



DATA COLLECTION METHODS AND TOOLS FOR PERFORMANCE MONITORING

Introduction

This resource supplements [ADS Chapter 201: Program Cycle Operational Policy](#) (in particular, sections 3.4.10 and 3.5.2 – 3.5.7) by providing an introduction to Data Collection Methods and Tools for Performance Monitoring.

Data collection is the process of systematically gathering quantitative and/or qualitative data used for purposes of monitoring, evaluation, and/or learning (MEL). Performance monitoring data are used to reveal whether project and activity implementation is on track and whether expected results are being achieved. As a result, performance data collection is critical to activity evaluation, accountability, and learning. Performance data collection supports the Program Cycle principle to “Manage Adaptively through Continuous Learning.” (See the [USAID Discussion Note: Adaptive Management](#).)

Performance monitoring data about USAID’s activities may be collected by USAID staff, implementing partners, and/or MEL support contractors. USAID may also use data collected by independent sources, such as local government statistical offices, for the purposes of USAID performance monitoring, where appropriate. Regardless of who collects performance data for USAID, it is ultimately USAID’s responsibility to make sure that performance monitoring data collected are of sufficiently high quality to support management needs.

This resource covers the basics of data collection for performance monitoring, including: primary and secondary types of data sources, common data collection methods, and the process of identifying appropriate data collection tools. The primary audience is USAID Program Officers, Monitoring, Evaluation, and Learning Specialists, Technical Officers, implementing partners, and MEL Support Contractors.

Types of Data: Primary and Secondary

Data sources may be categorized as either primary or secondary. Primary data are data collected directly by an investigator, researcher, or other staff of an organization, usually for a specific purpose. Primary data include data collected by USAID staff and data collected at USAID’s instruction by implementing partners, including MEL support contractors and external evaluators. Secondary data are data that have already been collected by other organizations, such as government agencies, public international organizations, and independent research organizations. Examples of secondary data sources include government population census data and country-level indices, such as the World Bank’s [Ease-of-Doing Business Index](#). While this document is mainly about primary data collection, both types of data may be of use for performance monitoring, though secondary data is often more appropriate for context monitoring.

PRIMARY DATA

Primary data collection for performance monitoring is common throughout USAID activities and is essential for monitoring the outputs and intermediate outcomes of our programming. For example:

- Implementing partners collect primary data when they record the number of beneficiaries who attend training.
- USAID staff members collect primary data when they record information about activities that they observe during site visits.
- MEL support contractors collect primary data when they conduct focus groups or survey program beneficiaries to obtain their feedback on the quality of services they are receiving from a USAID activity.

There are both advantages and disadvantages of using primary data. Foremost among the advantages is that the data are collected specifically for USAID's purposes. Primary data collection allows USAID and its partners to focus data collection efforts on its own activity beneficiaries in the technical sectors and geographic regions in which USAID works. This helps ensure that performance data are measuring the results to which U.S. foreign assistance has actively contributed. Primary data collection also gives USAID and its partners control over *how* the data are collected and the flexibility to determine how much data needs to be collected to serve its purposes.

Supporting Local Data Collection for Self-reliance

USAID should work with other donors, government agencies, and international organizations to harmonize data collection efforts, where possible. It is important to align with the priorities of partner governments and strengthen local data collection systems and institutions rather than always defaulting to USAID led primary data collection efforts.

The main disadvantage of primary data collection is that it is often costly relative to secondary data collection. Focus groups, surveys, and many other methods of data collection require specialized skills and may take considerable time, effort, and resources. Starting a *new* primary data collection effort also inherently entails some risks to data quality, since it is impossible to foresee all of the possible problems that may arise during a new data collection effort. Primary data collection can also be a burden for those from whom the data is collected, especially if it is not well-coordinated or harmonized with other data needs of the USAID Mission, or those of other donors or local partners, leading to multiple overlapping efforts to collect data from the same populations. Finally, because primary data collection by implementing partners occurs concurrently with the implementation of USAID activities, such data rarely includes useful information about pre-baseline and post-endline trends.

SECONDARY DATA

Secondary data are often used at USAID for the performance monitoring of higher-level outcomes, such as Intermediate Results in USAID Country Development Cooperation Strategies. Secondary data are also commonly used for *context monitoring*, or the collection of information about conditions and external factors relevant to the implementation and performance of strategies, projects, and activities. This includes information on local conditions or external factors that may directly or indirectly affect implementation and performance of USAID efforts. Many sources of country level secondary data are available at [IDEA](#).

As with primary data, there are a number of advantages and disadvantages to using secondary data. One advantage is that the data have already been collected. As a result, USAID will typically expend fewer resources to obtain secondary data compared to collecting primary data. In non-permissive

DATA COLLECTION METHODS AND TOOLS FOR PERFORMANCE MONITORING

environments, where USAID or its partners may find it difficult to collect primary data, local sources of secondary data may be particularly useful. In addition, secondary data is often collected over a longer time horizon than most USAID strategies, projects, or activities. This longer time horizon allows for analysis of pre-baseline and post-endline trends. The main disadvantage of secondary data is that it is not collected for USAID’s purposes and therefore may not align with the timing, scope, or population focus of USAID programming. This often makes secondary data difficult to use for monitoring activity outputs and beneficiary level outcomes.

In terms of data quality, the use of secondary data exposes USAID to different risks than primary data collection. On the one hand, the data are already collected, so USAID has an opportunity to assess the quality of the data prior to choosing to rely on it for monitoring. Many sources of secondary data have developed extensive expertise in their area of focus leading to high-quality data. However, the quality of secondary data can vary widely and assessing the quality of secondary data can often be difficult. Moreover, as it is outside USAID’s control, data collected by other organizations may change over time (or end altogether) in ways that can lead to problems with trying to analyze trends.

Select Appropriate Data Collection Method

A data collection method refers to the procedure for how data are collected. Quantitative data collection methods produce countable or numerical results. Qualitative data collection methods produce non-numerical data, such as perceptions and descriptions. While performance monitoring is often associated with quantitative indicators, data collection methods for performance monitoring may be either quantitative or qualitative.

The table below presents a snapshot of some of the data collection methods for performance monitoring:

Commonly Used Data Collection Methods	
Recording Data Through Administrative Actions	<p><i>Recording data through administrative actions</i> in the course of implementing activities is one of the most common methods of data collection, particularly for our implementing partners. Examples include recording attendance at training courses, awarding of grants to local organizations, hours of technical assistance provided, and deliveries of food aid.</p> <p>Recording data through administrative actions is primarily a method of quantitative data collection.</p>
Electronic Data Harvesting	<p><i>Electronic data harvesting</i> encompasses data collection of electronically generated data. This could include a record of people’s actions in an online environment (e.g., number of downloads) via texts or apps on mobile devices, social media data (e.g., “tweets” on Twitter), or data generated from cell phones and other mobile devices (such as human mobility data)</p> <p>Electronic data harvesting is a method of quantitative data collection.</p>

Commonly Used Data Collection Methods	
Survey	<p>A <i>survey</i> comprises a structured series of questions that respondents are asked according to a standard protocol. A survey tends to include mostly closed-ended questions, but can also include open-ended questions (e.g., to explore or better understand responses to closed-ended questions).</p> <p>A survey is conducted on a total population of interest (census), on a representative sample of the population of interest that through statistical methods can be generalized to the total population, or a non-representative sample of the population of interest that is not generalizable.</p> <p>Surveys are primarily a method of quantitative data collection, though survey questions can be either quantitative or qualitative in nature, and can measure coverage (i.e., who received an intervention), satisfaction, perceptions, knowledge, attitudes, and reported actions or behaviors.</p>
In-depth Interview (IDI)	<p>An <i>in-depth interview</i> is usually conducted one-on-one by an interviewer who asks an interviewee about their knowledge, experiences, feelings, perceptions, and preferences on a certain topic. IDIs can also be conducted with a group though this may not always be appropriate or optimal. The interviewer relies on a structured, semi-structured, or unstructured question guide or list of themes/points to be discussed and often encourages a free flow of ideas and information from the interviewee.</p> <p>A <i>Key Informant Interview (KII)</i> is a type of IDI, whereby an interviewee is selected for their first-hand knowledge of the topic of interest or geographical setting (e.g., community).</p> <p>IDIs are a method of qualitative data collection.</p>
Focus Group Discussion (FGD)	<p>A <i>focus group discussion</i> involves a skilled moderator who stimulates discussion among a group of individuals to elicit experiences, feelings, perceptions, and preferences about a topic. The moderator uses a list of topics to be discussed, ensures all voices are represented, and keeps the discussion on track. Focus group data may include information about body language, group dynamics, and tone, in addition to what is said. Typically, groups comprise 6-12 purposively selected participants; however, size and selection techniques may vary. Focus groups differ from group interviews in format, how they are facilitated, who may be chosen to participate, and the types of data that come out of the process.</p> <p>FGDs are a method of qualitative data collection.</p>

Commonly Used Data Collection Methods	
Observation	<p>Direct observation entails a trained (human) observer who records data based on what they see, hear, or touch, often based on a guided protocol. Examples include observation of skills-based performance and observations of a physical environment or setting of an intervention.</p> <p><i>Participant observation</i> involves a researcher participating in an activity and making observations informed by their experience interacting with others during the activity.</p> <p><i>Remote observation</i> or <i>remote sensing</i> entails gathering observational data through observation at a distance with the assistance of technology (e.g., satellite or aircraft-based imagery). Remote data collection is particularly useful in non-permissive environments.</p> <p><i>Observation</i> is more often used as a method of qualitative data collection but can also be used for quantitative data collection, especially when focused on the number of occurrences of a specific item, event or action.</p>

CONSIDERATIONS TO GUIDE METHOD SELECTION

In selecting which data collection method to choose for performance monitoring, USAID and its partners should primarily be guided by the monitoring *purpose* to be achieved by the data collection effort. Knowing the purpose of the data collection effort will help determine if a quantitative or qualitative method is more appropriate. For example, if the purpose of data collection is to determine whether the expected number of beneficiaries received a particular service, then a quantitative method, such as survey, might be appropriate. If the purpose is to understand the quality of instruction in an implementing partner’s training program, then direct observation or a focus group of participants may be appropriate.

Additional points to consider when selecting a data collection method for performance monitoring include:

- **Frequency:** How often are the data needed and with what frequency is the data expected to change? For example, in-person surveys can take a long time to administer and analyze, therefore they are more appropriate for monitoring outcomes that won’t change rapidly.
- **Rigor:** How rigorous does the data collection need to be? Not all monitoring data need to be the same level of rigor. For example, observational data collected during site visits can be helpful in identifying implementation problems, even if it did not involve a rigorous selection of program sites. For measuring a key performance outcome, however, such as change in household income of beneficiaries, a rigorous household survey of a representative sample of beneficiaries would be more appropriate.
- **Analysis:** How will data be analyzed? Data collection and analysis are directly related.

Data Collection Methods for Performance Indicators

When data are being collected for a performance indicator, review the steps outlined in the [USAID’s Selecting Performance Indicators guidance document](#). If data are being collected for a Standard Foreign Assistance Indicator, then the data collection method should conform to the information provided in the [Standard Foreign Assistance Indicator reference sheet](#) for that indicator.

DATA COLLECTION METHODS AND TOOLS FOR PERFORMANCE MONITORING

Consider data analysis plans while selecting your data collection method, so that you have exactly the kind of data you want to analyze.

- **Personnel:** Who and how many people will be needed to collect, process, and analyze the data? Different data collection methods may require a different number of personnel with varying qualifications. For example, monitoring specialists who excel at tracking administrative data may not be well-suited for facilitating focus groups, and vice versa.
- **Local context:** Is the data collection method appropriate in the local context? For example, in non-permissive environments, household surveys might put survey enumerators at risk.
- **Cost:** Is the estimated cost of the data collection method reasonable? High- and low-cost methods both have their place in performance monitoring, but high-cost methods should be reserved for when rigor or timeliness is of greatest importance.

Develop a Data Analysis Plan

A data analysis plan should be started as early as possible in the process of making choices about data collection efforts and should be completed prior to data collection. A data analysis plan documents the purpose of the data collection, the questions that the data collection and analysis is expected to answer, the specific data that will be collected, and how the data will be analyzed and used.

Documenting data collection efforts helps ensure:

- Data collection descriptions, definitions, units of measure, and calculations are specific enough to make clear how data are collected, compiled, and analyzed;
- Data collection responsibility or oversight is assigned to a specific individual, office, or team;
- Data collection methods are consistent and comparable over time;
- Data collection tools are clearly documented; and
- Data limitations are identified and clearly documented along with plans for mitigation.

For performance monitoring data collected for performance indicators, details should be recorded in the Performance Indicator Reference Sheet (PIRS) for each indicator. (See [USAID PIRS Guidance and Template for additional information](#).) The very exercise of filling out the PIRS may lead to further consideration of the chosen data methods and tools.

Ethical and security issues should also be considered and documented in the Data Analysis plan. Issues to address may include:

- How informed consent will be obtained from those agreeing to provide the data;
- Whether the data can be reused for different purposes;
- How to minimize the unnecessary collection of personally identifiable data;
- How data will be stored and protected from unauthorized access; and
- Whether and when data will eventually be deleted.

This plan for an individual data collection effort should be aligned to the overall Activity MEL Plan.

IDENTIFY AND PILOT SELECTED DATA COLLECTION TOOL(S)

A data collection tool is an instrument used to collect data, such as a discussion guide, paper-based survey questionnaire, or computer-assisted interviewing system (e.g., tablets, mobile phones, or computers). After determining the specific data source(s), identifying the appropriate data collection method(s) for the data required, and starting a data analysis plan, USAID and its partners need to identify whether a tool needs to be developed or if an existing tool can be adapted to collect the data.

Many data collection instruments already exist and only need to be carefully reviewed to determine to what extent customization (including adapting to the local context) is appropriate. When possible, USAID and its partners should look to validated data collection instruments – such as the Early Grade Reading assessment (EGRA) – before developing a new data collection instrument. When an existing data collection instrument is not appropriate, a new instrument may need to be developed. The basic process of developing a new data collection instrument usually involves the following steps:

- **Review your purpose:** Review your previously identified data needs and the data collection method that you selected to achieve that purpose. Identify how data may need to be disaggregated. Determine the type of instrument that will yield the data needed.
- **Create instrument:** Formulate the content of your instrument to respond to your identified data needs. Format the instruments (e.g., discussion guide, questionnaire, observation checklist, etc.) and include guidelines and instructions to be followed by the data collectors (i.e., discussion moderators/facilitators, interviewers/enumerators, observers, etc.). Translate the instruments to the target language(s).
- **Test instrument:** Test and adjust your instrument with a convenience sample of targeted respondents so as to calibrate the content to respondents' context and cognitive processes.
- **Train personnel:** Identify personnel to collect and process the data (field supervisors, interviewers, enumerators, coders, and other relevant personnel). Determine how the data will be analyzed and by whom. Train personnel on implementation of the instrument and how to process the data.
- **Pilot data collection:** Pilot test the instrument with a small population, preferably in the local context where the instrument will be used to collect data. Revise the instrument and retest as needed.

COLLECT THE DATA

Thoughtful planning for data collection will help ensure that the actual collection of data goes smoothly. But good planning is not enough. Oversight and attention are necessary throughout the data collection process to ensure both responsible and high-quality data collection.

Responsible and Ethical Data Collection: A USAID monitoring principle is to be responsible and ethical. Collecting data responsibly includes protecting participants and human subjects from harm, obtaining informed consent to collect and use data, safeguarding individual privacy and security, protecting the sensitive information of vulnerable and marginalized populations, and applying high ethical standards to deciding what information to collect and how to use it (see ADS 201.3.5.2).

Data collection should be approached with cultural, religious, and gender sensitivity, recognizing that norms in our partner communities may be different than yours. For instance, in some countries, surveys are not viewed positively or respondents have been over-surveyed. USAID programs may also address

DATA COLLECTION METHODS AND TOOLS FOR PERFORMANCE MONITORING

sensitive health or other outcomes, and collecting data may put people at risk for stigmatization. Some communities may have a negative reaction to having their photo taken or not feel comfortable discussing some topics in a focus group setting. (See the [USAID How-To Note: Engendering Evaluation](#) for more information on gender-sensitive data collection.)

USAID and its implementing partners should also pay close attention to confidentiality concerns of respondents and safeguard access to government, respondent, and other partners' information so that data (including electronic information) are not improperly disclosed or linked back to the individuals from whom data were collected. For example, discuss informed consent with respondents before starting individual or group interviews, including guardians of any children under the country's legal age of consent. Conduct interviews in a safe space where respondents will not be observed or overheard. The use of unique identifiers (such as a number) instead of a name may be of use when confidentiality is a concern, particularly if person level data sets are expected to be shared. To the extent possible, adhere to national or universally-recognized ethical standards for collection, maintenance, and reporting of personal data and information. Consider if an ethics review by a qualified third-party is required.

For more information on protecting beneficiaries' data, please see [USAID Data Security Guidance: Protecting Beneficiaries](#).

Quality Data Collection: As noted in ADS 201, high-quality data are the cornerstone for evidence-based decision making. To ensure that the quality of evidence is adequate and sufficient for decision making, performance data should reasonably meet the five USAID standards of data quality: validity, integrity, precision, reliability, and timeliness.

During data collection, all data face threats to quality. Data can be distorted, systematically biased, or erroneous due to errors in sampling design, poor implementation of methods, or problems in recording data. For example, biases can occur in the data because the interviewer asks a question in a way that encourages one response over another. Or, measurement bias may occur because a survey question is too personal or too sensitive for respondents to answer truthfully. Even the simple act of transcribing responses from a paper survey to an electronic database can lead to errors that ultimately affect the overall quality of the data. While the specific procedures for ensuring high-quality data collection will differ according to each specific data collection method, oversight and validation are necessary components of any data collection process. For instance, random spot checks comparing indicator data to the original source of the data can help determine if high standards of data collection are being maintained.

For performance indicator data that is reported externally, Data Quality Assessments (DQAs) are a required procedure for helping to ensure that data quality is being maintained. (See ADS 201.3.5.7 and the [How-To Note on Conducting a Data Quality Assessment](#) for additional detail on Data Quality Assessments.) Where the data quality is a particular concern, for instance when an implementing partner is attempting a technically difficult data collection method, conducting a DQA should be considered even when it is not required by policy. Ensuring high-quality data collection will ultimately support credible analysis of performance for decision making.

REFERENCES

- › [ADS Chapter 201: Program Cycle Operational Policy](#)
- › [ADS Chapter 205: Integrating Gender Equality and Women's Empowerment in USAID's Program Cycle](#)
- › [ADS Chapter 579: USAID Development Data](#)
- › [ADS Mandatory Reference: USAID PIRS Guidance and Template](#)
- › [USAID How-To Note: Conduct a Data Quality Assessment](#)
- › [USAID Discussion Note about Adaptive Management](#)
- › [USAID Selecting Performance Indicators](#)
- › [USAID How-To Note: Engendering Evaluation at USAID](#)
- › [USAID Data Security Guidance: Protecting Beneficiaries](#)