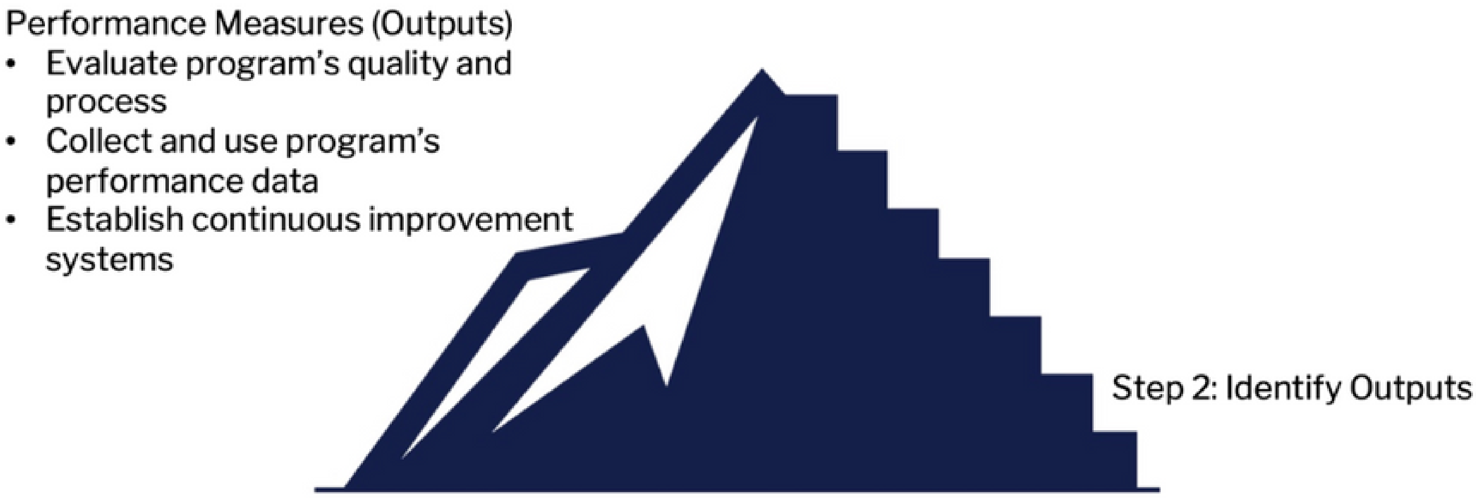
How will you include the voices of your intended target population?

Notes:

Once you feel comfortable that you have a well-defined Theory of Change, you are ready to proceed to Step 2.

Step 2: Identify and Measure Outputs

STEP 2: DETERMINE WHAT IS BEING DONE CONSISTENTLY AND EFFECTIVELY



Determine how well and consistently the program is being implemented. Implementation science is an important part of all of this work. For best practices in implementation science, please reference: [National Implementation Research Network (NIRN)]

What are your essential program activities?

How are you defining the successful implementation of your program?

What program activities and performance measures do you currently track/measure in order to assess how well your program is being implemented?

Do you have a system for tracking continuous quality improvement on these activities and performance measures (e.g., Google, Salesforce)? What does that look like?

Whom are you collecting the information for (internal partners, external partners, or both?)

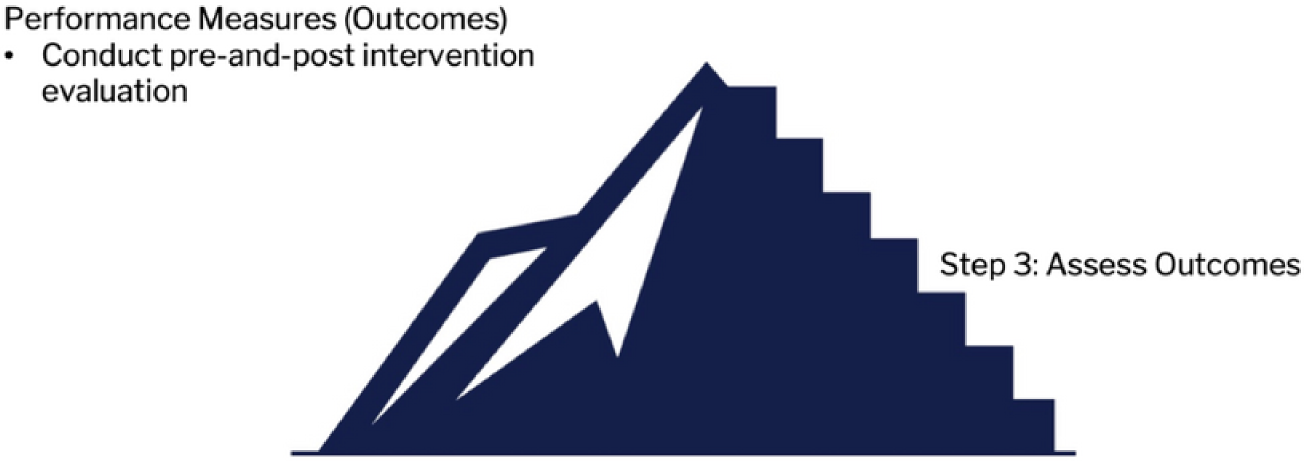
What are you doing with that information?

Notes:

Once your output measurement is telling you the program is being implemented as intended, you are ready to move to Step 3. If your program is not being implemented as intended, you may need to take a fresh look at Step 1 and modify your theory of change (e.g., change the underlying assumptions behind your program, add essential elements that are critical to quality implementation, and/or modify your program materials to facilitate consistent implementation).

Step 3: Assess Outcomes

STEP 3: MONITOR TRENDS IN DESIRED OUTCOMES OVER TIME



Monitor trends over time in desired outcomes-
– i.e., measures of what the program is meant to improve for its target population.

What do you currently measure to assess how program participants' lives have changed after receiving the program?

Have you conducted any pre-post assessments of the above outcomes to see whether they were different after the program than they were before it?

☐ **Yes** ☐ **No**

- **If yes:**
  - **How did you collect that information?**
  - **Can you report on trends?**

    ☐ **Yes** ☐ **No**

  - **Which trends can you report on?**
  - **Examples?**
- **If no, do you need help developing assessments?**

  ☐ **Yes** ☐ **No**

Do you collect demographic information on program participants?

☐ **Yes** ☐ **No**

- **If so, what information do you collect?**

- **Are you able to see if outcomes differ across different demographic groups, particularly for groups with the greatest barriers and challenges prior to the program?**

☐ **Yes** ☐ **No**

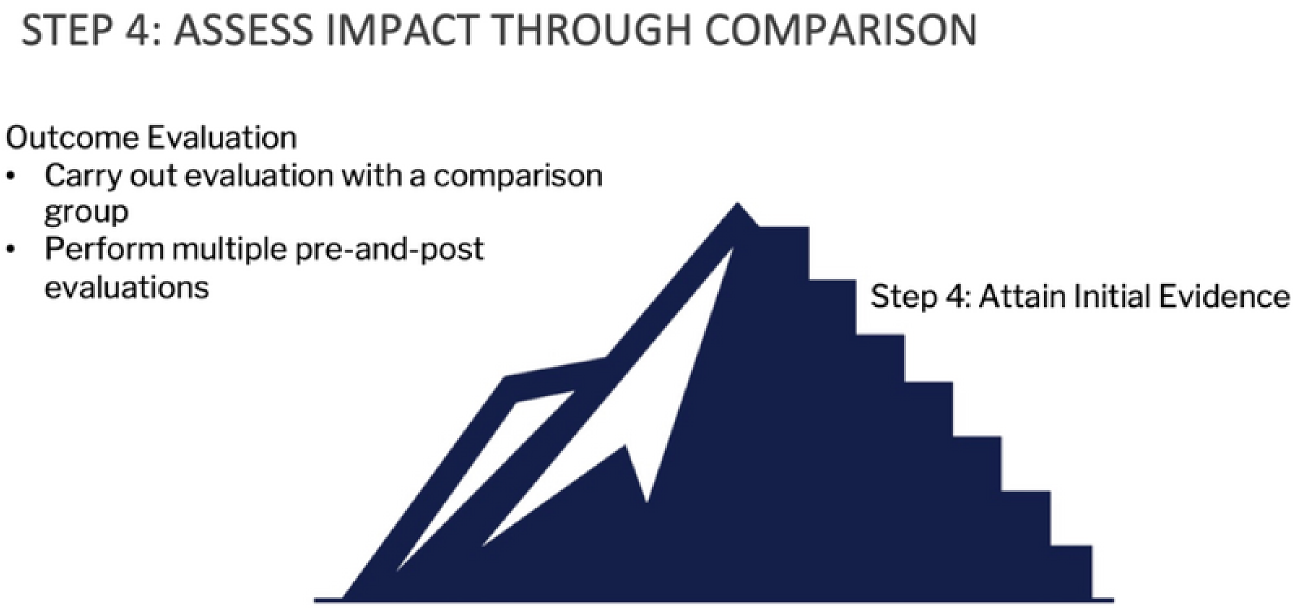
What additional systems do you need in place to measure outcomes?

Notes:

Once you have observed a positive trend in a pre-post study, you are ready to proceed to Step 4 to start measuring the impact of your program[1]. If you are not seeing a positive trend, you may need to go back to steps 1 and 2 to modify your program's design and/or implementation.

[1] Remember – a pre-post study tells you whether program participants' outcomes were better after the program than they were before, but it does not tell you that your program caused those changes (i.e., produced an impact). This is because many other factors could potentially explain program participants' change in outcomes besides just the program. Steps 4 and 5 are where you can start getting a sense of your program's impact through rigorous comparison group studies since these studies, when well-conducted, can rule out other possible explanations for improvements in program participants' outcomes.

Step 4: Attain Initial Evidence



Assess your program's effectiveness through either multiple pre-post outcome evaluations without a comparison group or one study with a rigorous comparison group.

If you are conducting additional pre-post studies, use the same checklist under Step 3 for these additional studies.

If you are conducting a comparison group study...

Have you identified a researcher or evaluator with experience carrying out comparison group studies, and who understands the importance of incorporating principles of diversity, equity, inclusion and justice (DEIJ) in evaluation design (e.g., in formulating research questions and selecting outcome measures)? Partnering with at least one such researcher is important, as they can help you determine the answers to the questions immediately below.

☐ Yes ☐ No

Can you identify a comparison group that is highly similar in observable characteristics to participants in your program (i.e., the program group)?[2]

☐ Yes ☐ No

Can you measure your target outcomes for both the program and comparison groups?

☐ Yes ☐ No

If not, what steps can you take to access the data you need to measure these outcomes?

---

[2] Importantly, the broader pool from which you draw your comparison group should be reasonably similar to the program group before statistical techniques are used to further equate the groups (this is to avoid a situation where any impact observed in the study is entirely due to the statistical techniques used to measure them as opposed to the program itself). It is also important to identify your program and comparison groups before measuring outcomes for your study sample in order to objectively estimate the program's impact.

Will your study have enough program and comparison group members (i.e., a big enough sample) to detect impacts that might be reasonably expected for your program? [3]

☐ **Yes** ☐ **No**

Can you collect demographic information for both your program and comparison groups?

☐ **Yes** ☐ **No**

- **If so, what information do you collect**

- **Are you able to measure how your program's impacts differ across different demographic groups, particularly for groups with the greatest barriers and challenges prior to the program?**

☐ **Yes** ☐ **No**

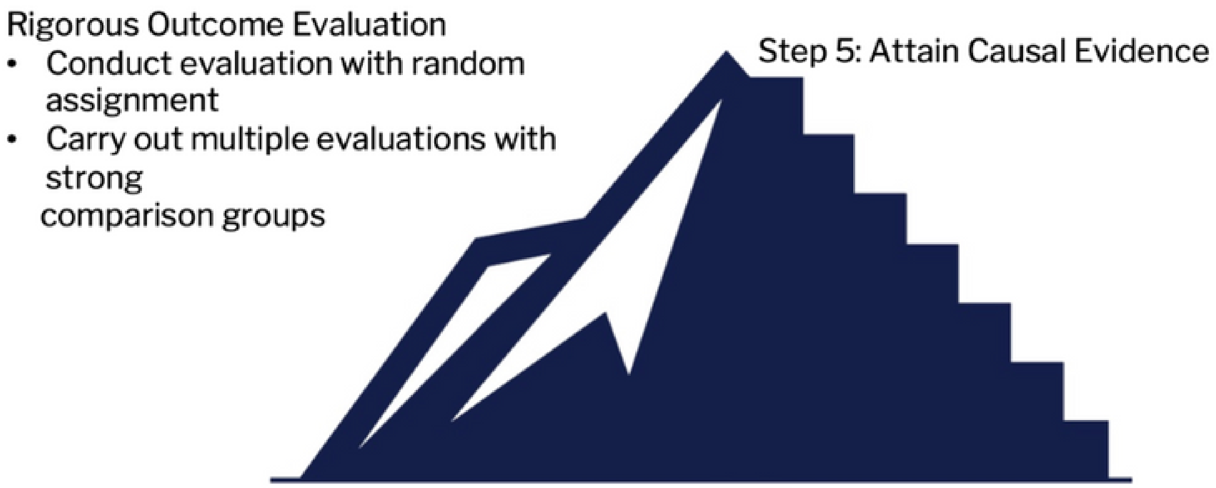
[3] Your research partner can help you with what is called a power analysis, which tells you how many people you will need in your study to confidently detect an impact that could reasonably be expected from your program.

Notes:

If you are observing consistent positive trends in multiple pre-post studies or a positive impact in a rigorous comparison group study, you are ready to move to Step 5. If not, you may need to go back to steps 1 and 2 to refine your theory of change and/or program implementation.

Step 5: Attain Causal Evidence

STEP 5: ASSESS IMPACT THROUGH COMPARISON (AND RANDOM ASSIGNMENT)



Assess impact through either (i) one randomized controlled trial (RCT) or (ii) multiple rigorous comparison group studies. For more information on which studies are most capable of producing valid evidence of a program's impact, see this guide from the Coalition for Evidence-Based Policy.

If you are undertaking additional comparison group studies that are not RCTs, follow the same comparison group checklist listed under Step 4.

If you are undertaking an RCT[4]...

Have you identified a researcher or evaluator with experience carrying out an RCT,[5] and who understands the importance of incorporating principles of diversity, equity, inclusion and justice (DEIJ) in evaluation design (e.g., in formulating research questions and selecting outcome measures)? Partnering with at least one such researcher is important, as they can help you determine the answers to the questions immediately below.

☐ **Yes** ☐ **No**

Are you able to randomly assign potentially eligible program participants to either a group that receives your program or a control group that does not? One scenario where this is often possible – and often the fairest way to allocate program slots – is when there are more people who could benefit from the program than the program can serve.

☐ **Yes** ☐ **No**

[4] Well-conducted RCTs are, when feasible, widely considered the most rigorous method for measuring a program's impact because they ensure that a study's program and control groups are highly similar on both observable characteristics (e.g., prior earnings/employment, demographics) and unobservable characteristics (e.g., motivation, psychological resilience). As a result, any difference in outcomes observed between the program and control groups can confidently be attributed to the program and not to any other factors.

[5] For a list of suggested and vetted evaluators, reference the *Additional Resources* page at the end of the document.

Are you able to measure your target outcomes for both the program and control groups?

☐ **Yes** ☐ **No**

- **If not, what steps can you take to access the data you need to measure these outcomes?**

Will your study have enough program and control group members (i.e., a big enough sample) to detect impacts that might be reasonably expected for your program?

☐ **Yes** ☐ **No**

Can you collect demographic information for both your program and control groups?

☐ **Yes** ☐ **No**

- **If so, what information do you collect?**

- **Are you able to measure how your program's impacts differ across different demographic groups, particularly for groups with the greatest barriers and challenges prior to the program?**

☐ **Yes** ☐ **No**

Notes:

RESEARCH, DATA, AND EVALUATION RESOURCES

These are all evaluators, research, and data teams that CEEMI has worked with in some capacity and can confidently recommend on this resource list. We are happy to make connections with any of the following:

Abt Associates

Actus Policy Research

Colorado Evaluation and Action Lab (Colorado Lab)

Coeffect

Economic Mobility Corporation

Mathematica

MDRC

Omni Institute

Results Lab

Pam Buckley, at the University of Colorado, Boulder

Christine Steeger, at the University of Colorado, Boulder

Oded Gurantz, at the University of Colorado, Boulder

Colorado Equitable Economic Mobility Initiative (CEEMI)

Building Evidence Questionnaire and Workbook

CEEMI COLORADO EQUITABLE ECONOMIC MOBILITY INITIATIVE

www.ceemi.org

Questions & Comments can be sent to Jessica Corvinus
Jessica@ceemi.org