M&E Attribute: Data Quality Assessment Checklist

STEPS:

1. Review the Performance Monitoring Plan (PMP)
2. Identify the critical information that needed data quality assessment
3. Identify if the source of information is from a survey, mission visits, routine visits, scheduled M&E visits, etc
4. Follow the checklist below

==========

Refer to this checklist when the team (or M&E Unit staff) conducts both initial and periodic data quality assessments. The full list does not have to be completed—the team may wish to identify the most critical data quality issues for formal or informal assessment.

Name of Project Purpose or Specific Objectives :

Name of Intermediate Result (if applicable):

Name of Performance indicator: PP1.1…. or ER1.11

Data source(s):

Partner or contractor or implementing agency who provided the data (if applicable):

Year or period for which the data are being reported:

Is this indicator reported in the Annual Report? *(circle one)* YES NO

Date(s) of assessment:

Location(s) of assessment:

Assessment team members:

For Office Use Only

SO team leader approval: X_____Date_____

Mission director or delegate approval:
X_____Date_____

Copies to:

Comments

1. VALIDITY—Do the data adequately represent performance?

	Yes	No	Comments
Face Validity			
➢ Is there a solid, logical relation between the activity or program and what is being measured, or are there significant uncontrollable factors?	❑	❑	
Measurement Error			
Sampling Error (only applies when the data source is a survey)			
➢ Were samples representative?	❑	❑	
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	
➢ Were response rates sufficiently large?	❑	❑	
➢ Has non-response rate been followed up?	❑	❑	
Non Sampling Error			
➢ Is the data collection instrument well designed?	❑	❑	
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	❑	
➢ Are definitions for data to be collected operationally precise?	❑	❑	
➢ Are enumerators well trained? How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?	❑	❑	
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	
Transcription Error			
➢ What is the data transcription process? Is there potential for error?	❑	❑	
➢ Are steps being taken to limit transcription error? (e.g., double keying of data for large surveys, electronic edit checking program to clean data, random checks of partner data entered by supervisors)	❑	❑	
➢ Have data errors been tracked to their original source and mistakes corrected?	❑	❑	
➢ If raw data need to be manipulated to produce the data required for the indicator:	❑	❑	

1. VALIDITY—Do the data adequately represent performance?

	Yes	No	Comments
➢ Are the correct formulae being applied?	❑	❑	
➢ Are the same formulae applied consistently from year to year, site to site, data source to data source (if data from multiple sources need to be aggregated)?	❑	❑	
➢ Have procedures for dealing with missing data been correctly applied?	❑	❑	
➢ Are final numbers reported accurate? (E.g., does a number reported as a "total" actually add up?)	❑	❑	
Representativeness of Data			
➢ Is the sample from which the data are drawn representative of the population served by the activity?	❑	❑	
➢ Did all units of the population have an equal chance of being selected for the sample?	❑	❑	
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	❑	❑	
➢ Is the sample of adequate size?	❑	❑	
➢ Are the data complete? (i.e., have all data points been recorded?)	❑	❑	

Recommendations for improvement:

2. RELIABILITY—Are data collection processes stable and consistent over time?

	Yes	No	Comments
Consistency			
➢ Is a consistent data collection process used from year to year, location to location, data source to data source (if data come from different sources)?	❑	❑	
➢ Is the same instrument used to collect data from year to year, location to location? If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	❑	❑	
➢ Is the same sampling method used from year to year, location to location, data source to data source?	❑	❑	
Internal quality control			
➢ Are there procedures to ensure that data are free of significant error and that bias is not introduced?	❑	❑	
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	❑	❑	
➢ Do these procedures provide for periodic sampling and quality assessment of data?	❑	❑	
Transparency			
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	❑	❑	
➢ Are data problems at each level reported to the next level?	❑	❑	
➢ Are data quality problems clearly described in final reports?	❑	❑	

Recommendations for improvement:

3. TIMELINESS—Are data collected frequently and are they current?

	Yes	No	Comments
Frequency			
➢ Are data available on a frequent enough basis to inform program management decisions?	❑	❑	
➢ Is a regularized schedule of data collection in place to meet program management needs?	❑	❑	
Currency			
➢ Are the data reported in a given timeframe the most current practically available?	❑	❑	
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	❑	❑	
➢ Are the data reported as soon as possible after collection?	❑	❑	
➢ Is the date of collection clearly identified in the report?	❑	❑	

Recommendations for improvement:

4. PRECISION—Do the data have an acceptable margin of error?

	Yes	No	Comments
➢ Is the margin of error less than the expected change being measured?	❑	❑	
➢ Is the margin of error is acceptable given the likely management decisions to be affected? (consider the consequences of the program or policy decisions based on the data)	❑	❑	
➢ Have targets been set for the acceptable margin of error?	❑	❑	
➢ Has the margin of error been reported along with the data?	❑	❑	
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	❑	❑	

Recommendations for improvement:

5. INTEGRITY—Are data are free of manipulation?			
	Yes	**No**	**Comments**
➢ Are mechanisms in place to reduce the possibility that data are manipulated for political or personal reasons?	❑	❑	
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	❑	❑	
➢ Has there been independent review?	❑	❑	
➢ If data is from a secondary source, is the management confident in the credibility of the data?	❑	❑	
Recommendations for improvement:			

For indicators for which no recent relevant data are available
If no recent relevant data are available for this indicator, why not?
What concrete actions are now being undertaken to collect and report this data as soon as possible?
On what date will data be reported?

Rebecca V Polestico revised with reference to USAID docs 2008