

Assessing an Organization’s Evidence of Effectiveness

The Edna McConnell Clark Foundation assesses an organization’s evaluations and other data to ascertain the quality and rigor of the evidence that its program is having a measurable impact on youth outcomes. We have developed a framework that categorizes a program’s evidence of effectiveness on one of three levels—a continuum from high apparent to proven. In the chart below, the first column defines each level and indicates what an organization should know about the effectiveness of its programs at that level. The second column specifies the kinds of information an organization must collect, and the types of evaluation activities required, to reach that particular level.

<p>Proven Effectiveness Experimental research has confirmed the program’s impact on participants.</p> <p><i>A program at this level should be able to answer the following question: Are there meaningful, positive, statistically significant outcomes for program participants that differ from outcomes for people in a randomized control group?</i></p>	<p>Key Characteristics of Data Collection and Evaluation Activities:</p> <ul style="list-style-type: none"> • A well-designed and well-executed experimental evaluation of program outcomes, created and conducted by an independent, external evaluator, establishes the most rigorous evidence of effectiveness. Ideally, participants in the study are randomly assigned to one of two groups—one that receives program services and a control group that does not. Outcome data for both groups is collected and compared in this randomized controlled trial. • The study concludes there are meaningful, positive, statistically significant differences between outcomes for youth served by the program and outcomes for youth in the control group. • At the highest level of proven effectiveness, a program has evidence of impact from multiple sites. • <i>Under some circumstances, a well-implemented program that has been proven effective elsewhere, or a third-party quasi-experimental evaluation that compares participants to a comparison group that has not been randomly assigned, may represent the highest proof point a program is capable of reaching.)</i>
<p>Demonstrated Effectiveness Systematically collected data comparing program participants with similar people not receiving a program’s services enables an organization to substantially conclude that youth are benefiting from program.</p> <p><i>A program at this level should be able to answer the following question: Are there meaningful, positive, statistically significant outcomes for program participants that differ from outcomes for people in a comparison group?</i></p>	<p>Key Characteristics of Data Collection and Evaluation Activities:</p> <ul style="list-style-type: none"> • A well-designed and well-executed quasi-experimental evaluation of program outcomes, created and conducted by an independent, external evaluator, measures outcomes for program participants against outcomes for a carefully chosen comparison group. People in both groups are at the same baseline on measured characteristics such as demographics and variables relevant to the study, and likely to be similar when it comes to unmeasured characteristics such as motivation at the start of the study. • This study, also called a comparison group evaluation, concludes there are meaningful, positive, statistically significant differences between outcomes for youth served by the program and outcomes for youth in the comparison group.
<p>High Apparent Effectiveness Systematically collected data indicates youth are probably benefiting as intended from participating in a specific program.</p> <p><i>A program at this level should be able to answer the following question: Who is accessing your services? What programs do they participate in? What outcomes do they achieve?</i></p>	<p>Key Characteristics of Data Collection and Evaluation Activities:</p> <ul style="list-style-type: none"> • Every program participant is given a unique identifier (such as a tracking or identification number). • The organization collects basic demographic data from program participants, such as address and contact information, age, gender, race/ethnicity, primary language, and socioeconomic status. • Initial data about program participants includes baseline data for measuring changes over time (outcomes). • The outcomes the organization intends for program participants are specified in a theory of change. • Outcomes are tracked for all program participants (or at least for a sample), and show meaningful, positive results, comparable to the results from similar well-implemented programs.

Many youth-serving nonprofits (perhaps even the majority of them) do not yet meet one of these three levels, as they are still in the process of collecting empirical evidence of their programs’ effectiveness. Nonetheless, many may gather basic information and/or have anecdotal evidence of a program’s beneficial outcomes even if they do not yet have the resources or capacity to systematically collect and analyze data (and thus meet High Apparent Effectiveness). Furthermore, an organization’s programs may not be mature enough operationally, or their performance management and measurement systems insufficiently developed, to evaluate outcomes rigorously. Although such programs may indeed benefit youth, this universe of organizations does not yet have systematically collected, empirical evidence that their programs are making an impact on young people’s lives.