Terms

Version: 6/13/2023

This document provides and explains terms that are directly relevant to CoLab's work. Many of these terms have many different definitions and applications in the literature and in practice. There is not a single right or wrong version of defining many of these terms, but we have created this document to help facilitate a shared language around CoLab's efforts.

Types of Evaluation

Term	What it is	When to use it	Example
Formative evaluation	A process of gathering and analyzing feedback during the development or implementation of a program.	When you want to collect information to inform the development or improvement of an intervention.	Conducting focus groups with first generation college students to inform the development of a program to support graduation.
Process evaluation and implementation evaluation	A process evaluation assesses whether program activities were implemented as intended (i.e., with fidelity) and the factors that support or inhibit high-quality implementation.	When you want to better understand the reasons for positive, negative, insignificant, or unexpected results of an intervention and gather information to help intervention developers and implementers improve a program's performance.	Evaluating the extent to which social workers providing home visiting carried out their home visiting activities as expected according to the program design and implementation manual.
Program monitoring	The systematic process of collecting, analyzing, and applying information as a program proceeds.	When an organization wants to track, report, and improve program implementation and outcomes.	A youth-serving organization monitors the implementation of its transitional living program's activities, such as case management and life skills training, and exit outcomes, such as housing, education, and employment status

Impact evaluation	An impact evaluation assesses the changes in outcomes that can be attributed to a specific intervention. Impact evaluations need to account for the counterfactual–what would have occurred without the intervention–through the use of an experimental or quasi-experimental design using comparison and treatment groups.	When you want to assess the intended and unintended changes in outcomes that are due to a specific intervention.	Researchers conducted a randomized evaluation comparing the effects of a cash transfer program for low-income mothers with babies to service-as-usual on early brain development and long-term health and education outcomes.
Cost-benefit analysis	An analysis comparing the costs and benefits of an intervention, where both are expressed in monetary units. This generally includes information on the costs of implementing the intervention and monetized benefits from improvements in outcomes, typically determined through impact evaluation.	When you want to show the return on investment of an intervention or understand whether the benefits of scaling an intervention outweigh the costs.	A cost-benefit analysis for a career training and placement program for formerly incarcerated people found a $3 return in future earnings for every $1 invested into the program.

Impact Evaluation Terms

Term	**What it is**	**When to use it**	**Example**
Experimental study and randomized controlled trial (RCT)	A type of impact evaluation that uses random assignment to establish a cause-and-effect relationship between an intervention (independent variable) and outcome (dependent variable). Usually, random assignment is done at the individual or household level, but	When you want to assess the intended and unintended changes in outcomes that are due to a specific intervention, and it is feasible and acceptable to use random assignment for establishing a counterfactual.	A city only had pilot funding for 100 housing vouchers when over 500 families needed housing. Researchers used a lottery to randomly assign the 100 vouchers and collected data over time to compare changes in outcomes

	researchers can also randomly assign schools, classes, organizations, neighborhoods, etc. to intervention and comparison study arms in a cluster-randomized controlled trial (C-RCT).		between families who did and did not receive vouchers.
Quasi-experimental study	A type of impact evaluation that uses approaches other than random assignment to establish a cause-and-effect relationship between an intervention and outcome, such as statistical matching of intervention group participants to similar people who did not participate in the intervention.	When you want to assess the intended and unintended changes in outcomes that are due to a specific intervention, and it is infeasible or unacceptable to use random assignment for establishing a counterfactual.	A city provided housing vouchers to all low-income families on a housing waiting list. Researchers used information available to compare changes in outcomes between voucher recipients and similar families in the community who were not enrolled in the waitlist.
Longitudinal impact evaluation	An impact evaluation in which researchers collect and analyze changes in outcomes for the same individuals at multiple points over time–not just before and after the program.	When you want to assess whether changes in outcomes sustain, fade-out, or change over time, especially for months or years after the program ends. Also when you want to know whether changes in short-term outcomes (e.g., educational achievement) lead to changes in long-term outcomes (e.g., income).	A longitudinal impact evaluation of a culturally-specific program to support school engagement among 9th grade students of color found that the program reduced absenteeism within 6 months, increased high school graduation within 4 years, and delayed childbearing and increased earnings within 10 years.
Service-as-usual comparison	An impact evaluation comparison group that involves people receiving usual services typically available in the organization, school, or community without	When you want to assess the impact of a specific intervention on outcomes in comparison to status quo services or supports they do or	Researchers compared the results of providing cash transfers and financial counseling to families in local homelessness

	receiving the additional intervention that is the subject of evaluation. With impact evaluations of social and economic interventions, we often describe the counterfactual comparison group as a "service-as-usual" group rather than a "control group," because unlike in clinical/pharmaceutical trials in laboratory environments, researchers do not control the conditions of the counterfactual in people's lives.	can already receive.	shelters to those families receiving services-as-usual provided by shelter programs, such as crisis services, case management, and housing assistance referrals.
Internal validity	Internal validity is defined as the extent to which the observed results represent the truth in the population we are studying and, thus, are not due to methodological errors.	When you want to consider how well a study was designed or carried out to mitigate the risk of potential bias or inaccurate findings.	Researchers decided not to compare changes in outcomes over time between a group that participated in an apprenticeship program and another group who chose not to because the difference in motivation between the two groups could play a major role in determining differences in outcomes–i.e., they questioned the internal validity of this study design.
External validity or generalizability	The extent to which you can generalize the findings of a study to other situations, people, settings, and measures. In other words, can you apply the findings of your study to a broader	When you want to consider how generalizable the findings of a study are to the population or context for which you would like to apply the findings.	An impact evaluation of a program to reduce teen pregnancy showed positive results with a small, mostly racially homogeneous sample in one

	context?		neighborhood. Practitioners questioned the study findings' external validity to other populations and geographic areas.

Theory of Change Terms

Term	What it is	When to use it	Example
Theory of change	A theory of change is essentially a comprehensive description and illustration of how and why a desired change is expected to happen in a particular context.	When you want to keep a program design focused on the intended impact, using the theory of change as a kind of road map to impact. And when you want an evaluation design to test and reveal hypotheses reflected by a theory of change.	A theory of change for a "cash plus" program for young adults experiencing homelessness captured expectations of how both cash and non-monetary program components would affect short-term housing and well-being outcomes and long-term social and economic mobility. It also included believed assumptions about what needed to be true for these results to occur.
Change theory	Change theories represent theoretical and empirically grounded knowledge about how change occurs that goes beyond any one project. Ideally, a theory of change for a specific program is informed by change theories.		A peer navigator program's change theory might reflect broader social theory and evidence suggesting that trained peers can be more effective in delivering a certain type of support than non-peers (e.g., traditional social workers).
Input	Inputs are things that a program can control		A peer navigator program's inputs

	directly, such as its resources, staffing, supplies, materials, technology, etc.		included resources such as funding to pay navigators and the staff supporting them and training for navigators.
Activity	Actions associated with delivering program goals.		A peer navigator program's activities included a semi-structured goal-setting conversation, an expected frequency of virtual and in-person follow-up meetings and communications, and peer support with specific topics and service connections.
Output	Numerical counts of a program's actions or products that were created or delivered, the number of people served, and the activities or services provided.		A peer navigator program's outputs included a certain number of contacts with participants, peer support groups organized, and service referrals made.
Outcome	Meaningful changes for the population served, such as anticipated changes in knowledge, skills, attitudes, behavior, condition, or status.		A peer navigator program's intended outcomes included improvements in participants' personal empowerment, use of services, housing stability, and education or employment.
Assumption	Assumptions are the necessary conditions for change, or the "underlying conditions or resources that need to exist for planned change to occur".		Developers of a peer navigator program assume that training and support of the peer navigators provide the necessary information and

			support for them to deliver the program as intended.

Other Research Terms

Term	What it is	When to use it	Example
Needs assessment	A needs assessment is a process for determining the needs, otherwise known as "gaps," between current and desired outcomes. Researchers can use both quantitative and qualitative methods to carry out this type of diagnostic study.	When you want to start by understanding the problems or gaps in an organization, system, or community before designing and evaluating specific solutions, such that the solutions emerge from a careful assessment of the need.	A needs assessment of a city's youth homelessness system found a lack of common points of entry to housing and services for young people and a lack of flexible, quickly accessible resources for exiting homelessness into acceptable housing.
Intervention	An intervention is a program, policy, or practice implemented with the intent to affect certain outcomes in a population.	When you want to put a solution into practice for improving specific outcomes for a population.	A city funded and delivered free, high-intensity tutoring for students falling behind in school to reduce achievement and graduation gaps.
Qualitative research	Qualitative research involves collecting and analyzing non-numerical data (e.g., text, video, or audio) to understand concepts, opinions, or experiences.	It can be used to gather in-depth insights into a problem, develop program theories, understand why a program did or did not have the intended effects on outcomes, and generate new ideas for research.	A process evaluation conducted alongside a quantitative impact evaluation included qualitative interviews with a subgroup of a home visiting program's participants to understand how they experienced the program, how they believed it influenced their outcomes, and how it might have worked better.
Quantitative research	Quantitative research is the process of collecting	It can be used to find patterns and	An impact evaluation is an example of

	and analyzing numerical data.	averages, make predictions, test causal relationships, and generalize results to wider populations.	quantitative research; it seeks to estimate the numeric effects of a specific intervention. A survey to estimate the prevalence of a problem in a population (e.g., depression or homelessness) or a relationship between two variables (e.g., childhood trauma and educational attainment) are other examples of quantitative research.
Community-based participatory research (CBPR)	An equitable approach to research in which researchers, organizations, and community members collaborate on all aspects of a research project. CBPR empowers all stakeholders to offer their expertise and partake in the decision-making process.	When you want to engage people closest to the problem studied in crafting the research questions and designing, carrying out, and interpreting the research.	A research initiative engaged current and former foster youth as collaborators in identifying research questions and programmatic solutions, developing and carrying out an evaluation of a program to support foster youth, and interpreting the policy and practice implications of the results.
Institutional Review Board (IRB)	An IRB is a committee within a university or other organization that reviews research proposals. The IRB reviews the proposals before a project begins (and sometimes before it is submitted for funding) to determine if the research project follows the ethical principles and federal regulations for the protection of human	Federal regulations require that research projects involving human subjects be reviewed by an IRB. The IRB must approve or determine the project to be exempt prior to the start of any research activities.	Before a research team began conducting an impact evaluation of an early childhood education program for low-income families, they submitted a detailed research protocol to the university's IRB for review and determination.

	subjects. The IRB has the authority to approve, disapprove or require modifications of these projects.		
Systematic review and meta-analysis	A systematic review answers a defined research question by collecting and summarizing all empirical evidence that fits pre-specified eligibility criteria. A meta-analysis is the use of statistical methods to summarize the results of these studies. Unlike a literature review, systematic reviews use transparent, predetermined criteria and methods for collecting, selecting, and synthesizing evidence to reduce bias in the review process.	Funders or policymakers might fund a systematic review to fully understand the state of the evidence on a particular intervention or topic before investing in it. Researchers might conduct a systematic review to understand the state of the evidence on an intervention or topic before conducting a new primary study, such as a new impact evaluation.	Before conducting a new impact evaluation of a financial coaching program for youth, researchers decided to conduct a systematic review of existing evidence on the effectiveness of financial coaching programs to inform whether and how to carry out the new study.

Contact: Matt Morton, Executive Director, CoLab, matt.morton@constellationlab.org