Corporation for NATIONAL & COMMUNITY SERVICE

Data Quality Review: Best Practices

Sarah Yue, Program Officer

Jim Stone, Senior Program and Project Specialist

Session Overview

- Why the emphasis on data quality?
- What are the elements of high-quality data?
- How should grantees and commissions assess data quality for their sites/subgrantees?
- What are some common “red flags” to look for when reviewing programmatic data?
- What corrective actions should grantees and commissions take if they encounter data quality issues?
- What resources are available to help with data quality?

Why Data Quality is Important

- Fundamental grant requirement
- Trustworthy story of collective impact for stakeholders
- Sound basis for programmatic and financial decision-making
- Potential audit focus

Elements of Data Quality

- Validity
- Completeness
- Consistency
- Accuracy
- Verifiability

Validity

- **Official definition:** Whether the data collected and reported appropriately relate to the approved program model and whether or not the data collected correspond to the information provided in the grant application
- **Plain language definition:** The data mean what they are supposed to mean
- How grantees/commissions should assess data validity for sites/subgrantees:
  - Review data collection tools; compare to objectives and PMs
  - Ask about data collection protocols
  - Request a completed data collection tool

Validity

- Examples of potential problems:
  - Invalid tools
    - Mismatch between tool and type of outcome
    - Incorrect approach to pre/post testing
    - Failure to use tools required by the National Performance Measure Instructions
  - Invalid data collection protocols
    - Assessing the wrong population
    - Implementing protocols in the wrong ways or times

Completeness

- **Official definition:** The grantee collects enough information to fully represent an activity, a population, and/or a sample.
- **Plain language definition:** Everyone is reporting a full set of data
- How grantees/commissions should assess data completeness for sites/subgrantees:
  - Keep a list of data submissions
  - Check source documentation
  - Request data at consistent intervals

Completeness

- Examples of potential problems:
  - Subgrantees/sites missing from totals
  - Non-approved sampling
  - Inconsistent reporting periods

Consistency

- **Official definition:** The extent to which data are collected using the same procedures and definitions across collectors and sites over time
- **Plain language definition:** Everyone is using the same data collection methods
- How grantees/commissions should assess data consistency for sites/subgrantees:
  - Request written data collection procedures; ask about dissemination/implementation
  - Cross-compare definitions and protocols

Consistency

- Examples of potential problems:
  - No standard definitions/methodologies
  - Lack of knowledge about required procedures and/or failure to follow them
  - Staff turnover

Accuracy

- **Official definition:** The extent to which data appear to be free from significant errors
- **Plain language definition:** The math is correct
- How grantees/commissions should assess data accuracy for sites/subgrantees:
  - Check your addition
  - Request data points from service locations
  - Check for double-counting in source data
  - Watch out for double-reporting

Accuracy

- Examples of potential problems:
  - Math errors
  - Counting individuals multiple times
  - Reporting the same individuals under different program streams

Verifiability

- **Official definition:** The extent to which recipients follow practices that govern data collection, aggregation, review, maintenance, and reporting
- **Plain language definition:** There is proof that the data are correct
- How grantees/commissions should assess data accuracy for sites/subgrantees:
  - Require source documentation
  - Review quality control plans

Verifiability

- Examples of potential problems:
  - Inexplicable data points
  - Estimated values

Common “Red Flags” in Reported Data

a) Large completion rates reported early in the program year
b) Actuals that are exactly the same as target values or consist solely of round numbers
c) Actuals that are substantially higher or lower than target values or are out of proportion to the MSY or members engaged in the activity
d) Substantial variation in actuals from one site to another
e) Substantial variation in actuals from one program year to the next
f) Outcome actuals that exceed outputs
g) Outcome actuals that are exactly the same as outputs

Example #1

Financial Literacy Education Program – Financial Literacy PM 1

Measure Type or Resource Type	Measure #	Target	Actual
Output	O1: Number of econ disadv individuals receiving financial literacy services	100	106
Outcome	O9: Individuals with improved financial knowledge.	40	106

Example #1 Continued

Financial Literacy Education Program – Financial Literacy PM 1

Measure Type or Resource Type	Measure #	Target	Actual
Output	O1: Number of econ disadv individuals receiving financial literacy services	100	106
Outcome	O9: Individuals with improved financial knowledge.	40	106

- The number of participants with increased knowledge (outcome) is exactly the same as the number of program participants (output)
- Possible explanations:
  - The program is extremely effective in achieving improved financial knowledge among program participants
  - The number of individuals with improved financial knowledge is not being reported correctly

Example #1 Additional Context

Financial Literacy Program Post-Assessment

Assessment of Knowledge Gains

Please respond to the questions below based on your experience participating in the Financial Literacy Corps' financial literacy training program.

1. How satisfied were you with the program?

Very Satisfied	Somewhat Satisfied	Neutral	Somewhat Dissatisfied	Very Dissatisfied
○	○	○	○	○

2. Do you feel that you learned something from your participation in the program?

○ Yes

○ No

3. What suggestions do you have for how to improve the program for future participants?

Done

Example #1 Additional Context

Financial Literacy Program Post-Assessment

Assessment of Knowledge Gains

Please respond to the questions below based on your experience participating in the Financial Literacy Corps' financial literacy training program.

1. How satisfied were you with the program?

Very Satisfied	Somewhat Satisfied	Neutral	Somewhat Dissatisfied	Very Dissatisfied
○	○	○	○	○

2. Do you feel that you learned something from your participation in the program?

- Yes
- No

3. What suggestions do you have for how to improve the program for future participants?

- **This assessment is primarily a customer satisfaction survey that does not objectively measure changes in knowledge**
- **To ensure validity, the grantee should use a pre-post assessment tool that asks content-based questions directly related to the subject matter**

Example #2

Demographics

Demographics Information	Value
Number of individuals who applied to be AmeriCorps members	116
Number of episodic volunteers generated by AmeriCorps members	389
Number of ongoing volunteers generated by AmeriCorps members	68
Number of AmeriCorps members who participated in at least one disaster services project	44
Number of disasters to which AmeriCorps members have responded	10
Number of individuals affected by disasters receiving assistance from members	2,000
Number of veterans serving as AmeriCorps members	2

Example #2 Continued

Demographics

Demographics Information	Value
Number of individuals who applied to be AmeriCorps members	116
Number of episodic volunteers generated by AmeriCorps members	389
Number of ongoing volunteers generated by AmeriCorps members	68
Number of AmeriCorps members who participated in at least one disaster services project	44
Number of disasters to which AmeriCorps members have responded	10
Number of individuals affected by disasters receiving assistance from members	2,000
Number of veterans serving as AmeriCorps members	2

- The number of individuals affected by disaster receiving assistance from members is an unusually round number
- Possible explanations:
  - Members served exactly 2,000 individuals over the course of the program year
  - The number of individuals receiving assistance from members is not reported correctly

Example #2 Additional Context

- Grantee Narrative
- Program Officer Analysis
- Program Officer Feedback
- Program Officer Clarification Items
- Grantee Clarification

Members responded to approximately one natural disaster per month over the course of the 10-month program year. We estimate that our cohort of members assisted about two hundred individuals during each disaster event, resulting in a total of 2000 individuals (10 events x 200 individuals per event).

Status

- [x] Clarification Required
- [ ] Clarification Resolved
- [ ] Review Complete

Example #2 Additional Context

- Grantee Narrative
- Program Officer Analysis
- Program Officer Feedback
- Program Officer Clarification Items
- Grantee Clarification

Members responded to approximately one natural disaster per month over the course of the 10-month program year. We estimate that our cohort of members assisted about two hundred individuals during each disaster event, resulting in a total of 2000 individuals (10 events x 200 individuals per event).

Status

- [x] Clarification Required
- [ ] Clarification Resolved
- [ ] Review Complete

- The grantee estimated the value of the demographic indicator rather than measuring it

- To ensure verifiability, the grantee should ensure that they have specific data collection procedures for all reported values and are maintaining source documentation for each number

Corrective Actions for Data Quality Issues

- Notify your CNCS Program/Grants Officer and discuss best way forward
- For issues related to invalid tools, incorrect protocols, or wrong definitions:
  - Switch to the correct tool/protocol/definition immediately if feasible; if not, switch in the next program year
  - Do not report data collected using incorrect tool/protocol/definition; document reasons for not reporting
- For issues related to incomplete reporting, math errors, or double-counting/double-reporting:
  - Correct the values
  - Put together quality control procedure

Corrective Actions for Data Quality Issues

- For issues related to missing/incomplete source documentation:
  - Develop a system for retaining source data
  - Do not report values for which there is little/no evidence. Document reasons for not reporting.

Resources for Data Quality Review

- Performance Measurement Core Curriculum (http://www.nationalservice.gov/resources/performance-measurement/training-resources)
  - Performance Measurement Basics
  - Theory of Change
  - Evidence
  - Quality Performance Measures
  - **Data Collection and Instruments**

More Resources for Data Quality Review

- Evaluation Core Curriculum: Implementing an Evaluation (http://www.nationalservice.gov/resources/evaluation/implementing-evaluation)
  - Basic Steps of an Evaluation
  - **Data Collection**
  - Managing an Evaluation
- CNCS Monitoring Tool: Data Quality Review Tab

CNCS Monitoring Tool

Monitoring Tool Data Quality Only 1June2015 [Read-Only] - Excel

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW

Reynolds, Barbara

F1 | Data Quality

Data Quality

		Not assessed	Yes	No		Comments
2					The Data Quality review is designed to assess the quality of reported data. The review entails engaging grantees in a dialogue about the underlying processes used to gather and handle data and will cover the following data quality elements: validity, completeness, consistency, accuracy, and verifiability.	
4					***Validity*** *addresses whether the data collected and reported appropriately relates to the approved program model and whether or not the data collected corresponds to the information provided in the grant application. In evaluating validity, CNCS staff will determine if the performance measure data being collected appear to measure the project's demographic data, performance measures, and/or outcomes.*	
6	1				Are the reported data consistent with the approved program design?	
8	2				Is the program measuring what it intended to measure?	
10					***Completeness.*** *Data reported to CNCS are considered complete when the grantee collects enough information to represent an activity, a population, and/or a sample. In evaluating completeness, CNCS staff will determine if the data collected and reported contains enough information to represent the activities or population as related to the performance measure.*	
12	3				Does the grantee collect all of the data relevant to the measure? *Consider :* *- Does it collect data from all sites? (if applicable to the grantee's program model)* *- Does it collect data with the same frequency from all sites? (if applicable to the grantee's program model)*	
14	4				Is any vital information (e.g. for a period of time or for a group of participants) missing from the data reported or collected? *Consider :* *- If data are missing, does the grantee have documentation to explain missing data?* *- If data are missing, what are the grantee's plans for resolving issues and collecting missing data in the future?*	

intro | data quality | Valid

READY 130%

8:26 AM 8/26/2015

Q&A

- *What questions do you have?*

Thank you!

- If you have questions or feedback about this presentation, feel free to contact us:

Sarah Yue: syue@cns.gov

Jim Stone: jstone@cns.gov