writing for green

Evaluation – Key Terms/ Definitions

Knowing whether your efforts are making a difference is important for both you, your team, your stakeholders, and your funder(s). Thus, every plan – whether the plan is a grant proposal, an operational document, or a strategic plan – should include an evaluation component.

An evaluation plan, however, need not be overly sophisticated or complicated. One of the most important considerations is that your plan be well-defined and developed, and provide a concrete means for the implementers/ planners to measure progress towards your defined goals. Below is a list of key evaluation terms, along with their descriptions, that you are likely to encounter in the process of applying for a funding opportunity.

Term	Definition
Efficacy	A measure of effectiveness, based on a desired or intended outcome (of a program, activity, or intervention). For example, if a vaccine has an "efficacy of 90%" (as determined in a trial), then it worked for 90% of the people who received the vaccine (they did not get sick, or did not test positive, or whatever the objective was).
Evaluation	An assessment of (something) to determine whether the defined goals and/or objectives were met. There are many different types of evaluation, and evaluations vary greatly in intensity, based on the purpose, scope, and available funding associated with the program, activity, product, service, etc. being evaluated.
Evaluator	The individual or entity responsible for an evaluation. Typically, when this term is used in a grant application, its use implies an external individual or entity. In the context of a grant application, a professional, independent evaluator may be identified within the application (if warranted based on the scope and nature of the program). This evaluator, usually a "contractor," will generally outline within the proposal a scope of work to be performed as part of the project evaluation, should the funding be awarded. If an organization has the in-house capacity to do so, they may identify an internal evaluator, unless the funding opportunity specifically requires otherwise.
Evidence	Objective data, often published in a peer-reviewed journal, that can be used to support a position, or an approach to implementing a program. Evidence can also be collected by the program implementers as a means to demonstrate whether a program is effective, or its level of its effectiveness. Evidence, or data, may be qualitative or quantitative and

	may be collected through a number of formats (devices, surveys, analysis of interviews, written reports, focus groups, etc.).
Experiment or "study"	A means to test a hypothesis. There are many different types of studies and such studies are generally not expected to be performed by small or even mid-sized organizations. Formal studies, usually performed by research institutions, require a high level of expertise and resources.
Fidelity	The level to which implementation of a program aligns with its original and intended design.
Focus group	A group of individuals meeting a specified criteria that are convened for the purpose of discussing a specific topic. Facilitating a focus group (and collecting/ analyzing the information collected) is an example of a qualitative method of collecting data. Depending on the level of sophistication of the researchers/ implementers, and their available resources, they may create a transcription and code/ evaluate the collected content using one of a number of statistical tools. In the context of meeting expectations of funders of small to midsize organizations, information/ feedback gathered from focus groups is used more informally as a means to better understand the impact of a program, or to gain information used to improve a program or activity.
Formative evaluation	Used to assess whether or not a program or activity is performing as expected, formative evaluation activities may be conducted at several points *during* the implementation of a program, with the goal of making improvements as informed by that evaluation. For example, whereas a "summative" assessment such as a final exam may be used to measure a student's knowledge or competencies at a given point in time, homework or classworks may be intended to provide real-time feedback to the instructor so they can assess a student's progress, and make adjustments to their methods, as appropriate. Similarly, a short survey of people who attended a community forum may provide insight about whether the organizer's approach is working – e.g., by asking whether the attendees found the forum interesting, educational, a good use of their time, etc. This information, in turn, will help the organizer to improve the next community forum. Such evaluation activities are legitimate to include in a grant proposal's evaluation plan and do not need to be overly complicated to indicate to reviewers that you have a well-developed approach to evaluation, and have potential to yield important results.
Goal	A target that is to be achieved. Typically, a goal is broader and more "big-picture" than an objective.

Indicator	Data that indicates progress toward a goal or objective.
Logic model	An organized depiction, usually in the form of a table or graph, of a problem and the path to a solution. Typical components of a logic model (also referred to sometimes as "Theory of Change") include the following: Problem --> Resources/Inputs --> Activities -- > Outputs --> Outcomes (short-term) --> Outcomes (long-term). Some grant solicitations explicitly request or require a logic model as part of the proposal/ evaluation plan. Even if it is not required, applicants may find the process of creating one useful in both planning and implementing the project.
Longitudinal data	Data that is tracked over a period of time that typically spans years, following the same subjects. In the context of applying for EJ grants, longitudinal data may not be highly relevant, as most projects will be completed before sufficient data can be collected to make assertions related to the project. However, while you may not propose to collect longitudinal data, you may find it useful to reference longitudinal studies who, for instance, studied the impact of a particular pollutant on human health over a long period of time.
Measures and Metrics	These terms are often used interchangeably. There is some slight difference in that a metric is a measure with more information or context attached to it. A metric can be thought of as a measure-of-performance. For example, the **"number of miles"** a subject rides their bike is a measure. More useful would be the **"number of miles they ride in a single day."** Metrics and measures are important ways to determine whether a program or activity is making progress towards (or has achieved) a goal or objective.
Methods	The way(s) in which evaluation activities (e.g., data collection) are carried out. Examples of evaluation methods include Interviews, focus groups, surveys/ polls, questionnaires, review of meeting minutes, analysis of data from air quality monitors, public health prevalence data from the CDC, collection and analysis of water quality data, etc.
Objective	A measurable action related to a goal, which can be achieved (often as a milestone) in a shorter period of time than a goal can be achieved.
Outcome	The desired or intended result of the proposed activities. In the context of an EJ project, the outcome may relate to the impact of a program on the quality of life of an individual (e.g., change in health measure, behavior, or competency), or an improvement to a community's well-being (e.g., reduced pollution in a defined geographic area). In

	most cases, the "outcome" should be associated with the final impact.
Outcome evaluation	An evaluation intended to measure the impact of a program/ intervention.
Output	The quantifiable results of program activities or interventions. Outputs typically demonstrate the extent of the work (e.g., how many town halls were held), reach (e.g., how many people viewed the social media post), or breadth or depth (e.g., the average length of activists' meetings with elected officials). Outputs typically indicate progress towards an objective, without directly measuring the intended outcome. For example, outputs of a program intended to teach community members to setup air quality monitors and collect and analyze the data could be: - # of trainings delivered; - # of people trained; - # monitors setup; - # of times data was collected or frequency of that collection. Outcomes of this same program could include: - % of those trained whose competencies in air quality monitoring improved; - Policy change (at any specified level) that was driven by this data; - Decrease in certain pollutants over time, as measured by monitors (long-term).
Pre/ Post Tests or Surveys	Surveys that are issued to participants before and then after a specific program activity or intervention. Such surveys can used to measure one or more changes experienced by participants as a result of the activity or intervention, and are often useful in the context of educational activities or trainings.
Process evaluation	An evaluation that measures the processes/ outputs of a program/ intervention, and is not designed to measure outcomes. Example of a process evaluation: interviews and review of publicly available information to determine whether a regional lobbying effort resulted in reforms to permitting regulations at the local level. Such an evaluation is not seeking to determine the impact (outcome evaluation) of the policy changes on the community, or individuals in that community.

Qualitative data	Describes qualities or characteristics, which may be subjective, based on the opinion or perspective of the actor. Qualitative data is often collected using questionnaires, interviews, or observation or focus groups. Qualitative data is still data and can be used in statistical analyses.
Qualitative interview	A systematic review of multiple sources/ studies (primarily qualitative), with conclusions drawn by the author.
Quantitative data	Data that expresses an amount or a range, and is not subjective/ dependent on an observer. Generally, units of measure are associated with quantitative data to indicate "how much" or the "rate of" whatever is being measured. Examples of quantitative data include: - # of people affected by a problem - # of people that attended a meeting - Incidence rate (#/ 100,000 ppl) in a specified geographic region exposed to a pollutant - # of times that unhealthy air-quality was recorded in a town
SMART	Used in reference to goals and objectives. S = specific M = measurable A = achievable R = relevant T = timely (or time-bound)
Statistically significant	Generally, an indication that determinations expressed by an analyst (based on data) are not based on chance alone. The caveat is that there is no 'absolute' in statistics. The statistical probability that something will happen, or is attributable to a certain intervention or cause, is quantified using terms such as "confidence level," "confidence interval," "odds ratio," and "p-value." While planners, grant writers and project implementers/ managers should be familiar with these terms, it is unlikely that CBOs applying for EJ-related funding will need to propose complex statistical analysis as part of their evaluation plans. One of the reasons it is useful to understand these terms, however, is that it will help to better understand and interpret technical reports and data that may be useful to incorporate into a grant narrative. To illustrate use of these terms, consider the following (fabricated) statement: "Our analysis concluded that the level of PM2.5 measured in Community X between March and July 2022 exceeded 30-35 $\mu g/m^3$ on 33 days (95%, CI: 27-36 days)."

	Simple translation: The researchers can say with 95% certainty (level of confidence) that their result (33 days) is within the expected range (27-36 days). How does a researcher arrive at such a determination? By using/ inputting data/ variables into statistical analysis software. For context, 95% is typical for a confidence level. If, considering the extreme example, a confidence level is particularly low (50%) or the confidence level is particularly high (in our case, for example 5-50 days) our finding of "33 days" of excessive PM2.5 would be meaningless.
Summative evaluation	An evaluation that seeks to determine the final result. In school, a final exam could be considered a "summative evaluation." Similarly, to assess the impact of a training on participant knowledge, a summative assessment could be administered after completion of the intervention (post-test), and compared to a pre-test.
Theory of Change	Sometimes used interchangeably with "logic model."