

Sample Evaluation Plan

Best Practice Resources – Program Evaluation

The purpose of an evaluation is to demonstrate if your program is working and how. To run a sustainable mentoring program you need to capture program data in an objective and if possible empirical format to demonstrate your program's efficacy to outsiders. An evaluation plan can help you develop, adapt and test your services through an on-going process. There are many different types of evaluation activities. For example some elements of evaluation help you monitor the delivery and quality of your services whereas others allow you to track participant outcomes.

A robust on-going evaluation process would have your program examine:

- **Participant outcomes** – *This helps you assess the impact of your services on participants. This kind of evaluation activity is helpful if you want to know if your program works, and if for example it should be continued, replicated or scaled.*
- **Service Quality** – *This helps you understand the usefulness of various aspects of program services. These kinds of evaluation activities help you understand potential problems and make adjustments to your model to improve the overall quality and impact of your services.*
- **Program/Process Fidelity** – *This helps you determine if your program is being delivered in a manner consistent with your model. These kinds of evaluation activities help you validate to program stakeholders that you're doing what you say you do.*

Participant outcomes

To examine participant outcomes you first need to answer this question:

- What specific outcomes do we aim to achieve for program participants?

Here are some examples of a community-based mentoring program's goals for youth participants:

- Increase a youth's self-esteem
- Increase a youth participant's positive connection to school
- Increase a youth's healthy peer relationships

Once you have identified your program's desired participant outcomes, you'll need to develop tools and processes to demonstrate progress towards these goals overtime. You should be able to demonstrate progress towards every outcome that your program aims to influence. For example, you should not state that your program aims to influence the healthy diets of youth participants if your program has no element of diet or health education incorporated into mentoring services and you have no intention of testing your influence on the dietary decisions of youth participants.

There are many ways to collect program data. Below are some examples of the most common data collection methods used by youth development organizations.

- **Surveys** – Usually a self-administered questionnaire
- **Interviews** – Typically an extensive, one-on-one meeting between an interviewee (participant) and the evaluator or program staff.
- **Focus Groups** – Small, informal group discussion among similar participants around a common topic.

- **Observations** –A standardized qualitative or quantitative assessment of a situation from a trained observer typically based on the observer’s audio and/or visual senses.
- **Documentation** –Information that is not recorded first-hand but is available for review and analysis.
- **Journal/Logs** –Detailed records of key events over a pro-longed period of time.

For a detailed overview of the strengths and weaknesses of these data collection methods, and tips for implementation check out the Northwest Health Foundation’s Evaluation Handbook.

Sherril B. Gelmon, Anna Foucek and Amy Waterbury. Program Evaluation: Principles and Practices. 2nd ed. Portland: Northwest Health Foundation, 2005.

Just as there are many types of evaluation, there are also many types of research design. Research design refers to how you employ your data collection tools to demonstrate your program’s influence on a particular youth outcome. Below you’ll find descriptions of commonly used research design methods in social services.

Randomized Controlled Study: The strongest way to demonstrate your program’s influence on a specific youth outcome like self-esteem is to compare the effects of your service on a treatment group to the absence of services on a control group.

Pre-Test/ Post-Test: A pre and post-test design allows you to collect information on your participants prior to and after involvement in your program. This type of research design allows you to measure change in your outcome indicators, but it does not allow you to control for other factors that may have influenced your outcomes.

One Shot Case Study/Post-Test: A post-test allows you to paint a basic picture of participant outcomes after program participation. Using this type of research design a group of participants would engage in your program and then you would conduct a post-test or case study to observe or document the experience of these participants. Because there is no pre/baseline test there is no way to determine whether any changes occurred and whether they were a result of your program.

Many youth development organizations do not have the infrastructure or resources to commit to a randomized controlled study due to the long-term and intensive nature of this process. If you’re mentoring program is well-established and you’re looking to expand or replicate services, you might consider applying for funding specifically for this type of evaluation.

Once you have identified a tool and or process to collect information from participants and you’ve selected a research design method you can:

- Collect relevant process, outcome, and record data.
- Analyze and interpret the data
- Disseminate findings to program stakeholders

Let’s take a closer look at how to do this with the first sample goal listed above.

1. Increase the self-esteem of youth participants.

To demonstrate your program’s influence on a youth’s self-esteem overtime we recommend utilizing an instrument to measure a youth’s self-esteem at the beginning of the program and then comparing this information to their level of self-esteem after treatment/ program participation. We have some

examples of validated survey instruments that assess a youth's self-esteem. With authorization from the instrument's author you could incorporate any of the instruments below into program services to utilize as pre and post-tests.

- Hare Area-Specific Self Esteem Scale
- Low Self-Esteem –Weinberger Adjustment Inventory
- Modified Rosenberg's Self-Esteem Inventory (a)
- Modified Rosenberg's Self-Esteem Inventory (b)
- Self-Concept –Individual Protective Factors Index
- Self-Esteem –Rochester Youth Development Study

Service Quality

Before conducting this kind of evaluation activity you need to answer the question below.

- How do we examine the quality and functionality of our services?

The following are some example measures of program quality:

- Participant feedback from initial and on-going volunteer trainings
- Parent and volunteer survey feedback regarding quality and consistency of match support services/case management.
- Parent and volunteer survey feedback regarding quality of overall program related services.

Information regarding program quality is typically collected in surveys administered during and after program participation. Survey questionnaires regarding program quality should be brief and easy to administer. We have some examples of basic parent and volunteer participant feedback surveys. Below you'll find a question from a sample survey testing a mentoring program's overall quality.

- I. I feel that I received adequate support and supervision from program staff

Strongly Agree

Agree

Disagree

Strongly Disagree

Program/Process Fidelity

This kind of evaluation activity will help you discern how well you adhere to your program model. Before conducting this kind of evaluation activity you need to answer the question below.

- What are core components of our model? Based on our model, how can I ensure that staff and mentors are operating as I think they are?

To demonstrate complete fidelity to your program model you can identify a measure to test the implementation of each level of program service.

A traditional mentoring program might test its fidelity by periodically reviewing a random selection of files to ensure proper documentation and fidelity. Here are some examples of the types of evidence to examine under each programmatic area, and the example of the types of files that would be useful:

1. Recruitment
 - A comprehensive recruitment plan and documentation of its implementation can serve as a fidelity measure here. Are you conducting regular recruitment activities and are these specific activities helping you reach your volunteer recruitment goals?
2. Orientation
 - Evidence of orientations conducted and data tracking the number of participants successfully transitioned from orientation to screening can serve as a program fidelity measure.
3. Screening
 - A review of volunteer files can serve as a fidelity measure here. Do you have documentation for every volunteer's completed screening protocol?
4. Training
 - Do all matches participate in the mandatory minimum for pre-match training? Documentation of a participant's engagement and feedback can serve as a fidelity measure. If training is not mandatory, can you determine what percentage of mentors are utilizing your training?
5. Matching
 - Do program coordinators make matches in concert with your program's matching policy? A review of volunteer and youth files can serve as a fidelity measure here. Do all volunteer files include a Match Assessment Form that details the coordinator's matching rationale?
6. Case Management/ Match Support
 - Do program coordinators stay in consistent communication with program participants including youth, families and mentors? A review of case management hours and relevant case-notes can serve as fidelity measure. What percentage of matches receive the full amount of case management?
7. Activities
 - Does the program offer consistent activities to engage participants? And does the program follow program safety standards/policies for group activities? Documentation of the number of activities offered and the number of participants in attendance can serve as fidelity measure. What percentage of matches participates in your activities?
8. Evaluation
 - A program report that summarizes your program's participant outcomes, quality of services, and fidelity to model can be used as evidence of a thorough evaluation process.
9. Closure
 - Do coordinators perform a healthy closure ritual for all matches graduating from services? This fidelity measure can be demonstrated by reviewing documentation of a closure form and final match case-note for all youth files.