

# Archived Information

---

## Data Quality (DQ) Standards

---

These standards are intended to assist ED managers as they collect, analyze, and report data about Federal education programs. Although no single document can anticipate the entire range of data uses, we designed these standards to have broad applicability.

Program managers can use the standards as a tool when monitoring grantees and evaluating the quality of the reported data and preparing submissions for the GPRA annual report.

To fully evaluate the quality of program data, data managers must ask themselves:

- Have I selected the appropriate measures? (Standard 1: Validity)
- Am I collecting sound data on those measures? (Standards 2 through 5: Accurate Description, Editing, Calculation, and Timeliness)
- Am I reporting the measures accurately? (Standard 6: Reporting)

These standards are an attempt to provide criteria against which to evaluate these aspects of program DQ. Not every example or item on the checklist is relevant to every indicator, or appropriate for every program.

### Relationship to the ED Strategic Plan

To ensure that program indicator performance data are of the highest quality, and to assist program managers in reaching toward this quality, the ED has included in its *Strategic Plan for 1998-2002* (September 30, 1997) a performance indicator on DQ:

**Performance Indicator 30. “By 2000, all ED program managers will assert that the data used for their program’s performance measurement are reliable and valid or will have plans for improvement.”**

This idea is further developed in ED’s Annual Plans. The *Fiscal Year 2000 Annual Plan* describes a five-part improvement strategy to ensure the quality of performance indicator information. The five parts are as follows:

- *Develop Department-wide standards for performance indicator measurement.* These standards have been developed as part of this strategy.
- *Programs systematically review the quality of their data collection systems.* Program managers use these standards to review performance indicators for their program.
- *Training of ED program data managers in the application of data standards for performance measurement.* Department-wide training on DQ standards for program data managers includes identification of the limitations of key performance indicators and development of concrete plans for their improvement.
- *Monitor DQ.* Program managers, division directors, and assistant secretaries examine the GPRA indicators and data for their programs to determine their accuracy and validity, and to develop plans for their improvement. In

addition, the Office of Inspector General (OIG) independently monitors the reliability of its data quality (DQ) in high-priority areas.

- *Managers attest to the reliability and validity of their performance measures or submit plans for data improvement.* Based on the information and improvement plans provided by the program managers and division directors, assistant secretaries will assert that they are aware of any DQ limitations and concur with the recommendations and plans for improvement.

## **Relationship to Legislative Requirements for DQ**

Some programs have specific criteria specified in the legislation for judging the quality of the performance indicators and other data provided by the program. For example, the Perkins Vocational and Applied Technology Act requires that data be “reliable, complete, and accurate.” Other programs’ data are expected to be “reliable and valid.” These standards are an attempt by the ED to develop a single set of DQ standards for all programs and to be concrete about what constitutes high-quality data in words understandable to nonstatisticians.

## **How this Document Is Structured**

The document consists of six standards for judging program performance DQ: validity, accurate description, editing, calculation, timeliness, and reporting. Associated with each standard are definitions, examples, and possible checks for application of that standard.

The DQ checklist for each standard consists of a series of questions that both primary data providers and secondary data managers should ask themselves as they evaluate the quality of the data. A primary data provider is a person or organization who carries out all or part of the study design, data collection, data processing, and initial reporting. A secondary data manager is someone who sponsors or requests a primary data collection or who uses data from the reports for other purposes. Not every item on the DQ checklist might be appropriate for every study; however, each represents a step that can be taken to ensure the quality of program data.

Some standards require a distinction between *sample* and *universe* data. Universe data are data collected from every unit in the population. For example, universe data for school districts include responses from *every* school district in the nation. By comparison, representative sample data are collected from only a fraction of the population. But this fraction is chosen in such a way that it describes *all* school districts. The distinction between sample and universe data is important. Sample data have measures associated with them related to their sampling variability (e.g., confidence intervals) typically not associated with universe data.

# Standard One: Validity— Data adequately represent performance.

DEFINITION	Examples Meeting the Standard and Failing to Meet the Standard
<ul style="list-style-type: none"> <li>• The performance indicators actually measure the goal or objective of interest.</li> <li>• The data “match” the performance indicator because they measure the same phenomenon.</li> <li>• The indicators and data are a useful guide to policy decision making.</li> </ul>	<p><b>Meeting:</b></p> <ul style="list-style-type: none"> <li>• The performance objective, indicator, and data all describe the same phenomenon.</li> <li>• The performance objective, indicator, and data all focus on the phenomenon of interest.</li> </ul> <p><b>Failing:</b></p> <ul style="list-style-type: none"> <li>• The program objective does not have a realistic, measurable performance indicator.</li> <li>• The data used to measure the performance indicator actually measure a somewhat different phenomenon.</li> <li>• The indicators provide data about aspects of the program that are unrelated to policy questions.</li> </ul>

<b>Validity Data Quality (DQ) Checklist</b> <i>(Most important checks are in italics)</i>
<p><b>For Primary Data Providers</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Is the way in which the data are being used to measure performance on the indicator appropriate given what you know about how they were collected?</li> </ul> <p><b>For Secondary Data Managers</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Do the indicators cover aspects of the program that are useful and important for policy decision making?</li> <li><input type="checkbox"/> <i>Have the objective, performance indicator, and data been scrutinized to be sure that they all describe the phenomenon of interest?</i></li> <li><input type="checkbox"/> <i>Is a realistic plan in place to collect data on all performance measures?</i></li> <li><input type="checkbox"/> If the available data do not appropriately match the indicator or objective, is the mismatch clearly stated in reporting, and are plans under way to properly align them?</li> </ul>

## Standard Two: Accurate Description— Definitions and counts are correct.

<b>DEFINITION</b>	<b>Examples Meeting the Standard and Failing to Meet the Standard</b>
<p><b>Generally for All Data</b></p> <ul style="list-style-type: none"> <li>All data providers use the same agreed-upon definitions.</li> <li>The phenomena being measured occurred or existed at the time for which they were reported.</li> </ul> <p><b>Specifically for Universe Data</b></p> <ul style="list-style-type: none"> <li>All data providers use the same agreed-upon definition.</li> <li>All instances of a phenomenon are counted, and no instances are omitted.</li> </ul>	<p><b>Generally for All Data</b></p> <p><i>Meeting</i></p> <ul style="list-style-type: none"> <li>Key terms are clearly defined and definitions followed by data providers.</li> <li>Data are provided for the correct time period.</li> <li>Counted program recipients actually received services.</li> </ul> <p><i>Failing</i></p> <ul style="list-style-type: none"> <li>Each data provider uses different definitions of key terms.</li> <li>Certain types of phenomena (e.g., private schools, disabled students) are overlooked or not consistently included in counts.</li> </ul> <p><b>Specifically for Universe Data</b></p> <p><i>Failing:</i></p> <ul style="list-style-type: none"> <li>Phenomena are double counted or undercounted.</li> <li>Estimates are provided instead of actual counts.</li> </ul>
<p><b>Specifically for Sample Data</b></p> <ul style="list-style-type: none"> <li>Population from which the sample is drawn is clearly defined.</li> <li>Samples are taken at the appropriate unit of analysis.</li> <li>Samples are of sufficient size to enable generalizations to the population to be made.</li> </ul>	<p><b>Specifically for Sample Data</b></p> <p><i>Meeting</i></p> <ul style="list-style-type: none"> <li>Data are provided for the correct time period.</li> <li>Those who are sampled are actually members of the population being studied.</li> </ul> <p><i>Failing</i></p> <ul style="list-style-type: none"> <li>Certain types of phenomena (e.g., private schools, disabled students) are not reported in proper proportion to their occurrence in the population.</li> </ul>

## Accurate Description DQ Checklist

*(Most important checks are in italics)*

### Generally for All Data

#### *For Primary Data Providers*

- Have definitions and time periods been followed or discrepancies clearly documented?

#### *For Secondary Data Managers*

- Have clear written definitions of key terms (including exclusions/inclusions) been communicated to data providers?*
- Do reporting forms provide spaces for data providers to report deviations from definitions and uses of estimation at the time they provide the data?*
- Have you invited feedback from data providers about data collection issues and possible problems?
- Have data been compared with other databases reporting similar statistics?
- Have entities for which counts have changed more than 10 percent since the previous report been double-checked?

### Specifically for Universe Data

#### *For Primary Data Providers*

- Have estimates been used for no more than 10 percent of the phenomena counted, and are estimates clearly differentiated from actual counts?
- Have counts been tallied at least twice, and do totals agree?

#### *For Secondary Data Managers*

- Have respondents been involved in setting definitions for key terms?
- Have definitions been communicated in sufficient time for data providers to prepare their system to properly implement them?
- Are independent undercount and overcount checks in place for counts associated with major program funding?

### Specifically for Sample Data

#### *For Primary Data Providers*

- Has sample been drawn from the most up-to-date population lists available?
- Have weights been properly applied?

#### *For Secondary Data Managers*

- Are data reported with weights properly applied?

## Standard Three: Editing—

Data are clean.

DEFINITION	Examples Meeting the Standard and Failing to Meet the Standard
<ul style="list-style-type: none"> <li>• Data are correct, internally consistent, and without mistakes.</li> <li>• Response rates are adequate across items and across responding units.</li> </ul>	<p><b>Meeting:</b></p> <ul style="list-style-type: none"> <li>• Data elements are accurately entered from original sources.</li> <li>• Data are internally consistent (e.g., totals, ratios, and products match).</li> </ul> <p><b>Failing:</b></p> <ul style="list-style-type: none"> <li>• Entries are out of range.</li> <li>• The total of the percentages does not sum to 100 percent.</li> </ul>

<b>Editing Data Quality (DQ) Checklist</b> <i>(Most important checks are in italics)</i>
<p><b>For Primary Data Providers</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Have the data, after they were entered, been systematically reviewed by a different person who is familiar with the data?</li> <li><input type="checkbox"/> <i>Has follow-up been done on non-response?</i></li> <li><input type="checkbox"/> Has an electronic edit checking program been used to clean the data?</li> <li><input type="checkbox"/> Have data errors been traced back to their original source and mistakes corrected?</li> </ul> <p><b>For Secondary Data Managers</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Have you “eyeballed” the data to see if they are reasonable given what you know about earlier years and other respondents?</li> <li><input type="checkbox"/> <i>Have you discussed large changes or unusual findings with the primary data providers to see if they might be due to editing errors?</i></li> </ul>

# Standard Four: Calculation—

The math is right.

DEFINITION	Examples Meeting the Standard and Failing to Meet the Standard
<p><b>For All Data</b></p> <ul style="list-style-type: none"> <li>Measured amounts (e.g., numbers, percentages, and ratios) are accurately computed using the right numbers and formulas.</li> </ul>	<p><b>Specifically for Universe Data</b></p> <p><i>Meeting:</i></p> <ul style="list-style-type: none"> <li>Rounding is done correctly.</li> <li>Percentages sum to 100.</li> <li>The denominator used accurately reflects the population of interest.</li> <li>Missing data are distinguished from true zeros.</li> </ul> <p><i>Failing:</i></p> <ul style="list-style-type: none"> <li>The increase from 3 percent to 13 percent is reported as a “10 percent increase” rather than an increase of 10 percentage points.</li> <li>Nonresponse and scores of “0” are aggregated.</li> </ul> <p><b>Specifically for Sample Data</b></p> <p><i>Meeting:</i></p> <ul style="list-style-type: none"> <li>Tests of significance are properly chosen and calculated.</li> <li>Correlations and other statistical formulae are properly applied.</li> </ul> <p><i>Failing:</i></p> <ul style="list-style-type: none"> <li>The wrong significance test is used.</li> </ul>

<b>Calculation Data Quality (DQ) Checklist</b> <i>(Most important checks are in italics)</i>
<p><b>For Primary Data Providers</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Has the application of the mathematical formula been double-checked by a colleague?</li> <li><input type="checkbox"/> Have procedures for dealing with missing data been correctly applied?</li> </ul> <p><b>For Secondary Data Managers</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> For sample data, has the data analysis plan been reviewed by outside experts to ensure that appropriate formulae and procedures are applied?</li> <li><input type="checkbox"/> <i>Have you “eyeballed” the data to see if they seem reasonable given what you know about earlier years and other respondents.</i></li> <li><input type="checkbox"/> Have the “+ or –” confidence intervals been reported for sample data?</li> </ul>

## Standard Five: Timeliness— Data are recent.

DEFINITION	Examples Meeting the Standard and Failing to Meet the Standard
<ul style="list-style-type: none"> <li>Automated electronic processes are used to speed data collection, analysis, and reporting.</li> <li>Data are reported in time to inform policy action.</li> </ul>	<p><b>Meeting:</b></p> <ul style="list-style-type: none"> <li>Data are collected and reports are forwarded as soon as possible after the close of the data collection period.</li> </ul> <p><b>Failing:</b></p> <ul style="list-style-type: none"> <li>Paper records are copied out by hand.</li> <li>Data from 1997 are used to determine whether or not programs have met their 1999 targets.</li> </ul>

<b>Timeliness Data Quality (DQ) Checklist</b> <i>(Most important checks are in italics)</i>
<p><b>For Primary Data Providers</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> <i>Are data from a time before the policy period of interest?</i></li> <li><input type="checkbox"/> Are data reported as soon as possible after collection?</li> </ul> <p>Are the data entered and processed in electronic machine-readable form?</p> <p><b>For Secondary Data Managers</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Are resources being invested in creating a modern automated electronic data system?</li> <li><input type="checkbox"/> <i>Is a regularized schedule of data collections in place to meet policy information needs?</i></li> <li><input type="checkbox"/> Are review processes designed to ensure that findings are made public in a timely fashion?</li> <li><input type="checkbox"/> Are respondents involved in setting time schedules for providing data?</li> <li><input type="checkbox"/> Are time schedules for providing data enforced with clear and frequent reminders?</li> <li><input type="checkbox"/> Is the year of the data collection clearly identified in the report?</li> </ul>



## Standard Six: Reporting— Full disclosure is made.

DEFINITION	Examples Meeting the Standard and Failing to Meet the Standard
<ul style="list-style-type: none"> <li>• What was done is clearly explained.</li> <li>• Limitations are clearly stated.</li> <li>• Findings are fairly and impartially summarized.</li> <li>• Graphics/report exhibits are properly documented, including complete legends, scales, sources, and time frames.</li> </ul>	<p><b>Meeting:</b></p> <ul style="list-style-type: none"> <li>• Data collection processes are documented.</li> <li>• Limitations of the data are clearly described in the report.</li> </ul> <p><b>Failing:</b></p> <ul style="list-style-type: none"> <li>• Sources of data and dates of collection are omitted from charts.</li> <li>• “+” or “-” confidence intervals are omitted when describing sample data.</li> <li>• A number is based on a sample when the actual number is known.</li> </ul>

<b>Reporting Data Quality (DQ) Checklist</b> <i>(Most important checks are in italics)</i>
<p><b>For Primary Data Providers:</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Are data collection, cleaning, and analysis procedures documented in writing?</li> <li><input type="checkbox"/> Is each step in the data collection process required to report deviations and problems in DQ?</li> <li><input type="checkbox"/> Are good graphics techniques used (e.g., axes begin at 0)?</li> <li><input type="checkbox"/> <i>Are DQ problems at each level reported to the next level?</i></li> </ul> <p><b>For Secondary Data Managers:</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> <i>Are DQ problems clearly described in the final reports?</i></li> <li><input type="checkbox"/> Are the data collection method and sample size mentioned at least briefly when findings are presented?</li> <li><input type="checkbox"/> Is the year that the data were collected clearly stated in the report?</li> <li><input type="checkbox"/> Have the types of exclusions and amount of nonresponse been clearly described?</li> <li><input type="checkbox"/> Are DQ problems reported together with the findings?</li> <li><input type="checkbox"/> Has the report of findings been edited by someone with expertise in DQ issues?</li> <li><input type="checkbox"/> Are reports widely announced and effectively disseminated to intended users?</li> <li><input type="checkbox"/> If there have been significant changes in program definitions that might break trend lines, have they been noted?</li> </ul>