

January 2003

A Framework for Assessing the Quality  
of  
Education Statistics

WORLD BANK  
Development Data Group  
and  
UNESCO Institute for Statistics

**Points of contact for this framework are:**

**Sulekha Patel Masako Hiraga, and Lianqin Wang**

Development Data Group  
The World Bank  
1818 H Street., NW  
Washington, D.C. 20433  
Fax No. (202) 522 3669  
Emails : [Spatell@worldbank.org](mailto:Spatell@worldbank.org), [Mhiraga@worldbank.org](mailto:Mhiraga@worldbank.org) and [Lwangw@worldbank.org](mailto:Lwangw@worldbank.org)

And

**Doug Drew and Doug Lynd**

UNESCO Institute for Statistics  
C. P. 6128  
Succursale Centre-Ville  
Montreal, QC, Canada H3C 3J7  
Fax No. (514) 343 6872  
E-mails: [D.Drew@unesco.org](mailto:D.Drew@unesco.org), [D.Lynd@unesco.org](mailto:D.Lynd@unesco.org)

# CONTENTS

	Page
Introduction.....	ii
A. Purpose of the Framework .....	ii
B. Organization of the Framework .....	ii
C. Structure of the Framework.....	iii
0. Pre-requisites of quality .....	1
0.1 Legal and institutional environment .....	1
0.2 Resources .....	3
0.3 Quality awareness .....	4
1. Integrity.....	6
1.1 Professionalism.....	6
1.2 Transparency.....	7
1.3 Ethical standards .....	8
2. Methodological soundness.....	10
2.1 Concepts and definitions.....	10
2.2 Scope.....	10
2.3 Classification/sectorization .....	11
2.4 Basis for recording.....	12
3. Accuracy and reliability.....	13
3.1 Source data.....	13
3.2 Statistical techniques.....	18
3.3 Assessment and validation of source data .....	20
3.4 Assessment and validation of intermediate data and statistical outputs .....	20
3.5 Revision studies .....	21
4. Serviceability .....	22
4.1 Relevance.....	22
4.3 Consistency .....	23
4.4 Revision policy and practice.....	24
5. Accessibility.....	25
5.1 Data accessibility .....	25
5.2 Metadata accessibility .....	26
5.3 Assistance to users .....	27

# Data Quality Assessment Framework for Education Statistics

## Introduction

### A. Purpose of the Framework

The main purpose of the Framework is to provide a flexible structure for the qualitative assessment of education statistics.

A general Data Quality Assessment Framework has been developed by the IMF and applied to statistics in a number of different subject matters – including poverty statistics. The World Bank in collaboration with the UNESCO Institute for Statistics has undertaken the application of the framework to education statistics. The approach followed has been to describe the general framework in a manner that is independent of the subject matter being applied in the main text of the document, and to highlight those elements which are specific to education statistics in boxes embedded within the report. This allows both a view of the general framework, while highlighting how it has been applied to education statistics.

The Framework covers all aspects of the statistical environment or infrastructure in which data are collected, processed, and disseminated, by integrating aspects of the quality of institution and of the quality of products.

The Framework could be used in a variety of contexts, including

- (i) the data module of the Reports on the Observance of Standards and Codes (ROSCs)
- (ii) reviews performed in the context of technical assistance programs and new initiatives related to statistical capacity building and data collection;
- (iii) self-assessments performed by data producers such as Education Management and Information Systems (EMIS) unit in the ministry of education, and national statistical offices; and
- (iv) assessments by other groups of data users.

### B. Organization of the Framework

The Framework is organized in a cascading structure that progresses from the abstract/general to the more concrete/specific.

The first-digit level defines the prerequisites of quality and five dimensions of quality: integrity, methodological soundness, accuracy and reliability, serviceability, and accessibility. The first-digit level is sub-divided by elements (two-digit level) and indicators (three-digit level).

At the next level, focal issues that are specific to the compilation of statistics for the subject area are addressed. Below each focal issue, key points describe quality features that may be considered in assessing the focal issues. The key points are meant to be suggestive, not exhaustive.

**Box A** provides a view of the cascading structure employed in the Framework.

### **C. Structure of the Framework**

The elements and indicators within their respective dimensions are described below.

“Statistical program” refers to the “ education statistical program” and  
“statistical series” refers to the “education data series.”, etc.

**Prerequisites of quality:** Although not itself a dimension of quality, this group of “pointers to quality” includes elements and indicators that have an overarching role as prerequisites, or institutional preconditions, for quality of statistics. Note that the focus is on the **umbrella agency** involved in the statistical work. These prerequisites cover the following elements:

- (i) legal and institutional environment, including coordination power within MOE and across different ministries and departments
- (ii) resources available for statistical work, and
- (iii) quality awareness informing statistical work.

**Integrity:** This dimension captures the notion that statistical systems should be based on adherence to the principle of objectivity in the collection, compilation, and dissemination of statistics. The dimension encompasses institutional arrangements that ensure professionalism in statistical policies and practices, transparency, and ethical standards. The three elements for this dimension of quality are

- (i) professionalism,
- (ii) transparency, and
- (iii) ethical standards.

**Methodological soundness:** This dimension covers the idea that the methodological basis for the production of statistics should be sound and that this can be attained by following internationally accepted standards, guidelines, or good practices. This dimension is necessarily dataset-specific, reflecting different methodologies for different datasets. This dimension has four elements, namely

- (i) concepts and definitions,
- (ii) scope,
- (iii) classification/sectorization, and
- (iv) basis for recording.

**Accuracy and reliability:** This dimension covers the idea that statistical outputs sufficiently portray the reality of the economy. It relates to the notion that source data provide an adequate basis to compile statistics that statistical techniques are sound, and that source data, intermediate data, and statistical outputs are regularly assessed and validated, inclusive of revision studies. The five elements of this dimension cover

- (i) source data,
- (ii) statistical techniques,
- (iii) assessment and validation of source data,
- (iv) assessment and validation of intermediate data and statistical outputs, and
- (v) revision studies.

**Serviceability:** This dimension relates to the need that statistics cover relevant information on the subject field, that they are disseminated in a timely fashion, with an appropriate periodicity, are consistent internally and with other major datasets, and follow a regular revision policy. The four elements for this dimension are

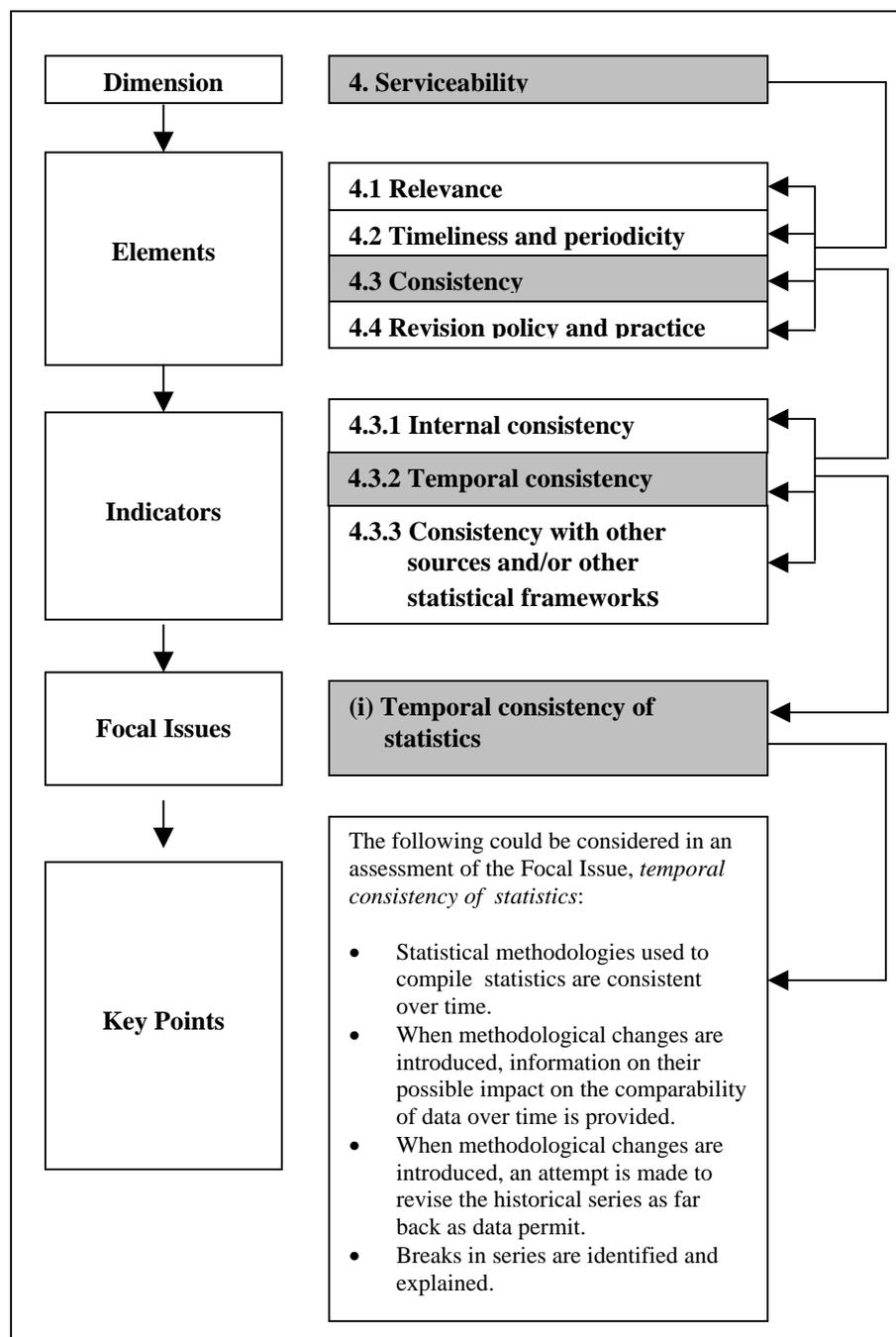
- (i) relevance,
- (ii) timeliness and periodicity,
- (iii) consistency, and
- (iv) revision policy and practice.

**Accessibility:** This dimension relates to the need to ensure that data and metadata are presented in a clear and understandable manner on an easily available and impartial basis, that metadata are up-to-date and pertinent, and that a prompt and knowledgeable support service is available. This dimension has three elements, namely

- (i) data accessibility,
- (ii) metadata accessibility, and
- (iii) assistance to users.

**Box A: An Example of the Cascading Structure of the Data Quality Assessment Framework of the Education Statistics:**

Using serviceability as the example of a dimension of quality, the box below shows how the framework identifies four elements that point toward quality. Within consistency, one of those elements, the framework next identifies three indicators. Specifically, for one of these, temporal consistency, quality is assessed by considering specific key points.



## **0. Pre-requisites of quality**

### **0.1 Legal and institutional environment**

*- The environment is supportive of statistics.*

#### **0.1.1 The responsibility for collecting, processing, and disseminating statistics is clearly specified.**

*i. The primary responsibility for compiling and disseminating statistics is clearly established.*

- A law, such as a statistical law, or other formal provision (e.g., inter-agency protocol or executive decree) assigns primary responsibility to an agency (agencies) and provides the authority to the agency (agencies) for the collection, processing, and dissemination of the statistical series.
- Working arrangements are consistent with this assignment of responsibility.

#### **0.1.2 Data sharing and coordination among data producing agencies are adequate.**

*i. Arrangements or procedures exist to facilitate data sharing and cooperation between the agency (agencies) with the primary responsibility for compiling the statistical series and other data producing agencies.*

- Arrangements or procedures are in place to ensure the efficient and timely flow of source data between the agency (agencies) with primary responsibility for compiling the statistical series and other data producing agencies.
- Arrangements are in place to ensure the consistency of methods and results.
- Contacts (e.g., regular meetings and/or workshops) are maintained with other data producing agencies to ensure proper understanding of data requirements, to avoid duplication of effort, and to take into account respondent burden.

#### **0.1.3 Respondents' data are to be kept confidential and used for statistical purposes only.**

*i. The **confidentiality** of individual respondent's data is guaranteed and that guarantee is widely known.*

- In collecting data, whether using administrative data or surveys, a law or other formal provision clearly states that individual responses are to be treated as confidential, and shall not be disclosed or used for other than statistical purposes unless disclosure is agreed to in writing by the respondent.

- In collecting survey data, respondents are informed of their rights and obligations with regard to the provision of information, and they are informed that the information they provide will be used for statistical purposes only.

***ii. Procedures are in place to prevent disclosure of individual data.***

- Rules and regulations to prevent disclosure include penalties against staff who disclose confidential data.
- Special aggregation rules have been developed to ensure that residual disclosure does not occur when aggregations of survey or other confidential information are disseminated.
- Staff reviews all data prepared for dissemination for possible indirect disclosure of confidential data and devise tables and outputs in a way that prevents disclosure.
- Access to individual data is restricted to staff who require the information in the performance of their duties.
- Steps are taken to secure the premises of the data producing agency and its computer systems to prevent unauthorized access to individual data. Confidentiality of data is appropriately guarded during storage and during the process of the destruction of records.

**0.1.4 Statistical reporting is ensured through legal mandate and/or measures to encourage response.**

***i. A law or other formal provisions are adequate to mandate reporting of information to compile statistics.***

- The data-producing agency has the legal authority to collect data required to compile the statistical series.
- Conflicts or potential conflicts between the legal authority to collect data that are required to compile the statistical series and other laws or provisions (e.g., access to information law) have been successfully resolved or reconciled with no major impairment to data production.
- There are penalties for noncompliance with reporting requirements (including misreporting), even if such provisions rarely need to be employed.

***ii. Other mechanisms are in place to ensure adequate reporting of data for compiling statistics.***

- The data producing agency considers carefully respondent burden and provides assistance to respondents in completing and submitting forms (e.g., by providing a point of contact)
- The data-producing agency seeks to secure the cooperation of respondents by creating goodwill (e.g., by informing respondents of measures to limit respondent burden, raising their awareness of the importance of good quality statistics, and providing them with data upon request).

## **0.2 Resources**

- *Resources are commensurate with needs of statistical programs.*

### **0.2.1 Staff, financial, and computing resources are commensurate with statistical programs of the agency.**

#### *i. Staff resources for compiling statistics are adequate to perform required tasks.*

- Overall, the number of the staff is adequate to perform the required tasks.
- The qualifications of the staff are adequate. They are provided formal (using internal and external experts) and on-the-job training in methodology and compilation methods relevant to the statistical series, including internationally accepted statistical standards (see 2.1.1), guidelines and procedures to compile data.
- Efforts are made to ensure the retention at any point of time of a core contingent of trained staff.

#### *ii. Computing resources for compiling statistics are adequate to perform required tasks.*

- Overall, sufficient resources are allocated and best efforts are made to exploit the full potential of modern computing technology for compiling and disseminating the statistical series.
- Software utilized for compiling and analyzing the statistical series is modern, continually updated, and well adapted to perform existing and emerging tasks.
- Hardware installation is distributed adequately to ensure efficient processing of data and management of the databases.

#### *iii. Financial resources for compiling statistics are adequate to perform required tasks.*

- Overall, financial resources for compiling the statistical series are adequate to perform required tasks and commensurate with the overall resource availability for the agency.
- There are forward plans that allocate budgetary resources to future statistical development based upon identified statistical needs for compiling the statistical series.

### **0.2.2 Measures to ensure efficient use of resources are implemented.**

#### *i. Processes and procedures are in place to ensure that resources are used efficiently.*

- Managers in the data-producing agency promote a policy vision and a direction that is shared with the staff.
- Efficiencies are sought, e.g., by encouraging consistency in concepts, definitions and methodologies across the different units within the data-producing agency.

- Data collection instruments are carefully designed to avoid duplication of information, and lengthy process in compiling data.
- Data compilation procedures are managed to minimize processing errors such as coding, editing, and tabulation errors.
- Internal processes exist within the data-producing agency to measure resources used to compile the statistical series and to compare the resource usage of the statistical program vis-à-vis other statistical programs.
- Periodic reviews of working processes are undertaken to ensure that they are improved upon.
- Periodic reviews of budgeting procedures are undertaken to ensure that scarce resources are best employed in addressing major data problems or meeting new data priorities.
- The data producing agency strives to make the best use of newly emerging opportunities, such as computing technology for data processing/dissemination, to effect resource savings. When necessary, the data-producing agency seeks outside expert assistance to evaluate statistical methodologies and compilation systems.

### **0.3 Quality awareness**

– *Quality is a cornerstone of statistical work.*

#### **0.3.1 Processes are in place to focus on quality.**

*i. There is recognition throughout the umbrella organization that quality builds trust and thus is a cornerstone of statistical work.*

- Managers are sensitive to all dimensions of data quality: integrity, methodological soundness, accuracy and reliability, serviceability, and accessibility.
- The umbrella agency has implemented processes or activities that focus on quality (e.g., Total Quality Management, ISO 9000, and external audits).
- The umbrella organization provides an organizational infrastructure for data quality (e.g., mission statements emphasizing quality and data banks that permit cross-checking) in awareness of the economies of scale and interrelationships between datasets.

#### **0.3.2 Processes are in place to monitor the quality of the collection, processing, and dissemination of statistics.**

*i. Measures are in place to conduct quality review at the various statistical stages.*

- Reviews are undertaken to identify problems at the various stages of collecting, processing, and disseminating data.
- There is a body distinct from the data-producing agency that provides guidance on the quality of the statistical series and on strategies for improving data production.
- Periodic users' surveys or other systematic processes exist to obtain feedback from users on data quality issues.

**0.3.3 Processes are in place to deal with quality considerations, including tradeoffs within quality, and to guide planning for existing and emerging needs.**

*i. Processes are in place at the level of the umbrella organization to deal with quality considerations, including implicit and explicit tradeoffs among the dimensions of quality and the reviews are used to inform planning.*

- There is recognition by the management of the umbrella organization of the tradeoffs among the dimensions of quality (for example between timeliness, completeness and accuracy/reliability).
- The significance of tradeoffs among the dimensions of quality is communicated to users and their views are taken into consideration.
- Improvement of data quality is taken into account in planning the forward work program.

*ii. Mechanisms exist for addressing new and emerging data requirements.*

- Meetings are periodically convened with policy makers and other data users to review the existing statistical series and statistical reports, and to identify any emerging data requirements.
- Users' feedback on the statistical series and statistical reports are encouraged.

## 1. Integrity

*The principle of objectivity in the collection, processing, and dissemination of statistics is firmly adhered to.*

### 1.1 Professionalism

*– Statistical policies and practices are guided by professional principles.*

#### 1.1.1 Statistics are compiled on an impartial basis.

*i. The terms and conditions under which statistics are produced guarantee the professional independence of the data producing agency.*

- A law or other formal provision
  - addresses the general need for the professional independence of the data producing agency,
  - prohibits interference from others, including other government agencies, in the compilation and/or dissemination of statistical information, and
  - ensures that the choice, tenure, and reporting arrangements of the umbrella agency's head are supportive of the professional independence of the statistical agency.
- If there is no law or formal provision to ensure professionalism, traditions or cultures of professionalism are encouraged and made known.

*ii. Professionalism is actively promoted and supported within the data producing agency.*

- Professional competency plays a key role in the recruitment and promotion practices.
- Professionalism is promoted by the publication of methodological papers and by encouraging participation in or organizing lectures, conferences, and meetings with other professional groups, etc.
- Research and analysis undertaken by the data-producing agency for publication are subject to internal review and other processes to maintain the agency's reputation for professionalism.

#### 1.1.2 Choices of sources and statistical techniques are informed solely by statistical considerations.

*i. The choices of data sources and statistical techniques are informed solely by statistical considerations.*

- The choice of source data (e.g. between **administrative records and surveys and among surveys**) and statistical techniques (e.g., processing and validation techniques) is based solely on statistical considerations.

- Staffs are encouraged to present their reasoning for the choice of methodologies in documents that are made public.

Sources include administrative data arising from school censuses, and data from household survey and population censuses (see 2.3.1 and 3.1).

**1.1.3 The appropriate statistical entity is entitled to comment on erroneous interpretation and misuse of statistics.**

*i. The data producing agency comments when its statistics are misinterpreted or misused.*

- The data-producing agency seeks to build trust in its work by commenting publicly on erroneous interpretations or misuse of the statistical series in the media and in other fora.
- The data-producing agency seeks to prevent misinterpretation or misuse of statistics by providing explanatory materials and briefings (e.g. to the press), and by following closely the press and other media (by means of a clipping service).

**1.2 Transparency**

– *Statistical policies and practices are transparent.*

**1.2.1 The terms and conditions under which statistics are collected, processed, and disseminated are available to the public.**

*i. Information is available to the public about the terms and conditions under which the statistical series are collected, compiled, and disseminated, including the obligation to compile and disseminate the statistics, the confidentiality of individual responses, and other key features.*

- Agency publications and/or website reproduce material about the terms and conditions under which official statistics are compiled and disseminated (e.g., the statistical law, the Fundamental Principles of Official Statistics, mission statements, and codes of conduct under which official statistics are compiled and disseminated).
- Statistical publications identify where more information about the data producing agency and its products can be found.

**1.2.2 Internal governmental access to statistics prior to their release is publicly identified.**

*i. The public is made aware of internal government access to statistics prior to their release to the public.*

- Access to statistics prior to release is made public in terms of who has access, and at what point of the compilation process access is given.
- The public is made aware that the approval processes for the publication of the statistical series rests entirely with the data-producing agency.

**1.2.3 Products of statistical agencies/units are clearly identified as such.**

*i. Statistical products are clearly identified so that the public is aware of what the data-producing agency takes responsibility for.*

- Data released to the public are clearly identified as the data producing agency's product (e.g., by name, logo, and insignia).
- In the case of joint publications, the part attributable to the data-producing agency is identified.
- The data-producing agency requests attribution when its statistics are used or reproduced.

**1.2.4 Advance notice is given of major changes in methodology, source data, and statistical techniques.**

*i. Users of statistics are made aware in advance of major changes in methodology, source data, and statistical techniques.*

- Advance notice is given when major changes in methodology, sources, and statistical techniques are introduced.

**1.3 Ethical standards**

– *Policies and practices are guided by ethical standards.*

**1.3.1 Guidelines for staff behavior are in place and are well known to the staff.**

*i. A clear set of ethical standards has been prepared.*

- There are clear guidelines outlining correct behavior when the agency or its staff is confronted with potential conflict of interest situations (e.g., with respect to avoiding holding up the data for collecting inappropriate fees).
- There are clear guidelines that make the connection between ethics and staff work (e.g., with respect to guarding against misuse and misrepresentation of statistics (see also 1.1.3)
- The reputation of the head of the umbrella agency and its management for the maintenance of ethical standards assure the autonomy from political interference.

*ii. Staff are made aware of the ethical standards.*

- Management acknowledges its status as role model and is vigilant in following the guidelines.
- New staff members are made aware of the guidelines when they join the organization.
- Staff members are reminded periodically of the guidelines.

## 2. Methodological soundness

*The methodological basis for the statistics follows internationally accepted standards, guidelines, or good practices.*

The methodological soundness is assessed on the basis of a hybrid of internationally accepted standards, guidelines and good practices, including, but not limited to:

UNESCO's International Standard Classification of Education 97  
UNESCO Institute for Statistics (UIS) technical guidelines and manuals  
Methodology used in the 1994 estimation and projection of adult illiteracy;  
Indicators developed by UIS and for OECD for countries participating in the  
World Education Indicators (WEI) programme.  
World Bank PRSP Source Book; and  
Other regional programs

For all UNESCO and UIS documentation above see [www.uis.unesco.org](http://www.uis.unesco.org).

### 2.1 Concepts and definitions

*– Concepts and definitions used are in accord with standard statistical frameworks.*

**2.1.1 The overall structure in terms of concepts and definitions follows internationally accepted standards, guidelines, or good practices.**

*i. The concepts and definitions follow internationally accepted standards, guidelines, or good practices.*

Concepts and definitions follow methodologies used by UNESCO Institute for Statistics, and those developed by the UIS and OECD for countries participating in the World Education Indicators project. Internationally accepted standards include

- Deviations from the above concepts and definitions are kept under review (see also 5.2.1).

### 2.2 Scope

*– The scope is in accord with internationally accepted standards, guidelines, or good practices.*

**2.2.1 The scope is broadly consistent with internationally accepted standards, guidelines, or good practices.**

*i. The scope of statistics is consistent with internationally accepted standards, guidelines, or good practices.*

- Statistics are sufficiently comprehensive in scope and in terms of conceptual development of concepts to adequately describe the subject area in question

Sufficient data on measuring education system that cover all levels and all types of education are available. The data should cover all major areas of education, including the following:

- Structure of education system,
- Human and financial resources invested in education

Performance of the education system (including access, enrollment, progression, completion, and student learning achievements)

- Scope of statistics is adequate in terms of other relevant variables for analytical purposes

- geographical boundaries (local, regional, central),
- socioeconomic groups (e.g. male and female, public and private, trained and untrained, full-time and part-time),
- the time period for which estimates are required (temporal coverage).

### 2.3 Classification/sectorization

– *Classification and sectorization systems are in accord with internationally accepted standards, guidelines, or good practices*

#### 2.3.1 Classification/sectorization systems used are broadly consistent with internationally accepted standards, guidelines, or good practices.

ii. *The classification of statistics complies with internationally accepted standards, guidelines, or good practices.*

Classification of education is based on UNESCO Institute for Statistics' International Standard Classification of Education 97(ISCED97) and technical guidelines and manuals (e.g., level of education, public and private, trained and untrained, full-time and part-time, trained and untrained.)

Classification of educational expenditure is based on [UNESCO Institute for Statistics' technical guidelines and manuals](#) as well as System of National Accounts (SNA).

For All UNESCO and UIS documentation see [www.uis.unesco.org](http://www.uis.unesco.org).

- iii. The *delimitation* of units of collection and analysis is consistent with internationally accepted standards, guidelines, or good practices

Delimitation of students, teachers and educational institutions is based on UNESCO Institute for Statistics' technical guidelines and manuals. See [www.uis.unesco.org](http://www.uis.unesco.org).

- iv. Classification systems are applied consistently across different units of collection and analysis

*ISCED and others UIS standards and guideline are applied consistently to statistics on the educational system, students, teachers and educational institutions, and educational expenditure.*

- v. Deviations from the above classification/sectorization are kept under review (see also 5.2.1).
- Trainings are available for respondents who need assistance.

## 2.4 Basis for recording

*– Data are recorded according to internationally accepted standards, guidelines, or good practices.*

### 2.4.1 Recording system follows internationally accepted standards, guidelines, or good practices.

- i. *Questionnaires are in accordance with internationally accepted standards, guidelines, or good practices.*

Questionnaires are in accordance with UNESCO Institute for [Statistics guidelines based on ISCED 97](#).  
Data are recorded according to classification/sectorization systems mentioned above.

### 3. Accuracy and reliability

*Source data and statistical techniques are sound and statistical outputs sufficiently portray reality.*

#### 3.1 Source data

*– Source data available provide an adequate basis to compile statistics.*

##### 3.1.1 Source data are collected from comprehensive data collection programs that take into account country-specific conditions.

###### *i. Statistics describing structure and normative characteristics of the education system*

- *Data collected include data on:*

Compulsory education:

- beginning and end age

School year:

- starting month
- ending month
- hours of instruction per year

Policy on promotion and repetition

- is there automatic promotion ?

Structure of education system

- name of programme (or certificate awarded for tertiary), by ISCED level and orientation
- existence of private education institutions, by level of education (yes or no)
- part-time pupils/students by level? (yes/no). If yes...What is the conversion factor for calculation of FTE by level?

Teachers

- part-time teachers by level (yes/no). If yes, What is the conversion factor for calculation of FTE by level?
- number of hours spent teaching per year, per full-time teacher (as distinct from other teacher functions such as preparation of material, marking, etc)

**ii. *Statistics on supply of education collected through a regular administrative school census program.***

- Administrative school census should collect information on structure of educational system, students, teachers, and educational expenditure.
- Coverage is comprehensive in terms of
  - geographic areas (local, regional, central),
  - relevant sub-groups of units of collection (e.g. male and female students and teachers, public and private schools, trained and untrained teachers, full-time and part-time students and teachers).
- Data collected are detailed and include, in particular, data on:

Students at pre-primary, primary, secondary, and post-secondary non-tertiary level

- Number of students (by gender, by level and orientation (general or trade/vocational) of education, by private and public, by full-time, part-time and full-time equivalents)
- Number of repeaters (by gender, by level of education, by age, private and public, by full-time and part-time)
- Number of graduates (by gender, by level of education, by age, private and public))
- Number of students in early childhood education (by age, gender, and by pre-primary vs. other type of programme)
- Number of students by grade (by age, and by level and orientation of education)
- Number of students in trade vocational education at secondary and post-secondary non-tertiary levels (by gender and field of study)

Students at tertiary level of education

- Number of students enrolled (by level, full-time part-time and full-time equivalents, by public, government dependent private, and independent private institutions)
- Number of students enrolled by age (by gender and by level of education)
- Number of students enrolled by field of education (by level of education and program duration, and by gender)
- Number of graduates (by field of education, by level of education and program duration, and by gender)
- Number of foreign students by country of origin

Teachers

- Number of teachers (by gender, by level and orientation of education, by private and public, by full-time, part-time, and full-time equivalents, by trained and untrained)
- Number of non-teaching staff (by gender, private and public, full-time and part-time)

#### Educational Institutions

- Number of institutions by level and orientation of education, by private and public.

#### Education Finances

- Public expenditure (by level of education, by level of government, by current and capital, and by public and private institution)
- Private expenditure (by level of education, by households and other private entities, by public and private institution)
- Expenditure from international sources (by level of education, by public and private institution)
- Direct institutional expenditure (by level of education, by public, government dependent private and independent private institutions, by capital, current compensation of personnel and other current)
- Educational revenues (by level of education, by source, by public, government dependent private and independent private institution)

### ***iii. Statistics on demand for education collected through household surveys and population censuses.***

- Survey modules should include school attendance, attainment, literacy, access, and educational expenditure.
- Data collected are detailed and include, in particular, data on:

#### School enrollment

- Enrollment (by level of education, public and private)
- Attendance (by level of education, public and private)
- 
- Repetition (by level of education, public and private, by cause)
- Drop out (by level of education, public and private, by cause)

#### Educational attainment

- Literacy level
- Highest educational level attained

#### Access to school

- Distance to school
- Transportation

#### Educational expenditure

- Educational expenditure by households (by level of education, by source, by purpose)

Adult education and training

- By purpose (job-related vs. personal)
- By sponsorship (employer sponsored or not)
- By hours of participation

***iv. Statistics on the quality of learning outcomes collected through assessments of student achievement***

- There is a regular programme assessment of student achievement, at one or more ages/or levels of education.
- Assessments vary in scope, and may be sub-national, national or international.
- Sub-national and national assessments are guided by standards of acceptable practice as developed and applied to international student assessments such as PISA, TIMSS, PIRLS, SAQMEQ, and others.
- Assessments cover key competencies in areas of learning including reading and writing, mathematics, and science.
- Assessment results, at a minimum, indicate the percentage of students in the age/grade level and topic area assessed meeting and not meeting expected levels of achievement.
- Assessments include background questionnaires of students and school administrators (principals), for the purpose of being able to study the relationships between family, socio-economic, and school factors contributing to learning outcomes.

***v. Statistics on the environment within schools that impact on quality of education collected via school surveys.***

School Characteristics:

- Governance structure
- Community in which the school is located
- Funding sources
- Estimated family income relative to the local region
- Drop-out, repetition, and “program completion without delay” rates
- Multi-grade teaching
- Single or multiple shifts

Teacher Characteristics:

- Gender
- Age
- Experience as a teacher
- Education
- Pre-service teacher training
- In-service training

**3.1.2 Source data reasonably approximate the definitions, scope, classifications, valuation, and time of recording required.**

*i. Source data approximate the definitions, scope, classifications, valuation, and the time of recording required in statistics in the subject area.*

- Source data are consistent with the definitions, scope, and classifications of statistics in the subject area.
- Source data are consistent with the time of recording, reference periods, and valuation of statistics.
- Source data are consistent with censuses and other data sources.
- Compilers are aware of differences in concepts and definitions used in the source data from those required of statistics in the subject area.

*ii. Specific procedures are used to improve the coverage, classification, valuation, and timing of information received by the data producing agency from various data sources.*

- Specific procedures have been developed to adjust data from various data sources to improve coverage, classification, and valuation and conform to internationally accepted standards, guidelines or good practices.

*iii. Information is available on the extent to which the coverage, classification, valuation, and timing of information contained in secondary data sources differ from international statistical guidelines.*

- Compilers are aware of differences in practices used in compiling source data for statistics.

**3.1.3 Source data are timely.**

*i. Data collection system provides for the timely receipt of source data and detailed data.*

- Respondents are made aware of the deadlines set for reporting.
- The producing agency employs rigorous follow-up procedures to ensure the timely receipt of respondents' data.
- If respondents fail to submit due to lack of knowledge or lack of resource, appropriate adjustments are made.

**3.1.4 Instances where statistical information for the subject area in question make use of data pertaining to other subject areas and produced by other data-producers are noted, and references are given to descriptions of their methodology and quality.**

*Education indicators are important in the area of education statistics. Indicator set are typically defined describing the education system in terms of some underlying model, for example in terms of education inputs, processes, and outcomes, as well as on the context of education are produced.*

- Indicators are measures derived from statistics that are relevant in describing diverse aspects of education systems. Indicators are generally normalized to facilitate comparisons across different units of analysis independent of their size (e.g., comparisons of sub-populations within a country, international comparisons across countries).
- Factors used in normalizing the indicators sometimes rely on non-education statistics such as population statistics, typically produced by National Statistical Agencies and internationally by the United Nations Population Division, and GDP or total government expenditures, typically produced by the SNA area of National Statistical Agencies.
- References are given to descriptions of methodology and quality of these external data, and where appropriate implications of quality limitations in these data in the interpretation of the education indicators are discussed.

## **3.2 Statistical techniques**

- *Statistical techniques employed conform to sound statistical procedures, and are documented.*

### **3.2.1 When sample surveys are used as a vehicle for collection of data, the sample design is appropriate:**

- Target population is defined.
- Sample frames are available for conducting surveys of statistical units (e.g., individual, household, community); minimize undercoverage and overcoverage; and are updated regularly.
- Scientific random sampling techniques are used.
- Sample size is appropriate.

### **3.2.2 Data compilation employs sound statistical techniques.**

#### *i. Data procedures are sound.*

- Data compilation procedures minimize processing errors such as tabulation errors and report generation.
- The report forms are designed in a way that makes them easy to complete and appropriate for computer processing, and they have also been pilot-tested with a sample of respondents.

### **3.2.3 Other statistical procedures (e.g., data adjustments and transformations, and statistical analysis) employ sound statistical techniques.**

#### *i. Data adjustments and transformations employ sound statistical techniques.*

- Imputation methods, estimation techniques (e.g., sampling weights, calibration weights), employ sound statistical techniques.

#### *ii. Internationally accepted statistical methods are used to handle non-sampling errors.*

- Problems regarding non-responses, recall errors, reporting errors, respondents effects, interviewer effects, and inappropriate instrument design are addressed.

#### *iii. Appropriate adjustments are made for inadequate coverage.*

- Imputation and estimation methods are appropriate:  
Proper imputation methods are used wherever feasible to handle missing, invalid or inconsistent responses. If there is a sizeable part of the population that is not covered by sources used for regular compilation of statistics, under-coverage adjustments are made, or if such adjustments are not feasible in terms of being statistically defensible, the limitation in the coverage of the statistics is described

Imputation methods and methods for dealing with missing data apply to data deriving from both administrative records, household surveys and population censuses, sample surveys of schools, and assessments of student achievement.
---

### 3.3 Assessment and validation of source data

- *Source data are regularly assessed and validated.*

#### 3.3.1 Source data – including censuses, sample surveys and administrative records are routinely assessed, e.g., for coverage, sample error, response error, and non-sampling error; the results of the assessments are monitored and made available to guide planning.

##### *i. Accuracy of information is routinely assessed.*

- Administrative and survey data are audited to check the accuracy of source data (e.g., inspection of field collections, random post-enumeration checks).
- Information is compiled on coverage, sampling errors (where applicable), non-response errors (e.g., non response rates for various socioeconomic groups), and the percentage of missing and/or imputed data by methods of imputation.
- For surveys, sampling standard errors of survey estimates in order to form confidence intervals for population values are provided, especially when the estimates are based on a small sample.

##### *ii. Appropriate measures are taken to validate data sources.*

- Training is provided to improve the accuracy.
- For Surveys, sample selection is adjusted when sampling errors become large.

##### *iii. Accuracy of source data is routinely assessed.*

- Accuracy of data from all sources used to compile statistics is routinely assessed in terms of monitored events, population coverage, and the time frames.

### 3.4 Assessment and validation of intermediate data and statistical outputs

- *Intermediate results and statistical outputs are regularly assessed and validated.*

#### 3.4.1 Main intermediate data are validated against other information where applicable.

##### *i. Aggregates from surveys are compared against independent data sources and statistical frameworks.*

- Data from different sources but measuring the same or closely related phenomena are compared against each other.

Administrative school census data, sample school survey data, and results from household surveys and population censuses are compared against each other (e.g. – comparison of common variables between administrative school censuses and surveys such as enrolments, teachers; and comparison of common variables between household surveys and censuses, e.g. educational attainment, literacy), enrolment by socio-demographic characteristics.

**3.4.2 Statistical discrepancies in intermediate data are assessed and investigated.**

*i. Statistical discrepancies are routinely assessed and investigated.*

- Post-survey data analysis is conducted to monitor statistical discrepancies.
- Provision is made for immediate follow-up to reconcile data inconsistencies.

**3.4.3 Statistical discrepancies and other potential indicators of problems in statistical outputs are investigated.**

*i. Errors and omissions are investigated and made public.*

- Systematic processes are in place to monitor errors and omissions, and address data problems.
- Results are checked against demographic data, and other survey/census results.

**3.5 Revision studies**

*– Revisions, as a gauge of reliability, are tracked and mined for the information they may provide.*

**3.5.1 Studies and analyses of revisions are carried out routinely and used to inform statistical processes (See also 4.4.3).**

*i. Revision studies are undertaken on a regular basis.*

- Revisions to methodology are assessed regularly.
- Analysis of preliminary versus revised data is conducted for major data series to assess the reliability of the preliminary data.
- The findings from these investigations are taken into account when compiling data for subsequent periods.
- The findings are made accessible to the data users and compilers.

## 4. Serviceability

- *Statistics are relevant, timely, consistent, and follow a predictable revisions policy.*

### 4.1 Relevance

- *Statistics cover relevant information on the subject field.*

#### 4.1.1 The relevance and practical utility of existing statistics in meeting user's needs are monitored and measures are taken to as appropriate and feasible to better address user needs.

*i. Specific actions are taken to ensure statistics collected adequately respond to users' needs.*

- An established process of consultation takes place periodically with policy departments/ministries, with a user advisory group that includes representatives from the private sector and academia, and/or through other forms of communication with users of statistics (e.g., users' surveys).
- An established process of review takes place periodically to assess whether the program meets the needs of users of the statistical series.
- The data producing agency regularly participates in international statistical meetings and seminars organized by international and regional organizations.

There is a regular dialogue between the area within the education ministry responsible for statistics (including the administrative school census, and where applicable other elements such as assessments of student achievement) and the policy area of the ministry on the statistical information needed for policy purposes, on the preparation and implementation of plans for improvements as needed, and on the use of statistical information for policy purposes.

### 4.2 Timeliness and periodicity

- *Timeliness and periodicity follow internationally accepted dissemination standards.*

#### 4.2.1 Timeliness follows dissemination standards.

*i. The timeliness of statistics follows internationally accepted good practices.*

Statistics derived from the administrative school census are disseminated within 6-12 months after the beginning of school year.  
The timeliness of other education statistics follow internationally accepted good practices.

#### **4.2.2 Periodicity follows dissemination standards.**

*i. The periodicity of statistics follows internationally accepted good practices.*

Education statistics derived from the administrative school census are disseminated annually.  
The periodicity of other education statistics follows internationally accepted good practices.

#### **4.3 Consistency**

– *Statistics are consistent within a dataset and over time, and with other major data sets.*

##### **4.3.1 Statistics are consistent within the dataset.**

*i. Statistics are internally consistent.*

- Accounting identities between aggregates and their components are observed for all involved data.
- Accounting identities and relationships between different variables are observed.

Relationships between enrollments, repeaters, dropouts, and survivors are consistent, as are relationships between these data and demographic data population of school age.

- Statistics are cross-checked within the survey, across geographic areas and sub-groups of population.

##### **4.3.2 Statistics are consistent or reconcilable over a reasonable period of time.**

*i. The statistical series is consistent over time.*

- Consistent time series data are available for an adequate period of time (at least five years)
- When changes in source data, methodology, and statistical techniques are introduced, historical series are reconstructed as far back as reasonably possible.
- Detailed methodological notes identify and explain the main breaks and discontinuities in time series, their causes, as well as adjustments made to maintain consistency over time.
- Any unusual changes in economic and demographic trends are explained in the analytical text included in the publication and in the database accessible to users.

**4.3.3 Statistics are consistent or reconcilable with those obtained through other data sources and/or statistical frameworks.**

- i. Statistics are consistent or reconcilable with those obtained through other surveys, data sources and/or statistical frameworks.*

Education statistics are reasonably reconciled with administrative data, census data, and socio-demographic data from other sources.
--

**4.4 Revision policy and practice**

– *Data revisions follow a regular and publicized procedure.*

**4.4.1 Revisions follow a regular, well established, and transparent schedule.**

- i. The practice of revisions (of provisional estimates, base years, weight updates and methodology) follows a predictable pattern of which users of statistics are informed.*

- Adequate documentation of revisions is included in the publication of the statistical series and in the database accessible to users.

**4.4.2 Preliminary data are clearly identified.**

- i. Preliminary data or first estimates are clearly identified in statistical releases.*

- Users are alerted that the initially published data are preliminary and subject to revision.
- The revised data are disseminated with the same level of detail as previously published for the data being revised.

**4.4.3 Studies and analyses of revisions are carried out routinely and made public (See also 3.5.1).**

- i. Users are informed of causes of revisions to the statistical series.*

- Revisions to methodology are assessed and explained in the publication of the statistical series and in the database accessible to users.
- Analysis of preliminary versus revised data is published for major data series to allow assessment of the reliability of the preliminary data.

## 5. Accessibility

*Data and metadata are easily available and assistance to users is adequate.*

### 5.1 Data accessibility

*– Statistics are presented in a clear and understandable manner, forms of dissemination are adequate, and statistics are made available on an impartial basis.*

#### 5.1.1 Statistics are presented in a way that facilitates proper interpretation and meaningful comparisons (layout and clarity of text, tables, and charts).

*i. The presentation of the statistical series is commensurate with users' needs.*

- Data are published in a clear manner; charts and tables are disseminated with the data to facilitate the analysis.
- Data offer adequate details and time series.
- Analysis of current period estimates is available.
- Depending on the intended audience and purposes, data of different degree of aggregation, sub-components and additional data are made available.

Examples: data of different degree of aggregation (e.g. school, region), sub-components (e.g., by gender, by level of education, by age, private and public, full-time and part-time) and additional data (e.g., demographic, socioeconomic, geographic information).
---

#### 5.1.2 Dissemination media and formats are adequate.

*i. The dissemination means are commensurate with users' needs.*

- Data are first time series can be accessed through an electronic database maintained by the data producing agency.
- Annual statistical yearbook can be made available and disseminated.

#### 5.1.3 Statistics are released on a pre-announced schedule.

*i. A schedule for data release is announced in advance.*

- The statistical series is released according to a pre-announced schedule.

#### 5.1.4 Statistics are made available to all users at the same time.

*i. The statistics are made available to all users at the same time.*

- The statistical series is released simultaneously to all interested users on the date and/or time specified in the pre-announced schedule.
- If the press is briefed in advance, measures are taken to avoid release to the public in advance of the regular schedule.

### **5.1.5 Non-published (but non-confidential) sub-aggregates are made available upon request. (see also 5.1.1)**

- i.* Non-published and non-confidential detailed data are made available to the users upon request.
  - Non-published (but non-confidential) specialized tabulations (e.g., sub-aggregates of units of analysis) are made available upon request.
  - Non-confidential micro-data files (e.g., with information permitting the identification of individual respondents removed) are available to permit analytical use by researchers and other users.
  - The availability of non-published statistics and data, and the terms and conditions on which they are made available are publicized.

## **5.2 Metadata accessibility**

– *Up-to-date and pertinent metadata are made available.*

### **5.2.1 Documentation on concepts, scope, classifications, basis of recording, data sources, and statistical methodologies and techniques is available, and differences from internationally accepted standards, guideline, or good practices are annotated.**

- i.* *The metadata for the statistical series provides users with an adequate information about what the data mean and about the methodology used to collect and process them.*
  - Metadata, including information on concepts, definitions, classification and other methodology, data sources, and statistical techniques are prepared and disseminated to the public. The metadata also provides information on:
    - biases in the data,
    - information about response rates to the main surveys used,
    - main linkages with other major data systems, highlighting important differences from these data systems,
    - other information users may need to assess the data.
  - Deviations from internationally accepted standards, guidelines, or good practices are well documented in the metadata.
  - The metadata is disseminated in a manner that facilitates its access (e.g., websites, statistical publications) and its availability is well publicized (e.g., in catalogs). The GDDS summary methodologies and other related metadata are regularly reviewed and updated.
  - Instances where statistical information for the subject area in question make use of data pertaining to other subject areas and produced by other data-producers are noted, and references are given to descriptions of their methodology and quality.

Where education indicators rely on non-education statistics such as population statistics, typically produced by National Statistical Agencies and internationally by the United Nations Population Division, and GDP or total government expenditures, typically produced by the SNA area of National Statistical Agencies, references are given to descriptions of methodology and quality of these external data, and where appropriate implications of quality limitations in these data in the interpretation of the education indicators are discussed.

### ***5.2.2 Levels of detail are adapted to the needs of the intended audience.***

- i.* Different levels of detail are made available to meet users' requirements.
  - A brochure has been prepared to inform general users about the statistical series.
  - A comprehensive sources and methods document produced to inform analysts and other users of statistics about how statistics are compiled. This document is updated regularly.

## **5.3 Assistance to users**

*– Prompt and knowledgeable support service is available.*

### **5.3.1 Contact person for each subject field is publicized.**

- i.* There are provisions to provide adequate assistance to users.
  - Prompt and knowledgeable service and support are available to users of statistics. All statistical releases identify specific individuals who may be contacted by mail, telephone, facsimile, or by email.
  - Documentation has been developed (e.g., brochures) to educate users of related datasets.
  - Assistance to users is monitored through periodic surveys of users.

### **5.3.2 Catalogues of publications, documents, and other services, including information on any charges, are widely available.**

- i.* Catalogues of publications and other services are available to users of statistics.
  - A catalogue of publications, documents, and other services to users is available and updated each year.
  - The prices of the statistical products and services are clearly disclosed and assistance is provided in placing orders.