Orientation to Data Collection

How to Create Systems to Yield High Quality Data

Michael Long
Andrew MacDonald

2/27/2018

Texas Evaluation Institute

Austin, Texas

ICF

Learning Objectives

By the end of this presentation, you will understand:

- What CNCS means by high-quality data systems and to document these in a data collection plan
- How to establish processes to ensure your data are complete and accurate
- How to use data for process and outcome evaluations

Why Build Systems for Collecting Data

- **Strong data systems are essential to maximizing your points during applicant review.**
  - The 2018 AmeriCorps NOFO requires that applicants "data collection systems are sufficient to yield high quality process and outcome data."
  - First time applicants must include a data collection plan.
  - Re-competing applicants must *describe* their data systems as part of their evaluation plan.
    - In other words, you should include a data collection plan as part of their evaluation plan
- **Well-developed data collection systems improve your evaluation capacity**
  - Building data collection systems requires stakeholders to agree on what indicators would provide answers your research questions.
  - Specifying when and how you will collect data can help you determine whether your evaluation plan is feasible to carry out given your resources.

Elements of a Strong Data Collection System

- **A strong data collection system should document how you will obtain, store, and use data. You should consider:**
  - *What* data will be collected
  - *What* are your sources of data
  - *Who* will collect data
  - *When* data will be collected
  - *What* data collection instruments and techniques you will use
  - *How* data will be cleaned and checked for consistency and accuracy
  - *How* data will be stored and protected
  - *How* data will be analyzed and used for program improvement

ICF

What Data Will Be Collected

- **Quantitative data:** Numerical information that can be counted, quantified, and mathematically analyzed (e.g., test scores, ratings)
- **Qualitative data:** Narrative information that describes the study subject(s) and context (e.g., transcripts of interviews and focus groups, field notes from observation of certain activities)
- **Both qualitative and quantitative data are systematically collected, recorded and analyzed**
  - Individual anecdotes and testimonials are NOT qualitative data if they are not systematically collected, recorded, and analyzed

What Data Will Be Collected

- **Indicators are the specific, observable data you collect to answer your research questions.**
  - Indicators may be qualitative or quantitative.
  - You can have more than one indicator for each research question

What Data Will Be Collected

- **Indicators for Process Evaluations**
  - Quantitative indicators for process evaluations typically measure inputs, activities, and/or outputs. For example:
    - The number of students who received tutoring
    - The number of families who attended seminars on first-time home-buying
    - The number of AmeriCorps members who obtained certification to teach ESL
  - Qualitative indicators for process evaluations describe feedback from stakeholders about inputs and activities. For example:
    - Students' perceptions about what did or did not work well in their program
    - Attributes of participants who complete all program components and those who drop out
    - AmeriCorps members' opinions about what they liked and did not like about their training

ICF

What Data Will Be Collected

- **Indicators for Outcome Evaluations**
  - Outcome indicators should measure changes in beneficiaries' knowledge, attitudes, behavior, and conditions.
  - Quantitative indicators are necessary to perform statistical analysis of changes. Examples of quantitative indicators are:
    - –Students' reading skills before and after they complete a year of tutoring
    - –Household savings before, immediately after, and one year after they take a seminar on budgeting
  - Qualitative indicators can also be used in outcome evaluation, even though they cannot show statistical differences across time. Examples of qualitative indicators are:
    - –Students reflections about how their knowledge and skills have changed after they participate in tutoring
    - –Families perceptions about how they manage their household finances differently after participating in a seminar on budgeting

ICF

What Are Your Sources of Data

- **Your data sources depend on your evaluation design and your resources.**
  - If you are undertaking an impact evaluation, you need to collect data from your program participants as well as a comparison group.
  - If you are not undertaking an impact evaluation, you need only collect data from program participants.

Sampling

- If you have a large number of participants, you may consider collecting data from a sample of participants.
- A sample is subset of individuals who are representative of your population of interest.
  - For example, if your program serves 1,000 people, it might make sense to collect data from 100 people. The important point is that your sample be intentionally drawn (e.g., random sample or stratified sample) and representative of the total population.
  - There is no hard-and-fast rule for minimal sample size. Talk with your external evaluator about carrying out a power analysis to estimate how large your sample needs to be to detect your expected effect size.

What Are Your Sources Of Data

- **"Secondary" data are data that already exists**
  - Sign-in sheets, participant records such as test scores or grades, program instructions, other administrative data
  - Your data collection plan should specify who has these data and if you have an agreement to access these data, and if not, how you plan to obtain one.
- **"Primary" data are data you need to go out and collect**
  - Surveys, tests, interviews, focus groups, observations
  - Your data collection plan should describe whether you are using an existing instrument(s) or creating your own instrument(s) to collect data.

Who will collect data

- **Your should specify who is responsible for collecting data**
  - You may have different individuals collect different data. Consider whether the person collecting data may introduce bias (or perception of bias) while collecting data.
    - For example, AmeriCorps members may collect sign-in sheets or administer pre/post tests, while an external evaluator may carry out observations and/or facilitate focus groups
- **You also should describe how data collectors will be trained, such as:**
  - An overview of what data to be collected and the data collection schedule
  - A description of the data collection instruments
  - Practice with data collection instruments
  - Instructions for submitting data and protecting personally identifiable information

When Data Will Be Collected

Your should specify what points in time, including the date, when you will gather data. You may want to collect data:

- **Before your intervention**
  - You can collect baseline data before your intervention so you have a comparison point.
- **During your intervention**
  - Some data are easier to collect during an intervention, such as sign-in sheets or observations.
- **Immediately after your intervention**
  - Collecting data immediately after your intervention can help you measure the short-term effects of your intervention, such as changes in knowledge.
- **After a period of time has elapsed after your intervention**
  - The medium and long-term outcomes in your logic model may not manifest immediately after your intervention.
  - You may also be interested in whether short-term outcomes persist over time.

What Data Collection Instruments To Use

- **Your should specify what *instruments* you will use to collect data, e.g., a survey, an assessment, or an interview protocol.**
- **Look to see if there are existing, validated instruments you can use. You may be able to make slight modifications to an existing instrument to fit your program model.**
  - Look on the CNCS Evidence Exchange to see if there are evaluation reports in your focus area that used an instrument that would be relevant to you.
- **If you cannot find an existing instrument, then you can create a new instrument.**
  - If you are using a new instrument, consider pilot testing it with a subset of participants. A pilot test can help you determine whether participants understand the instrument (e.g., the language or terminology), how long it takes to administer, and whether the data it yields are useful.

As we discuss later today, there are different considerations depending on your data collection technique (e.g., survey, assessment, focus group, etc).

How data will be cleaned and checked

- **Describe a systematic process for ensuring high quality data.**
- **Check for:**
  - Completeness—Surveys are completely filled out, sites have provided all their data, etc
  - Consistency—Sites and/or data collectors are using the same instruments and methods to collect data. For example, if you are administering pre/post tests, make sure everybody has the same amount of time to complete the tests.
  - Accuracy—There are no typos or illogical answers. For example, if a survey asks respondents to use a 1-5 point scale, check that nobody accidentally entered 6.
  - Verifiability—Describe how you will be able to confirm data, for example, by maintaining a paper copy of surveys.

- **In your application, be clear about who is responsible for checking data quality and when and how they will do so.**

How Data Will Be Stored and Protected

- **You should specify how you will store and protect data.**
  - If you are collecting data in hard copy (e.g., paper surveys), note how you will transfer to electronic form.
  - If you are collecting data from multiple sources, then you should describe how you will merge data.
  - If you are collecting personally identifiable information (PII), then make a note how you plan to protect these data, such as by using password protected files.
  - Your should also note how long you intend to store data, and when and how you plan to destroy data after your evaluation is complete.

How Data Will Be Analyzed And Used For Program Improvement

- **Your should describe the types of analysis you will carry out. These may include:**
  - Comparing output data to your targets
  - Measuring changes in knowledge, attitude, or behavior at different points in time and/or among different subsets of participants
  - Identifying themes or patterns among stakeholders' perceptions about program components.

- **You should also describe how you will use data for program improvement**
  - Identify what decisions may be influenced by data obtained through your evaluation. These may be related to topics such as:
    - Member training
    - Intervention dosage
    - Beneficiary recruitment
    - Other topics unique to your program

Resources

- **Building Effective Data Collection Plans presentation from 2017 ASN Symposium: https://www.nationalservice.gov/sites/default/files/resource/Building-Effective-Data-Collection-Plans.pdf**
- **Data Collection Plan Outline: https://www.nationalservice.gov/sites/default/files/resource/Building-Effective-Data-Collection-Plans.pdf**
- **Data Collection for Program Evaluation: https://www.nationalservice.gov/sites/default/files/resource/May_21_Data_Collection_for_Program_Evaluation.pdf**
- **Social Innovation Fund Evaluation Plan Guidance: https://www.nationalservice.gov/sites/default/files/documents/SIF%20Evaluation%20guidance%208%205%202014.pdf**